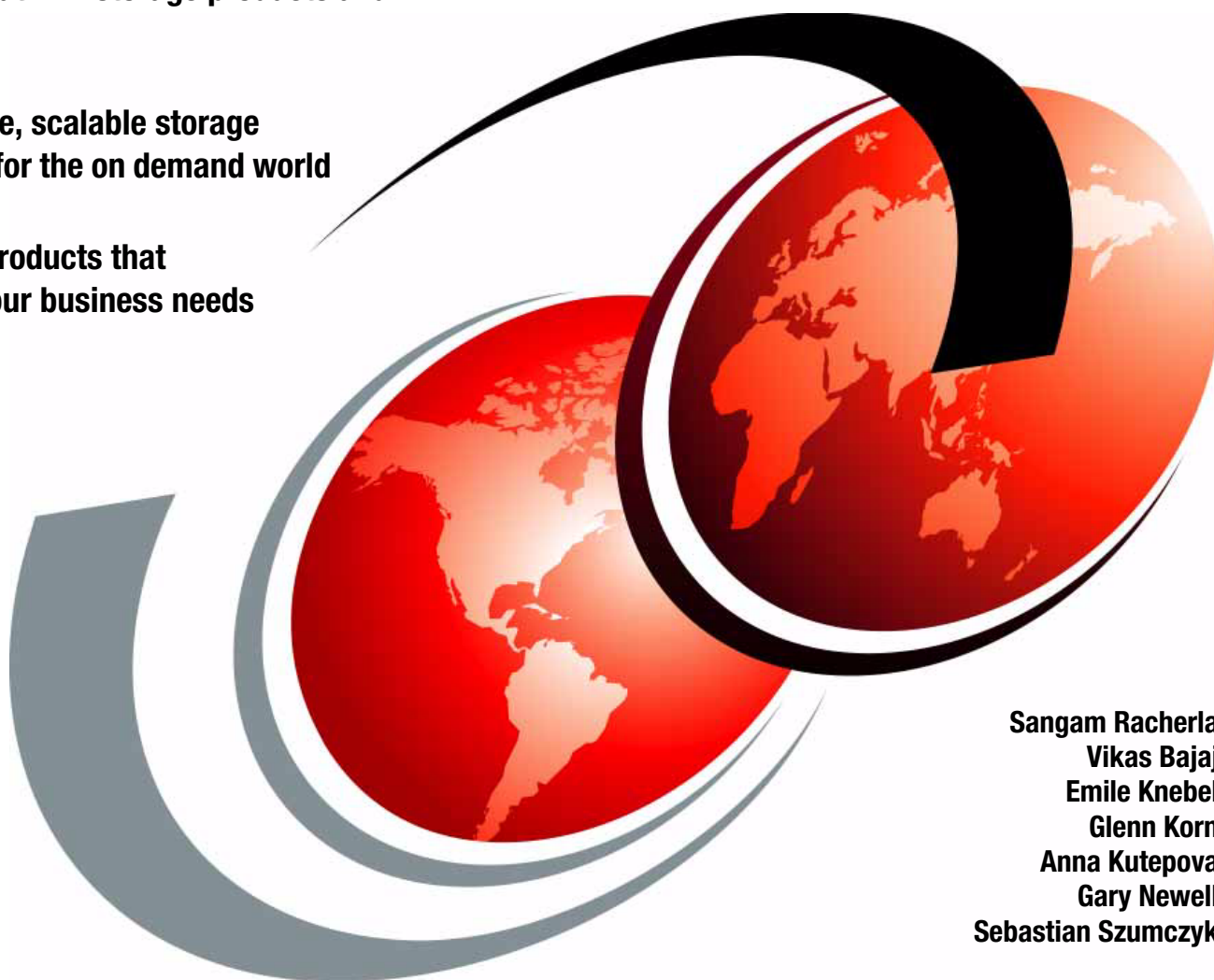


IBM System Storage Solutions Handbook

Learn about IBM storage products and solutions

Get flexible, scalable storage solutions for the on demand world

Find the products that answer your business needs



Sangam Racherla
Vikas Bajaj
Emile Knebel
Glenn Korn
Anna Kutepova
Gary Newell
Sebastian Szumczyk

Redbooks



International Technical Support Organization

IBM System Storage Solutions Handbook

May 2010

Note: Before using this information and the product it supports, read the information in “Notices” on page xvii.

Eighth Edition (May 2010)

This edition applies to IBM System Storage products portfolio as of November 2009.

© Copyright International Business Machines Corporation 1999-2010. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices	xvii
Trademarks	xviii
Preface	xix
The team who wrote this book	xx
Now you can become a published author, too!	xxii
Comments welcome	xxii
Stay connected to IBM Redbooks publications	xxii
Chapter 1. Introduction: Dynamic Infrastructure and IBM Smarter Planet Strategy ..	1
1.1 What a dynamic infrastructure is	3
1.2 Delivering a dynamic infrastructure	3
1.2.1 Information infrastructure	4
1.2.2 Business continuity and resiliency	5
1.2.3 Virtualization	5
1.2.4 Energy efficiency	5
1.2.5 Security	6
1.2.6 Service management	6
1.3 IBM System Storage Disk Systems	7
1.3.1 High-end and enterprise disk systems	7
1.3.2 Midrange disk storage	7
1.3.3 Entry-level disk	8
1.4 IBM Tape systems	9
1.4.1 Enterprise tape systems	9
1.4.2 Midrange tape systems	9
1.4.3 Entry tape systems	10
1.5 Storage Area Networks	11
1.5.1 Enterprise SAN directors	11
1.5.2 Midrange SAN switches	12
1.5.3 Entry SAN switches	13
1.5.4 Multiprotocol routers	13
1.5.5 New Ethernet switches	14
1.6 IBM NAS systems	16
1.6.1 IBM N series hardware	16
1.6.2 IBM N series software	17
1.7 Storage software	19
1.7.1 IBM SAN Volume Controller	19
1.7.2 IBM SVC Entry Edition	20
1.7.3 IBM Tivoli Storage Productivity Center	20
1.7.4 IBM Tivoli Storage Manager	20
1.7.5 IBM General Parallel File System	21
1.7.6 IBM Multilevel Grid Access Manager Software	21
1.7.7 z/OS storage management tools	21
1.7.8 IBM System Storage Productivity Center	22
Part 1. Disk systems	23
Types of storage	23
Enterprise storage	24
Mid-level storage	24

Near-line storage and reference data	25
Entry-level storage	25
Network attached storage	25
IBM System Storage point-in-time copy comparison	25
Meaning of indicators	26
FlashCopy function definitions	26
IBM System Storage disk mirroring comparison	27
Meaning of indicators	27
Disk mirroring function definitions	28
Chapter 2. Entry-level disk storage	29
2.1 Overview of DS3000 series models	30
2.2 General specifications	32
2.2.1 IBM System Storage model DS3200	32
2.2.2 IBM System Storage model DS3300	34
2.2.3 IBM System Storage model DS3400	35
2.3 IBM System Storage EXP3000	37
2.4 IBM DS Storage Manager	38
2.4.1 Host server	38
2.4.2 Management station	38
2.5 Premium features	39
2.5.1 DS3000 Partition Expansion License	39
2.5.2 DS3000 FlashCopy Expansion License	40
2.5.3 DS3000 VolumeCopy License	40
2.6 Advanced features	40
2.7 More information	41
Chapter 3. Midrange disk systems	43
3.1 Highlights: What's new	45
3.2 Overview of IBM System Storage DS Models	45
3.2.1 IBM System Storage DS4000	45
3.2.2 IBM System Storage DS4700 Express	46
3.2.3 IBM System Storage DS5000	46
3.2.4 IBM System Storage Model DCS9900	46
3.3 IBM System Storage Model DS4700 Express	46
3.3.1 IBM DS4700 Express Model features	47
3.3.2 IBM DS4700 Express specifications	48
3.4 IBM System Storage model DS5020	49
3.4.1 DS5020 description	49
3.4.2 DS5020 features	50
3.4.3 DS5020 specifications	51
3.5 IBM System Storage Models 5100 and 5300	52
3.5.1 DS5100/5300 storage subsystems	52
3.5.2 DS5100 and DS5300 features	53
3.6 IBM System Storage Model DCS9900	56
3.6.1 DCS9900 storage system	57
3.6.2 DCS9900 features	58
3.6.3 DCS9900 specifications	60
3.7 IBM System Storage DS4000 EXP810 Expansion Unit	61
3.7.1 EXP810 Expansion Unit features	61
3.7.2 EXP810 Expansion Unit specifications	62
3.8 IBM System Storage model EXP520 and EXP5000	63
3.8.1 EXP520 and EXP5000 Drive features	63

3.8.2 EXP520 and EXP5000 Drive support	63
3.9 IBM System Storage EXP5060	64
3.9.1 EXP5060 overview	64
3.9.2 Key prerequisites	64
3.10 Solid State Disk and Full Disk Encryption	65
3.10.1 Solid State Disk	65
3.10.2 Full Disk Encryption	67
3.11 IBM System Storage DS Storage Manager	70
3.11.1 New features: October 2009	72
3.11.2 Previous announcements: 2004	75
3.11.3 Other features	77
3.12 Introduction to IBM Midrange Storage Copy Services	77
3.12.1 FlashCopy premium feature	78
3.12.2 VolumeCopy	79
3.12.3 Enhanced Remote Mirroring	80
3.13 IBM Remote Support Manager for Storage	82
3.14 More information	83
Chapter 4. Enterprise Disk Systems	85
4.1 Overview	86
4.1.1 Infrastructure simplification	87
4.1.2 Business continuity	88
4.1.3 Performance	89
4.2 DS6000 and DS8000 technology	90
4.2.1 Multi-tiered storage environments	91
4.2.2 Enhancements of the DS6000 and DS8000 series	91
4.3 The DS6000 series	94
4.3.1 The DS6800 controller enclosure	95
4.3.2 The DS6000 expansion enclosure	97
4.3.3 DS6800 Model 1750-511	99
4.3.4 DS6000 Model 1750-EX1	100
4.3.5 More information	100
4.4 The DS8000 series	101
4.4.1 Optimizations to simplify and increase efficiency	101
4.4.2 Introduction to the DS8000 series features and functions	103
4.4.3 The DS8700: Latest member of the DS Family	104
4.4.4 The DS8000 design keys	105
4.4.5 The DS8000 architecture	109
4.4.6 DS8000 models	120
4.4.7 IBM Standby Capacity on Demand offering for the DS8000	125
4.4.8 More information	125
4.5 Storage management	126
4.5.1 DS Management Console	126
4.5.2 IBM TotalStorage Productivity Center	129
4.6 Copy Services	129
4.6.1 IBM System Storage FlashCopy: Point-in-time Copy feature	130
4.6.2 Remote Mirror and Copy feature	131
4.7 Additional information	135
4.8 IBM XIV Storage System	137
4.8.1 System models and components	137
4.8.2 Key design features	141
4.8.3 IBM XIV Storage System software	142
4.8.4 Parallelism	146

4.8.5 Full storage virtualization	148
4.8.6 Reliability, availability, and serviceability.	161
4.9 Comparison of products	169
4.10 More information	171
Chapter 5. IBM System Storage N series	173
5.1 IBM N series storage systems functionality	174
5.1.1 Drive flexibility	175
5.1.2 Near-line storage solution	175
5.2 IBM System Storage N series hardware	176
5.2.1 Maximum number of network interfaces	176
5.2.2 IBM System Storage N series A models hardware reference.	177
5.2.3 IBM System Storage N series A and G models hardware quick reference.	178
5.2.4 IBM System Storage N3000 introduction	179
5.2.5 IBM System Storage N6000 introduction	183
5.2.6 IBM System Storage N7000 introduction	185
5.3 IBM System Storage N series Gateways	187
5.3.1 IBM System Storage N series Gateway highlights	190
5.3.2 IBM System Storage N7700 and N7900 Gateway models	190
5.4 N series expansion units	192
5.4.1 Intermixing EXN units with N series A models	193
5.4.2 EXN1000.	193
5.4.3 EXN4000.	194
5.5 IBM System Storage N series standard software features	195
5.6 Optional software	196
5.7 Performance Accelerator Module	198
5.8 References	199
Chapter 6. IBM Scale Out Network Attached Storage	201
6.1 Overview	202
6.2 How SONAS works	202
6.3 SONAS features	204
6.3.1 Centrally managed, centrally deployed	204
6.3.2 Automated, integrated tiered storage and information life-cycle management.	204
6.3.3 Protecting data and replicating between sites.	206
6.3.4 Utilizing unique capabilities and modular hardware	206
6.3.5 Driving a dynamic infrastructure	207
6.4 Components within a SONAS solution	208
6.4.1 Interface node	208
6.4.2 Management node	209
6.4.3 Storage node	210
6.4.4 Infiniband switches	212
6.5 Summary of features and benefits	213
6.6 More information	214
Chapter 7. IBM Information Archive	215
7.1 Data retention and compliance	216
7.1.1 Characteristics of retention managed data	216
7.1.2 IBM strategy and positioning.	217
7.2 IBM Information Archive offering.	217
7.2.1 IBM Information Archive	218
7.2.2 IBM Information Archive components	220

Part 2. Tape systems	225
---------------------------------------	------------

Chapter 8. IBM TotalStorage and System Storage Tape Drives	227
8.1 LTO technology	228
8.1.1 Interleaved recording	229
8.1.2 Servo tracks	229
8.1.3 Track following	230
8.1.4 Linear density	234
8.2 IBM Ultrium 1, 2, 3, and 4 highlights	234
8.2.1 WORM tape format	235
8.2.2 Reliability	235
8.3 IBM LTO Tape Drives	236
8.3.1 IBM System Storage TS2230 Tape Drive Express Model	236
8.3.2 IBM System Storage TS2240 Tape Drive Express Model	237
8.3.3 IBM System Storage TS2340 Tape Drive Express Model	238
8.3.4 IBM TotalStorage 3580 Tape Drive	239
8.3.5 IBM 3592 Tape Drive	240
8.4 IBM System Storage TS1100 Tape Drive Family	240
8.4.1 The 1 TB background	241
8.4.2 Recording format	242
8.5 IBM System Storage TS1120 Tape Drive	243
8.5.1 Highlights	244
8.5.2 Media support	244
8.5.3 More information	245
8.5.4 IBM System Storage TS1120 Tape Controller Model C06	245
8.6 IBM System Storage TS1130 Tape Drive	246
8.7 IBM Midrange and Enterprise Storage Media	247
8.7.1 LTO cartridge	248
8.7.2 IBM 3590 and 3590E Cartridge	250
8.7.3 IBM 3592 Cartridge	250
Chapter 9. IBM Tape Automation Products	253
9.1 Overview	254
9.1.1 IBM System Storage TS2900 Autoloader	254
9.1.2 IBM System Storage TS3100 Tape Library	254
9.1.3 IBM System Storage TS3200 Tape Library	254
9.1.4 IBM System Storage TS3310 Tape Library	254
9.1.5 IBM System Storage TS3400 Tape Library	254
9.1.6 IBM System Storage TS3500 Tape Library	254
9.1.7 IBM TotalStorage 3494 Tape Library	255
9.2 IBM System Storage TS2900 Tape Autoloader Express	255
9.2.1 TS2900 features	256
9.2.2 TS2900 specifications	256
9.3 IBM System Storage TS3100 Tape Library Express Model	257
9.3.1 TS3100 features	258
9.3.2 TS3100 specifications	258
9.4 IBM System Storage TS3200 Tape Library	259
9.4.1 TS3200 features	260
9.4.2 TS3200 specifications	261
9.5 IBM System Storage TS3310 Tape Library	262
9.5.1 TS3310 features	263
9.5.2 TS3310 specifications	264
9.6 IBM System Storage TS3400 Tape Library	265
9.6.1 TS3400 features	266
9.6.2 TS3400 specifications	266

9.7 IBM System Storage TS3500 Tape Library	267
9.7.1 TS3500 features	268
9.7.2 TS3500 specifications	268
9.7.3 Advanced Library Management System	269
9.7.4 Virtual I/O slots	270
9.8 IBM TotalStorage 3953 Tape System for System z	270
9.8.1 Attachment methods	271
9.8.2 3953 Tape System features	272
9.9 Discontinued products	273
Chapter 10. IBM Tape Virtualization products	275
10.1 Introduction to Tape Virtualization	276
10.2 TS7650 ProtecTIER Deduplication Appliance	279
10.2.1 TS7650 ProtecTIER Deduplication Appliance features	280
10.2.2 TS7650 ProtecTIER Deduplication Appliance configurations	280
10.3 TS7650G ProtecTIER Deduplication Gateway	281
10.3.1 TS7650G ProtecTIER Deduplication Gateway features	282
10.3.2 TS7650G ProtecTIER Deduplication Gateway description	282
10.4 IBM System Storage ProtecTIER Enterprise Edition V2.3	282
10.4.1 Data deduplication	283
10.4.2 ProtecTIER Native Replication	284
10.5 IBM Virtualization Engine TS7700	285
10.5.1 TS7700 Virtualization Engine description	285
10.5.2 TS7700 Virtualization Engine highlights	286
10.5.3 IBM Virtualization Engine TS7700 Release 1.5 and 1.6	286
10.5.4 Characteristics of components comprising the TS7700	287
10.5.5 IBM Virtualization Engine TS7740	287
10.5.6 IBM Virtualization Engine TS7720	288
10.6 IBM Data Facility Storage Management Subsystem	288
10.6.1 The IBM Data Facility Storage Management Subsystem Removable Media Manager	288
10.6.2 DFSMSrmm highlights	289
10.7 More information	289
Chapter 11. Complementary storage products	291
11.1 IBM TotalStorage 7206 Model 336 External DDS Gen 5 Tape Drive	292
11.1.1 IBM 7206 Model 336 highlights	292
11.1.2 More information	292
11.2 IBM DDS Generation 5 USB Tape Drive	293
11.2.1 IBM DDS Generation 5 USB Tape Drive highlights	293
11.2.2 More information	294
11.3 IBM DDS Generation 6 USB Tape Drive	294
11.3.1 IBM DDS Generation 6 USB Tape Drive highlights	294
11.3.2 More information	295
11.4 IBM TotalStorage 7207 External Tape Drive	295
11.4.1 IBM TotalStorage 7207 External Tape Drive highlights	295
11.4.2 More information	296
11.5 IBM System Storage 7212 Storage Device Enclosure Express Model	296
11.5.1 IBM System Storage 7212 Storage Device Enclosure highlights	296
11.5.2 More information	297
11.6 IBM External Tape Enclosures	297
11.6.1 IBM System Storage 7214 Storage Device Enclosure	297
11.6.2 Half High Tape Drive External Enclosure	298

11.6.3 1U Tape Drive External Enclosure	299
Part 3. Storage networking	301
Chapter 12. Introduction to storage networking	303
12.1 Overview	304
12.2 Storage Area Network	305
12.2.1 Fibre Channel	306
12.2.2 SAN topologies	307
12.2.3 Physical components of the SAN infrastructure	311
12.2.4 Naming and addressing	314
12.2.5 Types of ports	315
12.2.6 FICON	315
12.3 IP storage networking technologies	315
12.3.1 iSCSI	316
12.3.2 FCIP	317
12.3.3 iFCP	318
12.3.4 Comparison of iSCSI, FCIP, iFCP	319
12.3.5 FCoE	321
12.4 Network Attached Storage	321
12.4.1 NAS benefits	322
12.4.2 NAS gateways	322
12.5 Wavelength Division Multiplexing	323
12.6 Selecting the best alternative	324
12.7 More information	324
Chapter 13. Storage Area Network products	325
13.1 SAN switches and connectivity products	326
13.1.1 Common characteristics	326
13.1.2 Other switch features	326
13.2 Entry SAN switches	328
13.2.1 IBM System Storage SAN24B-4 switch	329
13.2.2 Cisco MDS 9124 Express for IBM System Storage	330
13.3 Midrange SAN switches	331
13.3.1 IBM System Storage SAN40B-4	331
13.3.2 IBM System Storage SAN80B-4	332
13.3.3 Cisco MDS 9134 for IBM System Storage	333
13.4 Enterprise SAN switches	335
13.4.1 IBM TotalStorage SAN256B	335
13.4.2 IBM System Storage SAN384B	337
13.4.3 IBM System Storage SAN768B	338
13.4.4 Cisco MDS 9506 for IBM System Storage	340
13.4.5 Cisco MDS 9509 for IBM System Storage	342
13.4.6 Cisco MDS 9513 for IBM System Storage	344
13.5 Multiprotocol routers	346
13.5.1 IBM System Storage SAN04B-R multiprotocol router	347
13.5.2 IBM System Storage SAN06B-R multiprotocol router	349
13.5.3 Cisco MDS 9222i for IBM System Storage	349
13.6 SAN solutions	351
13.6.1 Infrastructure simplification solutions	351
13.6.2 Business continuity solutions	352
13.7 More information	354
Chapter 14. Data Center Networking	355

14.1 IBM b-type Ethernet switches and routers	356
14.1.1 IBM Converged Switch B32	356
14.1.2 IBM Ethernet Switch B24X	357
14.1.3 IBM r-series Ethernet Switches	358
14.1.4 IBM c-series Ethernet Switches	359
14.1.5 IBM g-series Ethernet Switches	360
14.1.6 IBM m-series Ethernet/IP routers	360
14.1.7 IBM s Series Ethernet Switch	362
14.2 IBM j-type Ethernet switches and routers	363
14.2.1 IBM Ethernet Switch J08E and IBM Ethernet Switch J16E	363
14.2.2 IBM Ethernet Switch J48E	365
14.2.3 IBM Ethernet Router J02M/J06M/J11M	366
14.3 Cisco Nexus 5000	367
14.3.1 Cisco Nexus 5000 overview	367
14.3.2 Cisco Nexus 5010/5020	367

Part 4. IBM System Storage software 369

Chapter 15. IBM System Storage Virtualization	371
15.1 What is storage virtualization?	372
15.2 The need for storage virtualization	372
15.2.1 Inefficiencies in sharing storage	373
15.2.2 Benefits of implementing storage virtualization in your SAN	373
15.3 Levels of storage virtualization	374
15.3.1 Server level	374
15.3.2 Network level	374
15.3.3 Storage subsystem level	374
15.4 Server, network, or storage subsystem virtualization	374
15.4.1 Implementing virtualization at a fabric or network level	375
15.4.2 Fabric level virtualization models	375
15.5 IBM System Storage approach to virtualization	376
15.5.1 IBM block aggregation characteristics	377
15.5.2 IBM file aggregation characteristics	377
15.6 IBM System Storage Virtualization family	377
15.6.1 IBM SAN Volume Controller	378
15.6.2 IBM XIV Storage System	378
15.6.3 IBM System Storage DS8000 Turbo	378
15.6.4 The IBM Virtualization Engine	378
15.6.5 IBM System Storage TS7650 ProtecTIER Appliance	379
15.6.6 IBM System Storage TS7650G ProtecTIER Deduplication Gateway	379
15.6.7 IBM Tivoli Storage Productivity Center	379
15.6.8 IBM System Storage N series	380
15.6.9 IBM General Parallel File System	380
15.7 More information	382
Chapter 16. IBM System Storage SAN Volume Controller	383
16.1 The need for storage virtualization	384
16.2 IBM System Storage SAN Volume Controller	385
16.2.1 Overview	385
16.2.2 Design intent	385
16.3 New features in recent SAN Volume Controller versions	387
16.3.1 SVC 5.1: New SVC engine hardware	388
16.3.2 SVC 5.1: New SVC software features	388
16.3.3 What was new with SVC 4.3.1	388

16.4	SAN Volume Controller architecture	389
16.5	SSD support in the SAN Volume Controller	391
16.6	SAN Volume Controller virtualization	392
16.6.1	Metro Mirror and Global Mirror	395
16.6.2	FlashCopy	397
16.6.3	Reverse FlashCopy	397
16.6.4	Mirrored VDisks	398
16.6.5	Space-Efficient VDisks	399
16.6.6	Space reclamation	400
16.7	SAN Volume Controller iSCSI host attachment	400
16.7.1	Supported iSCSI host attachments	403
16.7.2	Additional Ethernet changes	403
16.8	SAN Volume Controller software licensing	403
16.9	SAN Volume Controller solutions	406
16.9.1	Business continuity	406
16.9.2	Infrastructure simplification	407
16.10	SAN Volume Controller compatibility	409
16.11	SAN Volume Controller commonly used terms	410
16.12	More information	411
Chapter 17.	IBM Tivoli Storage Manager	413
17.1	New for IBM Tivoli Storage Manager V6.1.2	414
17.1.1	Server updates	414
17.1.2	Reporting and monitoring feature	418
17.1.3	Backup-archive client updates	418
17.1.4	Application program interface updates	419
17.1.5	Space Management for UNIX and Linux client updates	420
17.1.6	HSM for Windows updates	420
17.1.7	IBM Tivoli Storage FlashCopy Manager V2.1	420
17.1.8	Data Protection for Snapshot Devices updates	421
17.1.9	Data Protection for Mail updates	421
17.1.10	Data Protection for SAP updates	421
17.1.11	Storage agent updates	422
17.1.12	Products and components not updated in V6.1	422
17.1.13	Information center updates	422
17.2	IBM Tivoli Storage Manager overview	423
17.3	IBM Tivoli Storage Manager architecture	423
17.3.1	IBM Tivoli Storage Manager architectural components	425
17.3.2	How client data is stored	429
17.3.3	How the server manages storage	431
17.3.4	Tivoli Storage Manager backup and archive concepts	435
17.3.5	Progressive incremental backups	435
17.3.6	Tivoli Storage Manager reporting and monitoring	435
17.3.7	Tivoli Storage Manager Administration Center	438
17.3.8	Tivoli Storage Manager API	439
17.3.9	Tivoli Storage Manager security	439
17.4	IBM Tivoli Storage Manager family of products	440
17.4.1	IBM Tivoli Storage Manager Basic or Standard Edition	440
17.4.2	IBM Tivoli Storage Manager Extended Edition	441
17.5	IBM Tivoli Storage Manager complementary products	444
17.5.1	IBM Tivoli Storage Manager for Mail	445
17.5.2	IBM Tivoli Storage Manager for Databases	445
17.5.3	IBM Tivoli Storage Manager for Microsoft SharePoint	445

17.5.4 IBM Tivoli Storage Manager HSM for Windows	446
17.5.5 IBM Tivoli Storage Manager for Advanced Copy Services	446
17.5.6 IBM Tivoli Storage Manager for Copy Services	447
17.5.7 IBM Tivoli Storage Manager for SAP V6.1	447
17.5.8 IBM Tivoli Storage Manager for Space Management	448
17.5.9 IBM Tivoli Storage Manager for Storage Area Networks	448
17.5.10 IBM Tivoli Storage Manager for System Backup and Recovery	449
17.5.11 IBM Tivoli Storage Manager FastBack	449
17.5.12 IBM Tivoli Storage Manager FastBack for Microsoft Exchange	449
17.5.13 IBM Tivoli Storage Manager FastBack for Bare Machine Recovery	450
17.5.14 IBM Tivoli Continuous Data Protection for Files	450
17.6 Client backup and restore operations	451
17.6.1 Traditional LAN and WAN backup topology	451
17.6.2 SAN backup topology: LAN-free	452
17.6.3 Split-mirror/point-in-time copy backup using SAN	453
17.6.4 Network-attached storage backup and restore	454
17.6.5 Image backup	455
17.6.6 VMware Consolidated Backup	456
17.7 Tivoli Storage Manager server-to-server communications	456
17.7.1 Enterprise configuration	457
17.7.2 Using server-to-server virtual volumes for disaster recovery	458
17.8 Tivoli Storage Manager and high availability clustering	459
17.8.1 IBM High Availability Cluster Multi-Processing	459
17.8.2 Backup-archive client support with HACMP	461
17.8.3 Tivoli Storage Manager server and Microsoft Cluster Server (MSCS)	462
17.8.4 Tivoli Storage Manager backup-archive client support with MSCS	465
17.9 Tivoli Storage Manager and tape vaulting	466
17.9.1 Electronic tape vaulting	466
17.9.2 Remote disk mirroring and tape vaulting solutions	468
17.9.3 Collocation considerations for off-site vaulting	469
17.9.4 Reclamation considerations for off-site vaulting	469
17.10 More information	471
Chapter 18. IBM Tivoli Storage Productivity Center	473
18.1 New with IBM Tivoli Storage Productivity Center V4.1.1.	474
18.1.1 Name and licensing changes	474
18.1.2 Integration features	474
18.1.3 Storage Resource agents	476
18.1.4 SQL access to Tivoli Storage Productivity Center database	476
18.1.5 Performance management improvements	477
18.1.6 Storage Optimizer	477
18.1.7 Advanced Brocade Discovery	477
18.1.8 Discovery Library Adapter	478
18.1.9 Storage resource groups	478
18.1.10 IBM Support Assistant	478
18.1.11 IBM General Parallel File System	479
18.1.12 Installation changes	479
18.1.13 New device and application support	479
18.1.14 Features supported in a future release	481
18.2 Tivoli Storage Productivity Center overview	482
18.3 Tivoli Storage Productivity Center architecture	483
18.3.1 Tivoli Storage Productivity Center components	483
18.3.2 Tivoli Storage Productivity Center structure	486

18.4	Standards and protocols used in IBM Tivoli Storage Productivity Center	488
18.4.1	Common Information Model and Web-Based Enterprise Management	488
18.4.2	Storage Management Initiative - Specification	489
18.4.3	Service Location Protocol	489
18.5	IBM Tivoli Storage Productivity Center management applications	490
18.5.1	Data Manager	491
18.5.2	Disk Manager	497
18.5.3	Disk Manager Overview	497
18.5.4	Fabric Manager	501
18.5.5	Tape Manager	505
18.6	Performance analysis with Storage Optimizer	506
18.6.1	Storage Optimizer overview	507
18.6.2	How the Storage Optimizer works	508
18.6.3	Important performance metrics	509
18.6.4	Supported subsystems or applications	509
18.7	IBM Tivoli Provisioning Manager	510
18.7.1	Tivoli Provisioning Manager overview	510
18.7.2	Tivoli Provisioning Manager workflow	511
18.8	Tivoli Storage Productivity Center components and packages	512
18.8.1	IBM Tivoli Storage Productivity Center Basic Edition	513
18.8.2	IBM Tivoli Storage Productivity Center for Data	514
18.8.3	IBM Tivoli Storage Productivity Center for Disk	515
18.8.4	IBM Tivoli Storage Productivity Center For Replication	516
18.8.5	IBM Tivoli Storage Productivity Center Standard Edition	524
18.8.6	IBM Tivoli Storage Productivity Center with Advanced Provisioning	525
18.9	IBM Tivoli Storage Productivity Center licenses	525
18.10	IBM System Storage Productivity Center	527
18.10.1	New with IBM System Storage Productivity Center Version 1.4	527
18.10.2	System Storage Productivity Center overview	528
18.10.3	System Storage Productivity Center hardware components	532
18.10.4	System Storage Productivity Center licensing	532
18.10.5	System Storage Productivity Center support matrix	533
18.11	Interoperability	534
18.12	More information	535
Chapter 19.	IBM Grid Access Manager	537
19.1	GAM overview	538
19.2	Performance and availability	538
19.3	Data protection and business continuity	539
19.3.1	Data integrity	539
19.3.2	Multi site operation	539
19.4	Simplified management and optimized storage utilization	540
19.4.1	Intelligent information life-cycle management	540
19.4.2	Operation across facilities and tiers	540
19.4.3	Flexible hardware configurations	540
19.4.4	Automated data migration	540
19.4.5	Open, standards-based access	540
19.4.6	Proactive monitoring and management	541
19.5	Summary of the key features and benefits	541
19.6	More information	541
Chapter 20.	IBM GPFS and SoFS	543
20.1	IBM General Parallel File System	544

20.2	GPFS overview	544
20.3	Benefits and highlights of GPFS	545
20.3.1	Scalable high-performance shared disk file systems	545
20.3.2	Increased data availability	546
20.3.3	Simplified storage management	546
20.3.4	Simplified administration: Features and benefits of GPFS	547
20.4	SoFS overview	548
20.5	How SoFS works	549
20.6	Why use SoFS	550
20.6.1	The scale-up approach	550
20.6.2	Difference between SoFS and other clustered file systems	550
20.7	Components within an SoFS solution	551
20.7.1	SoFS hardware components	551
20.7.2	SoFS software components	552
20.8	Data protection and disaster recovery	553
20.8.1	Data protection	553
20.8.2	Simple SoFS single site solution: No disaster recovery	553
20.8.3	SoFS with synchronous replication on-site	554
20.8.4	SoFS with cross site replication	555
20.8.5	SoFS with a cross cluster mount	556
20.8.6	Snapshot technology	556
20.9	More information	557
Part 5.	z/OS storage management	559
Chapter 21.	Data Management with DFSMS and Data Sorting with DFSORT	561
21.1	DFSMS components	562
21.2	DFSMSdfp (Data Facility Product)	562
21.2.1	DFSMS SDM Copy Services	562
21.2.2	Features	563
21.2.3	z/OS NFS	563
21.2.4	XRC Performance Monitor	563
21.2.5	IBM TotalStorage Expert	563
21.3	DFSMSdss (Data Set Services)	564
21.3.1	Concurrent Copy	564
21.3.2	FlashCopy	565
21.3.3	DFSMSdss Stand-Alone Services	565
21.4	DFSMSHsm (Hierarchical Storage Manager)	565
21.4.1	The storage hierarchy	565
21.4.2	Availability management	565
21.4.3	Space management	565
21.4.4	Disaster recovery	566
21.4.5	List, report, query, audit	566
21.5	DFSMSrmm (Removable Media Manager)	566
21.5.1	Management of the tape library	566
21.5.2	DFSMSrmm volume retention and movement policies	566
21.5.3	Report generator	567
21.5.4	Full function removable media management system	567
21.5.5	Compatible with existing tape management systems	567
21.6	DFSMSStvs	567
21.7	More information	567
21.8	DFSORT	568
21.8.1	Product highlights	568

21.8.2 DFSORT/VSE	569
21.8.3 Features	569
Chapter 22. z/OS Storage Management Tools	571
22.1 IBM Tivoli Omegamon XE	572
22.2 IBM DB2 Cloning Tool	572
22.3 IBM Tivoli Advanced Catalog Management for z/OS	573
22.4 IBM Tivoli Advanced Backup and Recovery	573
22.5 IBM Tivoli Advanced Audit for DFSMSHsm	574
22.6 IBM Tivoli Automated Tape Allocation Manager	575
22.7 IBM Tivoli Advanced Allocation Management on z/OS	575
22.8 IBM Tivoli Storage Optimizer for z/OS	576
22.9 IBM Tivoli Tape Optimizer for z/OS	576
22.10 DFSMS Optimizer	576
Chapter 23. Smart Business Storage Cloud	579
23.1 Overview	580
23.1.1 Public cloud versus private cloud	580
23.1.2 Public cloud	580
23.1.3 Private cloud	580
23.2 Smart Business Storage Cloud components	581
23.2.1 How it works	582
23.2.2 The difference between SBSC and other clustered file systems	582
23.3 The benefits of using an SBSC	583
23.3.1 Continuous availability	583
23.3.2 Online transparent data migration	584
23.3.3 Summary of benefits	585
23.4 Components within an SBSC solution	586
23.4.1 Host attachment	586
23.4.2 Hardware	587
23.4.3 Software components	589
23.5 Data protection and disaster recovery	590
23.5.1 Data protection	590
23.5.2 Simple SBSC single site solution: No disaster recovery	590
23.5.3 SBSC with synchronous replication on-site	590
23.5.4 SBSC with cross site sync or async replication	591
23.5.5 SBSC with a cross cluster mount	591
23.5.6 Snapshot technology	592
23.6 IBM General Parallel File System (GPFS)	593
23.7 More information	593
Appendix A. Internet Small Computer System Interface (iSCSI)	595
Introduction to iSCSI	596
History	596
Overview	596
Benefits	597
Appendix B. Redundant Arrays of Independent Disks (RAID)	599
Appendix C. Introduction to SMI-S and CIM/WBEM	609
Related publications	613
IBM Redbooks publications	613
Other publications	614

Online resources	614
How to get Redbooks publications.....	614
Help from IBM	614
Index	615

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the web at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

1350™	HyperFactor®	Redbooks (logo)  ®
AIX 5L™	HyperSwap®	RS/6000®
AIX®	i5/OS®	S/390®
AS/400®	IBM®	SANergy®
BladeCenter®	IMS/ESA®	Smarter Planet™
Calibrated Vectors Cooling™	Informix®	SysBack®
CICS®	iSeries®	System i®
DB2 Universal Database™	Lotus®	System p5®
DB2®	Netfinity®	System p®
Domino®	OMEGAMON®	System Storage™
DS4000®	OS/390®	System Storage DS®
DS6000™	OS/400®	System x®
DS8000®	Parallel Sysplex®	System z9®
Dynamic Infrastructure®	Passport Advantage®	System z®
ECKD™	Power Systems™	TDMF®
Enterprise Storage Server®	POWER5™	Tivoli Enterprise Console®
ESCON®	POWER5+™	Tivoli®
eServer™	POWER6®	TotalStorage®
Express Storage™	PowerHA™	WebSphere®
FICON®	PowerPC®	XIV®
FlashCopy®	POWER®	z/OS®
GDPS®	ProtecTIER®	z/VM®
Geographically Dispersed Parallel Sysplex™	pSeries®	z9®
GPFS™	RACF®	zSeries®
HACMP™	Redbooks®	
	Redpaper™	

The following terms are trademarks of other companies:

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

Snapshot, SecureAdmin, RAID-DP, Network Appliance, FlexShare, FlexCache, VFM, SyncMirror, SnapVault, SnapValidator, SnapRestore, SnapMover, SnapMirror, SnapManager, SnapLock, SnapDrive, NearStore, MultiStore, FlexVol, FlexClone, FilerView, Data ONTAP, NetApp, and the NetApp logo are trademarks or registered trademarks of NetApp, Inc. in the U.S. and other countries.

Java, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows NT, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Celeron, Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Preface

IBM® System Storage solutions provide the capabilities needed to help organizations reduce cost and mitigate risks while managing continued information growth and service level demands. Dynamic environments, such as Cloud Computing and Service Oriented Architecture, require scalable, always on information services delivered by a robust storage infrastructure.

This IBM Redbooks® publication provides overviews and pointers for information about the most current IBM System Storage™ products, showing how IBM delivers the right mix of products for nearly every aspect of business continuance and business efficiency. IBM System Storage products can help you store, safeguard, retrieve, and share your data.

We cover these topics:

- ▶ First, we introduce the IBM Dynamic Infrastructure® and the Smarter Planet™ Strategy.
- ▶ Part 1 describes disk products that include IBM System Storage DS® Series (entry-level, midrange, and enterprise offerings), IBM XIV® Storage System, IBM System Storage N Series, and the IBM Information Archive.
- ▶ Part 2 presents an overview of the IBM TotalStorage® and System Storage Tape Drives, IBM Tape Automation Products, and IBM Tape Virtualization solutions Products.
- ▶ Part 3 describes storage networking infrastructure and presents the switches and directors to form SAN solutions. Also provided is an overview of the recently introduced IBM Ethernet switches and routers
- ▶ Part 4 discusses the IBM System Storage software portfolio, which includes IBM System Storage Virtualization, IBM Tivoli® Storage Manager, and IBM Tivoli Storage Productivity Center (TPC) Suite. We also introduce the IBM System Storage Multilevel Grid Access Manager Software (GAM), IBM General Parallel File System (GPFS™), and the IBM scale out file services (SOFS)
- ▶ Part 5 describes the z/OS® storage management software (DFSMS and DFSORT) and z/OS storage management tools, including an overview of the IBM Smart Business Storage Cloud solution.
- ▶ Finally, in the Appendixes, we provide information about the iSCSI, RAID and standards such as Storage Management Initiative Specification (SMI-S), Common Information Model (CIM), and Web Based Enterprise Management (WBEM).

This book is intended as a reference for basic and comprehensive information about the IBM System Storage products portfolio. This book provides a starting point when establishing your own enterprise storage environment.

Note: This book, previously published in 2006, has since been extensively updated for the 2010 edition.

The team who wrote this book

This book was produced by a team of specialists from around the world working at the IBM International Technical Support Organization (ITSO), San Jose Center.



The Team: Gary, Sebastian, Sangam, Anna, Emile, Glenn, and Vikas

Sangam Racherla is an IT Specialist and Project Leader working at the International Technical Support Organization (ITSO), San Jose Center. He holds a degree in electronics and communication engineering and has nine years of experience in the IT field. He has been with the ITSO for the past six years and has extensive experience installing and supporting the ITSO lab equipment for various Redbooks® publication projects. His areas of expertise include Microsoft® Windows®, Linux®, AIX®, System x®, and System p® servers, and various SAN and storage products.

Vikas Bajaj is a Senior Associate IT Specialist (Storage). He has been working with IBM for last three years and has worked in the domain of IBM storage, Brocade, and Cisco. He is responsible for maintaining the storage infrastructure for India Software Lab (ISL), which includes administering IBM DS8000®, DS6800, DS4700, Brocade Fibre Channel switch, Cisco Fibre Channel switch, and IBM N series NAS boxes. As a Senior Associate, he provides day to day support to all the customers on IBM storage and also helps design and implement IBM storage solutions for internal customers. In the area of TPC, he has provided training to partner labs and has published articles about TPC implementation.

Emile Knebel is a Senior IT Specialist at IBM the Netherlands. He joined IBM twelve years ago in the IBM Business Partner technical marketing support line. After this period, he gained eight years of experience in open systems platforms within STG. During this period he built up his technical pre-sales expertise within Server Consolidation and Enterprise Resource planning (ERP) projects. After working as a people manager in the IBM ISC Deal Hub

organization, he returned to STG the Netherlands, to become Storage Product Manager for the IMT Benelux, where he has been practicing this assignment for almost two years.

Glenn Korn is a member of the World-Wide Technical Support Marketing group for the IBM Systems and Technology group, where he specializes in SAN Storage solutions. Glenn has been working with computer hardware and networks for over 18 years. He has extensive experience in IT support of both hardware and software. He holds a B.S. degree in Industrial Technology and is a Microsoft Certified Systems Engineer (MCSE) and Novell CNE.

Anna Kutepova is a Customer Quality Engineer, doing problem analysis of the installation process for DS8000 data storage servers in the IBM DSS plant in Vac, Hungary. She is a leader of the Customer Engineer Workshop project for DS8000. She has a connection with the Budapest University of Technology and Economics and leads students' projects as a consultant from manufacturing.

Gary Newell is a Senior Support Specialist in the UK working for the last 10 years on open systems storage, SAN, and virtualization products. He has been with IBM for 10 years and has worked in the UK Hardware Support Centre and EMEA PFE teams in that time frame. He has more than 20 years experience in the IT industry, having worked in European and worldwide support roles for Honeywell, StorageTek, and Sequent before joining IBM.

Sebastian Szumczyk is an Accredited IT Architect for IBM Global Technology Services from Warsaw, Poland. He has twelve years experience in the IT industry and has focused primarily on infrastructure architectures in the last four years. His areas of expertise include Tivoli Storage and System Management, pSeries®, AIX, virtualization, storage, and high availability solutions. He holds several certifications within technical fields, including a prestigious IBM System p5® Certified Advanced Technical Expert certificate. He holds an engineering degree in Administering Operating Systems and Databases from Warsaw Information Technology under the auspices of the Polish Academy of Sciences.

Thanks to the following people for their contributions to this project:

Bertrand Dufrasne
Mary Lovelace
Ann Lund
Alex Osuna
Jon Tate
IBM ITSO, San Jose Center

Juergen Deicke
Roger L. Haskin
John Hoogenboom
Sven Oehme
Ronnie Sahlberg
Andrew Tridgell
Barry Whyte
IBM

Thomas Gruber
LSI Corporation

Many thanks also to the authors of the previous editions of this book.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author - all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

ibm.com/redbooks

- Send your comments in an email to:

redbooks@us.ibm.com

- Mail your comments to:

IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks publications

- Find us on Facebook:

<http://www.facebook.com/IBMRedbooks>

- Follow us on twitter:

<http://twitter.com/ibmredbooks>

- Look for us on LinkedIn:

<http://www.linkedin.com/groups?home=&gid=2130806>

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks publications weekly newsletter:

<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>

- Stay current on recent Redbooks publications with RSS Feeds:

<http://www.redbooks.ibm.com/rss.html>



Introduction: Dynamic Infrastructure and IBM Smarter Planet Strategy

In this chapter we present an introduction to the book as a whole. Subsequent chapters are organized under parts according to specific topics.

The IBM vision is to bring a new level of intelligence to the way the world works, and how every person, business, organization, government, natural system, and man-made system interacts. Each interaction represents a chance to do something better, more efficiently, and more productively. But, more than that, as the systems of the planet become smarter, we have a chance to open up meaningful new possibilities for progress. By harnessing our increasingly digitized world, we can help to solve intractable social problems, bring organizations closer to our customers, and vastly shrink decision windows that help executives drive better business outcomes.

Yet, as the pace of business continues to accelerate and the planet becomes smarter, the physical and digital foundations on which progress depends are straining to keep up. They are too complex, too inefficient, and too inflexible, and the current explosion of network connected devices and data will only intensify the demands that we continue to place on IT.

Within our own organizations, we face a growing realization that our physical and IT infrastructure is not built for what's coming. Going forward, the most successful organizations will be those with technology that enables informed, real-time decision making at the point of greatest business impact without requiring analytical expertise.

Consider what might it mean to your organization if you were always able to access critical business data at the exact moment you needed it. Likewise, imagine that you can improve service and reduce costs by delivering IT services when the customer requests those services and automatically reclaim the services when they are no longer needed.

Storage is as important to a dynamic infrastructure as servers and networks. Because of the explosion of information in most data centers, storage can be more challenging to manage than server resources, but the payoff is potentially more beneficial.

IBM is a market leader in the storage industry. IBM System Storage solutions offer excellent performance, innovative technology, and open standards. IBM System Storage solutions can help you manage information growth more effectively, dramatically improve service levels, lower costs associated with energy consumption and footprint, mitigate risks, and extract new intelligence for business insight.

IBM Storage solutions provide the capabilities needed to help organizations reduce cost and mitigate risks while managing continued information growth and service level demands. Dynamic environments, such as Cloud Computing and Service Oriented Architecture, require scalable, always on information services delivered by a robust storage infrastructure.

One of the aims of the IBM Smarter Planet strategy is to enable the dynamic infrastructure in order to deliver this improved service at a reduced cost and with better managed risk.

1.1 What a dynamic infrastructure is

The cornerstones of a dynamic infrastructure (Figure 1-1) are clear:

- ▶ Improving service
- ▶ Reducing cost
- ▶ Managing risk

In this book we look at several IBM system storage offerings available today and explore how they can be used to help deliver the benefits of a dynamic infrastructure.



Figure 1-1 The dynamic infrastructure

A dynamic infrastructure integrates the management of business and IT assets and aligns them with the overall goals of the business while taking a new, smarter, and more streamlined approach to help improve service, reduce cost, and manage risk. In today's ever more competitive and growing business environment, information is an increasingly valuable, but costly organizational asset. The volume of information is growing very rapidly in most organizations, and with this, the need to protect and manage information also continues to increase. Organizations are seeking to minimize risk, reduce costs, and increase flexibility by aligning IT investments according to information value and business needs.

1.2 Delivering a dynamic infrastructure

A dynamic infrastructure is designed for today's instrumented and interconnected world, helping clients to integrate their growing intelligent business infrastructure with the necessary underlying design of a flexible, secure, and seamlessly managed IT infrastructure.

IBM System Storage products can help deliver the solutions required to implement a dynamic infrastructure. Such solutions include:

- ▶ Information infrastructure
- ▶ Business continuity and resiliency
- ▶ Virtualization
- ▶ Energy efficiency
- ▶ Security
- ▶ Service management

1.2.1 Information infrastructure

IBM information infrastructure solutions are designed to help you manage the information explosion and address challenges around information compliance, availability, retention, and security. This can lead your company toward improved productivity, service delivery, and reduced risk while streamlining costs:

- ▶ **The goal: Continuous and reliable access to information**
Your organization's success can hinge on information availability. Providing access to critical information when and where it is needed, has an immediate impact on employee productivity, client satisfaction, and your bottom line. Even one hour of downtime can have dramatic effects and cause significant expense. Organizations must take steps to ensure information availability so users have reliable access to critical information in real time.
- ▶ **The advantage: Managing information availability for efficiency and cost effectiveness**
IBM can help. IBM offers comprehensive services, hardware, and software to help you build and implement an information availability strategy. IBM information infrastructure helps you deliver quality information services, even as demand increases or if components fail.
- ▶ **The benefits: Information and business resiliency**
IBM information infrastructure solutions can help you establish a solid foundation for keeping your information available. By optimizing the software, servers, storage, and networks, we can help you create a secure, resilient information infrastructure that allows you to exploit your data, maximize its value, and deliver the required information to the right people at the right time.

Here is how IBM system storage products can help you to develop your information infrastructure:

- ▶ **Storage virtualization:** Reducing SAN disk cost by increasing utilization
- ▶ **Data deduplication:** Lowering storage acquisition costs while reducing energy, cooling, floor space, management requirements, and maintenance costs
- ▶ **Solid-state storage architectures:** Improving drive response time up to 800% without application tuning
- ▶ **Enterprise Storage:** Choosing optimized enterprise storage solutions for Linux, UNIX®, Windows, and z/OS environments
- ▶ **Information Archiving:** Optimizing application performance and simplify application administration, while lowering total costs
- ▶ **Next Generation Storage Systems:** Revolutionary disk systems that can help deliver Tier 1 functionality at Tier 2 costs
- ▶ **Encryption and Security Management:** Encryption at the drive level that improves security with little or no performance impact
- ▶ **Storage Management:** Improving storage utilization and simplifying administration

- ▶ Continuous Data Protection: Getting applications and users up and running within minutes following a data loss

For more details about information infrastructure, see the website:

http://www.ibm.com/systems/information_infrastructure

1.2.2 Business continuity and resiliency

Business continuity and resiliency addresses the following topics:

- ▶ Business continuity planning: Evaluating, planning, and mitigating the business impact of various types of risks
- ▶ Regulatory compliance: Responding to regulatory requirements associated with business resiliency
- ▶ High availability: Assessing, planning for, creating, and running an infrastructure that supports the ongoing availability and performance objectives of the business
- ▶ Information protection: Ensuring that data is available and accessible in the event of a disruption
- ▶ Disaster recovery: Ensuring that human capital, processes, and systems are able to predictably recover and respond to catastrophic disruptive events that might require recovery at a remote location

For more information about business continuity and resiliency, see the website:

http://www.ibm.com/systems/business_resiliency

1.2.3 Virtualization

Virtualization and consolidation can reduce IT complexity to help your data center become more resilient and secure while reducing costs. The ability to deploy capacity and server images virtually increases speed of deployment roughly by a factor of 30 times¹.

Virtualization of physical storage resources is an essential part of a dynamic infrastructure environment. Such resources are more flexible because the logical files and disks can be manipulated more easily than numerous individual physical disk and tape devices. Virtualization can further improve resource utilization, simplify storage management, and accelerate application deployment with improved resource sharing. These capabilities can help organizations enact a tiered storage environment, where information is stored at a cost befitting its value at that point in time.

For more information about virtualization, see the website:

<http://www.ibm.com/itsolutions/virtualization>

1.2.4 Energy efficiency

Energy efficiency solutions can help lower data center energy costs by up to 40% or more through facilities design, power, and cooling infrastructure, active energy management, efficient, scalable IBM systems and, most importantly, better utilization of existing resources.

The challenges have become clear: the need for clean water and air; affordable and reliable delivery of energy; the dwindling supply of fossil fuels; the reality of climate change and its implications for future generations.

¹ Tom Bittman, Gartner 27th Annual Data Center Conference. December 2008

At IBM, our approach is two-fold: we are working to make our existing products and processes more efficient for both the environment and for business, while also developing new innovations that can accelerate the adoption of products and services that have lesser environmental impact.

Today's energy- and climate-related issues are at the top of our strategic agenda. We recognize that information technology plays an extremely important role in helping solve the myriad of ecological challenges faced by the global society, such as conserving our scarce resources even as global demand skyrockets, reducing pollution, minimizing the environmental impact of our activities, and enabling safe and renewable alternative sources of energy.

In the past year, IBM conducted more than 30 energy assessments around the world and found the average data center infrastructure efficiency (DCiE) index was 43 percent. This means that 43 cents of every energy dollar are spent to power IT, while 57 cents are spent on the physical infrastructure for conditioned power and cooling.

For more information about energy efficiency, see the website:

<http://www.ibm.com/itsolutions/energy-efficiency>

1.2.5 Security

Security is one critical component of a dynamic infrastructure. A flexible, dynamic infrastructure can enable real business innovation by integrating security with other key IT components so they function together to help you reduce cost, improve service, and manage risk.

IT security used to rely on a perimeter-based approach. But in today's environment, it is hard to define the perimeter. Web-based applications improve the productivity of your employees, partners, and customers, but they also expose you to risk. A service oriented infrastructure, offers extraordinary flexibility and efficiency, but it forces you to re-think data security. It used to be possible to protect your data with a perimeter, but now you have to secure that data wherever it resides and while it is in motion.

For more information about IBM data and information security solutions, see the website:

<http://www.ibm.com/software/tivoli/governance/security/information.html>

1.2.6 Service management

A smart planet is more instrumented, interconnected, and intelligent, creating new opportunities for business to deliver innovative, differentiated services across a dynamic infrastructure. IBM Service Management provides visibility, control, and automation across all stages of the service life cycle. Organizations can build, run, and manage services across both IT and business environments to support rapidly changing business requirements. We back this with years of industry expertise and worldwide experience.

For more information about Service management, see the website:

http://www.ibm.com/systems/dynamicinfrastructure/news/announcements/20091020/service_management.html

1.3 IBM System Storage Disk Systems

The IBM System Storage Disk Systems products and offerings provide storage solutions with superior value for all levels of business from small and medium business (SMB) to high-end and enterprise systems.

1.3.1 High-end and enterprise disk systems

High-performance disk storage meets the needs of business-critical enterprise workloads.

IBM System Storage DS8000

The IBM System Storage DS8000 Turbo offers high-performance, high-capacity, secure storage systems that are designed to deliver resiliency and total value for the most demanding, heterogeneous storage environments.

See 4.4, “The DS8000 series” for more information.

IBM XIV Storage System

The IBM XIV Storage System is a ground breaking, high-end, open disk system designed to support business requirements for a highly available information infrastructure. The XIV architecture is a grid of standard Intel/Linux components connected in any-to-any topology by means of massively paralleled, non-blocking Gigabit Ethernet, providing outstanding enterprise-class reliability, performance, and scalability.

See 4.8, “IBM XIV Storage System” for more information.

IBM System Storage DS6000

The IBM System Storage DS6000™ series is designed to provide high availability and high performance in a small, modular package. This series, along with the DS8000 series, offers an enterprise-class continuum of storage systems with shared replication services and common management interfaces.

See 4.3, “The DS6000 series” for more information.

1.3.2 Midrange disk storage

Midrange disk storage products provide convenient solutions for an on demand world.

IBM System Storage DS5020 Express

In a continuing effort to provide storage solutions that are designed to provide low total cost of ownership, high performance, robust functionality, and unparalleled ease of use, IBM offers the IBM System Storage DS5020 Express. As part of the DS series, the DS5020 Express offers high-performance 8 Gbps capable Fibre Channel connections, optional 1 Gbps iSCSI interface for less demanding applications and lower cost implementation, up to 50.4 TB of Fibre Channel physical storage capacity, up to 112 TB of SATA physical storage capacity.

See 3.4, “IBM System Storage model DS5020” for more information.

IBM System Storage DS5000 Series

The IBM System Storage DS5000 series is designed to meet today's and tomorrow's demanding open-systems requirements while establishing a new standard for life cycle longevity with field-replaceable host interface cards.

See 3.5, “IBM System Storage Models 5100 and 5300” for more information.

IBM System Storage DS4700 Express

The IBM DS4700 Express offers high-performance 4 Gbps capable Fibre Channel connections, up to 33.6 TB of Fibre Channel physical storage capacity, 112 TB of SATA physical storage capacity, and powerful system management, data management, and data protection features. The DS4700 Express is designed to expand from workgroup to enterprise-wide capability with up to six Fibre Channel expansion units with the DS4000® EXP810 Expansion Unit.

See 3.3, “IBM System Storage Model DS4700 Express” for more information.

1.3.3 Entry-level disk

Designed to deliver advanced functionality at a breakthrough price, these systems provide an exceptional solution for workgroup storage applications such as email, file, print, and web servers, as well as collaborative databases and remote boot for diskless servers.

IBM System Storage DS3200

The IBM System Storage DS3200 Express is an external storage enclosure specifically designed for the SMB. With SAS host interface, next generation SAS back-end technology, and SAS and SATA drive intermix support, The DS3200 addresses the top concerns of these businesses managing increasing demand for capacity, data availability, and functionality.

See 2.2.1, “IBM System Storage model DS3200” on page 32 for more information.

IBM System Storage DS3300

The IBM System Storage DS3300 Express is an external storage system specifically designed for a range of organizations’ data requirements. With iSCSI protocol, next-generation SAS back-end technology, and SAS and SATA drive intermix support, the DS3300 Express storage system provides businesses with robust, reliable, and cost-effective networked storage.

See 2.2.2, “IBM System Storage model DS3300” on page 34 for more information.

IBM System Storage DS3400

The IBM System Storage DS3400 Express is a FC host-based external storage system specifically designed for a wide range of organizations. With FC interface technology, next-generation SAS back-end technology, SAS and SATA drive intermix support, and DS3000 Storage Manager software, the DS3400 storage system provides businesses with robust, reliable, cost-effective FC networked storage.

See 2.2.3, “IBM System Storage model DS3400” on page 35 for more information.

IBM System Storage EXP3000

The IBM System Storage EXP3000 features 3 Gbps SAS interface technology. It has support for up to 5.4 TB of SAS disk physical storage capacity (450 GB SAS drives) in a single enclosure.

See 2.3, “IBM System Storage EXP3000” on page 37 for more information.

1.4 IBM Tape systems

IBM Tape storage, backup, and information management systems address a range of needs.

1.4.1 Enterprise tape systems

IBM enterprise class tape products are designed to offer high performance, availability, reliability, and capacity to help meet the needs of customers seeking enterprise class storage solutions for mass storage, data archiving, enterprise backup, and disaster recovery. IBM enterprise class storage products provide on demand solutions for business.

TS7700 Virtualization Engine

The IBM Virtualization Engine TS7700 (TS7700 Virtualization Engine) is a family of mainframe virtual-tape solutions that are designed to optimize tape processing. With one solution, the implementation of a fully integrated tiered storage hierarchy of disk and tape utilizes the benefits of both technologies to help enhance performance and provide the capacity needed for today's tape processing requirements. Deploying this innovative subsystem can help reduce batch processing time, total cost of ownership and management overhead.

See 10.5, "IBM Virtualization Engine TS7700" on page 285 for more information.

TS7650G ProtecTIER Deduplication Gateway

The IBM System Storage TS7650G ProtecTIER® Deduplication Gateway offers industry leading inline deduplication performance and scalability up to 1 petabyte (PB) of physical storage capacity per system that can provide up to 25 PBs of storage capacity. Combined with IBM storage, the ProtecTIER Gateway solution provides a powerful disk-based repository to improve the retention and availability of backup and archive data.

See 10.3, "TS7650G ProtecTIER Deduplication Gateway" on page 281 for more information.

TS3500 Tape Library

The IBM TS3500 Tape Library supports System z® through the IBM 3953 Tape System (3953 tape system) or the IBM Virtualization Engine TS7740, which enable System z hosts to access the TS3500 Tape Library cartridge inventory and allow connection to TS1130, TS1120 and IBM TotalStorage 3592 Model J1A tape drives.

See 9.7, "IBM System Storage TS3500 Tape Library" on page 267 for more information.

TS1130 Tape Drive

The IBM TS1130 Tape Drive offers high-performance flexible data storage with support for data encryption. The TS1130 Tape Drive supports IBM System Storage TS3400 and TS3500 Tape Libraries, IBM TotalStorage 3494 Tape Libraries, IBM Virtualization Engine TS7700, IBM racks that enable stand-alone installation, and IBM 3952 Tape Frames Model C20 (3952C20 frame) attached to a Sun 9310 library.

See 8.6, "IBM System Storage TS1130 Tape Drive" on page 246 for more information.

1.4.2 Midrange tape systems

Whether a small/medium size business is expanding operations or experiencing rapid growth, IBM midrange tape products are designed to help meet the customer needs of data backup, archive and management. Each tape product is designed to offer reliability, performance and flexibility for today and the future.

TS3200 Tape Library Express Model

The IBM TS3200 Tape Library Express Model is an entry-level tape library designed to provide reliable, high capacity, high performance tape backup. It is designed to support up to two IBM LTO Ultrium Full Height Tape Drives or up to four IBM LTO Ultrium Half Height Tape Drives.

See 9.4, “IBM System Storage TS3200 Tape Library” on page 259 for more information.

TS7650 ProtecTIER Deduplication Appliance

The IBM System Storage TS7650 ProtecTIER Deduplication Appliance is a preconfigured solution of IBM storage, IBM server, and IBM revolutionary ProtecTIER data deduplication software designed to improve backup and recovery operations, making it easy to deploy data deduplication for your information infrastructure.

See 10.2, “TS7650 ProtecTIER Deduplication Appliance” on page 279 for more information.

TS3310 Tape Library

The IBM System Storage TS3310 Tape Library is a modular, scalable tape library designed to address the tape storage needs of rapidly growing companies who find themselves space and resource constrained with tape backup and other tape applications. It is scalable from 2 to 18 drives and from 35 to 401 slots.

See 9.7, “IBM System Storage TS3500 Tape Library” on page 267 for more information.

1.4.3 Entry tape systems

IBM offers entry-level tape products that are designed to provide backup and protection of client data at an appealing cost to a company's budget.

TS2900 Tape Autoloader

The IBM System Storage TS2900 Tape Autoloader Express Model offers low profile, entry-level IBM LTO tape storage supported in open system platforms.

See 9.2, “IBM System Storage TS2900 Tape Autoloader Express” on page 255 for more information.

TS2230 Tape Drive Express

The IBM System Storage TS2230 Tape Drive Express Model offers high capacity and performance half-height Linear Tape-Open Tape Drive, with double the storage capacity of existing HH LTO2 drives in the market.

See 8.3.1, “IBM System Storage TS2230 Tape Drive Express Model” on page 236 for more information.

TS3100 Tape Library Express

The IBM TS3100 Tape Library Express Model offers large-capacity, high-performance tape backup for critical data. Designed to support one IBM LTO Ultrium Full Height Tape Drive or up to two IBM LTO Ultrium Half Height Tape Drives, to help increase capacity and performance. Designed to offer outstanding capacity, performance, and reliability for midrange and network tape storage environments in a 2U form factor with 24 data cartridge slots and a mail slot.

See 9.3, “IBM System Storage TS3100 Tape Library Express Model” on page 257 for more information.

TS2240 Tape Drive Express

The IBM TS2240 Tape Drive Express Model offers next generation, high capacity, half-high Linear Tape-Open Tape Drive. The TS2240 has a physical storage capacity of up to 1.6 TB (with 2:1 compression) in conjunction with the new IBM System Storage LTO Ultrium 800 GB data cartridge, which provides up to double the capacity of Ultrium 3 cartridges.

See 8.3.2, “IBM System Storage TS2240 Tape Drive Express Model” on page 237 for more information.

7214 Storage Device Enclosure Express

The IBM 7214 Tape and DVD Enclosure Express Model is a low-profile, entry-level IBM tape and DVD storage device enclosure supported in IBM Power Systems™ and open system platforms.

See 11.6.1, “IBM System Storage 7214 Storage Device Enclosure” on page 297 for more information.

1.5 Storage Area Networks

IBM SAN products and solutions provide integrated SMB and enterprise SAN solutions with multi-protocol local, campus, metropolitan, and global storage networking.

1.5.1 Enterprise SAN directors

The Enterprise SAN directors are designed for the highest availability and scalability enterprise solutions.

IBM System Storage SAN768B

The IBM System Storage SAN768B fabric backbone is designed to be the premier platform for consolidation of your data center connectivity, providing high-performance and high-availability data networking. It provides new levels of performance with industry-leading 8 Gbps per second (Gbps) Fibre Channel (FC) and 10 Gbps Fibre Channel over Converged Enhanced Ethernet (FCoCEE) technologies.

See 13.4.3, “IBM System Storage SAN768B” on page 338 for more information.

IBM System Storage SAN384B

The IBM System Storage SAN384B fabric backbone is designed to be the premier platform for consolidation of your data center connectivity, providing high-performance and highly available data networking. It provides new levels of performance with industry-leading 8 Gbps Fibre Channel (FC) and 10 Gbps Fibre Channel over Converged Enhanced Ethernet (FCoCEE) technologies.

See 13.4.2, “IBM System Storage SAN384B” on page 337 for more information.

IBM TotalStorage SAN256B

The IBM TotalStorage SAN256B is designed to provide outstanding performance, enhanced scalability, and a design ready for high-performance 4 Gbps, 8 Gbps, and 10 Gbps capable hardware and expanded capability features. The SAN256B is well suited to address enterprise SAN customer requirements for infrastructure simplification and improved business continuity.

See 13.4.1, “IBM TotalStorage SAN256B” on page 335 for more information.

Cisco MDS 9513

The Cisco MDS 9513 for IBM System Storage provides 12 to 528 Fibre Channel ports, with 4 and 8 Gbps support and a high-availability design. It offers 4 to 44 10 Gbps ports for ISL connectivity across metro optical networks. It includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN “islands” on a single physical fabric. The Cisco MDS 9513 provides network security features for large-enterprise SANs deployment. The director also offers intelligent networking services to help simplify mainframe FICON® and Fibre Channel SAN management and reduce total cost of ownership (TCO).

See 13.4.6, “Cisco MDS 9513 for IBM System Storage” on page 344 for more information.

Cisco MDS 9509

The Cisco MDS 9509 for IBM System Storage provides 1, 2, 4, 8, and 10 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance, and manageability required to consolidate geographically dispersed storage devices into a large enterprise SAN. The 9509 has from 12 to 336 Fibre Channel ports, which are 1, 2, 4, and 8 Gbps autosensing ports.

See 13.4.5, “Cisco MDS 9509 for IBM System Storage” on page 342 for more information.

Cisco MDS 9506

The Cisco MDS 9506 for IBM System Storage supports 1, 2, 4, 8, and 10 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance, and manageability required to consolidate geographically dispersed storage devices into a large enterprise SAN. The 9506 has from 12 to 192 Fibre Channel ports, which are 1, 2, 4, and 8 Gbps autosensing ports.

See 13.4.4, “Cisco MDS 9506 for IBM System Storage” on page 340 for more information.

1.5.2 Midrange SAN switches

IBM midrange SAN switches are scalable and affordable solutions ranging from SMB to an enterprise level.

IBM System Storage SAN80B-4

The IBM System Storage SAN80B-4 SAN fabric switch provides 48, 64, or 80 active ports, and is designed for high performance with 8 Gbps link speeds and backward compatibility to support links running at 4, 2, and 1 Gbps link speeds. High availability features make it suitable for use as a core switch in midrange environments or as an edge-switch in enterprise environments.

See 13.3.2, “IBM System Storage SAN80B-4” on page 332 for more information.

IBM System Storage SAN40B-4

The IBM System Storage SAN40B-4 is a simple-to-use midrange and enterprise SAN fabric switch for IBM POWER® Systems, System x, System z, and other server environments. They offer high port density design with up to 40 ports in an efficient, space saving 1U format, and non-disruptive capacity activation from 24 to 32 to 40 ports with Ports on Demand scalability.

See 13.3.1, “IBM System Storage SAN40B-4” on page 331 for more information.

Cisco MDS 9134

The Cisco MDS 9134 for IBM System Storage is designed to address the needs of medium-sized businesses and large enterprises with a wide range of Storage Area Network (SAN) capabilities. It can be used as part of a high performance simple SAN with single-switch or stacked switch configurations for business-class customers in support of IT simplification and business continuity solutions. It can also be used as an edge switch for device aggregation with 10 Gbps core director configurations for large enterprise customers.

See 13.3.3, “Cisco MDS 9134 for IBM System Storage” on page 333 for more information.

1.5.3 Entry SAN switches

Scalable, flexible storage infrastructure IBM SAN products and solutions provide integrated, simple, affordable SMB solutions with multi-protocol local, campus, metropolitan and global storage networking.

IBM System Storage SAN24B-4 Express

The IBM System Storage SAN24B-4 Express fabric switch is designed specifically to address the needs of small to medium-sized SAN environments. It can be used to create a wide range of high-performance SAN solutions, from simple, single-switch configurations to larger, multi-switch configurations which support fabric connectivity and advanced business continuity capabilities.

See 13.2.1, “IBM System Storage SAN24B-4 switch” on page 329 for more information.

Cisco MDS 9124 Express

The Cisco MDS 9124 Express for IBM System Storage is designed to address the needs of small and medium-sized businesses with a wide range of SAN capabilities. It can be used as part of SAN solutions from simple single-switch configurations to larger multi-switch configurations in support of fabric connectivity and advanced business continuity capabilities.

See 13.2.2, “Cisco MDS 9124 Express for IBM System Storage” on page 330 for more information.

1.5.4 Multiprotocol routers

Multiprotocol routers are designed to connect two Storage Area Networks (SANs) over a wide distance using the Internet as the interconnection fabric (upgrades to enterprise-level functions are available). They are intended to support business continuity solutions between supported servers at one site and support servers or IBM System Storage disk or tape devices at a distant location.

IBM System Storage SAN04B-R

The IBM System Storage SAN04B-R is designed for high performance with 4 Gigabit per second (Gbps) Fibre Channel (FC) ports and hardware-assisted traffic processing for line-rate performance across Ethernet Internet Protocol (IP) ports. It utilizes existing Internet, IP-based Metropolitan Area Network (MAN) or Wide Area Network (WAN) infrastructures for metro and global SAN extension for business continuity solutions.

See 13.5.1, “IBM System Storage SAN04B-R multiprotocol router” on page 347 for more information.

IBM System Storage SAN06B-R

The IBM System Storage SAN06B-R is designed for high performance with 8 Gigabit per second (Gbps) Fibre Channel (FC) ports and hardware-assisted traffic processing for line-rate performance across 1 GbE Ethernet Internet Protocol (IP) ports. It utilizes existing Internet, IP-based Metropolitan Area Network (MAN) or Wide Area Network (WAN) infrastructures for metro and global SAN extension for business continuity solutions.

See 13.5.2, “IBM System Storage SAN06B-R multiprotocol router” on page 349 for more information.

Cisco MDS 9222i

The Cisco MDS 9222i for IBM System Storage is designed to address the needs of medium-sized businesses and large enterprises with a wide range of Storage Area Network (SAN) capabilities. It can be used as a cost-effective high performance SAN extension over IP router switch for midrange SMB customers in support of IT simplification and business continuity solutions. It can also provide remote site device aggregation and SAN extension connectivity to large customer data center directors.

See 13.5.3, “Cisco MDS 9222i for IBM System Storage” on page 349 for more information.

1.5.5 New Ethernet switches

Expanded network technology choices with the new IBM Ethernet switches and routers provide an integrated end-to-end resiliency and security framework. The big product range from IBM makes data center networking a cost efficient solution.

IBM b-type Ethernet switches and routers

IBM offers a versatile switch that supports both Fibre Channel (FC) and Ethernet LAN data in one fabric using FCoE to help data centers simplify their growing infrastructures.

IBM Converged Switch B32

The IBM Converged Switch B32 offers outstanding performance with eight Fibre Channel (FC) ports concurrently active at 8 Gbps and 24 Converged Enhanced Ethernet (CEE) ports concurrently active at 10 Gbps link speeds. It has a high density design with 32 ports in a 1 U enclosure.

See 14.1.1, “IBM Converged Switch B32” on page 356 for more information.

IBM Ethernet Switch B24X

The IBM Ethernet Switch B24X is a compact, high-performance, highly-available, and high-density 10 Gigabit Ethernet (GbE) solution designed for mission critical data centers, large enterprises, and High-Performance Computer (HPC) requirements.

See 14.1.2, “IBM Ethernet Switch B24X” on page 357 for more information.

IBM r-series Ethernet switches

The IBM r-series Ethernet switches are available in four chassis models, and allow network designers to standardize on a single product family for end-of-row, aggregation, and backbone switching.

See 14.1.3, “IBM r-series Ethernet Switches” on page 358 for more information.

IBM c-series Ethernet switches

The IBM c-series Ethernet switches are powerful enablers of advanced converged enterprise backbones. Featuring state-of-the-art Quality of Service and wire-speed unicast/multicast routing, they enable the efficient rollout of converged backbones, providing reliable transport of Voice over IP (VoIP), video services, and mission critical data.

See 14.1.4, “IBM c-series Ethernet Switches” on page 359 for more information.

IBM g-series Ethernet switches

The IBM g-series Ethernet access switches provide enterprise organizations with a flexible and feature-rich solution for building a secure and converged network edge. Upgradeable with 10-Gigabit Ethernet, PoE, and IronStack stacking technology, the switches provide enterprises with the cost and operational benefits of a “pay-as-you-grow” architecture.

See 14.1.5, “IBM g-series Ethernet Switches” on page 360 for more information.

IBM m-series Ethernet/IP routers

The IBM m-series Ethernet/IP routers are high performance multi-service IP/MPLS routers that help virtualize data center core routing and are available in four high density chassis. They feature state-of-the-art QoS and wire-speed unicast/multicast routing for IPv4, IPv6, and MPLS services.

See 14.1.6, “IBM m-series Ethernet/IP routers” on page 360 for more information.

IBM Ethernet Switch B08S and IBM Ethernet Switch B16S

The IBM s-series Ethernet Switch modular PoE capable switches provide an industry-leading scalable, low-latency, fault-tolerant infrastructure for cost-effective deployment of Voice over IP (VoIP), wireless, and high-capacity data services throughout the enterprise.

See 14.1.7, “IBM s Series Ethernet Switch” on page 362 for more information.

IBM j-type Ethernet routers and switches

The IBM j-type Ethernet switches and routers are designed to deliver the performance, scalability, and high availability required for today’s high-density data center and cloud computing environments. The IBM j-type family includes the following switches and router.

IBM Ethernet Switch J08E/J16E

The IBM Ethernet Switch J08E/J16E are high-performance 8-slot (J08E) and 16-slot (J16E) switches that support data center as well as campus LAN core and aggregation layer deployments.

See 14.2.1, “IBM Ethernet Switch J08E and IBM Ethernet Switch J16E” on page 363 for more information.

IBM Ethernet Switch J48E

The IBM Ethernet Switch J48E was designed for high-performance server access deployments, running Juniper Networks’ JUNOS Software operating system. A single switch can be deployed initially; as requirements grow, Virtual Chassis technology allows up to nine additional switches to be interconnected over a 128 gigabit-per-second (Gbps) backplane and managed as a single device.

See 14.2.2, “IBM Ethernet Switch J48E” on page 365 for more information.

IBM Ethernet Router J02M/J06M/J011M

The IBM Ethernet Routers J02M, J06M, and J11M deliver high-port density as well as performance of up to 960 Gbps throughput, scalability and reliability in a space-efficient package. The routers offer fully redundant hardware that includes a redundant Switch Control Board (SCB) and Routing Engines (REs) plus fan trays and power supplies designed to increase system availability.

See 14.2.3, “IBM Ethernet Router J02M/J06M/J11M” on page 366 for more information.

Cisco Nexus 5000 family

The Cisco Nexus 5010 for IBM System Storage and Cisco Nexus 5020 for IBM System Storage FCoE switches help reduce costs through data center infrastructure simplification. A unified fabric over 10 Gigabit Ethernet (GbE) for server LAN and SAN traffic enables consolidation of server adapters, cables, and top-of-rack (TOR) switches by up to 50 percent.

Cisco Nexus 5020 provides up to 56 ports: 40 fixed 10 GbE and FCoE ports and 16 optional Fibre Channel, FC ports

Cisco Nexus 5010 provides up to 28 ports: 20 fixed 10 GbE and FCoE ports and 8 optional FC ports.

See 14.3.2, “Cisco Nexus 5010/5020” on page 367 for more information.

1.6 IBM NAS systems

IBM Network attached storage (NAS) products provide a wide-range of network attachment capabilities to a broad range of host and client systems.

1.6.1 IBM N series hardware

IBM N series unified system storage solutions can provide customers with the latest technology to help them improve performance, virtualization manageability, and system efficiency at a reduced total cost of ownership.

IBM System Storage N3000 Express

The IBM System Storage N3000 Express series Modular Disk Storage Systems. Advanced system design enables outstanding, cost-effective deployment versatility.

The N3300 Express squeezes 12 TB of internal raw capacity into a 2U enclosure and optional external expansion that can increase total system raw capacity to 68 TB. The N3600 Express scales up to 20 TB of internal raw capacity and can scale up to 104 TB by supporting up to 104 disk drives.

See 5.2.4, “IBM System Storage N3000 introduction” on page 179 for more information.

IBM System Storage N series Gateway

The IBM System Storage N series Gateway for SAN Storage Environments allows you to unify NAS, FC SAN, and IP SAN under a common architecture.

This gateway is designed to enable organizations to consolidate UNIX, Linux, Windows, and web workloads with existing SAN storage, thereby helping to increase storage utilization.

See 5.3, “IBM System Storage N series Gateways” on page 187 for more information.

IBM System Storage N6000

The IBM System Storage N6000 series systems can be used to simultaneously meet your diverse needs for SAN and NAS, and for primary and secondary storage. You get outstanding value as well: These flexible systems offer excellent performance and impressive scalability at a low total cost of ownership.

This highly available and scalable storage solution includes the following features:

- ▶ Up to 840 TB raw storage capacity
- ▶ 4 GB to 32 GB cache memory
- ▶ 512 MB to 4 GB nonvolatile memory

See 5.2.5, “IBM System Storage N6000 introduction” on page 183 for more information.

IBM System Storage N7000

The IBM System Storage N7000 series systems are intended to help IT organizations tackle the challenge of effective data management using virtualization technology and a unified storage architecture. The N7000 series is designed as a reliable, powerful, scalable modular storage system, able to deliver high-end enterprise storage and data management value with midrange affordability. Built-in enterprise serviceability and manageability features help support your efforts to increase reliability, simplify, and unify storage infrastructure and maintenance, and deliver exceptional economy. It is designed for non-disruptive expansion to more than 1.1 petabytes (1.1 PB, or 1100 TB) storage capacity.

See 5.2.6, “IBM System Storage N7000 introduction” on page 185 for more information.

1.6.2 IBM N series software

The IBM System Storage N series also provides a selection of features and functions designed to provide a comprehensive set of robust management and operational tools as well as high availability features, disaster recovery, and data copy services that help the system administration provide a high level of support for environments requiring IP attached storage solutions.

See 5.6, “Optional software” on page 196 for more information.

SnapManager for Hyper-V

SnapManager® for Microsoft Hyper-V is an extensive data management technology designed to simplify storage management, backup, restore, and disaster recovery processes in a Microsoft Hyper-V environment.

SnapManager for SAP

The foundations of many of today’s leading companies rest on SAP. IBM System Storage N series SnapManager for SAP integrates closely with SAP BRTools to help automate and streamline SAP data management and provide fast, space-efficient, disk-based backup, rapid restore and recovery, and flexible cloning.

SnapManager for Virtual Infrastructure

SnapManager for Virtual Infrastructure is a comprehensive data management solution intended to streamline storage management in a VMware ESX environment. It can also help simplify backup, restore, and disaster recovery operations for virtual machines.

SnapManager for Oracle

IBM System Storage N series with SnapManager for Oracle is a comprehensive data management solution designed to help you streamline storage management while simplifying configuration, backup, and restore operations for Oracle databases. Deployed in combination with N series storage systems, SnapManager supports tight integration with Oracle Database to automate critical tasks such as backup, restore, database recovery, and cloning. SnapManager management of the underlying data layout simplifies routine data management tasks, allowing administrators to focus more effort on value-added tasks. Environments using real application clusters (RAC) and automatic storage management (ASM) are fully supported.

SnapManager for Microsoft Office SharePoint Server 2007

SnapManager provides high availability and data protection for Microsoft Office SharePoint Server data. As Microsoft Office SharePoint Server use increases, so has the number of discrete SharePoint instances on each physical server. This can lead to rapidly growing, fragmented deployments. The complexity of large Microsoft Office SharePoint Server environments can be further exasperated by the use of direct-attached storage subsystems. Managing numerous servers and the related data usually requires a lot of time and coordination across various departments within an organization.

SnapManager for Microsoft SQL server

IBM System Storage with SnapManager software is designed to help you achieve significant savings from a time and resource perspective by streamlining database storage while simplifying configuration, backup, and restore operations for SQL Server databases. SnapManager supports rapid SQL Server backup times, from hours to as little as seconds, and is designed to make each backup a complete and consistent copy of the original.

Backups are based on Snapshot™ copies, which are designed to require minimal disk space for each additional full backup. This capability can help you achieve availability, scalability, and reliability for SQL Server environments.

SnapManager for Microsoft Exchange

SnapManager software for Microsoft Exchange is a comprehensive data management solution designed to help you streamline storage management while simplifying configuration, backup, and restore operations for Exchange databases. Deployed in combination with N series storage systems, SnapManager software is designed to provide near-instantaneous hot backups and rapid restores to deliver a robust combination of availability, scalability, performance, and reliability for Exchange environments.

SnapLock

The SnapLock® function is designed to deliver high performance and high-security data function to disk-based nearline and primary IBM System Storage N series storage. The SnapLock function can help manage the permanence, accuracy, integrity, and security of data by storing business records in an inalterable form and allowing for their rapid online accessibility for long periods of time.

Protection Manager

Offering intuitive, elegant interface for visibility into, and control of, data protection processes, Protection Manager is designed to provide data center professionals with an innovative software solution to manage the enterprise's disk-based data protection environment. The application is designed to allow administrators to apply consistent data protection policies across the enterprise, automate complex data protection processes, and pool backup and replication resources to help improve utilization.

Operations Manager

With comprehensive monitoring and management for your N series environment, IBM System Storage N series with Operations Manager software offers comprehensive monitoring and management for N series enterprise storage and content delivery environments. Operations Manager is designed to provide alerts, reports, and configuration tools from a central control point, helping you keep your storage and content delivery infrastructure in-line with business requirements for high availability and low total cost of ownership. Operations Manager can effectively support your organization's efforts to rapidly deploy, provision and manage a comprehensive enterprise storage and content delivery network.

SnapMirror

As a global data availability and disaster recovery software solution, IBM System Storage N series systems with SnapMirror® technology can help you implement the disaster recovery and data distribution solution that your enterprise needs. By replicating data at high speeds over a LAN or a WAN, SnapMirror technology can help support high data availability and quick recovery for mission-critical applications.

SnapMirror technology is designed to mirror data to one or more N series filers. It supports constant updates to the mirrored data to keep it current and available for disaster recovery, off loading tape backup, and read-only data distribution, testing on non-production filers, online data migration, and more. If your enterprise is geographically dispersed and all locations need access to the same data set (such as training videos, CAD tools, and so on), SnapMirror can distribute the same data to all locations. By automatically updating this data and allowing local access to mirrored data, SnapMirror can help improve employee productivity and efficiency.

Note: For more comprehensive information about the full range of N series software products, see the website:

<http://www.ibm.com/systems/storage/network/software>

1.7 Storage software

Here we describe IBM Storage software and Storage technologies for a Smarter Planet.

1.7.1 IBM SAN Volume Controller

The IBM SAN Volume Controller (SVC) is designed to help simplify storage infrastructure, optimize storage utilization, and enable business to adapt quickly and dynamically to variable environments.

SVC is designed to combine storage capacity from multiple disk systems into a reservoir of capacity that can be managed more efficiently. It is designed to help increase storage utilization, thus reducing costs, by providing host applications with more flexible access to capacity. Designed also to help improve storage administrator productivity by automating provisioning and enabling management of heterogeneous storage systems using a simple common interface.

SVC supports improved application availability by practically eliminating storage-related causes of application downtime. SVC also enables a tiered storage environment in which the cost of storage can be better matched to the value of data. These are just a sample from the range of features and benefits provided by SVC.

1.7.2 IBM SVC Entry Edition

Small and medium businesses looking for the benefits of storage virtualization can consider the IBM SVC Entry Edition, which can simplify management of storage while improving storage utilization. It can grow with your business by converting easily to the full SVC offering.

See Chapter 16., “IBM System Storage SAN Volume Controller” for more information.

1.7.3 IBM Tivoli Storage Productivity Center

IBM Tivoli Storage Productivity Center (TPC) is an integrated suite that includes a single user interface to manage capacity utilization of storage systems, file systems, and databases, and to automate file system capacity provisioning in both physical and virtual environments. IBM Tivoli Storage Productivity Center components can be ordered as a bundled suite or ordered separately based on specific needs.

The individual components include:

- ▶ IBM Tivoli Storage Productivity Center Basic Edition
- ▶ IBM Tivoli Storage Productivity Center for Data
- ▶ IBM Tivoli Storage Productivity Center for Disk
- ▶ IBM Tivoli Storage Productivity Center for Replication.

The integrated packages include:

- ▶ IBM Tivoli Storage Productivity Center Standard Edition
- ▶ IBM Tivoli Storage Productivity Center with Advanced Provisioning
- ▶ IBM System Storage Productivity Center

See Chapter 18., “IBM Tivoli Storage Productivity Center” for more information.

1.7.4 IBM Tivoli Storage Manager

IBM Tivoli Storage Manager is the number one product of choice for an efficient and effective enterprise wide storage solution. It provides a data protection solution for backup, archiving, disaster recovery planning, space management, database and application protection, bare machine recovery, and record retention. More than 44 operating platforms are supported, using a consistent graphical user interface.

Tivoli Storage Manager provides:

- ▶ Centralized administration for data and storage management
- ▶ Fully automated data protection
- ▶ Efficient management of information growth
- ▶ High-speed automated server recovery
- ▶ Full compatibility with hundreds of storage devices, and local area network (LAN), wide area network (WAN), and storage area network (SAN) infrastructures
- ▶ Optional specifically designed backup solutions for major groupware, enterprise resource planning (ERP) applications, and database products.

Tivoli Storage Manager is the premier storage management solution for mixed platform environments.

See Chapter 17., “IBM Tivoli Storage Manager” for more information.

1.7.5 IBM General Parallel File System

IBM General Parallel File System (GPFS) is a scalable, parallel cluster file system. In addition to high-speed parallel file access, GPFS provides fault tolerance, including automatic recovery from disk and node failures.

IBM General Parallel File System (GPFS) currently powers many of the world's largest scientific supercomputers and commercial applications requiring high-speed access to large volumes of data.

See Chapter 20., “IBM GPFS and SoFS” for more information.

1.7.6 IBM Multilevel Grid Access Manager Software

IBM System Storage Multilevel Grid Access Manager Software (Grid Access Manager Software) is built on an open, high-performance grid architecture that delivers data protection, information life-cycle management, simplified storage management and multi-site data access based on open standards. The potential benefits derived from these features can help deliver important cost savings and operational efficiencies:

- ▶ Simplified management and improved storage utilization, with excellent performance
- ▶ Data protection and improved business continuity
- ▶ Support for global access, multi-site operation

See Chapter 19., “IBM Grid Access Manager” for more information.

1.7.7 z/OS storage management tools

The z/OS storage management tools are a collection of products designed to provide a comprehensive menu of solutions in order to meet the storage management requirements of z/OS customers. The tools are positioned to complement and extend the storage management capabilities of DFSMS.

These products allow IBM to provide a more complete suite of storage management solutions:

- ▶ IBM Tivoli Omegamon XE
- ▶ IBM DB2® Cloning Tool
- ▶ IBM Tivoli Advanced Catalog Management for z/OS
- ▶ IBM Tivoli Advanced Backup and Recovery
- ▶ IBM Tivoli Advanced Audit for DFSMSHsm
- ▶ IBM Tivoli Automated Tape Allocation Manager
- ▶ IBM Tivoli Storage Optimizer for z/OS
- ▶ IBM Tivoli Tape Optimizer for z/OS
- ▶ DFSMS Optimizer

See Chapter 22., “z/OS Storage Management Tools” for more information.

1.7.8 IBM System Storage Productivity Center

The IBM System Storage Productivity Center (SSPC) is an integrated hardware and software solution that can help you improve and centralize the management of your storage environment through the integration of products. It provides a single point from which to manage your storage systems.

See 18.10, “IBM System Storage Productivity Center” for more information.

Note: Only IBM can provide a complete portfolio of solutions and services, innovative servers, storage, and systems software that facilitate combined business and IT.

To learn more about the IBM Dynamic Infrastructure and Smarter Planet strategies, see the following websites:

<http://www-03.ibm.com/systems/dynamicinfrastructure>

http://www-03.ibm.com/systems/dynamicinfrastructure/smarter_planet/index.html



Part 1

Disk systems

In Part 1 we describe disk products that either can be directly attached to a server by Direct Attached Storage (DAS), used in a Storage Area Network (SAN) infrastructure, or used with Network Attached Storage (NAS).

The IBM System Storage disk products portfolio covers the needs of a wide spectrum of possible implementations, from entry-level to large enterprise.

The IBM System Storage DS family combines the high-performance of the IBM System Storage DS6000/DS8000 series and XIV enterprise servers with the IBM System Storage DS5000 series of midrange systems, and with DS3000 series low priced entry systems.

The family is further complemented by a range of expansion enclosures to expand the disk storage capacities of individual systems into hundreds of terabytes (TB), or even to a petabyte (PB).

Furthermore, a full range of IBM System Storage capabilities such as advanced copy services, management tools, and virtualization services are available to help protect data.

Types of storage

Figure 1 shows the extensive range of DS disk storage systems that we discuss in more detail in subsequent chapters.

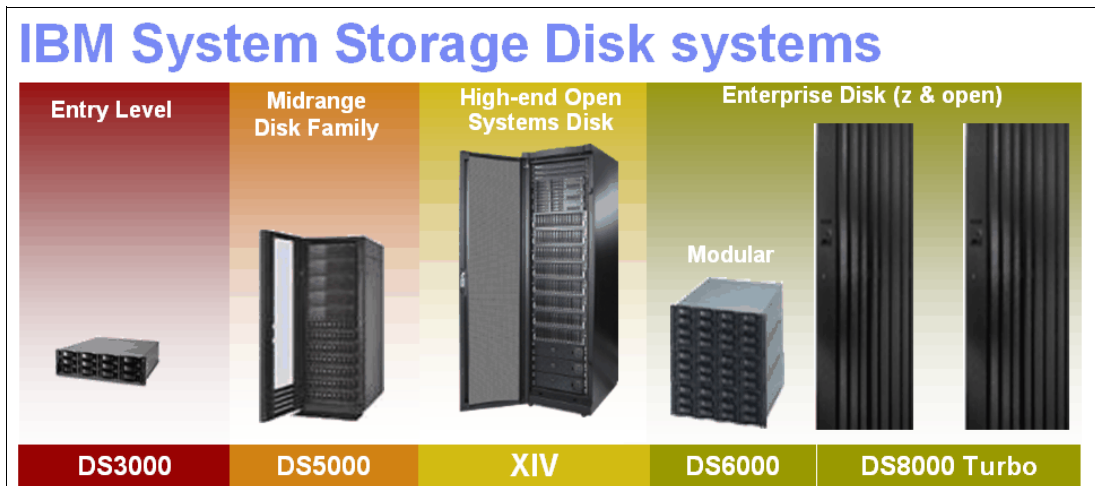


Figure 1 IBM System Storage Disk systems

Enterprise storage

Under Enterprise disk storage, you find disk systems with the following characteristics:

- ▶ Excellent external storage solution for medium to large scale deployments
- ▶ Medium and large data centers
- ▶ Multi-workload mission critical environments
- ▶ Long distance disaster replication
- ▶ High performance, highly scalable, high function storage systems
- ▶ Designed to avoid any single points of failure for high availability; highly reliable architecture with hot-swap, redundant components
- ▶ Designed for nondisruptive upgrades in capacity, performance, and cache; long-term, highly expandable storage solutions to protect investment
- ▶ Flexible management including SMI-S compatible API
- ▶ Advanced business continuance solutions
- ▶ Excellent price/performance/value

Mid-level storage

IBM System Storage mid-level storage disks systems are characterized as follows:

- ▶ Ideal external storage solution for smaller deployments
- ▶ Often used in remote branch office locations
- ▶ Storage systems designed to offer low cost, scalability, high performance

In particular, the DS5000 series incorporates these characteristics:

- ▶ High reliability with hot-swap redundant power supplies, controllers and drives
- ▶ Fully redundant pathing from host to drive
- ▶ Current models are prepared for new iSCSI Host interfaces developments (10Gb iSCSI)
- ▶ High performance dual ported 4Gbps FC drives as well as SATA drives
- ▶ Easy to use, common DS Storage Manager

Near-line storage and reference data

For near-line storage and reference data, the IBM System Storage disk products in that category offer the following features:

- ▶ Excellent external storage solution for small to large scale deployments
- ▶ Data Protection with local snapshot and remote mirrors to help support rapid recovery and zero-time backup windows
- ▶ Fixed Content Support for those unchanging files and blocks that have light to moderate access profiles
- ▶ A design with no single point of failure for high availability, which offers high reliability with hot-swappable redundant components
- ▶ Designed to be nondisruptive for upgrades in capacity and performance, helping to protect investment
- ▶ Flexible management, including SMI-S compatible FSM API

Entry-level storage

IBM System Storage entry-level disk systems offer the following benefits:

- ▶ Excellent external storage solution for smaller deployments:
 - Smaller data center
 - Remote branch office locations
- ▶ Storage subsystems designed to provide low cost, scalability, high performance
- ▶ A design that offers high reliability with hot-swap redundant power supplies, controllers, and drives
- ▶ Easy to use Storage Manager helps remove complexity from SAN deployment and management

Network attached storage

IBM N series (NAS) products provide a wide-range of network attachment capabilities to a broad range of host and client systems. The range includes NAS gateways, providing heterogeneous access to Fibre Channel attached storage arrays, and NAS expandable storage systems, which come with Fibre Channel disks.

IBM System Storage point-in-time copy comparison

Table 1-1 provides a quick reference to the point-in-time copy (FlashCopy®) functions available on the IBM System Storage SAN Volume Controller, DS5000, and DS6000-DS8000. The meaning of each function is stated in the definitions after the table.

Table 1-1 Point-in-time copy product function comparison

Function	SVC	DS5000	DS6000	DS8000
File Level	N	N	N	N
Block Level	Y	Y	Y	Y
Physical Copy	Y	Y	Y	Y
Logical Copy (NoCopy)	Y	Y	Y	Y

Function	SVC	DS5000	DS6000	DS8000
Transition NoCopy into Copy	Y	N	Y	Y
Logical Copy (space efficient)	Y	N	N	Y
Persistent Copies	Y	Y	Y	Y
Multiple Concurrent Copies	Y	Y	Y	Y
Source /Target Read/ Write capable	Y	Y	Y	Y
Incremental Copies	Y	Y	Y	Y
Consistency Groups	Y	Y	Y	Y
Across heterogeneous disk vendors	Y	N	N	N
Remote Pair FlashCopy	N	N	N	Y
Cascaded FlashCopy	Y	N	N	N

Meaning of indicators

Each box has one of the following indicators:

- **Y** – function is currently generally available
- **N** – function is not available

The table is intended as a guideline only; consult the appropriate product chapters and product documentation for details of the individual product implementations.

As product enhancements are continually made, see your IBM representative for the latest information.

FlashCopy function definitions

File Level: Is the point-in-time copy performed at the file level?

Block / LUN Level: Is the point-in-time copy performed at the block / LUN / volume level?

Physical Copy: Can a physical copy be made of the entire source to the target?

Logical Copy (no background copy): Does the capability exist to not copy the complete source LUN / volume?

Transition Nocopy → Copy: Does the capability exist to change no background copy to a full copy without doing a new point-in-time copy operation?

Logical Copy (space efficient): Does the capability exist to utilize less target space than occupied by the source?

Persistent Copies: Does the source / target relationship remain intact until explicitly withdrawn?

Multiple Concurrent Copies: Can multiple targets be made from one source at the same point in time?

Source / Target Read / Write Capable: Are both the source and target fully read / write capable?

Incremental Copies: Are changes to the source tracked so only changed data is copied to the target?

Consistency Groups: Can associated volumes be treated as one or more groupings and have the point-in-time operation performed across the grouping at the same point in time?

Heterogeneous disk subsystems: Can the point-in-time copy have source and target on various vendor disk subsystems?

Remote Pair FlashCopy: Can this function allow a significant reduction of the recoverability time when a FlashCopy target device is also a Metro Mirror primary device?

Cascaded FlashCopy: Can a VDisk be a source for one FlashCopy mapping and the target for another FlashCopy mapping?

IBM System Storage disk mirroring comparison

Table 1-2 provides a quick reference to the disk mirroring (Metro Mirror, Global Copy, Global Mirror) functions available on the IBM System Storage SAN Volume Controller, DS5000, and DS6000-DS8000. The meaning of each function is stated in the definitions after the table.

Table 1-2 Disk mirroring product function comparison

Function	SVC	DS5000	DS6000	DS8000
File Level	N	N	N	N
Block Level	Y	Y	Y	Y
Metro Mirror	Y	Y	Y	Y
Distance support Sync	300Km	300Km	300Km	300Km
Distance support Async	NA	Unlimited	Unlimited	Unlimited
Global Copy	N	Y	Y	Y
Global Mirror	Y	Y	Y	Y
Asynchronous Cascading 3 site	Y	N	Y	Y
Failover/Failback support	Y	N	Y	Y
Consistency groups	Y	N	Y	Y
Freeze / Thaw support	Y	N	Y	Y
Suspend / Resume	Y	Y	Y	Y
Source/Target can be FC source	Y	N	Y	Y
Source/Target same cluster	N	N	Y	Y
Heterogeneous disk Source/Target	Y	N	N	N

Meaning of indicators

Each box has one of the following indicators:

- **Y** – function is currently generally available
- **N** – function is not available

The table is intended as a guideline only; consult the appropriate product chapters and product documentation for details of the individual product implementations.

As product enhancements are continually made, see your IBM representative for the latest information.

Disk mirroring function definitions

File Level: Is the point-in-time copy performed at the file level?

Block / LUN Level: Is the point-in-time copy performed at the block / LUN level / volume level?

Metro Mirror synchronous copy - Host is notified of write complete after source and target I/O have completed.

Distance Support:

- ▶ SVC – 10 km Fibre Channel - channel extender IP/DWDM based solutions for longer distances.
- ▶ DS6000 / DS8000 – 103 km over ESCON®, 300 km over Fibre Channel or IP/DWDM based network; unlimited distance supported for asynchronous mirroring.
- ▶ DS5000 – 10 km over FCP; channel extender IP/DWDM based solutions for longer distances.

Global Copy asynchronous copy: Host is notified of write complete after I/O to source is completed. I/O to target is performed later; data broadcast where data integrity not insured.

Global Mirror asynchronous copy with consistency: Target I/O is applied in with data integrity; data is suitable for database restart at remote site.

Failover / failback support: Indicates capability to quickly re-establish target → source relationship in event of a primary site failure, copying only incremental changes back.

Consistency groups: Can associated volumes be treated as one or more groupings such that disk mirroring error triggers cause suspension and keep data integrity across an entire grouping of volumes?

Freeze / thaw (run) support: This function is linked to consistency groups. If an error occurs, I/O stops (freeze) to the consistency group, and an error is posted to the host for automation to act upon, and then host I/O resumes (thaw / run)?

Suspend / resume: Can the remote copy environment be suspended in the event of a planned or unplanned outage and then resumed? This means that only changes while the remote copy environment is suspended are copied to the targets versus entire remote copy environment must be copied in entirety.

Source / target can be FlashCopy source: Can the source or target of a disk mirror be the Source for FlashCopy?

Source / Target same cluster: Can the source and target of a remote copy pair reside in the same cluster / subsystem? This capability is primarily useful for testing purposes.

Heterogeneous disk source / target: Can the source and target of a remote copy pair reside on various vendor disk subsystems?



Entry-level disk storage

In this chapter we describe the IBM System Storage DS3000 family of external storage enclosures. The models are entry-level, cost-effective workgroup storage servers, and are flexible for use with the IBM System x, IBM system p, and IBM BladeCenter® servers, but also for selected third party products.

The main difference between the models is the way they are connected to the host. Model DS3200 is designed to use SAS host connectivity. Model DS3300 has iSCSI to connect over an existing Ethernet infrastructure, whereas the DS3400 is capable of using Fibre Channel connectors. This connection information is shown in Figure 2-1.

All models can make use of the EXP3000 Disk expansion cabinet, which allows them to grow in steps to a maximum of 48 Disks. Disk types are intermixed within the controller and in the Expansion cabinets.

With a wide range of Advanced functions such as Partitioning, FlashCopy, and VolumeCopy, this entry level disk solution is highly suitable for small to midsize companies.

Furthermore, the DS3000 series systems are easy to deploy and manage through the IBM System Storage DS Storage Manager (also known as IBM DS Storage Manager).

2.1 Overview of DS3000 series models

The IBM System Storage DS3000 series models are entry-level, low cost workgroup storage subsystems for IBM System x, IBM System p, IBM BladeCenter servers, and selected other parties servers. They come in either single or dual-controller configurations, and even Telco prepared versions. The DS3000 series offer a solution for workgroup storage applications, such as file, print, and web serving, as well as remote boot storage for diskless servers. The DS3200, DS3300, and DS3400 models are shown in Figure 2-1.



Figure 2-1 DS3000 Series

Model DS3200 will have direct 3 Gb SAS connectivity or be part of a BladeCenter SAS Storage Area Network through the BladeCenter connectivity. For even greater flexibility, the DS3200 allows simultaneous support of a range of operating systems for System x and BladeCenter servers, in either single or dual-controller configurations.

Utilizing standard Ethernet infrastructure by 1 Gb iSCSI connection to an IP switch, and low cost highly reliable SAS and SATA drives, the model DS3300 offers a simple, affordable storage solution that can be used in direct attached or network-attached environments for shared storage as shown in Figure 2-2.

The DS3400 offers two 4 Gbps FC host fabric ports and can therefore be connected directly to two servers for availability purposes, or to Fibre Channel switches, as shown in Figure 2-2.

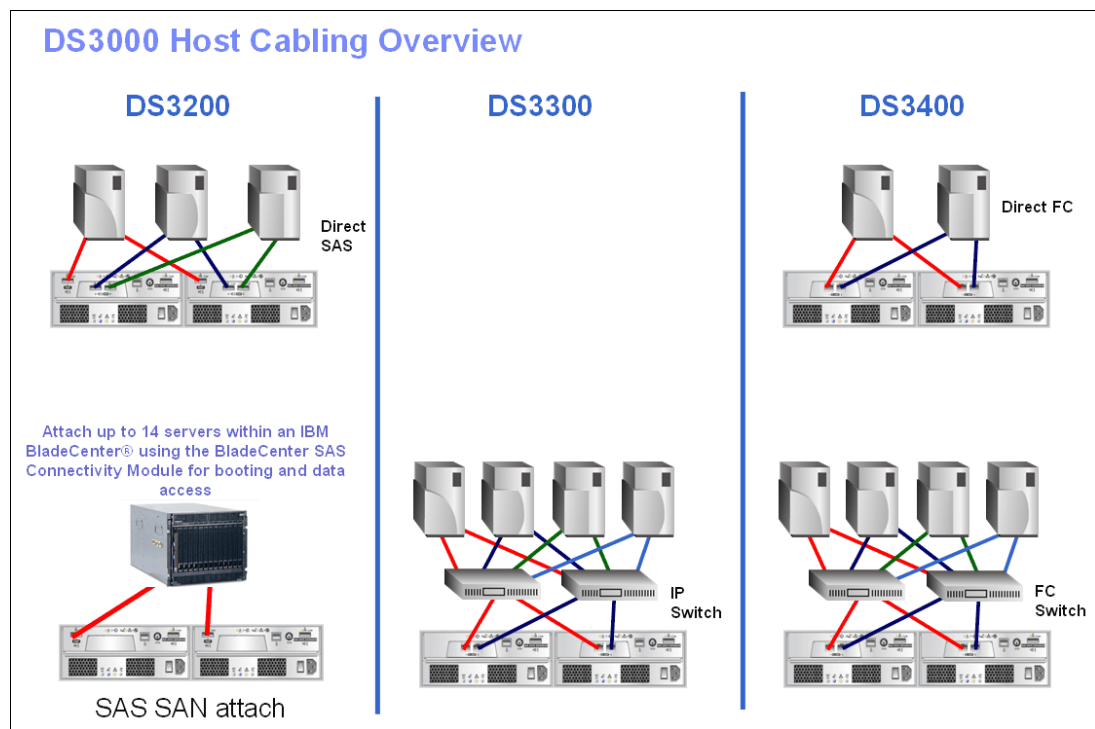


Figure 2-2 Host Connection options for DS3000 series

The modular and scalable design of the DS3000 Series can easily be upgraded to meet the growing needs of storage. Disk enclosures will enable flexible growth up to higher Disk space requirement regions.

Availability of data is secured by common high availability features:

- ▶ Dual active RAID controllers
- ▶ Redundant hot-swappable components (controllers, drives, cables, power/cooling)
- ▶ Automated IO path failover
- ▶ Online administration, expansion, configuration, RAID level migration, firmware upgrades¹
- ▶ Unlimited Global Hot-Spares
- ▶ Battery backed cache with cache mirroring

Additional advanced (Premium) software features are available on the DS3000 series:

- ▶ Storage partitions
- ▶ FlashCopy functionality
- ▶ VolumeCopy

Out of the box, the DS3000 series models include licenses for 4 Partitions, 2 FlashCopy images, an attachment license for 1-3 Expansion units and Windows and Linux on Intel® Host-attach license.

All DS3000 series models have a special Telco DC Power model that supports -48v DC power supplies and are NEBS level 3 certified.

¹ An exception to this is the DS3000 Storage Manager V10.35 with Firmware 7.35 upgrade, which is offline.

2.2 General specifications

In this section we take a closer look to the IBM System Storage models DS3200, DS3300, and DS3400.

2.2.1 IBM System Storage model DS3200

The model DS3200 is ideal for an internal DAS user to get started with external storage. Compared to internal storage, the DS3200 can increase performance, enhance functionality, increase data protection, and is easy to set up and administer. It is well suited for part-time administrators. You can start small but still be prepared for future growth. The modular design of the DS3200 can easily be upgraded to support increasing volumes of data, in a flexible and affordable way.

The DS3200 can also be enabled to share storage across three application servers, making it a very low cost shared DAS solution, as well as making sure that you get all the TCO benefits of consolidation, single management, increased efficiency and High Availability. As the model DS3200 is also supported by an IBM BladeCenter, it can support up to 14 blade servers. At the same time, it can serve as a boot device for these blade servers, making it a very cost effective total solution.

The DS3200 storage subsystem 1726-21x and 1726-22x (dual controller) consists of the following components (Figure 2-3):

- ▶ 2U rack-mountable enclosure with 12 easily accessible drive bays
- ▶ One or three 3 Gbps SAS host ports per controller
- ▶ Support for dual-ported, and hot-swappable SAS (15k rpm) and SATA (7.2k rpm) disks
- ▶ Scalability of up to 12 TB internal physical storage capacity using 1 TB hot-swappable SAS disks
- ▶ Expandability of up to 48 TB physical storage capacity with attachment of three EXP3000 expansion units
- ▶ RAID 0, 1, 3, 5, 6 and 10 support
- ▶ LED indicators (8)
- ▶ Minimum 512 MB RAID cache memory
- ▶ Maximum 2 GB system memory, battery backup
- ▶ Built-in reliability features with dual-redundant power-supplies standard

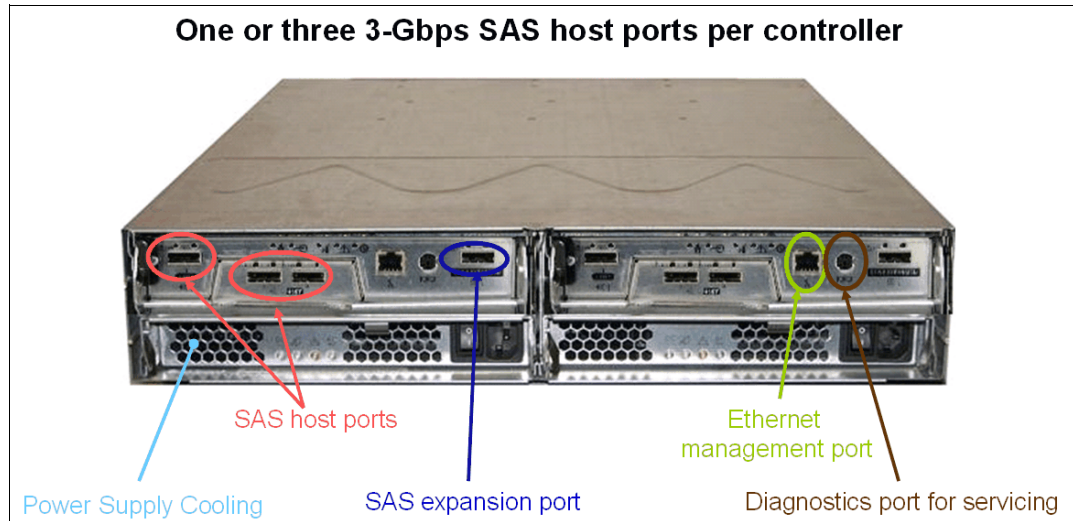


Figure 2-3 DS3200 (1726-22x) storage subsystem rear view

Specifications: Table 2-1 lists the DS3200 specifications.

Table 2-1 IBM System Storage DS3200 specifications

Model	1726-21x single controller 1726-22x dual controller 1726-22T dual controller Telco,
RAID controller	Dual active
Cache per controller	512 MB battery-backed cache with 1 GB upgrade option
Host interface	One or three host ports per controller, Serial Attached SCSI (SAS) 3 Gbps
Drive interface	Serial Attached SCSI (SAS)
Supported drives	3 Gbps SAS: 73 GB, 146 GB, 300 GB, and 450 GB SAS drives at 15,000 rpm speeds 3 Gbps SATA: 500 GB, 1.0 TB SATA drives at 7,200 rpm speed
RAID	RAID-0, -1, -3, -5, -6, -10
Storage partitions	4, 8, 16, 32
Maximum drives supported	48 drives (utilizing three EXP3000 Expansion Units)
Fans and power supplies	Dual-redundant, hot-swappable
Rack support	19-inch, industry-standard rack

For the latest specification information, including the DS3200 Interoperability Matrix, see the website:

<http://www.ibm.com/systems/storage/disk/ds3000/ds3200>

2.2.2 IBM System Storage model DS3300

The model DS3300 is ideally suited for SMB range customers who also transition from direct attached to external storage, but are hesitant to implement FC SAN due to perceived complexity. Making use of the customer's inhouse IP-network and expertise, it is an ideal and comfortable solution. With the model DS3300, the pervasiveness of IP networks are brought together, because it is easy to install and manage and has all the obvious TCO and availability advantages to go with it. Furthermore, it gives you the freedom of extensive reach with hardly any distance limitation.

The DS3300 storage subsystem 1726-31x and 1726-32x (dual controller) comes with the following components (Figure 2-4):

- ▶ 2U rack-mountable enclosure with 12 easily accessible drive bays
- ▶ Two 1 Gbps iSCSI host ports per controller
- ▶ Support for dual-ported, and hot-swappable SAS (15k rpm) and SATA (7.2k rpm) disks
- ▶ Scalability of up to 12 TB internal physical storage capacity using 1 TB hot-swappable SAS disks
- ▶ Expand ability of up to 48 TB physical storage capacity with attachment of three EXP3000 expansion units
- ▶ RAID 0, 1, 3, 5, 6 and 10 support
- ▶ LED indicators (8)
- ▶ Minimum 512MB RAID cache memory
- ▶ Maximum 2 GB system memory, battery backup
- ▶ Built-in reliability features with dual-redundant power-supplies standard

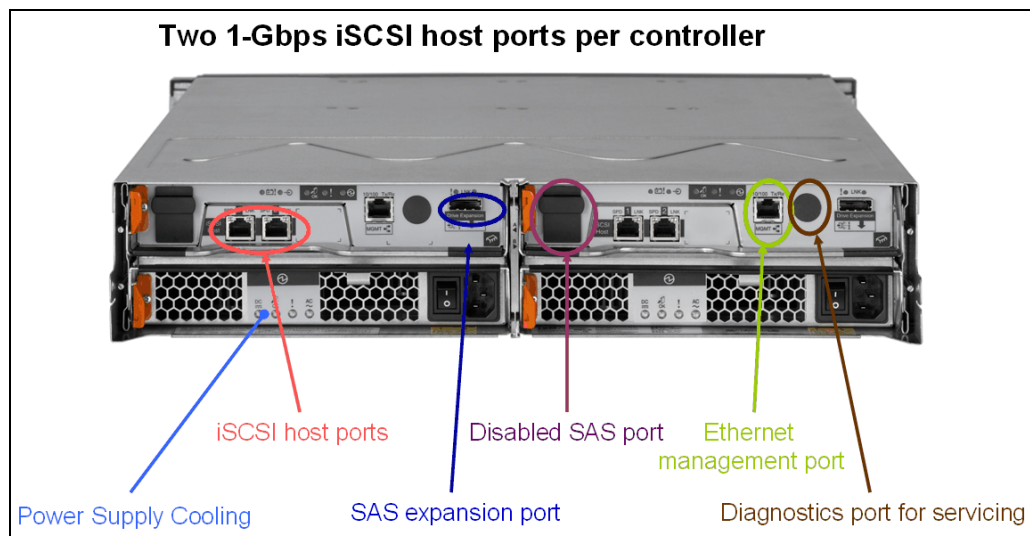


Figure 2-4 DS3300 (1726-32x) storage subsystem rear view

Specifications: Table 2-2 lists the DS3300 specifications.

Table 2-2 IBM System Storage DS3300 specifications

Model	1726-31X - single controller 1726-32X - dual controller 1726-32T - dual controller Telco
RAID controller	Dual active
Cache	512 MB battery-backed cache with 1 GB upgrade option
Host interface	Two host ports per controller, 1 Gbps iSCSI
Drive interface	Serial Attached SCSI (SAS)
Supported drives	Supports 3 Gbps SAS: 73 GB, 146 GB, 300 GB, and 450 GB SAS drives at 15,000 rpm speeds 3 Gbps SATA: 500 GB, 750 GB and 1.0 TB SATA drives at 7,200 rpm speed
RAID	RAID 0, 1, 3, 5, 6, 10
Storage partitions	4, 8, 16, 32
Maximum drives supported	48 SAS or SATA drives (using three EXP3000 expansion units)
Fans and power supplies	Dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack

For the latest specification information, including the DS3300 Interoperability Matrix, see the website:

<http://www-03.ibm.com/systems/storage/disk/ds3000/ds3300/index.html>

2.2.3 IBM System Storage model DS3400

Model DS3400 fits the SMB market very well. You can start with a simple direct-attached FC solution, and as your storage demands grow, seamlessly transition to an FC SAN.

Growth will often lead you to a Fiber Channel SAN solution. The DS3400 is scalable, extremely fast, very reliable, congestion free, and fit for long distances. It is one of the world's best price-performance entry FC storage servers. It can support your remote sites, offices, and departments, while you can be self-sufficient due to its simple administration. The advantages of shared-storage are high; enablement of consolidation (storage and server), increased efficiencies within the infrastructure, and centralized single management.

The DS3400 storage subsystem 1726-41x and the 1726-42x (Dual controller) contain the following components (Figure 2-5):

- ▶ 2U rack-mountable enclosure with 12 easily accessible drive bays
- ▶ Two 4 Gbps FC host ports per controller
- ▶ Support for dual-ported and hot-swappable SAS (15k rpm) and SATA (7.2k rpm) disks
- ▶ Scalability of up to 12 TB internal physical storage capacity using 1 TB hot-swappable SAS disks
- ▶ Expands up to 48 TB physical storage capacity with attachment of three EXP3000 expansion units LED indicators (8)
- ▶ RAID 0, 1, 3, 5, 6 and 10 support

- ▶ Minimum 512MB RAID cache memory
- ▶ Maximum 2 GB system memory, battery backup
- ▶ Built-in reliability features with dual-redundant power-supplies standard

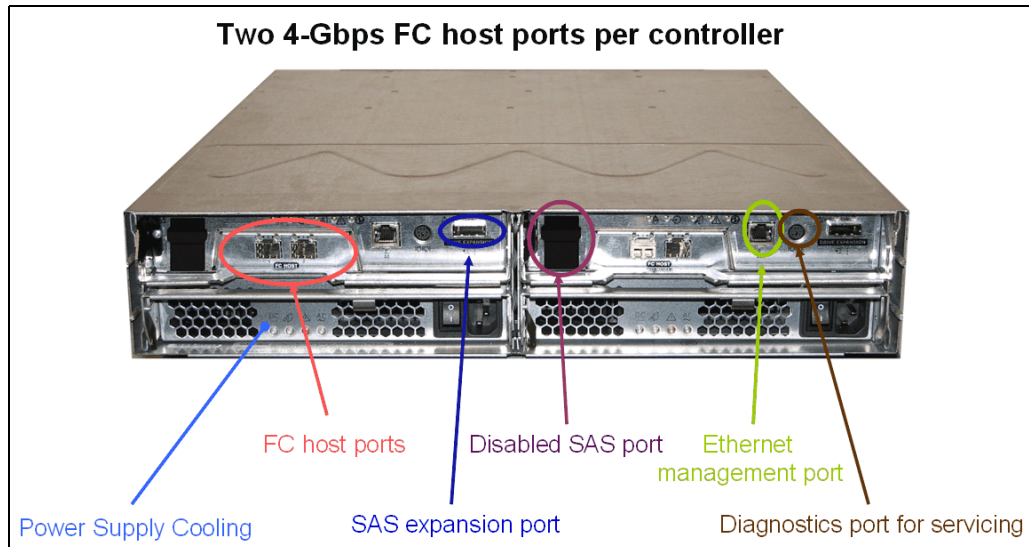


Figure 2-5 DS3400 storage subsystem model 1726-42x rear view

Specifications: Table 2-3 lists the DS3400 specifications.

Table 2-3 IBM System Storage DS3400 specifications

Model	1726-41X - single controller 1726-42X - dual controller 1726-42T - dual controller Telco
RAID controller	Dual active
Cache	512 MB battery-backed cache with 1 GB upgrade option
Host interface	Two host ports per controller, Fibre Channel (FC) 4 Gbps auto-sensing 1 Gbps/2 Gbps
Drive interface	Serial Attached SCSI (SAS)
Supported drives	Supports 3 G bps SAS 73 GB, 146 GB, 300 GB, and 450 GB SAS drives at 15,000 rpm speeds 3 Gbps SATA: 500 GB, 750 GB and 1.0 TB SATA drives at 7,200 rpm speed
RAID	RAID 0, 1, 3, 5, 6, 10
Storage partitions	4, 8, 16, 32
Maximum drives supported	48 SAS or SATA drives (using three EXP3000 Expansion Units)
Fans and power supplies	Dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack

For the latest specification information, including the DS3400 Interoperability Matrix, see the website:

<http://www.ibm.com/systems/storage/disk/ds3000/ds3400>

2.3 IBM System Storage EXP3000

The EXP3000 is a 2U rack-mountable SAS disk drive enclosure. It supports up to 12 SAS or SATA 3.5-inch disk drives. If you use 1 TB SATA disk drives, the maximum storage capacity in a single EXP3000 will be 12 TB. You can attach up to three EXP3000 enclosures to a DS3000 storage subsystem, and this gives you a maximum storage capacity of 48 TB. Additionally, the 50 GB SATA Solid State Disk drives are supported in EXP3000 expansion units that are directly attached to System x servers by the ServeRAID MR10M SAS/SATA controller adapter installed in the System x server.

The EXP3000 enclosures connect to the drive side SAS port of the DS3000 through a 3 Gbps SAS interface. The main components of an EXP3000 are one or two Environmental Services Modules (ESMs), two power supply modules, and up to 12 SAS disk drives. ESMs and power supply modules are in the rear of the enclosure (as shown in Figure 2-6), whereas the SAS drives are installed in the front.

The EXP3000 comes with one ESM in the standard configuration, and another ESM can optionally be installed. Both ESMs must be installed in order to connect the EXP3000 to a dual controller DS3000 storage subsystem.

Each EXP3000 is also equipped with two hot-swappable and redundant power supply modules. Each power supply module contains two fans for cooling redundancy. If a power supply or a fan fails, you can replace it while the EXP3000 remains fully operational.



Figure 2-6 EXP3000 storage subsystem rear view

Specifications: Table 2-4 lists the EXP3000 specifications.

Table 2-4 IBM System Storage EXP3000 specifications

Model	1727-01X - IBM System Storage EXP3000 1727-02T - IBM System Storage EXP3000 with DC power supplies
Environmental Services Module	Single ESM with dual ESM upgrade SAS interface: Mini-SAS connectors
Drives supported	SAS 15,000 rpm capacities: 73 GB, 146 GB, 300 GB, 450 GB SATA 7,200 rpm capacities: 500 GB, 750 GB, 1.0 TB Solid State Disk: 50 GB SATA (System x direct attach only by the ServeRAID MR10M SAS/SATA controller adapter)
Fans and power supplies	Two hot-swappable 530 watt (115 – 230 V AC) power supplies
Rack support	19-inch, industry-standard rack

For the latest specification information, including the EXP3000 dimensions, see the website:

<http://www-03.ibm.com/systems/storage/disk/exp3000/index.html>

2.4 IBM DS Storage Manager

The IBM DS Storage Manager is used to manage the DS3200, DS3300, and DS3400 storage subsystems. Along with your storage subsystem, you might get DS Storage Manager Software and Host Kit CDs; otherwise, you can download it from the IBM support website. Using DS Storage Manager software, you can perform tasks such as creating arrays and logical drives, assigning logical drives to the host servers, setting up FlashCopy and VolumeCopy, capturing logs for troubleshooting, and so on.

When discussing the DS Storage Manager, it is important to differentiate between the two terms, *host server* and *management station*.

2.4.1 Host server

The *host server* is a server attached to the DS3000 storage subsystem through the I/O path (SAS, iSCSI, or Fibre Channel). The host server has access to certain logical drives defined on the DS3000.

2.4.2 Management station

The *management station* is a system used for DS3000 management. The DS Storage Manager GUI runs on the management station. You need to establish a management connection between the management station and the DS3000 storage subsystem, which can be done in two ways:

- ▶ Out-of-band: The management station is connected to the Ethernet management port in each DS3000 RAID controller. We also call this method *direct-attached management*. The management station in this case only requires an Ethernet connection to the DS3000.
- ▶ In-band: This method utilizes the I/O path between a host server and the DS3000. In this case, the management station does not have a direct TCP/IP connection to the DS3000, but rather communicates with the DS3000 through an HBA, which acts as a gateway to the DS3000 storage subsystem. We also call this method *host-attached management*.

Figure 2-7 shows the “look and feel” of the Storage Manager.

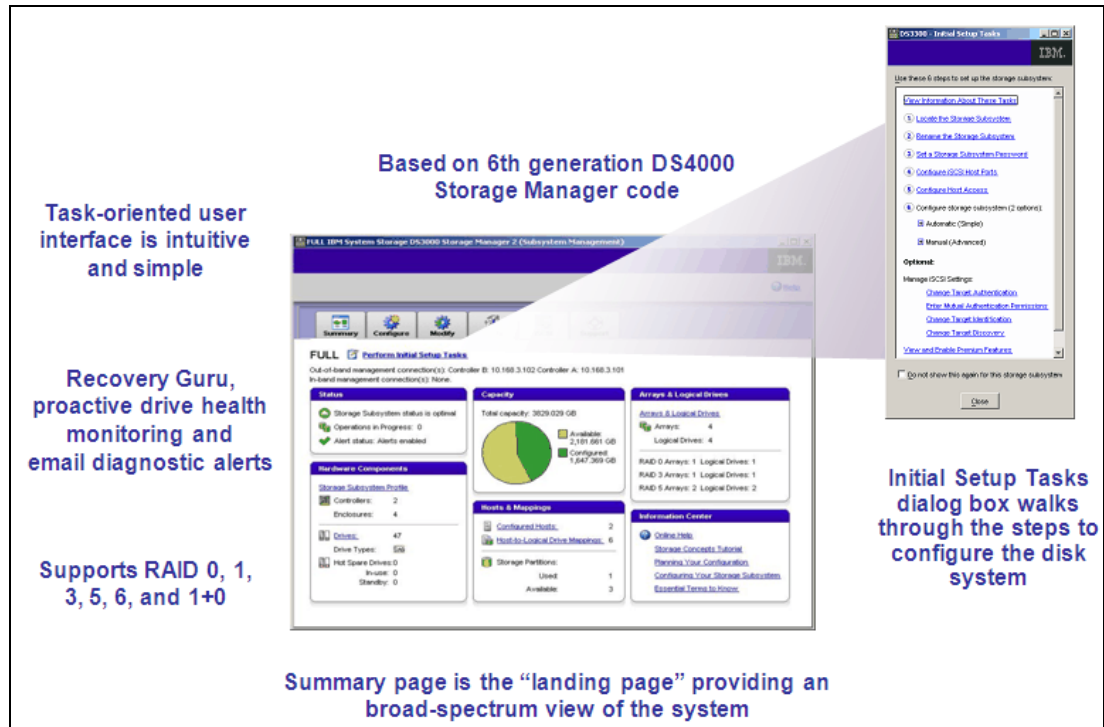


Figure 2-7 Storage Manager GUI

More information: For examples of typical configurations, and problem determination flowcharts, see the *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065.

2.5 Premium features

Standard configurations of DS3000 storage subsystems come with the following premium features.

2.5.1 DS3000 Partition Expansion License

As part of the standard configuration, four storage partitions are enabled on the DS3200, DS3300, and DS3400 with Windows and Linux on Intel host attach license (this can be optionally increased to 8, 16, or 32). The maximum number of storage partitions is 32 on all DS3000 products. Use the DS3000 Partition Expansion License to enable all 32 storage partitions.

2.5.2 DS3000 FlashCopy Expansion License

This feature enables FlashCopy. FlashCopy is a point-in-time copy of a source logical drive. The FlashCopy logical drive becomes available almost instantaneously. FlashCopy requires the use of a defined FlashCopy *repository*, which contains the original content of the data that has since been altered. FlashCopy logical drives are often used as a source for a backup operation. They can also be used to simply and quickly roll back to an original data state, thus providing a restore point. However, if the source logical drive is lost, point-in-time FlashCopy will be lost as well. As part of the standard configuration, two FlashCopies are enabled on every DS3000 storage subsystem and this premium feature enables up to 64 FlashCopies.

2.5.3 DS3000 VolumeCopy License

VolumeCopy is a way to provide a complete point-in-time copy of a source logical drive. As opposed to FlashCopy (where only the original values of changed data are copied to the repository), the whole source logical drive is copied to target. You can use this functionality for data replication, relocation, backup, or to restore snapshot data to the original logical drive.

The time required to establish a copy will depend on the size of the source data and the operation priority settings. While establishing the copy, the source logical drive will be in read-only state. After all the data is copied to the target, the target will remain available if the source logical drive is lost. The VolumeCopy premium feature allows for up to 128 VolumeCopies. Be aware that FlashCopy is a prerequisite for VolumeCopy.

2.6 Advanced features

The extended advanced features are listed in Figure 2-8, included with the part numbers. This diagram describes Out of the Box or premium features, and shows you all current advanced features that are optional and currently available on the DS3000 series.

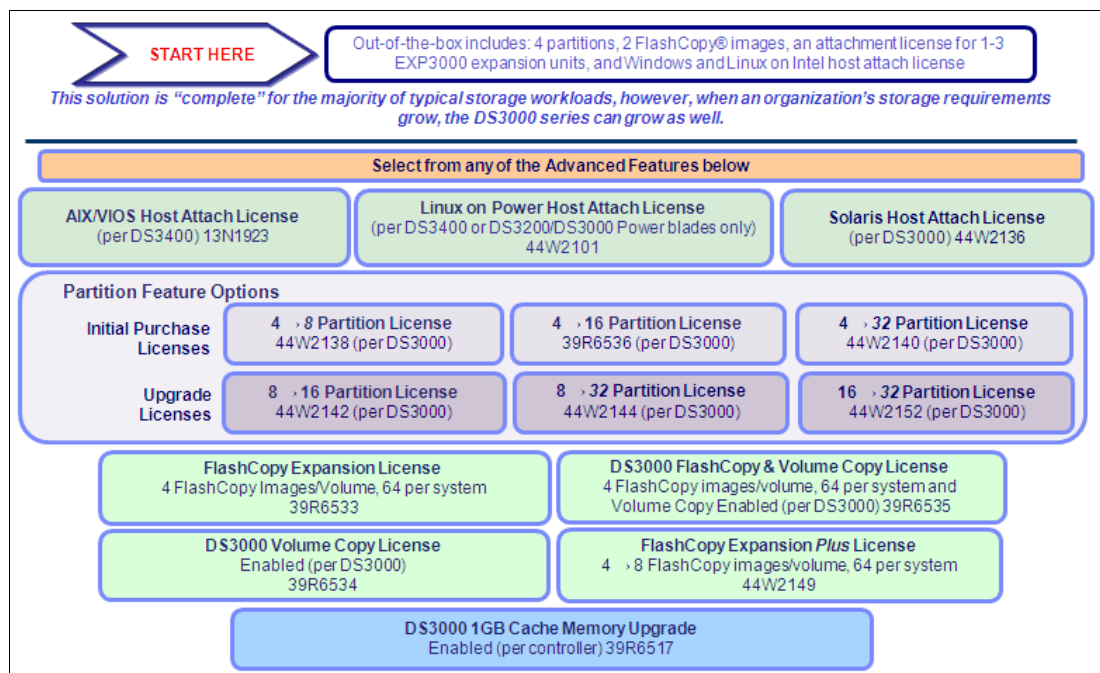


Figure 2-8 Advanced features

2.7 More information

For more information about the IBM System Storage entry-level disk systems, see the *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065.

You can also consult the following websites:

<http://www-03.ibm.com/systems/storage/disk/entry/index.html>

- IBM/LSI Internet site for IBM employees and Business Partners (after subscription):

<http://www.ibmdsseries.com>



Midrange disk systems

The introduction of the IBM Model DS5000 series offers a replacement path for the very successful range of DS4000 series. Whereas the midrange entry model DS5020 will be the natural replacement for the model DS4700, the midrange model DS5100 will be the natural replacement for the model DS4800.

In this chapter, we focus on the DS5000 models, also including the DS4000 series models for comparison when choosing an appropriate upgrade path.

The DS4000 series systems are RAID controller devices that contain Fibre Channel (FC) interfaces to connect the host systems and the disk drive enclosures. All DS4000 series systems have dual, hot swappable RAID controllers, thereby providing excellent system availability even if one of the parts malfunctions. Furthermore, hot-swap and redundant power supplies and fans can also help with the availability.

The next-generation DS5000 series are dual RAID controller devices that support intermixing Fibre Channel and iSCSI interfaces to the host systems and FC interfaces to the disk drive enclosures. They deliver a tremendous balanced performance for IOPS and MB/s application behavior, while having all the availability features to keep your data safe and reachable.

The IBM System Storage model DCS9900 is designed for applications with high-performance streaming data requirements served by deep computing servers, high performing computers, and clustered systems. With its parallel storage solution architecture, the DCS9900 is especially designed to address such needs.

In this chapter, we present the features and major characteristics of the various models that make up the current DS Midrange disk systems family, as known at the time of writing:

- ▶ IBM System Storage model DS4700 Express
- ▶ IBM System Storage model DS5020 Express
- ▶ IBM System Storage model DS5100
- ▶ IBM System Storage model DS5300
- ▶ IBM System Storage model DCS9900
- ▶ IBM System Storage DS4000 EXP810 Expansion Unit
- ▶ IBM System Storage DS5020 EXP520 Expansion Unit
- ▶ IBM System Storage DS5100 and DS5300 EXP5000 Expansion Unit
- ▶ IBM System Storage EXP5060 High Density Drawer

Technical descriptions and specifications are given for each model in the corresponding sections.

We also briefly discuss the IBM System Storage DS Storage Manager software that is used to manage the DS family series systems, and we give highlights of Full Disk Encryption and Solid State Disk. Then we close this chapter with a description of the Remote Support Manager product.

3.1 Highlights: What's new

This section summarizes the withdrawn products, new products, still available products, functions, supported platforms, and features of the DS midrange family of disk storage products, since the last edition of this book in July 2006.

Withdrawn (or communicated to be withdrawn) models:

- ▶ DS4100
- ▶ DS4300, and DS4300Turbo
- ▶ DS4500
- ▶ DS4700 Express models 70A, 70S, 72A, 72S
- ▶ DS4800 All Models
- ▶ EXP100 expansion
- ▶ EXP710 expansion

Newly introduced midrange product models:

- ▶ IBM System Storage DS5020 Express
- ▶ IBM System Storage DS5100
- ▶ IBM System Storage DS5300
- ▶ IBM System Storage model DCS9900
- ▶ IBM System Storage DS5020 EXP520 Expansion Unit
- ▶ IBM System Storage DS5100 and DS5300 EX5000 Expansion Unit
- ▶ IBM High Density Drawer model 5060

Updated general information after introduction DS5000 series:

- ▶ Storage Manager updated to Version 10.x
- ▶ Attachment support up to 448 Drives for DS5100 and DS5300
- ▶ 73 GB Solid State Disk support
- ▶ Full Disk Encryption for FC Disks (FDE)
- ▶ Power systems Virtual I/O support
- ▶ Native system Power I support for DS5100 and DS5300
- ▶ 1gb iSCSI Host support for DS5100 and DS5300
- ▶ 32MB and 64MB cache Memory features and MES
- ▶ High bandwidth OM3 fiber optic cables

3.2 Overview of IBM System Storage DS Models

In this section we present an overview of the IBM System Storage DS Models.

3.2.1 IBM System Storage DS4000

The IBM System Storage and System Storage DS4000 series are disk storage products using redundant array of independent disks (RAID) that contain the Fibre Channel (FC) interface to connect both the host systems and the disk drive enclosures. The DS4000 series of disk storage systems are an IBM solution for low to midrange/departmental storage requirements

The IBM System Storage DS4000 series now includes the only supported DS4700.

3.2.2 IBM System Storage DS4700 Express

IBM System Storage model DS4700 Express offers high performance 4 Gbps capable Fibre Channel connections and up to 112 TB of physical storage capacity. The DS4700 can expand from workgroup to enterprise-wide capability by attaching up to six EXP810 disk enclosures.

3.2.3 IBM System Storage DS5000

IBM System Storage and System Storage DS5000 series are disk storage products using redundant array of independent disks (RAID) that contain 4 or 8Gbps Fibre Channel (FC) and 1 Gbps Internet SCSI (iSCSI) interface to connect to the host systems, thus enabling host tiering, and 4 Gbps FC to the disk drive enclosures. The DS5000 series of disk storage systems are an IBM solution for low to high-range departmental (and even low Enterprise) storage requirements.

The IBM System Storage DS5000 series family consists of the following products:

- ▶ IBM System Storage DS5020 Express:

The IBM System Storage model DS5020 Express offers high performance 8 Gbps Fibre Channel and 1 Gbps iSCSI connections to the Host and 4 Gbps FC drive interfaces to the EXP520 expansion units. The DS5020 can support up to 112 disk drives by 6 EXP520 Expansion Drawers, offering 67.2TB FC or 112TB SATA of physical storage capacity.

- ▶ IBM System Storage DS5100 and DS5300:

The IBM System Storage model DS5100 and DS5300 offer high performance 4 Gbps and 8 Gbps Fibre Channel and 1Gbps iSCSI connections to the Host and 4 Gbps FC drive interfaces to the EXP5000 and EXP810 (for migration purposes only) expansion units. Both can support up to 448 disk drives by 6 EXP5000 Expansion Drawers offering 268 TB FC or 448 TB SATA of physical storage capacity. Furthermore, they can support up to 8 EXP5060 High Density Drawers, which offers 480 TB of physical storage.

The DS5100 can have the same performance as the DS5300 using the option of enhanced performance key, so technically there is no difference. The difference is within the price setting. The DS5100 starts at a much lower price than the DS5300, but after full transition to a DS5300, it results in just a slightly higher price than an original DS5300. In this chapter, we regard the DS5100 and the DS5300 as one.

3.2.4 IBM System Storage Model DCS9900

The IBM System Storage Model DCS9900 offers a dual active RAID controller containing 8 Gbps Fibre Channel (FC) or IB 4x IDR interface to connect to the host systems. The DCS9900 can support up to 1200 Serial Attached SCSI (SAS) or SATA disks. The DCS9900 is designed for applications with high-performance streaming data requirements served by Deep Computing servers, high performing computers, and clustered systems.

3.3 IBM System Storage Model DS4700 Express

The IBM System Storage Model DS4700 Express supports a high-performance 4 Gbps FC interface for data-intensive applications that demand increased connectivity. The DS4700 Express 4 Gbps technology is backward compatible with both 2 Gbps and 1 Gbps technologies. The DS4700 Express is available in two models: Model 70 has 2 GB of physical cache memory (1 GB per controller), and Model 72 has 4 GB of physical cache memory (2 GB per controller). The DS4700 Express supports up to 112 disk drives with attachment of six EXP810 disk enclosures, for over 112 TB of physical storage capacity. DS4700 Express supports up to 16 hard disk drives internally in the standard controller.

Figure 3-1 shows the DS4700 with 16 hard disk drives installed.

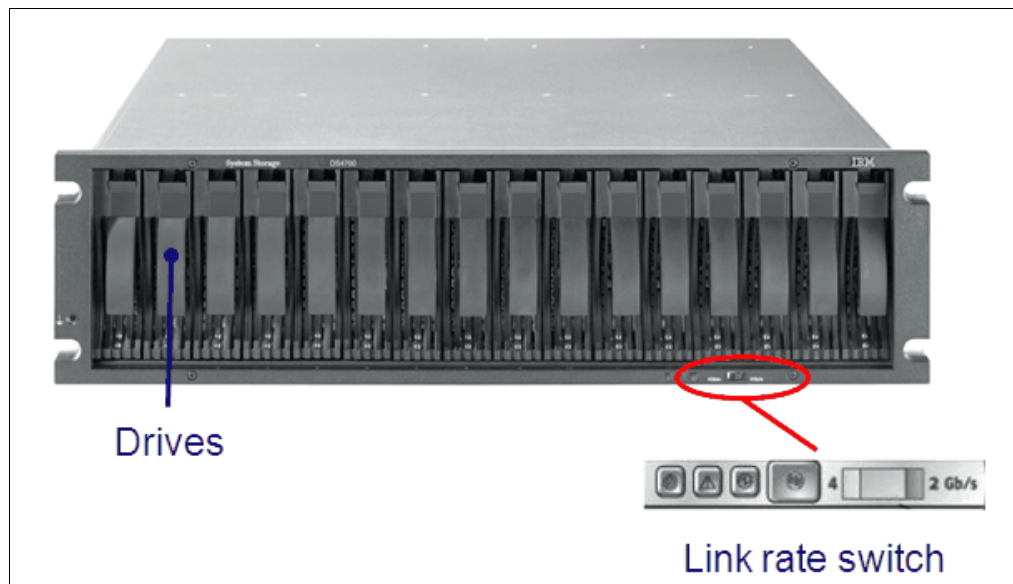


Figure 3-1 DS4700 Express front view

3.3.1 IBM DS4700 Express Model features

Following are the IBM System Storage DS4700 Express Model features.

- ▶ Dual RAID controllers: Dual 4 Gbps capable, redundant, and hot-swappable, battery backed, write cache mirroring
- ▶ Model 72 supports:
 - 4 GB battery backed cache, which is designed to be protected for at least three days
 - Eight Host/SAN connections side and four drive side ports for two controllers
- ▶ Model 70 supports:
 - 2 GB battery backed cache, which is designed to be protected for at least three days
 - Four Host/SAN connections side and four drive side ports for two controllers
- ▶ Supporting 2 dual redundant FC disk loops
- ▶ Eight or four host ports providing a cluster solution without using a switch
- ▶ Auto negotiates 1, 2, or 4 Gbps host connection speeds
- ▶ Supports up to 16 disk drives inside the controller, of the following types:
 - 4 Gbps FC: 15K rpm, 600 GB/300 GB/146 GB (E-DDM)
 - 4 Gbps SATA: 7.2K rpm, 1 TB/750 GB (E-DDM)
- ▶ Hot pluggable drives: Adds new drives for extra capacity without rebooting the server
- ▶ Supports up to 112 FC drives with 6 DS4000 EXP810 Expansion Units
- ▶ Includes System Storage DS Storage Manager to help manage the storage subsystem
- ▶ Allows the intermixing of expansion units (EXP710 and EXP810) and also the intermixing of FC and SATA disks in the same enclosure (only for EXP810) or storage subsystem

- Provides flexibility for multiplatform storage environments by supporting a wide variety of servers, operating systems, and cluster technologies, offering support for System p, System x, System i® w/ VIOS, Windows Server 2003 w/SP1, Windows Server 2003 w/SP2, Windows Server 2008 w/SP1, Windows 2000 Server & Advanced Server w/SP4, Novell NetWare 6.0 w/SP5 & 6.5 w/SP5, Red Hat Enterprise Linux 3.0 U7, Red Hat Enterprise Linux 4.0 U3, Red Hat Enterprise Linux 5.2 and 5.3, SUSE Linux Enterprise Server 8 SP4, SUSE Linux Enterprise Server 9 SP3, VMware ESX 3.0/3.5/3i, VMware ESX 2.5.2, AIX 5.1, 5.2, 5.3, 6.1, HP-UX 11i and 11.23, Solaris 8,9,10
- Offers advanced replication functions such as FlashCopy, VolumeCopy, and Remote Mirroring for high data availability and protection

Figure 3-2 shows the DS4700 model 72x rear view. Each controller has two Ethernet ports, independent slots battery, and two modules for power and cooling.

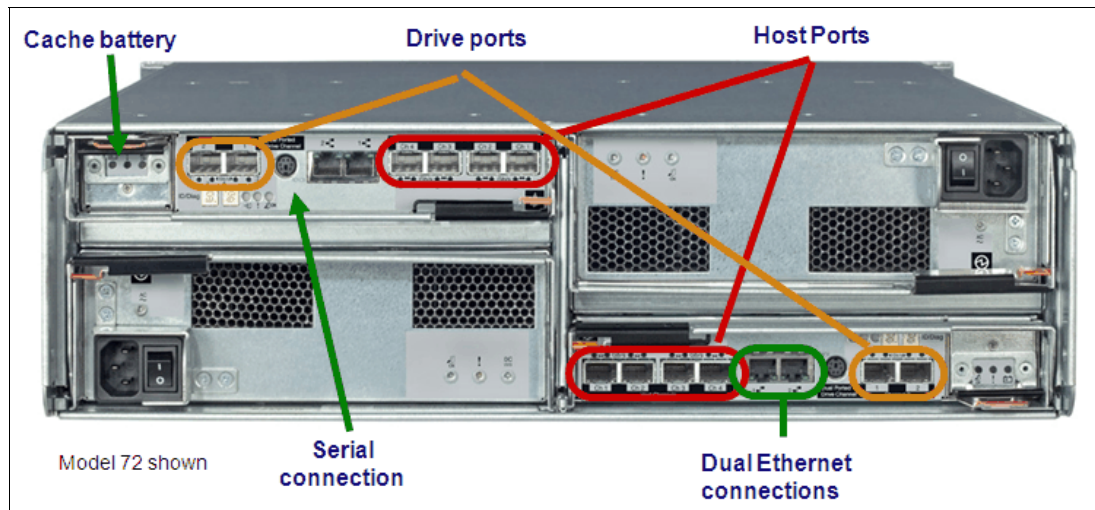


Figure 3-2 DS4700 mod 72x rear view

3.3.2 IBM DS4700 Express specifications

Table 3-1 shows the specifications of the DS4700.

Table 3-1 System Storage DS4700 Express specifications

Model	1814-72H	1814-70H
RAID controller	Dual active 4 GB RAID controllers	Dual active 4 GB RAID controllers
Capacity of base enclosure	16 HDD	16 HDD
Controller cache	Total of 4 GB, battery-backed	Total of 2 GB, battery-backed
Host interface	FC-AL	FC-AL
Connectivity	Eight host side and four drive side ports	Four host side and four drive side ports
Drive interface	FC-AL	FC-AL

Model	1814-72H	1814-70H
Supported drives	4 Gbps FC: 15K rpm, 600 GB/300 GB/146 GB/73 GB (E-DDM) 4 Gbps SATA: 7.2K rpm 1 TB SATA II*/750 GB SATA II / 500GBSATA (E-DDM)B * requires firmw. 6.23.05	4 Gbps FC: 15K rpm, 600 GB/300 GB/146 GB/73 GB (E-DDM) 4 Gbps SATA: 7.2K rpm 1 TB SATA II*/750 GB SATA II /500 GB (E-DDM) * requires firmw. 6.23.05
RAID levels	0, 1, 3, 5, 6 and 10	0, 1, 3, 5, 6 and 10
Storage partitioning	Up to 128	Up to 128
Maximum drives supported	Up to 112 FC or SATA drives	Up to 112 FC or SATA drives
Fans and power supplies	Dual redundant, hot-swappable	Dual redundant, hot-swappable
Rack support	Industry standard 19" rack - 3U	Industry standard 19" rack - 3U

For the latest specification information, see the website:

<http://www.ibm.com/servers/storage/disk/ds4000/ds4700/index.html>

3.4 IBM System Storage model DS5020

The DS5020 is the newest model within the DS5000 series and is designed to help address midrange or departmental storage requirements. Equally adept at IOPS and MB/s, the DS5020 is able to support applications with wide-ranging performance requirements. It is well suited for data warehousing, consolidation, and virtualization environments that have diverse workloads and application requirements, and is designed to concurrently support transactional intensive applications, such as databases and OLTP, and throughput intensive applications, such as HPC and rich media at entry price level.

The DS5020's design avoids over-configuration for an affordable entry-point, while offering seamless "pay-as-you-grow" scalability as requirements change. Its efficient storage utilization lowers raw capacity requirement, and support for intermixing high-performance and high-capacity drives enables enclosure-based tiered storage. These unique capabilities reduce the number of drives needed to meet performance and/or capacity demands, thus lowering acquisition and operational expenditures.

3.4.1 DS5020 description

The DS5020 is a 3U rack-mountable enclosure, has four 4 Gbps FC drive interfaces, and can comprise a maximum of six EXP520 expansion units for a total of up to 112 disk drives. Through a specific activation feature, six EXP810 expansions can be used in place of the EXP520s, thus enabling investment protection.

The following model is available: DS5020 Disk Controller (1814-20A).

Figure 3-3 shows the Model DS5020.

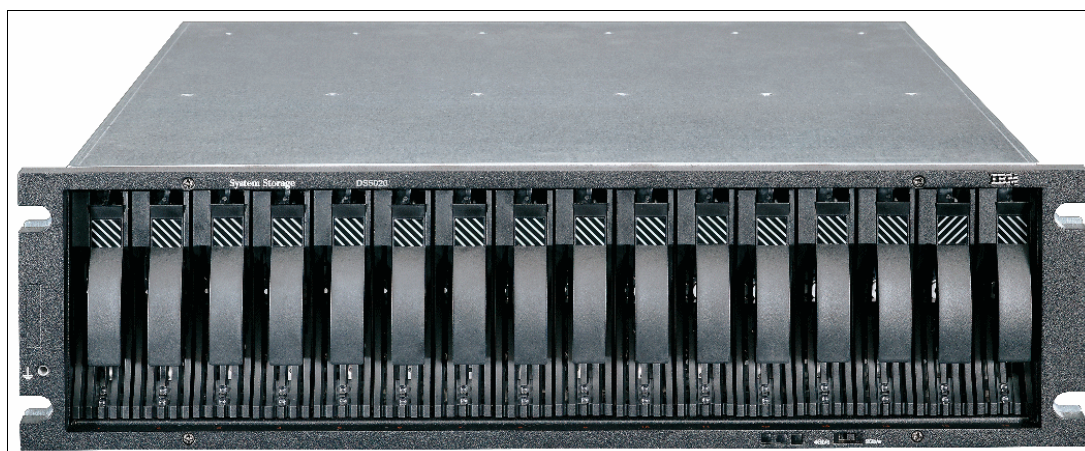


Figure 3-3 IBM System Storage DS5020

3.4.2 DS5020 features

Here we summarize the DS5020 features:

- ▶ Up to 112 TB physical storage capacity
- ▶ Comes with 2 GB cache memory (1 GB per internal RAID controller)
- ▶ 4 GB cache memory (2 GB per RAID controller) feature is available as an initial plant order feature; there are no cache memory upgrades available as field (MES) features for the DS5020
- ▶ Accommodates up to 16 disk drives installed within the DS5020 enclosure
- ▶ Attachment support for up to six EXP520 expansion enclosures
- ▶ Attachment support for EXP810; this is supported with the Attach EXP810 to DS5020 Activation feature and is available only for migration purposes
- ▶ Supports intermix of SATA drives, FC drives, and encryption-capable FC drives (FDE) within the DS5020 and EXP520 enclosures
- ▶ Provides SAN-attached 8 Gbps Fibre Channel (FC) host connectivity, as well as optional 1GbE iSCSI host connectivity
- ▶ All DS5020s have four 8 Gbps FC ports (two per controller)
- ▶ Additionally, you can order initially either of the following options:
 - 2-Dual 8Gbps Host Interface Cards (HIC)
 - 2-Dual 1Gbps iSCSI HIC
- ▶ Drive options:
 - FC disks without encryption:
 - 146.8 GB/15K 4Gbps FC DDM
 - 300 GB/15K 4Gbps FC DDM
 - 450 GB/15K 4Gbps FC DDM
 - 600 GB/15k 4Gbps FC DDM

- FC disk with encryption:
 - 146.8 GB/15K 4Gbps FC encryption-capable DDM
 - 300 GB/15K 4Gbps FC encryption-capable DDM
 - 450 GB/15K 4Gbps FC DDM encryption-capable DDM
 - 600 GB/15k 4Gbps FC DDM encryption-capable DDM
- SATA disks:
 - 750 GB/7.2K SATA DDM
 - 1000 GB/7.2K SATA DDM
- Supports RAID 0, 1, 3, 5, 6, and 10
- At the time of writing, one model is available: 1814-20A

3.4.3 DS5020 specifications

Table 3-2 shows the DS5020 specifications.

Table 3-2 IBM System Storage DS5020 specifications

Model	1814-20A
RAID controller	Dual active
Cache	2 GB battery-backed or 4 GB battery-backed initial option (no MES)
Host interface	Four 8 Gb/s FC or Eight 8 Gb/s FC; or Four 4 Gb/s FC and Four 1 Gb iSCSI
Drive interface	4 drive ports - Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard, Auto-sensing 2 Gbps/4 Gbps
Supported drives	Supports 4 Gbps FC/FDE: 15k – 600 GB, 450 GB, 300 GB, 146.8 GB, E-DDM Supports 4 Gbps SATA: 7.2K 750 GB and 1 TB E-DDM Disk drives
RAID	Levels 0, 1, 3, 5, 6, and 10
Storage partitions	4, 8, 16, 64, or 128 storage partitions
Maximum drives supported	112 FC or 112 SATA drives (using 6 EXP520 Expansion Units)
Fans and power supplies	Dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack - 4U

For the latest specification information and Interoperability Matrix, see the website:

<http://www.ibm.com/systems/storage/disk/ds5020/index.html>

3.5 IBM System Storage Models 5100 and 5300

The first models of the DS5000 family that have been released offer increased performance compared to the family's DS4000 predecessors. It offers "Pay-as-you-grow" scalability up to 448 drives for both the DS5100 and the DS5300, thus enabling the most demanding capacity requirements, by utilizing the EXP5000 expansion enclosure. With the introduction of the DS5060 Expansion Drawer, we also can deliver up to 480 TB of physical storage housed in one single Rack.

The DS5000 series is equally adept at supporting transactional applications such as databases and OLTP, throughput-intensive applications such as HPC and rich media, and concurrent workloads for consolidation and virtualization. With its relentless performance and superior reliability and availability, the DS5000 series storage system can support the most demanding service level agreements (SLAs). And when requirements change, you can add or replace host interfaces, grow capacity, add cache and reconfigure the system on the fly, ensuring that it will keep pace with your growing company.

3.5.1 DS5100/5300 storage subsystems

The following two models are available:

- ▶ DS5100 (1818-51A)
- ▶ DS5300 (1818-53A)

The two models differ in the number of host ports and amount of cache for data and performance.

Model DS5100 initially comes with eight host channels, 4 Gbps FC, and 8 GB to 64 GB cache for data. Upgrade path is available to DS5300 by using the optional full performance activation.

Model DS5300 comes with eight or sixteen host channels, 4 Gbps FC, and 8 GB to 64 GB cache for data and full performance activation.

Figure 3-4 shows the front view of the DS5000.



Figure 3-4 IBM System Storage DS5000

3.5.2 DS5100 and DS5300 features

Models DS5100 (1818-51A) and DS5300 (1818-53A) have the following features:

- ▶ Compact 4U rack-mountable enclosure
- ▶ Utilizes new 7th generation dedicated ZIP ASIC engines on RAID controllers
- ▶ Features Intel Xeon® 2.8 GHz processor
- ▶ Dual, redundant controllers
- ▶ PCI Express x8 bus technology
- ▶ Ability to attach up to Eight EXP5060 Enclosures feature
- ▶ New DS5000 IBM i Host Kit feature: The DS5000 IBM i Host Kit feature (#7735) is supported with DS5300 and DS5100 controllers at firmware level 07.60, or later operating with IBM Power Systems (rack/tower systems only, no Power Blade support) while hosting IBM iOS 6.1 with 6.1.1 machine code
- ▶ Dedicated cache for data (base model has 8 GB cache) with enhanced diagnostics; the architecture is designed to support a 64 GB cache
- ▶ Dedicated processor memory of 2 GB per controller
- ▶ Hot-swappable lithium-ion battery for backup and de-staging data from cache
- ▶ New flash memory to store data from cache during power outage
- ▶ Two dedicated PCI Express buses for cache mirroring
- ▶ Redundant, hot-swappable power supplies and fans
- ▶ Hot-swappable interconnect module acts as midplane
- ▶ Supports RAID 0, 1, 3, 5, 6, and 10 (RAID 1+0)
- ▶ Supports RAID 1 and 10 dedicated mirrored drive pair configurations
- ▶ Ability to create a RAID 10 or 0 group on all available drives to maximize performance for a LUN
- ▶ Flexible host interface modules that can be added, changed, or mixed as the infrastructure changes; quad-4 Gbps Fibre Channel (FC) Host Interface Cards (HIC) that can now be replaced with quad-8 Gbps FC HIC or a dual-1 Gbps iSCSI Host Interface Card; field replacements (MES) available
- ▶ Supports an unlimited number of Global Hot Spare drives with the ability to enable/disable the copy back function (important for SATA drives)
- ▶ Host-side connections support:
 - Fibre Channel Switched Fabric
 - Arbitrated Loop and Fibre Channel Direct Connections
 - Ethernet Direct Connection and Ethernet Switched Network (with iSCSI HIC)
- ▶ Supports sixteen 4 Gbps drive-side connections for both controllers, allowing a total of eight dual-redundant drive channel pairs to be implemented to support expansion enclosure additions
- ▶ Redundant drive-side connections designed to avoid any single-point of failure and maintain high availability
- ▶ Supports up to 28 EXP5000 (or a mix of EXP5000 and EXP810 for migration purposes) for a total of 448 disks, thus allowing you to install up to 448 TB of raw capacity with 1 TB SATA disks or 268 TB of raw capacity with FC drives

Figure 3-5 shows the ability and terms to attach an EXP810 directly to the DS5000.

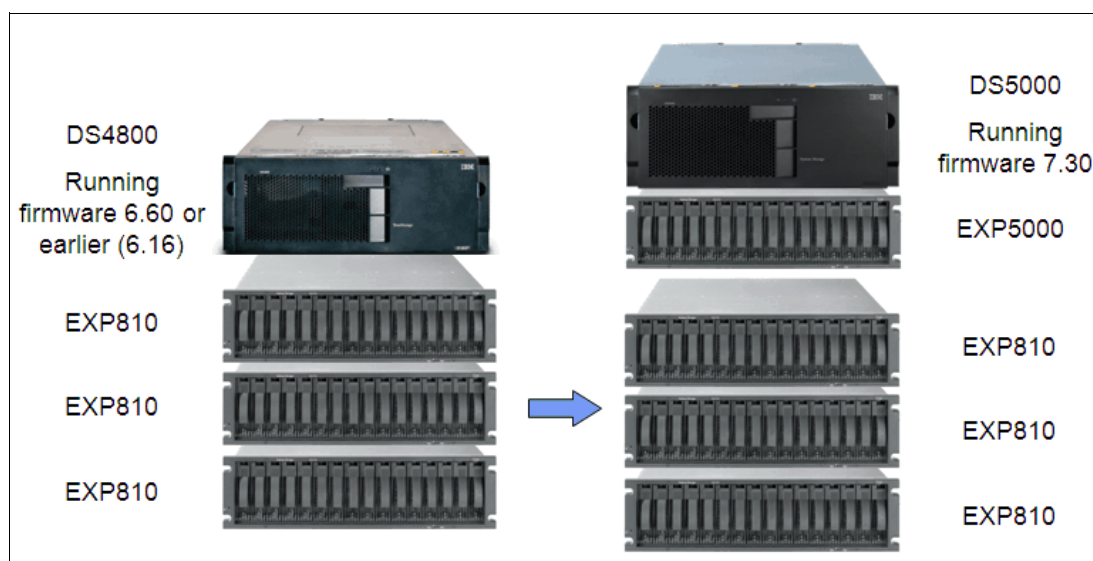


Figure 3-5 Example of moving EXP810 from DS4800 to DS5000

Note: DS4800 running firmware version 6.60 or earlier can migrate its EXP810s into a DS5000 with at least one EXP5000.

The following features are available:

- ▶ Supports a maximum of 20 solid state drives (SSDs) within EXP5000:

Fully supports Fibre Channel/SATA/SSD intermix (premium feature) by allowing the simultaneous usage of SATA, Fibre Channel, and SSD behind one DS5000 controller, allowing user flexibility and increased storage capacity utilization; also possible to mix disks of various size and technology inside one enclosure
- ▶ Supports up to 512 host storage partitions that isolate LUNs for various servers or groups of servers
- ▶ Supports up to 2048 logical volumes
- ▶ Supports up to 2048 host logins
- ▶ Supports 4096 command queue depth (maximum drive queue depth is 16)
- ▶ Supports logical volumes greater than 2 TB (when required and supported by the operating system)
- ▶ Supports shortwave Fibre Channel 4 and 8 Gbps host attachment
- ▶ Supports 1 Gbps copper iSCSI host attachment
- ▶ Multiple heterogeneous server and operating system support (host kits required)
- ▶ Powerful On Demand functions: Dynamic Volume Expansion, Dynamic Capacity Expansion, and Dynamic RAID Level Migration. Dynamic Segment Size allows users to modify storage configurations on-the-fly without incurring any downtime.
- ▶ Remote Support Manager notifies IBM if there is an issue; see 3.13, “IBM Remote Support Manager for Storage” on page 82.
- ▶ New dual 10/100/1000 Ethernet for out-of-band management to separate out-of-band management from service diagnostics for each controller

- FlashCopy (premium feature) for up to 16 copies per base volume. Two FlashCopies per default (without premium feature)
- VolumeCopy (premium feature)
- Remote Volume Mirroring: Metro Mirror, Global Mirror, and Global Copy (premium features) for up to 128 pairs
- Standard DB-9 serial connection for service purposes

Additionally, the DS5100 offers further value in flexibility, scalability, and investment protection by providing the ability to upgrade to a DS5300. The DS5100 provides excellent performance for dynamic storage consolidation solutions in the deployment stage, with the full assurance that this model can be upgraded to a DS5300 performance level if required.

Table 3-3 shows the DS5000 specifications.

Table 3-3 IBM System Storage DS5000 specifications

Model	1815-80A, 1815-82A
RAID controller	Dual active
Cache	Starts with 8 GB battery-backed Maximum 64 GB battery-backed per controller
Host interface	8 or 16 host ports - Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard, auto-sensing 4 GBps or 8 GBps
Drive interface	16x 4Gbps drive ports - Fibre Channel (FC) Switched and FC Arbitrated Loop (FC-AL) standard
Supported drives	146 GB, 300 GB, 450 GB, 600 GB 15,000 rpm (Fibre Channel) 146 GB, 300 GB, 450 GB, 600 GB 15,000 rpm (Fibre Channel) FDE Also full 16 Encryption capable Disk packs and FC Disk packs available 750 GB, 1 TB SATAII 7,200 rpm 73 GB Solid State Disk (SSD)
RAID	Levels 0, 1, 3, 5, 6, and 10
Storage partitions	8, 16, 32, 64, 128, 256, 512 storage partitions - choice required Upgrades in steps available: 8–16, 16–64, and so on
Maximum drives supported	All models: 448 FC/FDE/SATA drives (using 28 EXP5000 Expansion Units) Maximum of 480 SATA drives (using 8 EXP5060 High Density Drawers) Maximum of 20 Solid State Disks per Model
Fans and power supplies	Dual redundant, hot-swappable
Rack support	19-inch, industry-standard rack - 4U

For the latest specification information, including Interoperability Matrix, see the website:

<http://www.ibm.com/systems/storage/disk/ds5000>

3.6 IBM System Storage Model DCS9900

The IBM System Storage DCS9900 Storage System is designed for applications with high-performance streaming data requirements served by Deep Computing systems, IBM System p High Performance Computing (HPC) systems, System x 1350™ Linux Cluster systems, and Broadcast Video applications.

Applications such as those in HPC, rich media, life sciences, and government research require high-performance, reliable access to streaming data and extreme capacity and density to simplify management and reduce cost. Examples of such applications include weather forecasting, seismic processing, reservoir modeling, high definition (HD) creation and distribution, proteomics, structural analysis, fluid dynamics, and gene sequencing. With its parallel storage solution architecture, the DCS9900 is specifically designed to address those needs.

The DCS9900 brings many advantages:

- ▶ Performance: Industry leading streaming I/O performance (up to 5.7 GB/s)
- ▶ Density/Scalability: 1200 TB in two Racks
- ▶ Energy efficiency: Fewer Controller/Powersupplies/Fans and Sleep mode enabled
- ▶ Availability and Reliability: Hardware based RAID 6...8+2...with no performance penalty, Parity computed on every read – no SATA silent corruption errors, and all data remaining visible to all clients even in the case of a disk, enclosure or controller failure; parallel architecture at every level

Orders: All orders for the DCS9900 Storage System, which includes the DCS9900 Controller Model CP2, DCS9900 Storage Expansion Units, and features, must first be approved by the mandatory I-listed RPQ 8S0870. Ask your IBM representative how to order the I-listed RPQ.

Figure 3-6 shows the front view of the DCS9900.



Figure 3-6 IBM System Storage DCS9900

3.6.1 DCS9900 storage system

The following two models are available:

- ▶ DCS9900 Controller Model CP2 (1269-CP2)
- ▶ DCS9900 3S1 SAS/SATA Storage Expansion Unit (1269-3S1)

The DCS9900 Storage System initial offering consists of the DCS9900 Controller Model CP2 and the DCS9900 3S1 SAS/SATA Storage Expansion Unit. The DCS9900 Controller Model CP2 attaches to the DCS9900 Storage Expansion Units using twenty 3 Gbps SAS v1.0 disk expansion ports and is designed to scale up to 1200 disks, providing up to 1200 terabytes (TB) storage capacity in one DCS9900 controller system.

The DCS9900 Storage System consolidates and fully integrates a parallel, non-blocking architecture which enables industry leading data streaming performance. The DCS9900 is designed to provide optimal block level and file system performance with up to 6 gigabytes-per-second sustained data streaming bandwidth delivered to deep computing applications.

3.6.2 DCS9900 features

The DCS9900 controller model CP2 has the following features:

- ▶ The DCS9900 Controller Model CP2 (1269-CP2) comes in a 4U rack-mount enclosure containing dual controllers with 5 GB cache (2.5 GB cache per controller) and eight 8 Gbps Fibre Channel host ports. The DCS9900 controller attaches to DCS9900 3S1 SAS/SATA Storage Expansion Units using 3 Gbps SAS v1.0 disk expansion ports. The DCS9900 Model CP2 controller can support up to 1200 disks, yielding up to 1200 TB physical storage capacity, 64 thousand LUNs, and up to 4096 direct host logins per DCS9900 system.
- ▶ Infiniband host connectivity delivers high speed host connectivity option for attachment of host systems to the DCS9900 Storage System, it complements Fibre Channel host connectivity and offers host cables in several convenient lengths.
- ▶ Improved temperature reporting, fault handling, and ability to shut down all drives with one command are provided.
- ▶ The DCS9900 controller is designed to support managed Quality of Service to provide uninterrupted data delivery, as well as source all of its performance from multiple target LUNs or a single target “PowerLUN” reducing the need for host-based striping software.
- ▶ The DCS9900 controller incorporates enterprise-class data protection with on-the-fly parity checking on all read I/Os, as well as hardware-enabled RAID 6, which protects data in the event of double disk failure in the same redundancy group, without adversely affecting data availability or system performance.
- ▶ Additionally, the DCS9900 includes block level virtualization, to virtualize storage deployment and system management, through LUN aliasing, WWN masking/filtering, or port zoning.
- ▶ The DCS9900 Controller Model CP2 attaches to DCS9900 Storage Expansion Units to form the following supported, orderable configurations:
 - One DCS9900 Model CP2 Controller attached to five DCS9900 3S1 SAS/SATA Storage Expansion Units:
 - Range from 150 to 300 disk drives (in increments of 10 disk drives)
 - Scales from 21.9 to 300 TB physical storage capacity
 - One DCS9900 Model CP2 Controller attached to 10 DCS9900 3S1 SAS/SATA Storage Expansion Units:
 - Range from 150 to 600 disk drives (in increments of 10 disk drives)
 - Scales from 21.9 to 600 TB physical storage capacity
 - One DCS9900 Model CP2 Controller attached to 20 DCS9900 3S1 SAS/SATA Storage Expansion Units:
 - Range from 600 to 1200 disk drives (in increments of 10 disk drives)
 - Scales from 87.6 to 1200 TB physical storage capacity
- ▶ Each DCS9900 Storage System must be ordered with the minimum of number of disk drives, and then can be incremented an additional ten drives at a time, applied to the overall DCS9900 system configuration.

The DCS9900 3S1 SAS/SATA Storage Expansion Unit (1269-3S1) has the following features:

- ▶ The DCS9900 will support Intermix of SAS and SATA in the same enclosures.
- ▶ The DCS9900 will also support three Disk enclosure configurations.
- ▶ The DCS9900 3S1 SAS/SATA Storage Expansion Unit comes in a 4U, rack-mounted, 60 SAS or 60 SATA disk enclosure. The Model 3S1 internally accommodates up to 60 SAS or 60 SATA disk drives logically configured as a 1x60 or 2x30 connection to the DCS9900 Model CP2 controller, dependent upon whether the configuration will contain five 3S1 expansion units or ten/twenty 3S1 expansion units. The Model 3S1 includes hot-swappable power supplies and hot-swappable cooling fans.
- ▶ The Model 3S1 disk expansion unit integrates into a 45U 1050mm deep rack, rack feature (#2003). A single DCS9900 system can consist of a configuration made up of either one or two racks. The first rack can contain either five 3S1 expansion units or ten 3S1 expansion units. The second rack must contain ten 3S1 expansion units for a maximum configuration of up to twenty 3S1 expansion units and up to 1200 disk drives.
- ▶ An order for a configuration made up of five 3S1 disk expansion units must be ordered with a minimum of 150 disk drives (30 drives per 3S1 expansion unit), and scales in increments of ten up to a maximum of 300 disk drives (60 drives per 3S1 expansion unit).
- ▶ Each group of five 3S1 disk expansion units ordered must contain the same number and same type of disk drives and ordered in increments of ten drives applied across the entire configuration. For example, each of the five 3S1 expansion units in the rack can contain either 30-, 32-, 34-, 36-,.....54-, 56-, 58-, up to 60-disk drives. But all 3S1 expansion units must contain the same number and type of drives.
- ▶ Additionally, when field upgrading from five 3S1 expansion units to ten 3S1 expansion units, specify feature 9231, the resulting rack of ten 3S1 expansion units must contain the same number and type of disk drives in each expansion unit.
- ▶ An order for a configuration made up of ten 3S1 disk expansion units must be ordered with a minimum of 150 disk drives (15 drives per 3S1 expansion unit), and scales in increments of ten drives up to a maximum of 600 drives (60 disk drives per 3S1 enclosure).
- ▶ Each group of ten 3S1 disk expansion units ordered must contain the same number and type of disk drives and ordered in increments of ten drives applied across the entire configuration. For example, each of the ten 3S1 expansion units in the rack can contain either 15-, 16-, 17-, 18-,.....57-, 58-, 59-, up to 60-disk drives. But all expansion units must contain the same number and type of drives.
- ▶ A configuration made up of two racks of ten 3S1 expansion units attached to a DCS9900 controller can have a unique disk drive feature in each rack. Therefore, disk drive features must not be intermixed within each rack of five 3S1s or each rack of ten 3S1s, however, configurations made up of two rack of ten 3S1 expansion units can have a unique SAS or SATA disk drive feature in each rack.
 - 1000 GB SATA disk drive (#3203)
 - 450 GB/15K SAS disk drive (#3233)
- ▶ Every DCS9900 system location which has a DCS9900 system configuration that uses the 45U 1050mm deep Rack Unit (#2003), must have access to one Service Ladder (#2011). The Service Ladder is mandatory for installation, service, and maintenance of components within the upper drawers of the 45U 1050mm deep Rack Unit. Failure to have at least one Service Ladder available in a required location can result in delayed or prolonged maintenance times.

.Figure 3-7 shows the possible configurations of the DCS9900.

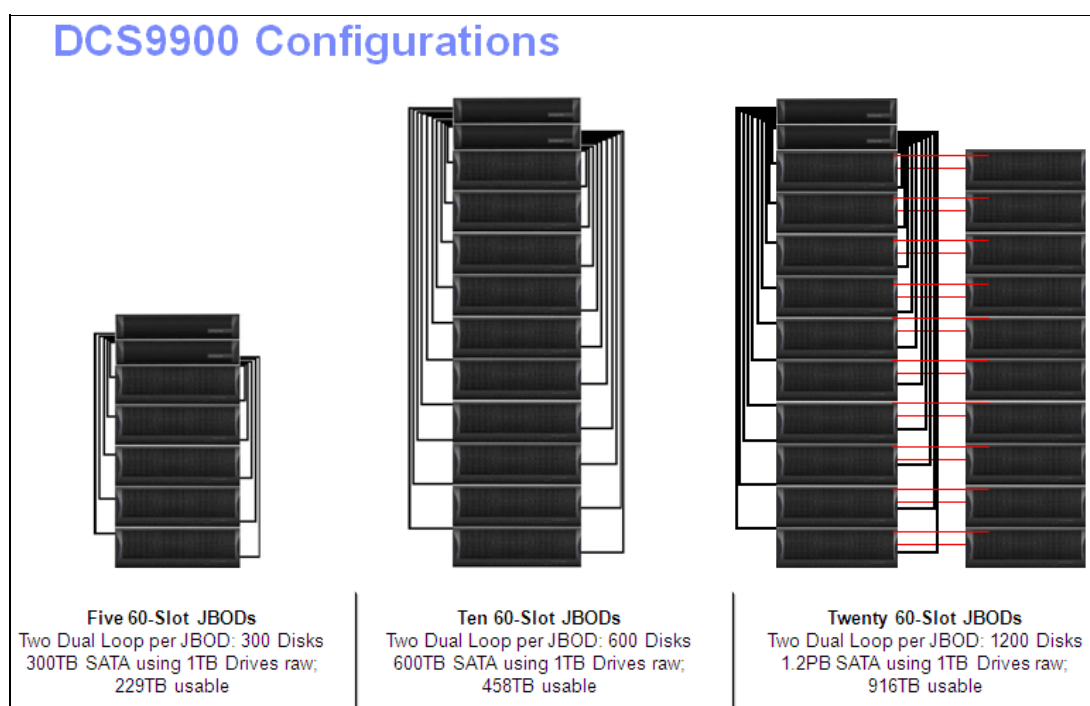


Figure 3-7 DCS9900 configurations

3.6.3 DCS9900 specifications

Table 3-4 shows the DCS9900 specifications.

Table 3-4 IBM System Storage DCS9900 specifications

Model	1269-CP2 and 1269-3S1
RAID controller	Dual active
Cache	Dual controller with 5 GB of RAID protected cache (2.5 GB cache per controller)
Host interface	Eight FC8 or IB 4X DDR host ports (4 ports per controller)
Drive interface	Twenty SAS 4-lane (3 Gbps) connections (10 per controller)
Supported drives	450 GB 15.000 rpm SAS and 1000 GB 7.200 rpm SATA
RAID	Hardware enabled RAID 6, which protects data in the event of double disk failure in the same redundancy group
Maximum # of supported LUN's	1024 LUNS, 512 concurrent logins (IB) and 1024 concurrent logins (FC) per DCS9900 couplet
Maximum drives supported	Maximum of 1200 SAS/SATA drives in 2 Racks
Fans and power supplies	2 redundant power supplies 3 redundant blowers with temp sensors, ensuring safe write-through and shutdown
Rack support	19-inch, industry-standard rack - 4U

For the latest specification information, see the website:

<http://www-03.ibm.com/systems/storage/disk/dcs/dcs9900/index.html>

3.7 IBM System Storage DS4000 EXP810 Expansion Unit

The EXP810 Expansion Unit is packaged in a 3U rack-mountable enclosure, supports up to 16 FC disk drives or E-DDM SATA drives. It contains 16 drive bays, dual-switched 4 Gbps ESMs, and dual power supplies and cooling components. Fully populated with 450 GB FC disk drive modules, this enclosure offers up to 7.2 TB of raw storage capacity or up to 16 TB when populated with the 1000 GB E-DDM SATA drives. This expansion unit is the only one that can be connected to every storage subsystem of the DS4000/DS5000 family. On the DS5000 family, the EXP810 can be attached for migration purposes only.

Through the proper firmware level, this expansion unit is able to host both FC and SATA Drives. Intermix of FC and SATA drives is supported within this expansion enclosure. The IBM System Storage DS4000 EXP810 storage expansion enclosure (Machine Type 1812, Models 81A and 81S) implements high-capacity, Fibre Channel disk storage. The expansion unit delivers fast, high-volume data transfer, retrieval, and storage functions for multiple drives to multiple hosts.

The EXP810 allows continuous, reliable service, using hot-swap technology for easy replacement without shutting down the system and support redundant, dual-loop configurations. Figure 3-8 shows the EXP810.

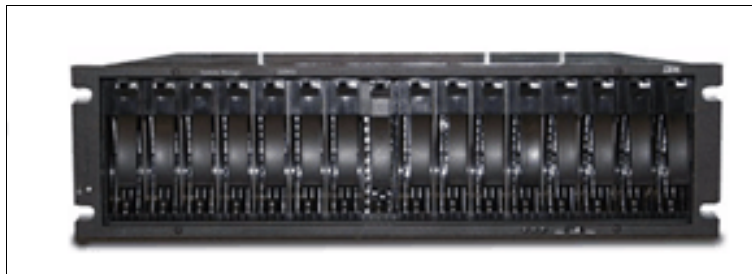


Figure 3-8 IBM DS4000 EXP810 Expansion Unit

3.7.1 EXP810 Expansion Unit features

The IBM System Storage DS4000 EXP810 expansion unit has the following features:

- ▶ Machine type: 1812, Model number: 81A/81S (Telco)
- ▶ 4 Gbps Fibre Channel and SATA drives currently available for the EXP810:
 - Supported drives 2 Gbps FC: 15K rpm, 146 GB/73 GB/36 GB (E-DDM)
 - Supported drives 2 Gbps FC: 10K rpm, 300 GB/146 GB/73 GB (E-DDM)
 - Supported drives SATA drives: 7.2K rpm, 1000 GB¹/750 GB¹/500 GB (E-DDM)
 - Supported drives 4 Gbps FC: 15K rpm, 600 GB²/450 GB²/300 GB/146 GB/73 GB (E-DDM)
- ▶ Attachment supported on models DS4300, DS4500, or DS4700 Express, DS4800 at V6.16 firmware and higher; attachment to DS5000 series for migration purposes only
- ▶ Can be intermixed with EXP710
- ▶ New design supporting up to 16 drives in 3U enclosure
- ▶ Up to 16 TB physical capacity per expansion unit using sixteen 1000 GB SATA disk drives
- ▶ Dual Environmental Service Modules (ESM) - 4 Gb FC
- ▶ Field replaceable components
- ▶ NEBS Level 3 certification and 48V dc Telco (model 1812-81S)

¹ Minimum required firmware level 06.23.05.00 or above

² Minimum required firmware level 06.23 and Storage Manager 9.23 or above

- Supported by IBM System p servers, IBM System x and selected Netfinity® servers, Selected Sun and HP UNIX servers, and Intel-based servers running Windows 2000, Windows Server 2003 SP2, Windows Server 2008 SP1, Linux, Novell NetWare, or VMware

For the latest configuration options, supported server models, supported operating system levels, supported system levels, and attachment capabilities, see the website:

<http://www-03.ibm.com/systems/storage/disk/ds4000/exp810/index.html>

Figure 3-9 shows the rear components of DS4000 EXP810.

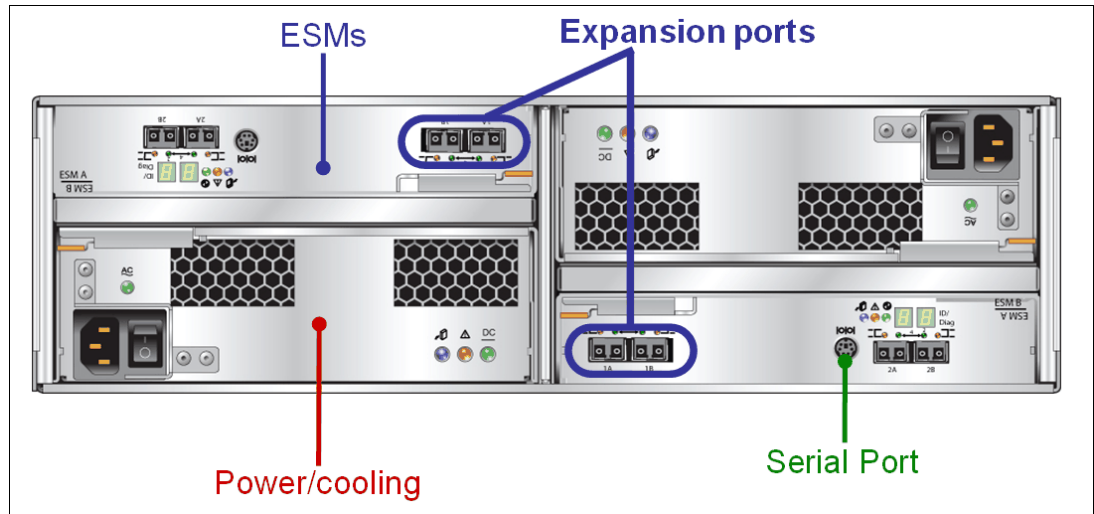


Figure 3-9 DS4000 EXP810 rear components

3.7.2 EXP810 Expansion Unit specifications

Table 3-5 shows the specifications of the IBM System Storage DS4000 EXP810 expansion unit.

Table 3-5 IBM System Storage DS4000 EXP810 Expansion Unit

Model	1812, -81A/81S
Number of drives supported	Up to 16 FC drives
Fans and power supplies	Dual redundant, hot-swappable
Dimensions	(With bezel) 12.95 cm H x 48.26 cm W x 57.15 cm D (5.1 in x 19 in x 22.5 in)
Weight	40 kg (88 lb)

3.8 IBM System Storage model EXP520 and EXP5000

The EXP520 and EXP5000 Drive enclosures are more than just-a-bunch-of-disks. They are designed to optimize performance, availability, and serviceability. Redundant 4 Gbps FC drive loops ensure complete accessibility to all drives in the event of a loop or cable failure, having redundant power supplies, cooling fans, and ESMs, while all primary components are hot-swappable CRUs and can be easily accessed and removed or replaced.

3.8.1 EXP520 and EXP5000 Drive features

The EXP5000 (1818-D1A) and EXP520 (1814-52A) storage expansion units are packaged in a 3U rack-mountable, high-capacity 16-drive bay enclosure containing dual switched 4 Gbps ESMs, dual power supplies, and redundant cooling. EXP5000s connect to DS5100 and DS5300 controllers through high-speed 4 Gbps FC disk expansion ports and have a physical storage capacity of up to 16 TB per enclosure, using 1000 GB SATA DDMs (disk drive modules) and up to 9.6 TB per enclosure, using 600 GB FC DDMs. It is also possible to mix SATA and FC drives in the same enclosure. The EXP520 uses the same disk types and sizes but only connects to DS5020 storage controllers.

The EXP5000 Expansion Unit (1818-D1A) and EXP520 (1814-52A) base model includes a 3U, rack-mount 16-bay disk enclosure, two SW SFP transceivers, dual power supplies, redundant cooling, rack mounting rails, soft copy documentation, and two rack PDU power cords (36L8886).

Figure 3-10 shows the EXP 520 unit.



Figure 3-10 EXP 520 front view

3.8.2 EXP520 and EXP5000 Drive support

Both expansion units support the following drives:

- ▶ 600 GB, 450 GB, 300 GB, 146 GB FC 4 Gbps, 15000 rpm (FDE and non-FDE)
- ▶ 1000 GB SATAII and 750 GB SATAII 4 Gbps 7,200 rpm
- ▶ 76 GB, FC, solid state drive (*only in EXP5000*)

For the latest specification information about the expansion units, see the website:

<http://www.ibm.com/systems/storage/disk/>

For the latest specification information, see the website:

<http://www.ibm.com/systems/storage/disk/ds5000/specifications.html>

3.9 IBM System Storage EXP5060

The IBM System Storage DS5000 series now offers an EXP5060 High Density Disk Enclosure. The EXP5060 can support up to 60 disk drives in a 4U unit, allowing a great reduction of footprint. Next to its lower infrastructure costs, with its high density format and amount of disks, it will deliver beneficial performance for HPC capacity intensive application environments, and with its efficient power supplies and variable speed cooling fans, it is a real energy saver.

Customers who will benefit most are the ones who:

- ▶ Have traditional HPC or GPC environments
- ▶ Have growing capacity needs
- ▶ Want to reduce power consumption
- ▶ Want to maximize density to reduce footprint in the data center

The EXP5060 comes in a 4U rack-mount enclosure and accommodates up to sixty 1000 GB 7.2K SATA DDMs. Up to eight EXP5060 Disk Enclosures are attachable to a DS5100 or DS5300 controller; providing up to 480 TB of physical storage.

3.9.1 EXP5060 overview

The IBM System Storage DS5000 series is expanded to include an EXP5060 High Density Disk Drive Enclosure designed to accommodate up to 60 disk drive modules (DDM) in a 4U rack-mount enclosure.

DS5100 and DS5300 models support attachment of up to eight EXP5060 enclosures. A configuration composed of a DS5100 or DS5300 controller attached to eight EXP5060 disk enclosures provide up to 480 TB of physical storage, all of which can be housed in a single IBM 2101 Model 200 Storage Solutions rack system.

To help support the integration of EXP5060 enclosures into 2101-200 rack systems, new power cord features are offered for the 2101-200 rack. These power cords provide connections from the rack power distribution units (PDUs) within the rack enclosure to applicable 240 V ac power sources in the data center.

3.9.2 Key prerequisites

DS5100 and DS5300 controllers must be at firmware level 07.60, or later to support new EXP5060. Initial plant orders (IPOs) for the EXP5060 High Density Disk Enclosure must include a minimum of two 10-Pack EXP5060 1000 GB/7.2K SATA DDMs feature (#4740).

When EXP5060 enclosures are to be plant or field integrated into the IBM 2101 Model 200 Storage Solutions Rack, the rack must be configured with PDU power cords rated for 240 V AC electrical service using either of rack power cord options 9491, 9492, 6491, or 6492. See the *2101 Model 200 Rack Sales Manual* for details.

Figure 3-10 shows the top view of the EXP5060 unit.

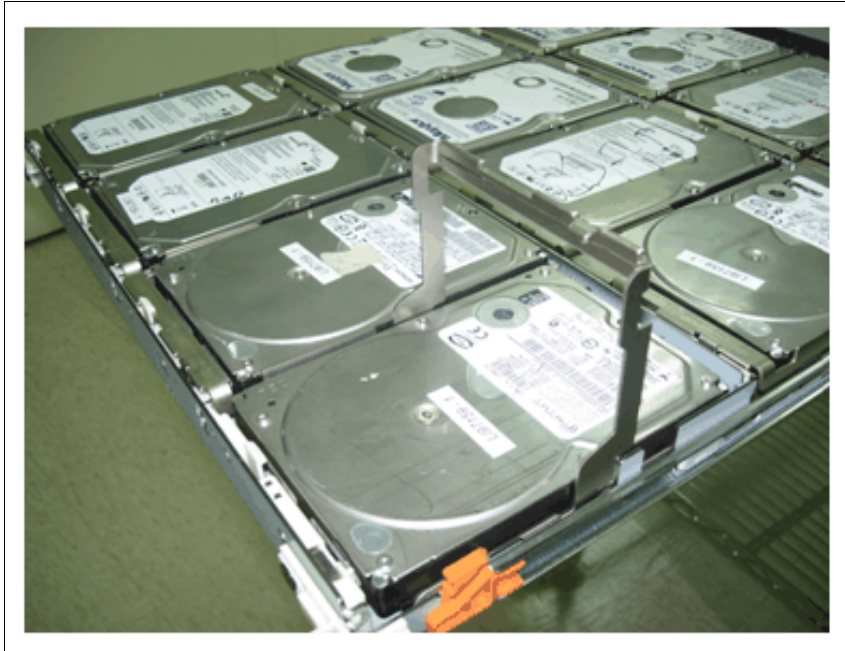


Figure 3-11 EXP5060 - High Density Drawer

For the latest specification information for the EXP5060, see the website:

<http://www.ibm.com/systems/storage/disk/>

3.10 Solid State Disk and Full Disk Encryption

In this section we explain the high-level concepts and advantages of the relatively new products, Solid State Disk and Full Disk Encryption.

3.10.1 Solid State Disk

This section summarizes the major advantages offered by Solid State Disk (SSD) when compared to spinning, magnetic hard disk drives.

Figure 3-12 shows an SSD and its dimensions.

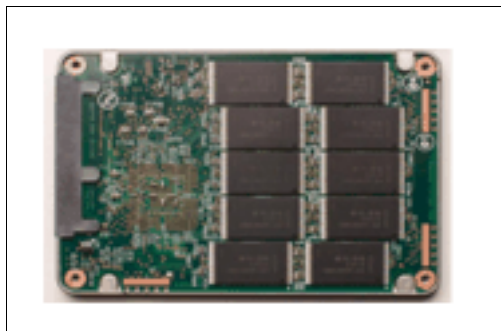


Figure 3-12 SSD

The simple fact that there are no moving parts (disk platters, magnetic heads, motor) in an SSD results in:

- **Faster data access and throughput:**

The access to the data with an SSD is faster because again, there is no read/write head to move and no magnetic platters need to spin up (no latency). On an SSD, the data can be read almost immediately. The SSD has up to 100x more throughput and 10x better response time than 15K RPM spinning disks. High transactional environments benefit most of SSD technology.

- **Better reliability:**

Again, the lack of moving and rotating parts almost eliminates the risk of mechanical failure. SSDs have the ability to tolerate extreme shocks, higher altitude, vibration, and extremes of temperature. However, they can still fail and must be RAID protected like traditional drives.

- **Less power consumption:**

Because there is no power for the motor required to spin up the magnetic platters and to move the heads, the drive uses less energy than a traditional hard disk drive. Each SSD uses about half of the power of a 15K RPM HDD. The savings can be substantial if a few SSDs can replace many HDDs to deliver the same performance. This is particularly true for applications that were forced, for performance reasons, to use large quantities of HDDs to get as many spindles as they can, while only using a small portion of the disk capacity. Besides power consumption savings, the overall system weighs much less because SSDs are already much lighter than HDDs.

Considerations for using SSDs

In this section we review special considerations and general usage guidelines to follow when SSDs are used:

- **Copy Services:**

Copy Services operations are not specifically affected by SSD. However, it is best (as is the case with HDDs as well) not to use a slower device as the target of a copy services function.

- **Sparing and rebuild time:**

The sparing algorithm is unchanged for storage systems equipped with SSDs. The rebuild time for an SSD is, however, much faster than with regular HDDs of the same capacity. For an idle 6+P SSD array, we have measured a rebuild time that is around 10-15% faster than a 6+P array with 15K RPM HDDs.

SSD usage

SSDs cost more per unit of storage than HDDs but the gap in price is expected to narrow over time. Also SSDs offer only relatively small capacity per drive in comparison to HDDs. You thus want to make sure that they are used (reserved) for the most business critical applications and that they are used under conditions that can guarantee optimum performance. Note that any update of the SSD firmware must be done in non-concurrent mode. This means that any application IO to the drives must be stopped prior to the upgrade.

Use the following guidelines when deciding:

- SSDs currently remain targeted at applications with high access workloads (IOPS). Applications with high IO rates and poor cache hit ratios are ideal candidates for SSDs while applications with a good cache hit ratios or low IO rates just remain on spinning drives.
- Also, sequential workloads (as opposed to random) stay on spinning disks.

- ▶ Given the current high cost per GB for SSDs, use a mix of SSDs and HDDs (not on the same DA), such that data from less critical applications can be stored on HDDs (you can use FC and SATA drives).
- ▶ Do not use SSDs as a source of a Remote Mirror or FlashCopy where the target is a regular HDD.
- ▶ If you run an I/O driver to a single SSD that is set up for direct attach, it can do up to 30k IOPS. The DS8000 for example can do about high IOPS across seven or eight disks on one device adapter. This is close to double on a device adapter pair. The throughput can scale up as more device adapter pairs are used for SSDs. Even the most active production data cannot exploit the range of 30k IOPS making use of 73 GB or 146 GB disk.

In other words, the SSDs are so fast that the Device Adapter becomes the performance bottleneck on random workloads. For sequential workloads the DA was already the bottleneck with traditional HDDs (hence using SSDs for sequential IO is not a good investment).

- ▶ Use optimum host attachment options. The functionality provided by a storage server such as virtualization layers that make the particularities of disk hardware transparent to applications is critical for the usability of the system. However, this functionality adds to the response time that you normally get from an SSD. It is important not to further impact the response time that you use the host attachment options that guarantee the best performance.

Attention: You need sufficient amount of SSD to gain the best performance benefit. Indeed, applications such as databases can have high access rates, which makes them good candidates for deployment on SSD. However, the same database might require a large storage capacity. In this case you thus need a sufficient amount of SSD capacity to get the best performance improvement. To help customers get insight of their current Storage Data behavior and clarifying which data is expected to benefit from SSD, IBM has developed “*Smart Data Placement Tools*.”

For specific information about SSD usage with DB2, see the *Ready to Access DB2 for z/OS Data on Solid-State Drives*, REDP-4537.

For more SSD performance related information, see the website:

<http://www.ibm.com/systems/storage/solutions/ssd/>

3.10.2 Full Disk Encryption

This section summarizes the major advantages offered by Full Disk Encryption (FDE) when compared to regular hard disk drives.

Preventing data breaches is becoming more and more important, as they appear more frequently, and not seldom make the headlines in newspapers or Internet. This can have a severe impact on companies.

The negative impact of data breaches are:

- ▶ Loss of customers and revenue
- ▶ Unplanned expenses of remedy a breach
- ▶ Legal implications, penalties and fines
- ▶ Negative press and tarnished reputation
- ▶ Lost goodwill and undermining other corporate relationships

The difficulty today is to keep track of “data at rest” because eventually, the data will be out of the administrator’s control, and therefore vulnerable to a breach. When the storage system leaves the building, frequently sensitive client/partner data or intellectual capital resides on it.

Security technology efforts are fragmented across point products, leaving holes in the security system. Also, other current methods fall short. Because drive disposal methods are subject to human error, deleted files can be reinstated and password protected data can still be accessed by skilled individuals.

Because data security is now a common component of the corporate landscape due to necessity and regulatory compliance, there is a need for new technologies that can enable companies to comply. FDE provides such a solution.

FDE value propositions

FDE offers the following benefits:

- ▶ Data is secure at the most basic level – the hard drive:
 - It covers most data exposures and vulnerabilities all at once.
 - Security planning is simplified.
- ▶ Unparalleled security assurance with government-grade encryption is provided:
 - It is National Security Agency (NSA) qualified.
 - Advanced Encryption Standard (AES) 128 encryption algorithm is approved by the United States government for protecting secret-level classified data.
- ▶ Data is protected through the hardware life cycle:
 - FDE enables return of defective drives for servicing.
 - You can quickly decommission or re-purpose drives with instant secure erase.
- ▶ Performance is preserved:
 - FDE will not slow a system down; the encryption engine matches the drive’s maximum port speed and scales as more drives are added.

In order for drive-level encryption to provide this value, the key that enables encryption must be protected and managed.

Figure 3-13 shows an FDE setup.

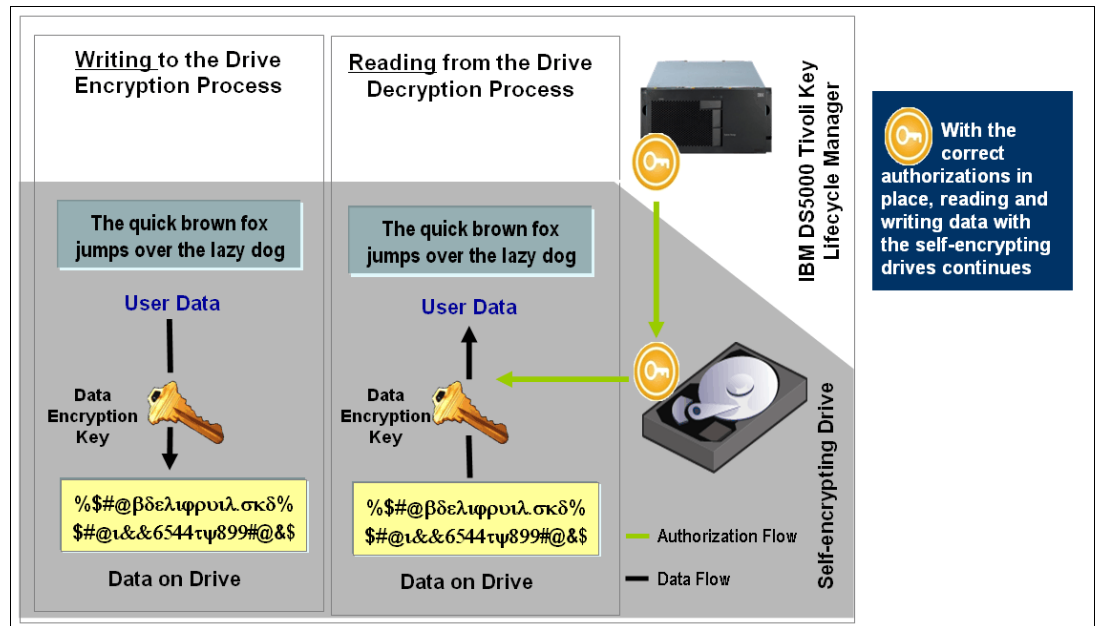


Figure 3-13 Locked, secured Data with self encrypting drives

Components of FDE solution

The IBM Full Disk Encryption (FDE) solution has two components:

- ▶ FDE Hard Disk Drives (HW):
 - Encryption is performed at HDD level.
- ▶ Tivoli Key Lifecycle Manager (TKLM SW):
 - With the FDE drive, it generates and encrypts a security key.
 - It creates a unique security key ID that is paired with the security key.
 - It adds its WW identifier and a randomly generated number
 - The security key ID is saved – this folder location will be needed whenever a security operation requires (such as when a drive powers up).
 - It creates a backup of the security key and the security key identifier.
 - A secure backup is provided in which the security key and the security key identifier are encrypted utilizing a user-selected pass phrase.

Figure 3-14 shows the classification of data levels.

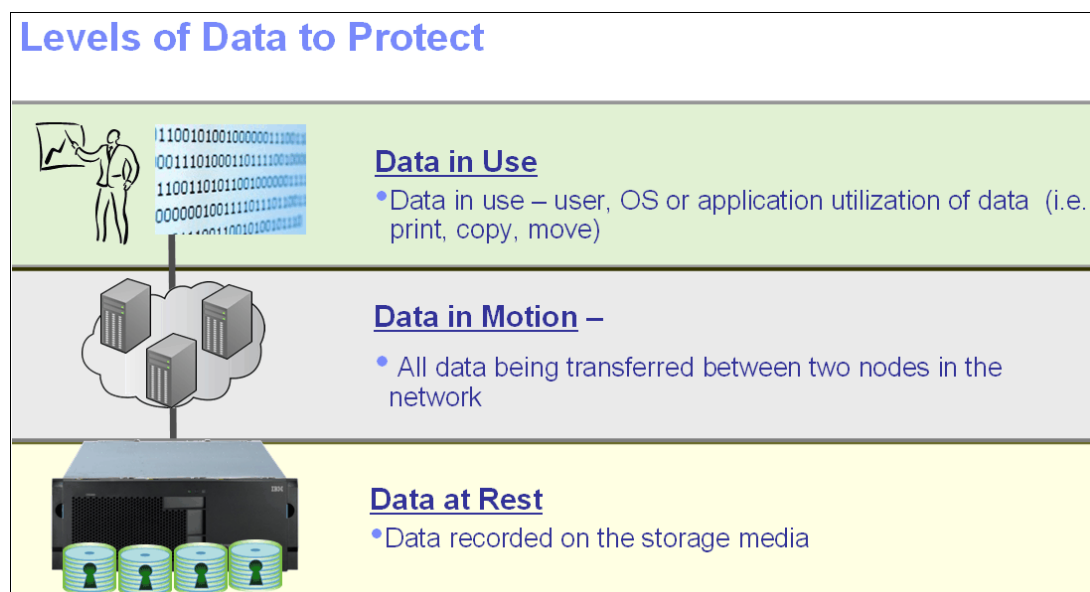


Figure 3-14 Levels of Data to protect

Reasons to classify data for encryption

Here are some reasons to classify data for encryption:

- ▶ Encryption of the entire data center requires a large investment:
 - Utilization of earlier investments of non-FDEs
 - Utilization of SATA drives
- ▶ Instead, utilize available FDEs for the most sensitive high-security data:
 - Determine the associated risk with the potential breach of that data.
 - Determine what data is regulated and must be encrypted.
 - Prioritize data accordingly.
 - Cut in FDE drives as storage is routinely replaced.
 - Enterprise businesses can choose to encrypt only their regulated data, because managing the keys and the overall encryption process remains complex.
 - Small and medium-size companies can consider encrypting just about everything to ease the management challenge of determining what to encrypt.

3.11 IBM System Storage DS Storage Manager

The IBM System Storage DS Storage Manager software is used to configure, manage, and troubleshoot the Disk System. Its front-end component, the Storage Manager client, is a Java™ based management tool that is available for various operating systems.

This tool is used primarily to configure RAID arrays and logical drives, assign logical drives into storage partitions, replace and rebuild failed disk drives, expand the size of the arrays and logical drives, and convert from one RAID level to another. It allows troubleshooting and management tasks, such as checking the status of the Disk System components, updating the firmware of the RAID controllers, and managing the Disk System. Finally, it offers advanced functions such as FlashCopy, Enhanced Remote Mirroring and VolumeCopy (these are premium features that need to be purchased).

IBM DS Storage Manager(SM)10.60 is the most recent version available at the time of writing this book. SM 10.60 supports controller firmware versions 4.01.xx.xx - 07.60.xx.xx. DS4000 Storage Manager 10.60 (with controller firmware 06.16.xx.xx and above) provides support for attachment of the DS4000 EXP810 storage expansion enclosures to DS4800 storage subsystems. IBM DS SM 9.16 and above supports automatic ESM firmware synchronization with the new EXP810 storage expansion enclosure. When you install a new ESM into an existing EXP810, in a DS4000 storage subsystem that supports automatic ESM firmware synchronization, this feature resolves firmware mismatch conditions by automatically synchronizing the firmware in the new ESM with the firmware in the existing ESM.

Figure 3-15 and Figure 3-16 show the Summary and the Logical tabs of the Storage Manager with a DS5000 storage subsystem.

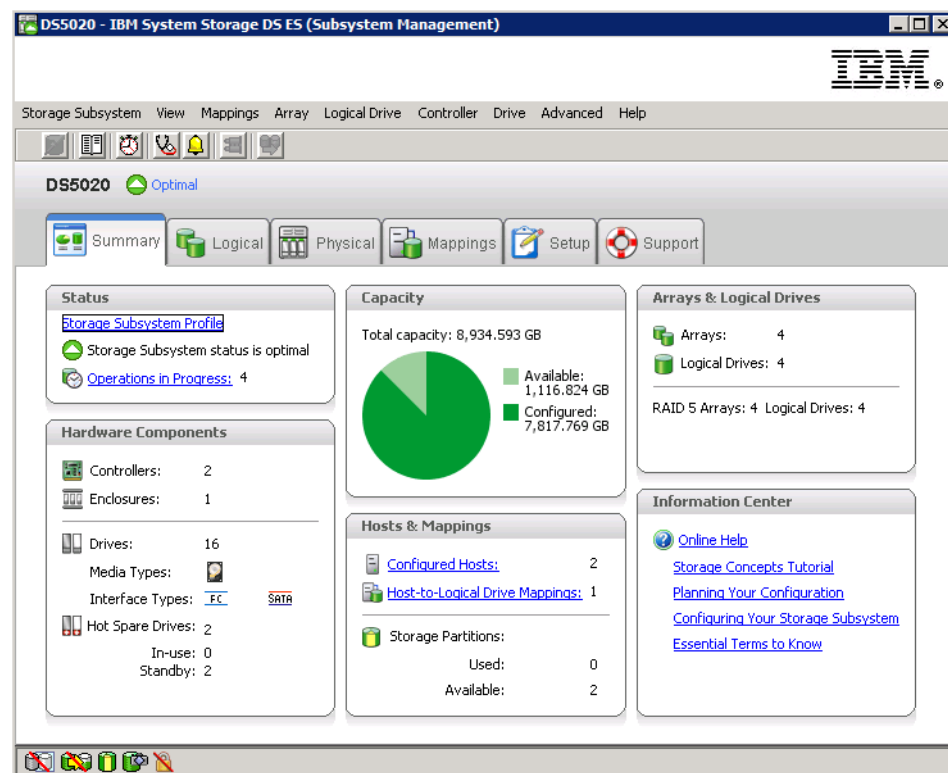


Figure 3-15 Storage Manager Example - Summary

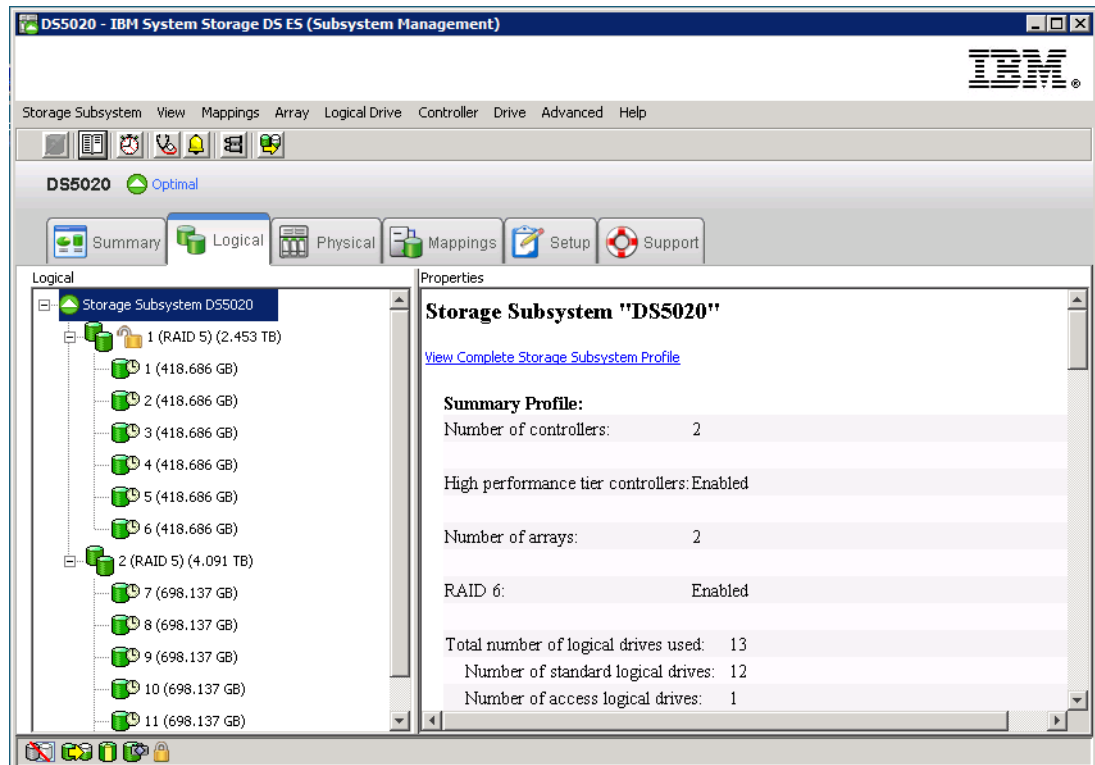


Figure 3-16 Storage Manager Example - Logical

Next we describe the latest enhancements to the Storage Manager.

IBM DS Storage Manager Version 10.60 (controller code Version 7.60) includes all the functions already available in previous Storage Manager releases, and also offers several significant new features. The previous enhancements are listed at the end of this section.

The controller code (firmware) Version 7.60 is based on a new layered code design that provides better scalability and flexibility for future code enhancements. As part of the new design, the configuration database (DACStor region) has also been restructured. This new structure provides the foundation for supporting new features such as greater than 2 TB logical drives, RAID-6, and an increased number of logical partitions.

The new 7.60 firmware is only supported on the DS4200, DS4700, DS4800, DS5020, and the DS5100/DS5300. Former DS4000 models (such as the DS4500 or DS4300) must remain at the latest compatible 6.xx firmware. Those former models can, however, be managed from Storage Manager V10.60 client.

3.11.1 New features: October 2009

New features discussed in this section were first introduced with Storage Manager V10.10 (firmware V7.10) and are applicable to Storage Manager V10.60 (firmware V7.60) as well.

The following features have already been described in this book:

- ▶ Memory upgrade for DS5000 series
- ▶ iSCSI support for DS5000
- ▶ SSD support
- ▶ EXP5060 support
- ▶ 8 Gb Host interface cards
- ▶ 448 HDD support on DS5100/5300

Here we describe other versions (up to October 2009):

- RAID-6 support:

RAID-6 technology is designed to provide improved data protection against multiple disk drive failure. RAID-6 uses a double parity check implementation (designated as p+q). Dynamic RAID-level migration is supported but might require capacity expansion prior to RAID migration.

Attention: RAID-6 is not supported on the DS4800. This feature is available only on DS4200 and DS4700. This is because various RAID-6 parity calculations rely on electronic components available only with the DS4700 and DS4200 systems. This hardware-based implementation offers better performance than the RAID-6 software-based implementation used by other storage vendors.

- Support for more than 30 drives for RAID-0 and RAID-1:

A logical drive can be created using all the drives in the storage system. Although not a guideline, you can now create a logical drive using a possible maximum of 224 drives for a DS4800 or 112 drives for a DS4700 and 448 for a DS5100/5300.

- Support for greater than 2 TB logical drives:

You can now create greater than 2 TB (terabyte) logical drives. If you combine this feature now also with the possibility to support more than 30 physical drives (assuming RAID-0 or RAID-1), and if you use 500 GB drives in a DS4800 with its maximum of 224 disk drives, you can create a 112 TB logical drive. Keep in mind, however, that the operating system (OS) making use of the logical drive can impose other limitations.

Important: Because of the design changes, upgrading the firmware to Version 7.15 is a non-concurrent upgrade and it must be done offline. See “Upgrading to firmware level7.xx” on page 668 in *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010 for details.

- Increased maximum number of partitions:

The maximum number of supported partitions for DS4200, DS4700, and DS4800 with firmware versions older than 7.10 was 64. This number is increased to 128 for DS4200, DS4700, and DS5020 and to 512 for DS4800 and DS5100/5300.

- Increased queue depth:

The controller queue depth is increased to 4096 for DS4800 and DS5100/5300 and to 2048 for DS4200, DS4700 and DS5020.

- Support bundle collected on critical events:

The event monitor on the client system, upon occurrence of a critical event, saves diagnostic data (support bundle) to the local hard drive of the client system in the same area already used for other recovery information. This information is stored and is not overwritten for 72 hours. Also, email notification can be set up for all critical events.

- Use diagnostic data capture for additional failure events:

The goal of this feature is to capture enough information about the controller state at the time of an unusual event and to store that diagnostic data for later retrieval.

- Increased number of GHS (Global Hot Spares):

The new controller code allows the creation of unlimited Global Hot Spares. There are now three ways to create GHS drives:

- Automatic assignment of hot spares
- Automatic assignment by specification of number of drives
- Explicit assignment by specification of a list of drives

More information: See “Hot spare drive” section in *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010 for details.

- Un-configuring Storage Subsystem and arrays:

Storage Manager 10.10 (controller code 7.10) and above allows you to clear the storage subsystem or previously created arrays if you must do so. Clearing the storage subsystem completely removes all configurations of the storage subsystem. Clearing the array configuration clears logical drives and group configuration in all arrays while leaving the remaining subsystem configuration intact.

- Drive migration warning:

This feature allows the user to manually intervene and permits safer migration scenarios. (When migrating configured logical drives to an existing storage subsystem, there is a warning given if the number of the logical drives being added will overrun the maximum number of drives allowed on this particular subsystem.) This feature thus prevents users from being allowed to import more logical drives than a particular subsystem can support.

- Configurable failed drive replacement:

You can now designate a hot spare to be a permanent member of a logical drive, thus eliminating the copyback portion after the failed drive has been replaced. The replacement drive in that case is in an unassigned state.

- Enhanced performance feature:

This is a premium feature and requires a premium key to enable higher-performance on a DS4800 Mod 80 and DS5100. Application of this key is an off-line operation (requiring a quiesce of the storage system and a subsequent reboot of the controllers).

- RAID redundant data verification prior to read:

This feature is supported on all RAID levels and enables verification of redundant RAID data consistency before returning the requested read data back to the host. This must be enabled on a logical drive basis. The function is automatically disabled when a logical drive becomes degraded. If a mismatch is detected, the system returns a media error (unrecoverable read error (03/11/00)). The system does not attempt to correct or recalculate parity automatically.

- 8 K cache block size:

Cache block size dictates the size of individual data buffers within the cache. In addition to the previously possible 4 K and 16 K cache block sizes, the new firmware V7.10 and later also adds the choice for an 8 K block size.

- Increased number of FlashCopies per logical drive:

The maximum number of FlashCopies per logical drive that is supported on DS4800 and DS5100/5300 is 16, instead of 4 with previous versions of the firmware. For DS4200, DS4700, and DS5020, the maximum number of FlashCopies is 8 per logical drive.

- Increased number of mirrors supported:

With firmware V7.10 and later, the number of mirrors supported on DS4800 and DS5100/5300 increases to 128 and, on DS4200, DS4700, and DS5020 to 64.

- ▶ Icon for unmanageable controllers:
A new icon is available in the DS Storage Manager GUI to differentiate between unresponsive and unmanageable controllers. An unmanageable controller means any storage subsystem that has no corresponding configuration file.
- ▶ New multipathing device driver support: SDDDSM:
SDDDSM is a multipath I/O solution based on Microsoft MPIO technology. The SDDDSM driver also supports other storage solutions from IBM such as IBM System Storage DS8000 and IBM SAN Volume Controller (SVC). It is now also available for the DS4000 and DS5000 systems.
- ▶ Windows RDAC no longer supported:
RDAC for Windows is not supported with this new version of Storage Manager (V10.10 and later). For multipathing and failover support, you must install the SM failover driver (MPIO/DSM) or the new, separately provided, SDDDSM driver (RPQ).
- ▶ Enhanced cache support:
The cache memory management has been improved. All cache is now read/write. Previously only the first 1 GB was read/write all other was read only. Additional statistics gathering by GUI and CLI Controller response and performance along with additional drive statistics by the GUI and CLI option is available now.
- ▶ Additional operating system support includes:
 - MS Vista client support
 - MicroSoft Windows Server 2008
 - Support Red Hat v5.2
 - HP-Ux v11.23 full solution
 - Solaris 10 x86 support
 - SLES 10 SP-2
- ▶ Enhancements in Version 10.30: In summary, with the Storage Manager 10.30, IBM introduced the following major new features:
 - 1,000 GB: 7.2K rpm, SATA (E-DDM) disk drive. Using 1,000 GB disk, the capacity can go up from 12 TB to 16 TB per enclosure.
 - DS4000 SNMP MIB file change to support SNMPv2.
 - IPv6 support on the Ethernet management ports for DS4800, DS4700, and DS4200.
 - Set the number of supported host logins on the DS4200, DS4700, and DS5020 controller to 640 and to 2048 for the DS5100/5300.
 - Windows Server 2008, RHEL 5.2, SLES 10 SP2, HP-UX 11.23, and Solaris 10 on x86 support.
 - IBM SDD support on Windows and HP-UX.

3.11.2 Previous announcements: 2004

In this section we summarize the previous enhancements:

- ▶ Support for mixed drive configurations:
It is now possible to intermix Fibre Channel (FC) drives and SATA drives in one Disk System. The flexibility to intermix various drive types (FC and SATA) in one Disk System gives you the ability to use the advantages of both drive technologies. For example, it is now possible to have the primary (production) storage on FC drives and the secondary (near line or archiving) storage on SATA drives without the need of having different, separate Disk Systems. Or the FlashCopy or VolumeCopy of an array made of FC drives can be created on cheaper SATA drives.

► Support for Microsoft VDS/VSS:

Microsoft has defined an interface layer for the purpose of integrating storage system functions with the Windows Server 2003 operating system. The ultimate aim is to integrate snapshot and capacity management functions into applications, such as Microsoft Exchange. within their applications. Storage Manager 9.12 or higher supports VDS/VSS.

– Volume Shadow Copy Service (VSS):

- Provides Point-in-Time (PiT) copies
- Works with applications to suspend for a few seconds
- Uses snapshot (FlashCopy) services from the storage subsystem

– Virtual Disk Service (VDS):

- Common disk interface for block storage
- Common storage management across vendors
- API for third party management tools

► Task assistants:

The task assistants provide a task-oriented starting point for carrying out common storage management operations, helping to simplify storage administration and reduce TCO.

With the Enterprise Management Window Task Assistant, you can add, name, or rename storage subsystems, set up an alert destination, and manage a storage subsystem. This last option launches the Subsystem Management Window Task Assistant. With this assistant, you can configure a storage subsystem, define hosts, create new storage partitions, map additional logical drives, and other common tasks.

► New RDAC driver architecture:

The new RDAC driver, sometimes referred to as the Multi-Path Proxy (MPP) driver is a new multipath driver architecture that is used in the new versions of multipathing drivers (also known as RDAC) for Windows 2000/2003 and Linux. The driver was completely rewritten from scratch to include improvements and to be able to accommodate future planned enhancements in the device driver.

► Full command line interface:

Any task that can be performed with the Storage Manager GUI can now also be performed through the Command Line Interface.

► Controller service mode:

In previous versions of Storage Manager, it was not possible to do any problem determination on a “live” controller, because the controller was either online (processing IO) or failed (not processing IO). Now, if a controller is suspected as the root cause of a problem, it can be set into service mode and IBM support can run low-level shell commands to further analyze the errors.

► Support for VERITAS ASL 4.0:

Previous versions of the DS4000 Storage Manager support a component called the VERITAS Array Support Library (ASL), which is essentially a vendor-supplied plug-in for the VERITAS Volume Manager (also known as VxVM) Dynamic Multipathing Driver (DMP). VxVM 4.0 has defined new capabilities of the ASL that are supported in Storage Manager 9.10, and later by the feature known as “VERITAS ASL 4.0”.

3.11.3 Other features

In this section we describe the dynamic expansion features of Storage Manager:

- **Dynamic Volume Expansion (DVE):**

DVE is a modification operation that you can use to increase the capacity of a standard logical drive or a FlashCopy repository logical drive. To increase the capacity, a DVE operation uses the free capacity that is available on the logical drive group of the logical drive. The operation is considered to be dynamic because you can continuously access data on logical drive groups, logical drives, and disk drives during the operation.

- **Dynamic Capacity Expansion (DCE):**

DCE is the ability to increase the available free capacity on an array without needing to restart the host system, and is a very important feature. In today's IT environment, the need for storage space grows constantly. Many customers exhaust their existing space sooner or later and have to expand their storage capacity. It is essential that this process be non-disruptive and not cause any downtime. With DS4000 Storage Manager, you can simply add new disk drives to the Disk System and start the expansion procedure while the system remains fully operational. After the procedure starts, you cannot stop it. Be aware that you might see a performance impact because the expansion process competes with normal disk access.

Further dynamic volume/array reconfiguration such as changing the logical drive segment size (DSS) and changing the RAID level (DRM) are discussed in more detail in *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010.

3.12 Introduction to IBM Midrange Storage Copy Services

This section provides a high-level overview of the difference in functionality between FlashCopy, VolumeCopy, and Enhanced Remote Mirror features, which are supported on all of the midrange storage servers mentioned in this Redbooks publication.

Customers in today's IT world require good archives of their data as well as the ability to create these archives with minimal interruptions. The IBM Midrange Storage Subsystem helps to fulfill these requirements by offering three premium features for Copy Services: *FlashCopy*, *VolumeCopy*, and *Enhanced Remote Mirroring (ERM)*.

All of these three premium features can be used individually, or can be combined together to help with providing a solution to customers for designing a disaster recovery answer for their critical data protection needs. All of these premium features require license keys for activation. FlashCopy and ERM can be purchased individually, however VolumeCopy requires the purchase of a combined FlashCopy and VolumeCopy license key.

In the following section, we present a brief overview of these premium features and their capabilities. For more information, see *IBM Midrange System Storage Copy Services Guide*, SG24-7822.

3.12.1 FlashCopy premium feature

When customers are in need of doing their backups, they frequently want to do so without requiring long application backup windows. With the IBM Midrange Storage Subsystem, this capability is handled by using the FlashCopy feature.

A FlashCopy can be used to create a virtual logical drive that is a point-in-time (PiT) image of a real logical drive. The FlashCopy is the logical equivalent of its complete physical source copy, at the time created, but it is created in seconds and requires much less physical resources than a full copy of the source to another logical drive might take. Flashcopies can be used for performing backup images, to create a temporary test image and to create a source image for a VolumeCopy operation that can be copied without extended impacts to the production operations.

Using the FlashCopy also requires the use of less physical disk space, because it makes use of a repository logical drive for storing only data change operations that have been made since its creation. Because the FlashCopy is not a full physical copy, if the source logical drive is damaged, the FlashCopy logical drive itself cannot be used to recover it.

Figure 3-17 shows how FlashCopy works.

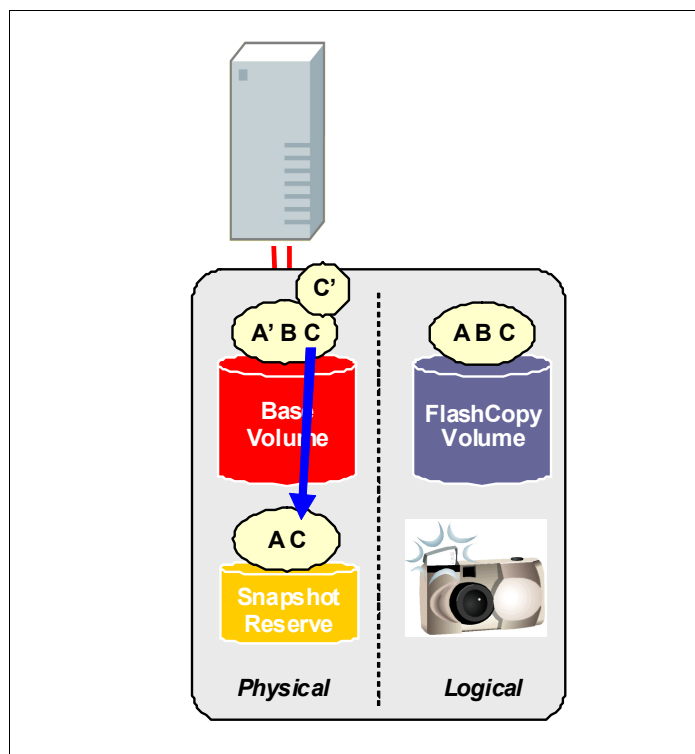


Figure 3-17 How DS4000/5000 FlashCopy works

The Storage Management software issues a warning message when the FlashCopy repository logical drive nears a user-specified threshold (a percentage of its full capacity; the default is 20 percent). When this condition occurs, you can use the Storage Manager software to expand the capacity of your FlashCopy repository logical drive from free capacity on the array.

3.12.2 VolumeCopy

The VolumeCopy premium feature is used to copy data from one logical drive (source) to another logical drive (target) to create a complete PiT image copy to a target logical drive within a single IBM Midrange Storage Subsystem. The target logical drive is an exact copy or clone of the source logical drive. This feature can be used to migrate data from one array to another. There is no requirement that the arrays be of a common RAID type, physical drive type or size, only that the target logical drive be of at least the same size in capacity as the source.

VolumeCopy is a full PiT replication of the source; which enables its use for the creation of a logical drive image to be used in analysis, data mining, and testing without placing additional overhead on the source production logical drives.

Figure 3-18 shows how VolumeCopy works.

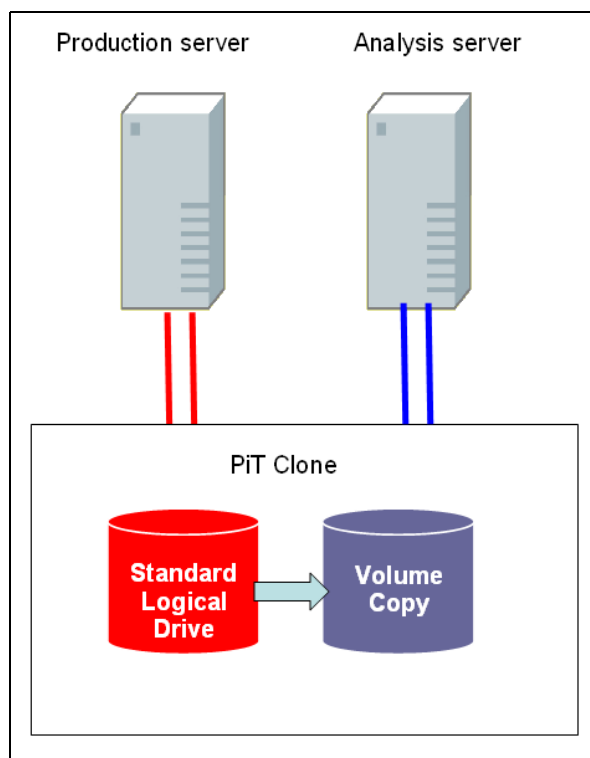


Figure 3-18 VolumeCopy

VolumeCopy does require the source logical drive to be quiesced while the target is being created so the image is not being changed while the copy is being done. To enable minimal impact here, this is best performed in conjunction with the FlashCopy feature.

Figure 3-19 shows how this process works.

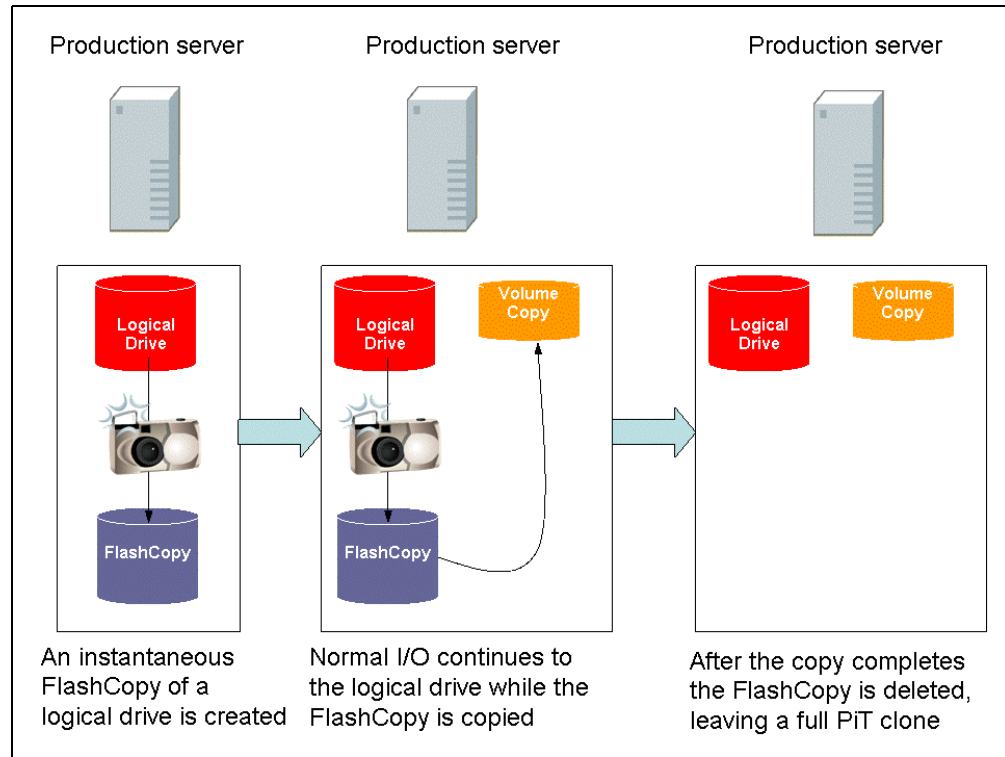


Figure 3-19 VolumeCopy integration with FlashCopy

3.12.3 Enhanced Remote Mirroring

Enhanced Remote Mirroring (ERM) or Remote Mirroring is the premium feature that enables the IBM Midrange Storage Subsystem to create a full image mirror of its logical drives on a second IBM Midrange Storage Subsystem, and keep both images synchronized with each other.

ERM is generally used by customers to migrate their data from one subsystem to another, or it is used as the remote mirroring engine for their disaster recovery solution (DR) between two site locations over varying distances.

The ERM option is used for online, real-time replication of data between storage subsystems over a remote distance (Figure 3-20). In the event of disaster or unrecoverable error at one storage subsystem, you can promote the second storage subsystem to take over responsibility for normal I/O operations. Currently, a high number of mirror relationships are supported per DS storage subsystem.

Figure 3-20 shows how ERM works.

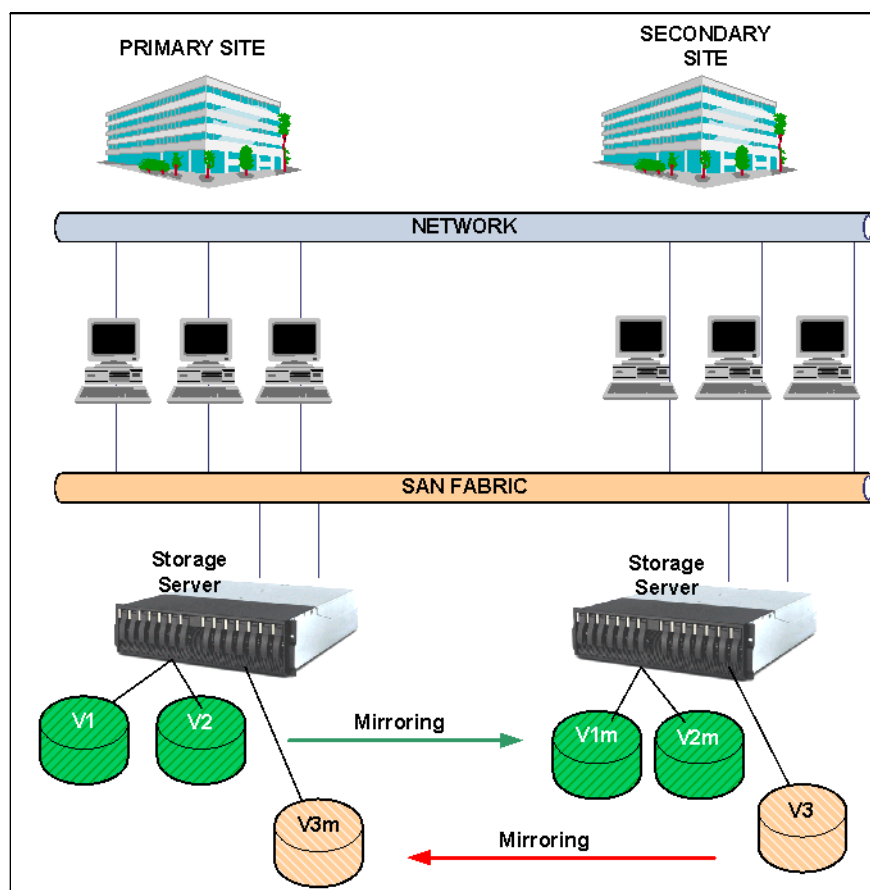


Figure 3-20 ERM

ERM supports three methods of creating the target images over various distances and networks. In the following topics, we describe the mirroring methods that are supported.

Metro Mirror method

Metro mirroring is a synchronous mirroring method, generally used for campus or local (up to 10 mile) distances. Any host write request is written to the primary (local) storage subsystem and then transferred to the secondary (remote) storage subsystem. The host application must wait until receiving acknowledgement that the write request has been executed on both (local and remote) storage controllers.

Global mirror method

Global mirroring is an asynchronous method of mirroring that also uses write order consistency (WOC) to ensure that the IO changes are sent to the secondary in the same sequence as they were processed on the primary. This method must be used when the host application has critical data to be mirrored, spread across several logical drives. This method writes all requests from a host to the primary (local) storage subsystem and immediately reports them as completed to the host. Doing this ensures that the inter-dependent write requests are carried out in the same order at the remote site as they were performed at primary (local) site so that it performs the mirroring writes in the same order that they were received. This is accomplished through the use of a consistency group that tracks the order in which the member logical drives' IOs were received and processed.

Global copy method

Global copy is a second asynchronous ERM write method. With this method all write requests from a host are written to its primary (local) storage subsystem and immediately reported as completed back to the host system as is done with global mirroring above. The secondary (target) is also written to at a later time however, there is no inter-dependency check done to any of the other logical drives to determine if there are any relationships there. This method of mirroring can be used to copy static logical drives which have no IO being performed to them; or to mirror LUNs which are totally independent from all other logical drives which are being mirrored.

For information about the Copy Services, see the following Redbooks publications:

- ▶ *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010
- ▶ *IBM Midrange System Storage Copy Services Guide*, SG24-7822
- ▶ *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065

3.13 IBM Remote Support Manager for Storage

The IBM Remote Support Manager (RSM) for Storage is designed to help manage problem reporting and analyze alert notifications for IBM's midrange disk storage subsystems. The DS RSM Model RS2 (1818-RS2) contains advanced processor and memory technology for enhanced processing capabilities. The DS RSM Model RS2 can attach to the LAN and collect alerts from IBM midrange disk systems within the data center and expedite forwarding of the alerts to an IBM service support center.

Additionally, the DS RSM Model RS2 is designed:

- ▶ To enable IBM to establish communication with the attached supported storage systems
- ▶ To provide faster response time on processing alerts
- ▶ To offer the ability to transfer detailed information to assist error analysis
- ▶ To facilitate enhanced problem resolution

The DS RSM can help improve and facilitate service and support for DS series disk storage systems. Additionally, for each data center with a DS employed, it is desirable to install at least one DS RSM that will be operational with all of the DS5000/DS4000/DS3000 controllers located at the site. Each DS RSM unit has the capability to manage up to 50 DS5000/DS4000/DS3000 storage systems.

Figure 3-21 shows the model IBM DS RSM model RS2.



Figure 3-21 DS RSM RS2 model

The DS RSM model RS2 contains an Intel Nehalem Quad-Core processor and improved RAM memory technology. The DS RSM RS2 delivers a rack-mount server in a 1U enclosure containing dual power supplies, SUSE Linux Enterprise Server (SLES) operating system, RSM code, RSM installation CD, RSM documentation, rack mounting rails, and rack power cords.

RSM requires a service contract. You can contact your IBM representative for further information regarding the service offering.

For more detailed information about the DS-RSM product, see the website:

<http://www.ibm.com/systems/storage/disk/rsm/index.html>

3.14 More information

More detailed information about the DS4000 and DS5000 families can be found in the following publications and websites:

- ▶ *Deployment Guide Series: IBM TotalStorage Productivity Center for Data*, SG24-7140
- ▶ *IBM i and Midrange External Storage*, SG24-7668
- ▶ *IBM Midrange System Storage Hardware Guide*, SG24-7676
- ▶ *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- ▶ *IBM Midrange System Storage Copy Services Guide*, SG24-7822
- ▶ *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065
- ▶ *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010
- ▶ *IBM System Storage FlashCopy Manager and PPRC Manager Overview*, REDP-4065
- ▶ Interoperability matrix:

<http://www.ibm.com/servers/storage/disk/ds4000/interop-matrix.html>

<http://www-03.ibm.com/systems/storage/disk/ds5000/interop-matrix.html>

- ▶ Servers and storage:

<http://www.ibm.com/servers/storage/disk/index.html>



Enterprise Disk Systems

The family of IBM Enterprise Disk Systems offers a broad range of scalable solutions to address various enterprise storage needs. By making use of IBM leading technology, the Enterprise Disk Systems provide a significant choice in functionality, performance, and resiliency.

In this chapter, we discuss the features and major characteristics of the following enterprise disk storage products:

- ▶ IBM System Storage DS6000 Series
- ▶ IBM System Storage DS8000 Series
- ▶ IBM System Storage XIV

4.1 Overview

The IBM Enterprise Disk Systems deliver high-performance, high-availability storage, with flexible configurations for various business requirements. The IBM System Storage DS6000, DS8000, and XIV deliver an enterprise storage continuum of systems with shared replication services and common management interfaces. The DS6000, DS8000, and XIV series systems are designed to help simplify the storage infrastructure, support business continuity, and optimize information lifecycle management. See Figure 4-1 for an overview of the family.

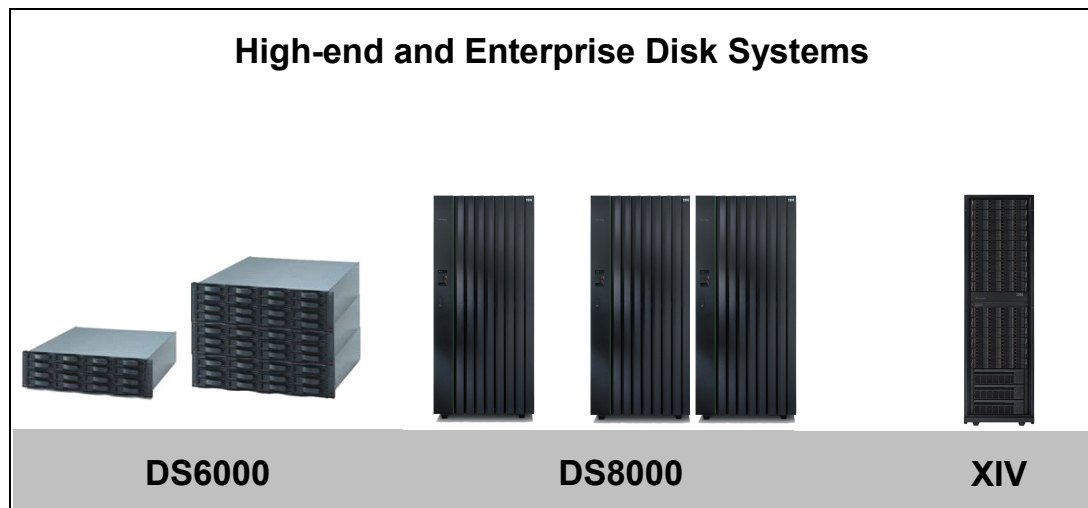


Figure 4-1 Enterprise Disk Systems

For a quick overview of storage solutions, see the *IBM System Storage Product Guide* at the website:

<http://www.ibm.com/systems/storage/resource/pguide/>

The DS6000 series

The DS6000 offers true enterprise-class functionality with modular design and an attractive entry-level pricing for medium and large businesses. The DS6000 series helps simplify data management and enables easy scalability, which allows accommodation of the continuing exponential data growth. By using the proven software function of the Enterprise Storage Server®, the DS6000 series brings proven enterprise class technology to a modular package.

The DS6000 series offers high scalability and excellent performance. With the DS6800 (Models 1750-511 and 1750-522), you can install up to 16 disk drive modules (DDMs). The minimum storage capability with 4 DDMs is 292 GB. The DS6000 series also offers 500 GB, 7,200 rpm Fibre Channel ATA (FATA) disk drives, so the maximum storage capability with 16 DDMs for the DS6800 model is 8 TB.

For more information about the DS6000, see 4.3, “The DS6000 series” on page 94.

The DS8000 series

The IBM System Storage DS8000 series is designed for the most demanding, mission critical environments requiring the highest level of availability. The DS8000 series is designed to set an entirely new industry standard for high performance and high capacity by delivering a dramatic leap in performance and scalability.

The DS8100 and the DS8300 use the IBM POWER5™ server technology, and the DS8700 uses the IBM POWER6® that is integrated with the IBM Virtualization Engine technology. All DS8000 series models consist of a storage unit and one or two management consoles, two being the preferred configuration.

For more information about the DS8000, see the following references:

- ▶ Section 4.4, “The DS8000 series” on page 101 in this book
- ▶ *IBM System Storage DS8000: Architecture and Implementation*, SG24-6786
- ▶ *IBM System Storage DS8700: Architecture and Implementation*, SG24-8786

The XIV series

The XIV Storage System architecture is designed to deliver performance, scalability, and ease of management while harnessing the high capacity and cost benefits of Serial Advanced Technology Attachment (SATA) drives. The system employs off-the-shelf products as opposed to traditional offerings that use proprietary designs and therefore require more expensive components. The main idea of this machine is the Modules architecture based on the massive parallelism: The system architecture ensures full exploitation of all system components.

The IBM XIV Storage System is a next-generation high-end open disk storage system. An integral part of the IBM broad spectrum of system storage and SAN offerings, the XIV system is based on a grid of standard, off-the-shelf hardware components connected in any-to-any topology by means of massively paralleled, non-blocking Gigabit Ethernet. Its innovative architecture is designed to deliver the highest levels of reliability, performance, scalability, and functionality at low overall cost, while providing unprecedented ease of management.

For more information about the XIV, see 4.8, “IBM XIV Storage System” on page 137 and *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

4.1.1 Infrastructure simplification

The IBM System Storage DS Family and XIV offers the opportunity to simplify your IT infrastructure through consolidation and streamlined storage management.

Consolidate storage assets

Consolidation begins with interoperability. The IBM System Storage DS Family and XIV can be connected across a broad range of server environment, including IBM z/OS, z/VM®, OS/400®, i5/OS®, and AIX, as well as Linux, HP-UX, Sun SOLARIS, Novell NetWare, UNIX, and Microsoft Windows. You can easily split storage capacity among the attached servers, reducing the total number of storage systems required.

Streamline storage management

The IBM System Storage DS Family and XIV incorporates streamlined management tools with an easy-to-use and straightforward GUI based on open Storage Management Initiative Specification (SMI-S) interfaces. The GUI allows users to manage multiple subsystems and controllers, perform logical configurations, and administer copy service management functions, all by a web browser.

The DS6000 series, the DS8000 series, and XIV share a common set of advanced functions, including FlashCopy, Metro Mirror, Global Copy, and Global Mirror. So there is only one set of skills necessary to manage the whole enterprise disk storage system.

There is also a set of common functions for storage management, including the IBM System Storage DS Command-Line Interface (DS CLI) and the IBM System Storage DS open application programming interface (API). The XIV Storage System offers a comprehensive set of Extended Command Line Interface (XCLI) commands to configure and monitor the system. All the functions available in the GUI are also available in the XCLI.

Advanced copy functions and storage management tools are described later in this chapter.

4.1.2 Business continuity

The IBM Enterprise Family offers the capability to increase your IT continual operation tolerance by data protection and compatible copying service.

Data protection

Many design characteristics and advanced functions of the IBM Enterprise Disk Systems contribute to protect data effectively:

- Fault-tolerant:

The IBM System Storage DS8000 and XIV series are designed with no single point of failure. They are fault-tolerant storage subsystems that can be maintained and upgraded concurrently with user operations. The DS6000 series is also designed and implemented with component redundancy to help reduce and avoid many potential single points of failure.

- RAID and mirroring protected storage:

The IBM DS Family Systems support RAID-5, RAID-10 configurations, or a combination of both. This gives more flexibility when selecting the redundancy technique for data protection. Also, RAID-6 is supported by the DS8000 with the latest code. RAID 6 protection provides more fault tolerance than RAID 5 in the case of disk failures and uses less raw disk capacity than RAID 10.

The XIV systems use remote mirroring. Asynchronous and synchronous remote mirroring provides protection against primary site outages, disasters, and site failures.

Copy Services

IBM FlashCopy point-in-time copy functions support higher application availability and continuity of operations because they are designed to shrink backup window time. Metro and Global Mirror functions allow for creation of duplicate copies of application data at remote sites for rapid recovery purposes.

The IBM System Storage DS Family also supports open standards. As a result, you can mirror a DS8000 series system with a DS6000 series system to help lower the total cost of the disaster recovery solution.

The DS8000 and XIV systems use the Thin Provisioning Copy Service. Thin Provisioning allows a more efficient capacity utilization in the storage subsystem. The automated capacity allocation (when required by applications) facilitates capacity management as well.

The Copy Services can be managed and automated with Tivoli Storage Productivity Center for Replication (TPC-R).

4.1.3 Performance

IBM Enterprise Disk Systems are designed for high performance that takes advantage of IBM leading technologies. In today's world, enterprises need business solutions that can deliver high levels of performance continuously every day, day after day. They also need a solution that can handle various workloads simultaneously, so they can run business intelligence models, large databases for enterprise resource planning (ERP), and online and Internet transactions alongside each other. In the following sections we discuss unique features that contribute to the overall high-performance design of the IBM Enterprise Disk Systems.

Server-based design

The design decision to use processor memory as I/O cache is a key element of the IBM storage architecture. Performance improvements can be traced to the capabilities of the processor speeds, the L1/L2 cache sizes and speeds, the memory bandwidth and response time, and the PCI bus performance (Figure 4-2).

With the DS6000 and DS8000 series, the cache access has been accelerated further by making the non-volatile storage (NVS) a part of the main memory. Part of the memory is used for the operating system and another part in each controller card acts as non-volatile storage (NVS), but most of the memory is used as cache. This design to use processor memory makes cache accesses very fast.

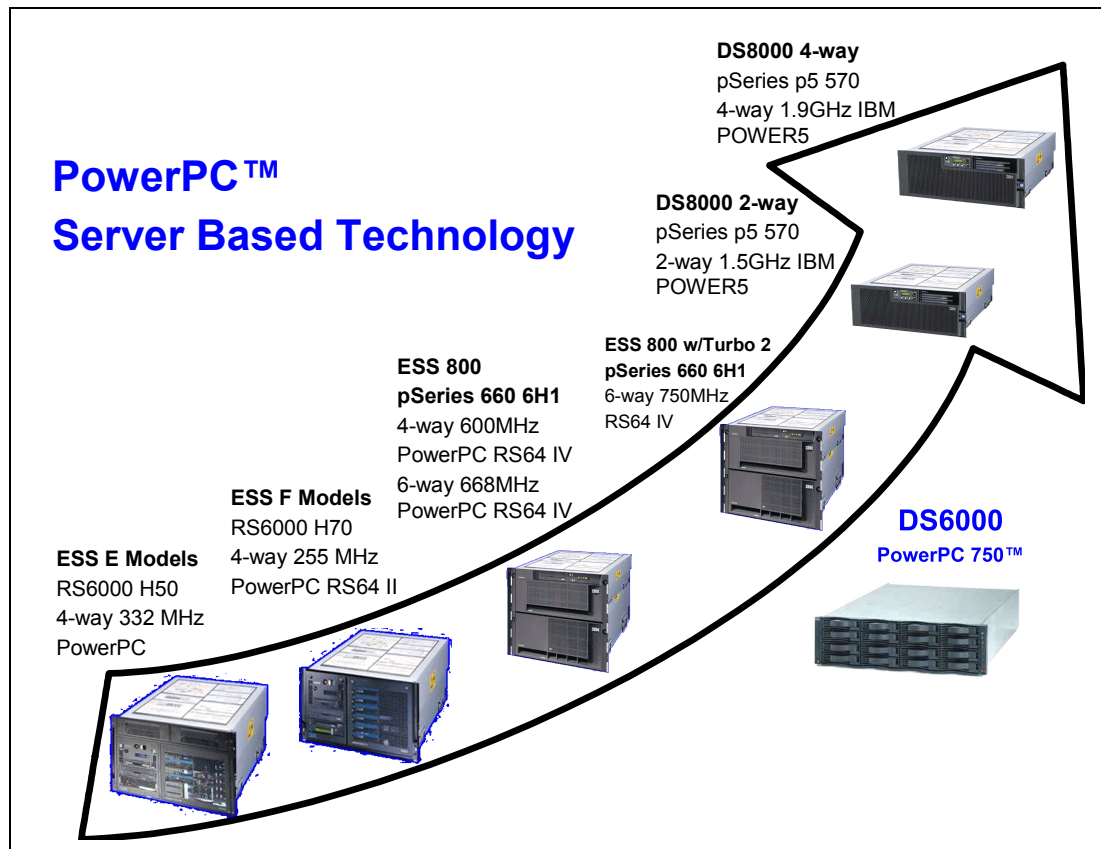


Figure 4-2 The server based architecture rings continuous performance improvements

Sequential Prefetching in Adaptive Replacement Cache

Another performance enhancer is the new self-learning cache algorithm, Sequential Prefetching in Adaptive Replacement Cache (SARC), which is described further in 4.2.2, "Enhancements of the DS6000 and DS8000 series" on page 91.

Multipathing software

IBM Multipath Subsystem Device Driver (SDD) provides load balancing and enhanced data availability in configurations with more than one I/O path between the host server and the storage server. Most vendors' priced multipathing software selects the preferred path at the time of initial request. IBM free of charge preferred path multipathing software dynamically selects the most efficient and optimum path to use at each data interchange during read and write operations.

Command Tag Queuing

Command Tag Queuing provides multiple AIX/UNIX I/O commands that can be queued to the DS6000, which improve performance through autonomic storage management, as compared to having the server queuing one I/O request at a time.

4.2 DS6000 and DS8000 technology

The maximum storage capability of the DS6800 controller with 16 DDMs is 8 TB. Optional DS6000 expansion enclosures provide a maximum storage capacity of 64 TB. The physical storage capacity of the DS8000 series systems can range from 1.1 TB to 512 TB (for DS8700 up to 1152 TB) and it has an architecture designed to scale up to a petabyte. The DS8000 series allows additions and upgrades from one model to another to adapt to changing business requirements (the exception is DS8700).

With this choice, you can build very cost efficient storage systems by adding expansion enclosures to the DS6800 controller or also grow horizontally by adding other DS6800 controllers. You have the option to easily grow into the DS8000 series by adding DS8000 systems to your environment or by replacing DS6000 systems. The expandability is shown in Figure 4-3.

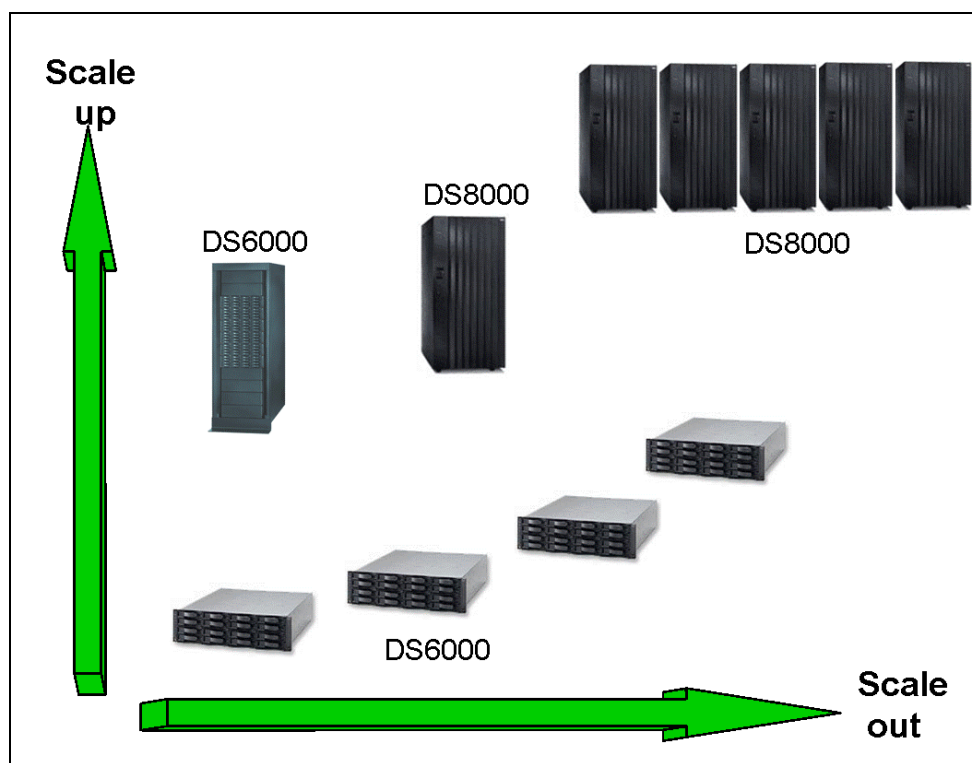


Figure 4-3 Scaling options

4.2.1 Multi-tiered storage environments

The DS6000/DS8000 series is ideally suited to a multi-tiered storage environment. This helps minimize storage costs by retaining frequently accessed or high-value data on higher performance DS8000 storage servers and archiving less frequently accessed or less valuable information kept on less-costly DS6000 systems.

4.2.2 Enhancements of the DS6000 and DS8000 series

The DS6000 and DS8000 series incorporate the following performance, resiliency, and scalability innovations:

Switched Fibre Channel Arbitrated Loop (FC-AL)

There is a switched FC-AL implementation in the DS Family (Figure 4-4). This switching technology uses a point-to-point connection to each disk drive and adapter, which allows maximum bandwidth for data movement, eliminates the bottlenecks of loop designs, and allows for specific disk drive fault indication and isolation in the same time. Each disk is attached to both switches. Whenever the device adapter connects to a disk, it uses a switched connection to transfer data. This means that all data travels by the shortest possible path.

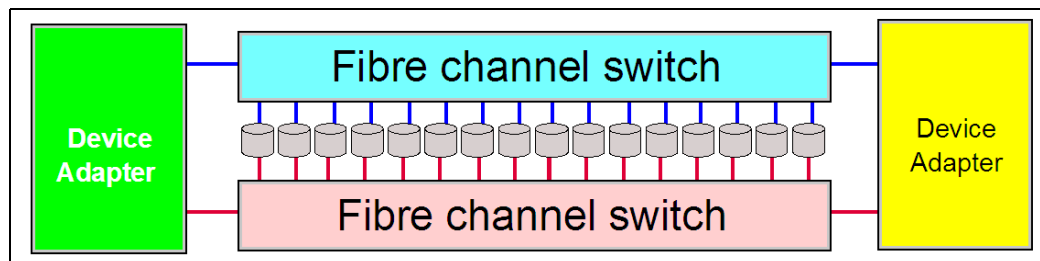


Figure 4-4 Disk enclosure switched connections

Key features of switched FC-AL technology are as follows:

- ▶ Standard FC-AL communication protocol from DA to DDMs
- ▶ Direct point-to-point links are established between DA and DDM
- ▶ Isolation capabilities in case of DDM failures, providing easy problem determination
- ▶ Predictive failure statistics
- ▶ Simplified expansion: No cable rerouting is required when adding another disk enclosure

The DS8000 architecture employs dual redundant switched FC-AL access to each of the disk enclosures. The key benefits of doing this are as follows:

- ▶ Two independent networks to access the disk enclosures
- ▶ Four access paths to each DDM
- ▶ Each device adapter port operating independently
- ▶ Double the bandwidth over traditional FC-AL loop implementations

Four path switched drive subsystem

There are four paths from the device adapters to each disk drive to provide greater data availability in the event of multiple failures along the data path (Figure 4-5). The four paths provide two FC-AL device interfaces, each with two paths such that either path can be used to communicate with any disk drive on that device interface (in other words, the paths are redundant). One device interface from each device adapter connects to a set of FC-AL devices such that either device adapter has access to any disk drive through two independent switched fabrics (in other words, the device adapters and switches are redundant).

In normal operation, however, disk drives are typically accessed by one device adapter and one server. Each path on each device adapter can be active concurrently, but the set of eight paths on the two device adapters can all be concurrently accessing independent disk drives. This avoids any contention between the two device adapters for access to the same disk, and means that all eight ports on the two device adapters can be concurrently communicating with independent disk drives.

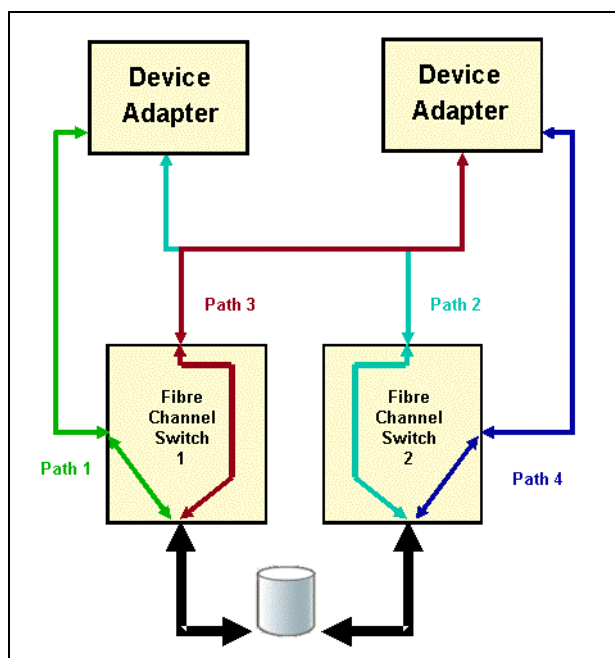


Figure 4-5 Four paths to each disk drive

Self-learning cache algorithms: SARC

The DS6800 and DS8000 use the patent-pending Sequential Prefetching in Adaptive Replacement Cache (SARC) algorithm, developed by IBM Storage Development in partnership with IBM Research. It is a self-tuning, self-optimizing solution for a wide range of workloads with a varying mix of sequential and random I/O streams. SARC is inspired by the Adaptive Replacement Cache (ARC) algorithm and inherits many features from it.

When a host performs a read I/O, the DS6000 and DS8000 fetch the data from the disk arrays through the high performance switched disk architecture. The data is then cached in volatile memory in case it is required again. The DS6000 and DS8000 attempt to anticipate future reads by the SARC algorithm. Data is held in cache as long as possible using this smart algorithm. If a cache hit occurs where requested data is already in cache, then the host does not have to wait for it to be read from the disks.

SARC provides the following benefits:

- ▶ Sophisticated, patented algorithms to determine what data is to be stored in cache based upon the recent access and frequency needs of the hosts.
- ▶ Pre-fetching, which anticipates data prior to a host request and loads it into cache.

Predictive Failure Analysis

The DS Enterprise Family uses Predictive Failure Analysis (PFA) to monitor disk drive operations. PFA takes preemptive and automatic actions before critical drive failures occur. The disk drives can anticipate certain forms of failures by keeping internal statistics of read and write errors. If the error rates exceed predetermined threshold values, the drive will be nominated for replacement. Because the drive has not yet failed, data can be copied directly to a spare drive. This avoids using RAID recovery to reconstruct all of the data onto the spare drive.

Configuration flexibility

The DS6000 and DS8000 use virtualization techniques to separate the logical view of hosts from the underlying physical layer. On an ESS there was a fixed association between logical subsystems (LSS) and device adapters. With the DS6000 and the DS8000, these limitations no longer apply (Figure 4-6).

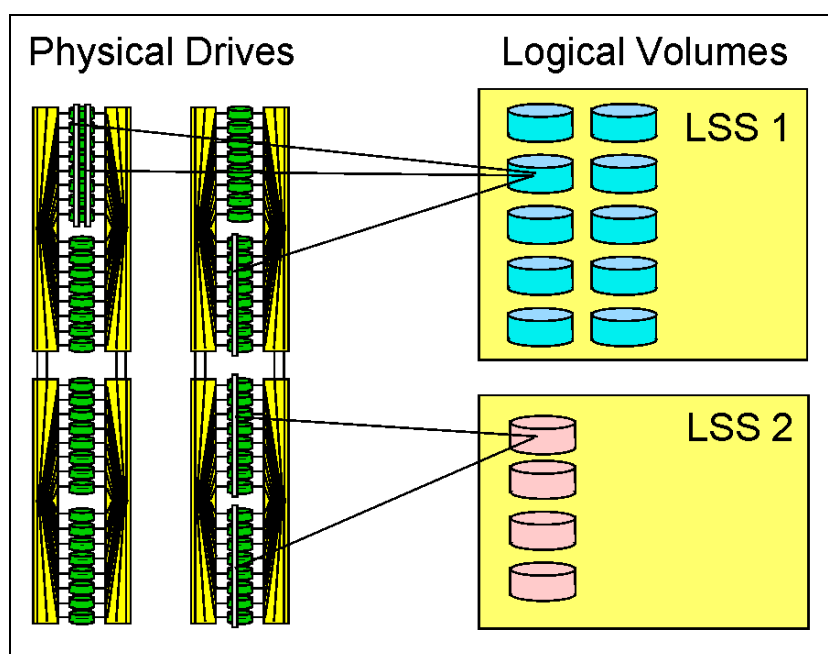


Figure 4-6 Grouping of volumes in LSSs

The DS6000 and DS8000 series provide a high configuration flexibility with the following virtualization techniques:

- ▶ Dynamic LUN/Volume creation and deletion:
LUNs can be created and deleted without having to reformat a whole array.
- ▶ Large LUN and large CKD volume support:
LUNs and volumes can be configured to span arrays, which allows for large LUN sizes.
- ▶ Flexible LUN to LSS association:
There is no predefined association of arrays to logical subsystems (LSS).
- ▶ Simplified LUN masking:
The access to LUNs by the host systems is controlled by volume groups. Hosts or disks in the same volume group share access to data.

IBM System Storage DS Open Application Programming Interface (API)

The DS Open API supports routine LUN management activities, such as LUN creation, mapping, and masking, and the management of point-in-time copy and remote mirroring. It supports these activities through the use of a standard interface as defined by the Storage Networking Industry Association (SNIA) Storage Management Initiative Specification (SMI-S). It is implemented through the IBM System Storage Common Information Model Agent (CIM Agent) for the DS series, a middleware application that provides a CIM-compliant interface. The DS Open API and CIM Agent are provided with the DS6800 and DS8000.

1. The DS API coordinates with Microsoft Volume Shadow (VSS) Copy Service and the DS6000 series FlashCopy function to create a consistent, point-in-time shadow copy of a Windows 2003 Server volume or group of volumes. This ensures consistency for VSS aware applications.
2. The DS API also coordinates with Microsoft Virtual Disk Service (VDS). Microsoft VDS is a software interface that manages block storage virtualization. This interface module makes it possible for the DS6000 series to interoperate with applications that utilize the VDS Application Programming Interface in a Windows 2003 environment.

See the individual server pages within this document for operating system level requirements and for DS6000 series LMC prerequisites. While the CIM Agent is available for only selected operating system environments, it can be used to manage all LUNs within a DS6000 series.

For additional information, see the *IBM System Storage DS Open Application Programming Interface Reference* at the following website:

<http://www.ibm.com/support/docview.wss?uid=ssg1S7001160>

4.3 The DS6000 series

The DS6000 series (Figure 4-7) offers high reliability and enterprise-class functionality for medium and large businesses. The DS6000 series supports connectivity with open systems and mainframe hosts. Copy Services are equivalent and interoperable with the DS8000 and the ESS. With the DS6000, enterprise class resiliency is now available in an incredibly small, modular, affordable package. The DS6000 series can grow up to 64 TB of physical storage.



Figure 4-7 The DS6000 series

The DS6000 consists of the DS6800 controller enclosure (1750-522), which has dual Fibre Channel RAID controllers. Capacity can be increased by adding up to seven DS6000 expansion enclosures (1750-EX2).

4.3.1 The DS6800 controller enclosure

The DS6000 series is based on IBM industry-leading PowerPC® architecture, which accelerates data response time and offers impressive performance. At the core of the DS6800 controller enclosure are two active/active RAID controllers (Figure 4-8).

The controller cards

Each controller card contains an integrated four-port host adapter and a high performance device adapter chip set. The device adapter ASIC on the DS6800 controller provides two disk expansion and two disk control ports. These ports are used to connect the expansion enclosures. Host adapter ports are used for host or SAN attachment.



Figure 4-8 DS6800 controller card

The RAID controller cards are the heart and soul of the system. Each card is the equivalent of a cluster node in an ESS. IBM has utilized its extensive development of the ESS host adapter and device adapter function to create a total repackaging. It actually uses DS8000 host adapter and device adapter logic, which allows almost complete commonality of function and code between the two series (DS6000 and DS8000).

The processors

The DS6800 utilizes two 64-bit PowerPC 750GX 1 GHz processors for the storage server and the host adapters, respectively, and another PowerPC 750FX 500 MHz processor for the device adapter on each controller card. The DS6800 has 2 GB memory in each controller card, for a total of 4 GB.

The disk subsystem

For the disk subsystem (Figure 4-9), each controller card has an integrated four-port Fibre Channel Arbitrated Loop (FC-AL) device adapter that connects the controller card to two separate Fibre Channel loops. Each switched loop attaches disk enclosures that each contain up to 16 disks. Each enclosure contains two 22-port Fibre Channel switches. Of these 22 ports, 16 ports are used to attach to the 16 disks in the enclosure and four ports are used to interconnect with other enclosures. The remaining two ports are reserved for internal use.

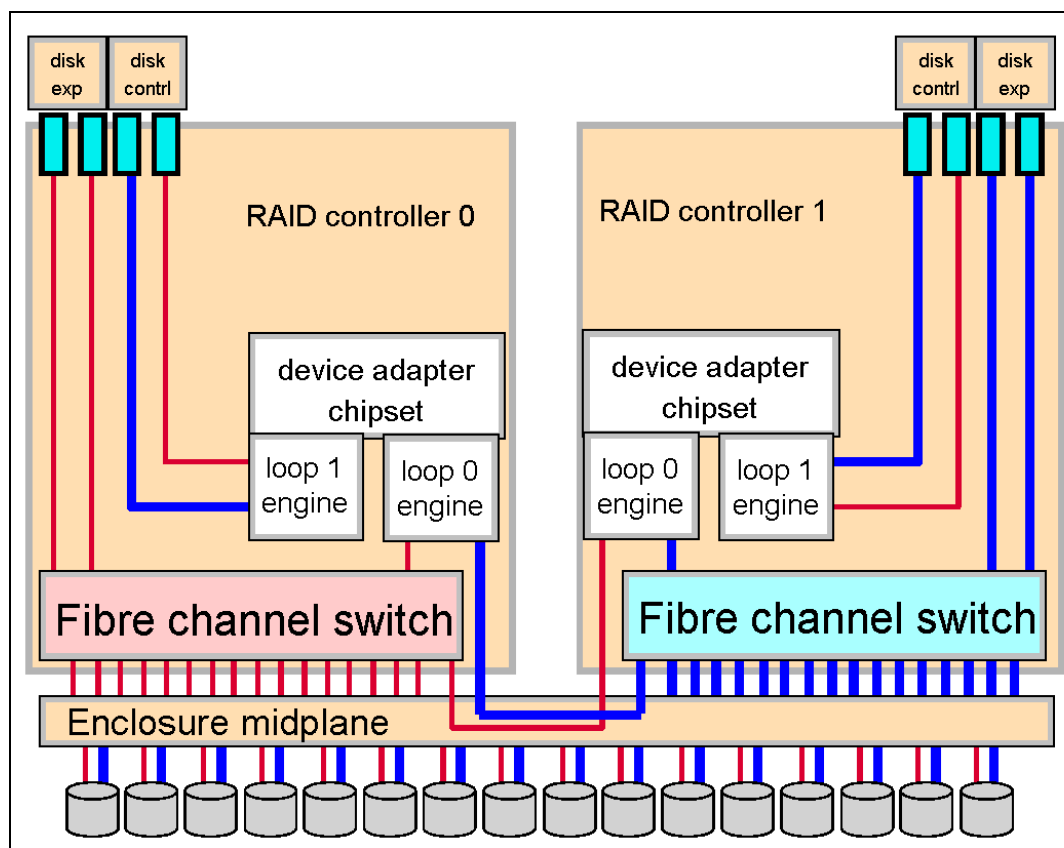


Figure 4-9 DS6000 disk architecture

The configuration process when forming RAID-5 or RAID-10 arrays requires that two global spares be defined in the DS6800 controller enclosure. If you have expansion enclosures, the first enclosure needs another two global spares. More spares can be assigned when drive groups with larger capacity drives are added.

Host adapters

The DS6800 has eight 2 Gbps Fibre Channel host ports. Each host port can be configured individually to operate in Fibre Channel or FICON mode. A host port cannot be both FICON and FCP simultaneously, but it can be changed as required. Ensure that host servers have paths to each of the two DS6800 RAID controllers to achieve best reliability and performance.

The attached hosts interact with microcode running on a Power PC chip set to access data on logical volumes. The microcode manages all read and write requests to the logical volumes on the disk arrays. For write I/O operations, the controllers use fast-write, which writes the data to volatile memory on one controller and to persistent memory on the other controller. The DS6800 then reports to the host that the write is complete before it has actually been written to disk, thus providing much faster write performance. Persistent memory is also called NVS or non-volatile storage (Figure 4-10).

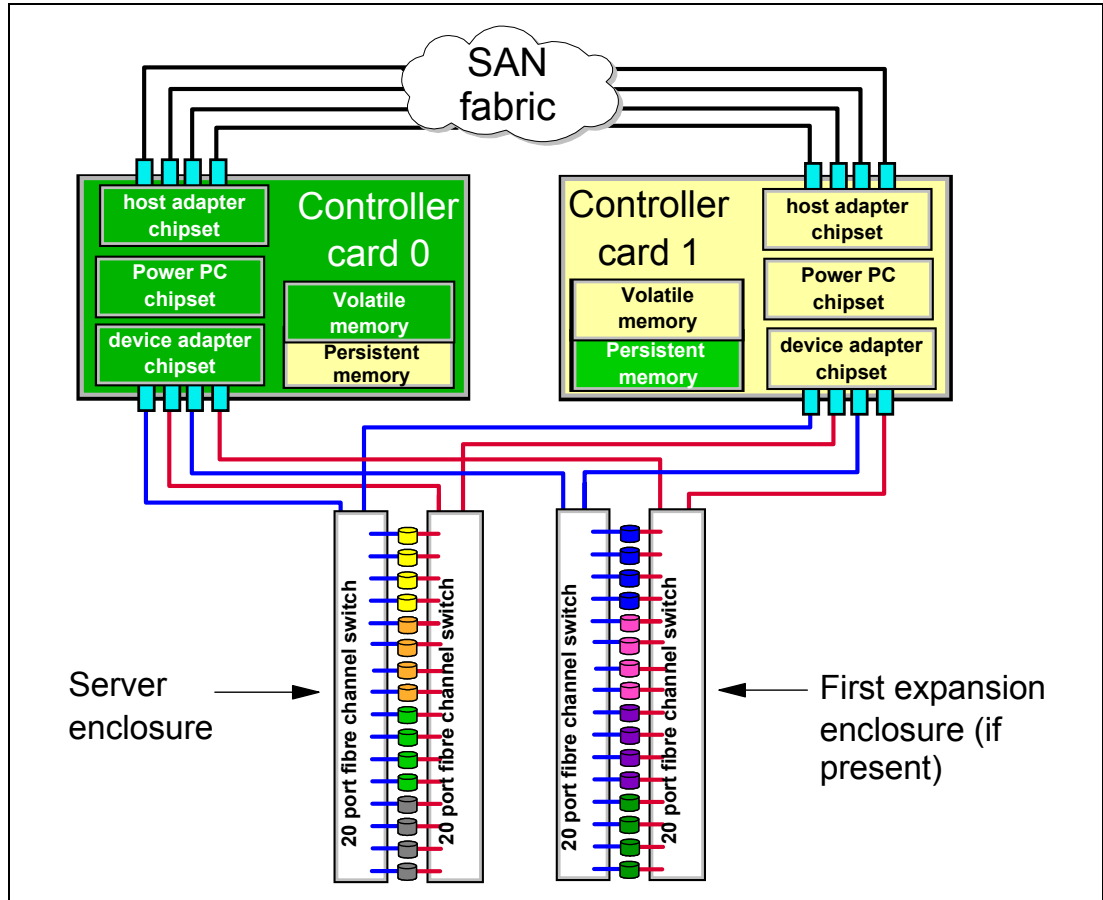


Figure 4-10 DS6800 controller architecture

Power subsystem

The power subsystem of the DS6800 has two redundant power supplies and two battery backup units (BBUs). DS6000 expansion enclosures contain power supplies but not BBUs. The DS6000 power supplies are hot-swappable and a single power supply can support the power requirements of an entire enclosure.

Dense packaging

Calibrated Vectors Cooling™ technology used in System x and BladeCenter to achieve dense space saving packaging is also used in the DS6800.

4.3.2 The DS6000 expansion enclosure

The DS6000 expansion enclosure is used to add capacity to a DS6800 server enclosure. Up to 7 DS6000 expansion enclosures can be added to a DS6800 controller enclosure.

Expansion enclosure SBOD controller card

The DS6000 SBOD controller card (Figure 4-11) is only found in the expansion enclosure. Each SBOD controller card contains an independent 22-port Fibre Channel switch. Of these 22 ports, 16 are used to attach to the 16 disks in the expansion enclosure. Four more are used to interconnect with other enclosures, with the remaining two ports reserved for internal use.



Figure 4-11 SBOD controller card

Dual redundant switched loops

The DS6800 supports two dual redundant switched loops. The first loop is for the DS6800 and up to three DS6000 expansion enclosures. The second switched loop is for up to four expansion enclosures. For connections to the previous and next enclosure, four inbound and four outbound 2 Gbps Fibre Channel ports are available.

The disks in the server enclosure are on the first disk loop (loop 0). The first two expansion enclosures are attached to the DISK CONTRL ports to start the second disk loop (loop 1). The DISK EXP ports are used to attach the third expansion enclosure. It joins the same switched loop as the disks in the server enclosure (loop 0), as shown in Figure 4-12.

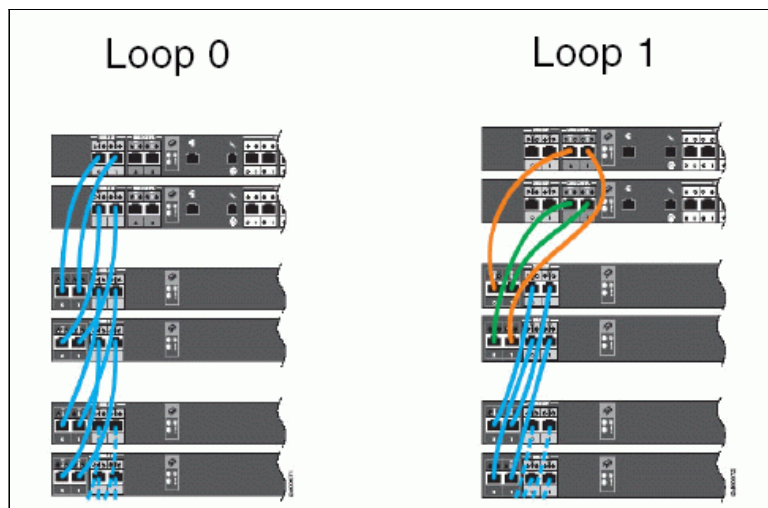


Figure 4-12 Example of expansion enclosure cabling

To distribute the storage enclosures equally on both storage loops, follow the installation sequence shown in Table 4-1.

Table 4-1 Install sequence of expansion enclosures

Expansion enclosure	Loop
1	1
2	1
3	0
4	1
5	0
6	1
7	0

4.3.3 DS6800 Model 1750-511

The 1750-511 (Figure 4-13) model contains control unit functions as well as a rich set of advanced functions and holds up to 16 disk drive modules (DDMs). It provides a minimum capacity of 584 GB with 8 DDMs and 73 GB per DDM. The maximum storage capacity with 450 GB 15k DDMs is 7.2 TB with 16 DDMs.



Figure 4-13 DS6800 Model 1750-511 rear view

The 1750-511 model has the following features:

- ▶ Two RAID controller cards
- ▶ PowerPC 750GX 1 GHz processors and one PowerPC 750FX processor on each RAID controller card
- ▶ Dual active controllers to provide continuous operations and back up the other controller in case of controller maintenance or an unplanned outage of a controller
- ▶ 4 GB of cache memory (2 GB in each controller card)
- ▶ Battery backed mirrored cache
- ▶ Two battery backup units (one per controller card)
- ▶ Two AC/DC power supplies with imbedded enclosure cooling units
- ▶ Eight Fibre Channel host ports that can be configured as a pairs of FCP or FICON host ports. The hosts ports auto-negotiate to either 2 Gbps or 1 Gbps link speeds
- ▶ Two Fibre Channel switches for disk drive connectivity
- ▶ Eight 2 Gbps device ports (for additional DS6000 expansion enclosures)
- ▶ Attachment to up to seven DS6000 Model 1750-EX1 expansion enclosures
- ▶ Choice of 2 Gbps Fibre Channel disk drives including:
 - 73 GB 15k RPM
 - 146 GB 15k RPM
 - 146 GB 10k RPM
 - 300 GB 10k RPM
 - 450 GB 15k RPM
- ▶ 3U in height and mountable in a standard 19-inch rack

4.3.4 DS6000 Model 1750-EX1

DS6000 can expand with the expansion enclosure model 1750-EX1 (Figure 4-14). Up to seven expansion drawers build up to the maximum configuration of 128 DDMs, which are comprised of 8 drawers x 16 DDMs.

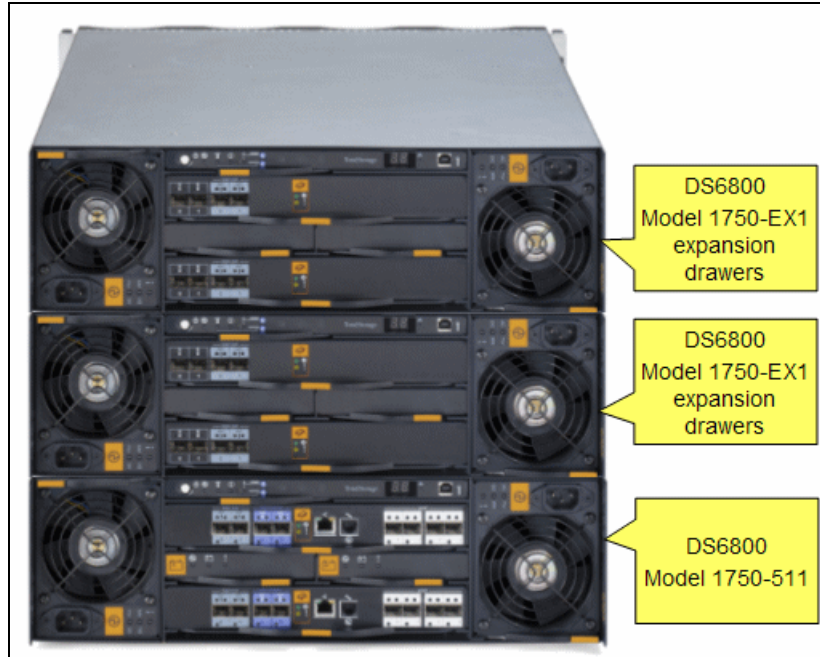


Figure 4-14 DS6000 with 2 expansion enclosures

Each expansion enclosure contains the following features:

- ▶ Two expansion controller cards, each with:
 - Two 2 Gbps inbound ports
 - Two 2 Gbps outbound ports
 - One Fibre Channel switch
- ▶ Disk enclosure, which holds up to 16 Fibre Channel DDMs
- ▶ Two AC/DC power supplies with embedded enclosure cooling units
- ▶ Supports attachment to DS6800 Model 1750-511
- ▶ The DS6800 Model 1750-EX1 can also be mounted in a standard 19-inch rack

4.3.5 More information

For more information about the DS6000 series, see the following Redbooks publications:

- ▶ *The IBM TotalStorage DS6000 Series: Concepts and Architecture*, SG24-6471
- ▶ *The IBM TotalStorage DS6000 Series: Implementation*, SG24-6781
- ▶ *The IBM TotalStorage DS6000 Series: Copy Services with IBM eServer zSeries*, SG24-6782
- ▶ *The IBM TotalStorage DS6000 Series: Copy Services in Open Environments*, SG24-6783
- ▶ *iSeries and IBM TotalStorage: A Guide to Implementing External Disk on eServer i5*, SG24-7120
- ▶ *IBM TotalStorage DS6000 Series: Performance Monitoring and Tuning*, SG24-7145

4.4 The DS8000 series

The DS8000 series (Figure 4-15) is designed to break through to a new dimension of on demand storage. Its innovative design incorporates a high-bandwidth and fault-tolerant Fibre Channel disk technology and hardware-based partitioning implementation. The DS8000 delivers robust, flexible, and cost-effective disk storage for mission-critical workloads and helps to ensure exceptionally high system availability for continuous operations in 24x7 environments.



Figure 4-15 DS8000 base frame with covers removed (left) and with Model 94E (right)

The hardware has been optimized to provide enhancements in performance, connectivity, and reliability. The DS8000 series features several models in a new, higher-density footprint, while re-using much of the fundamental architecture of the previous ESS models. This ensures that the DS8000 benefits from a very stable and well-proven operating environment.

During recent years, the DS8000 series were developed and improved; this process is still ongoing and brings new solutions and technologies to the DS Family. In the following sections we describe innovations and new implementations for the DS8300 and DS8700.

4.4.1 Optimizations to simplify and increase efficiency

With its high capacity, scalability, broad server support, and virtualization features, the DS8000 is well suited for simplifying the storage environment by consolidating data from multiple storage systems on a single system.

Virtualization

Through virtualization, organizations can allocate system resources more effectively and better control application quality of service. The implementation of storage system logical partitions (LPARs) allows the DS8300 Turbo Model 9B2 (not available for DS8700) to run two independent workloads on completely independent virtual storage systems. This unique feature of the DS8000 series delivers opportunities for new levels of efficiency and cost-effectiveness.

The DS8000 improves the cost structure of operations and lowers energy consumption through a tiered storage environment. The availability of Serial ATA (SATA) drives gives you the option to retain frequently accessed or high-value data on Fibre Channel (FC) disk drives and to archive less valuable information about the less costly SATA disk drives.

The dynamic volume expansion simplifies management by enabling easier, online volume expansion to support application data growth, and to support data center migrations to larger volumes to ease addressing constraints.

The Thin Provisioning feature in Licensed Machine Code 5.4.3x.xx provides over-provisioning capability for more efficient usage of the storage capacity and reduced storage management requirements, hence improving cost efficiency and reducing the Total Cost of Ownership (TCO). For more information, see *DS8000 Thin Provisioning*, REDP-4554.

The FlashCopy SE capability enables more Space Efficient utilization of capacity for FlashCopy operations, improving cost effectiveness. System Storage Productivity Center (SSPC) single pane control and management integrates the power of the Tivoli Storage Productivity Center (TPC) and the DS Storage Manager user interfaces into a single view.

Licensed Machine Code 5.4.25.xx provides support for Remote Pair FlashCopy (not in Licensed Machine Code 5.4.3x.xx), which enables more effective two-site business continuity with FlashCopy and Metro Mirror. See *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, REDP-4504 for more information.

The LDAP authentication support allows single sign-on functionality. This can simplify user management by allowing the DS8000 to rely on a centralized LDAP directory rather than a local user repository. For more information, see *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, REDP-4504.

General innovations

The DS8000 Release 3 microcode (Licensed Machine Code 5.3.xx.xx) introduced new features reinforcing the DS8000's world class performance:

- ▶ Storage Pool Striping helps maximize performance without special tuning.
- ▶ Adaptive Multistream Prefetching (AMP) is a breakthrough caching technology that can dramatically improve sequential performance, thereby reducing times for backup, processing for Business Intelligence, and streaming media.

The DS8000 Release 4 microcode (Licensed Machine Code 5.4.xx.xx) and higher brings additional innovations that gives the DS8000 increased flexibility and data protection with features, such as:

- ▶ Variable LPAR: This new function provides the ability for Dual Shark Images where one image has more processor and cache than the other for increased flexibility (except DS8700).
- ▶ RAID 6: This new function allows for additional fault tolerance by using a second independent distributed parity scheme (dual parity).
- ▶ IPv6 support: The IBM System Storage DS8000 has been certified as meeting the requirements of the IPv6 Read Logo program, indicating its implementation of IPv6 mandatory core protocols and the ability to interoperate with other IPv6 implementations. The IBM DS8000 can be configured in native IPv6 environments. The logo program provides conformance and interoperability test specifications based on open standards to support IPv6 deployment globally.

- ▶ **Extended Address Volume:** This new function extends the addressing capability of System z environments. Volumes can scale up to approximately 223 GB (262,668 cylinders). This capability can help relieve address constraints to support large storage capacity needs. This Extended Address Volumes are going to be supported by z/OS 1.10 or later versions.

The Release 4.2 microcode (Licensed Machine Code 5.4.20.xx) provides new innovations to increase the performance again:

- ▶ **Optional Solid State Drives** provide extremely fast access to data, energy efficiency, and higher system availability.
- ▶ **Intelligent Write Caching (IWC)** improves the Cache Algorithm for random writes. Specifically, database applications can benefit from the new IWC technology.
- ▶ **The Full Disk Encryption (FDE)** feature can protect business sensitive data by providing disk based hardware encryption combined with a sophisticated key management software (Tivoli Key Lifecycle Manager). The Full Disk Encryption support feature is available only as a plant order.
- ▶ A new, optional mobile computer-based HMC is available. It is mandatory with Full Disk Encryption systems.
- ▶ You now have the ability to enable the Network Time Protocol (NTP) service.

The Release 4.3 microcode (Licensed Machine Code 5.4.3x.xx) and higher introduces new features to improve efficiency and performance:

- ▶ **DS8000 Thin Provisioning:** This new feature allows the creation of over-provisioned devices for more efficient usage of the storage capacity.
- ▶ **Quick Initialization for open system (FB) volumes:** This new feature provides volume initialization up to 2.6 times faster and therefore allows creation of devices making them available as soon as the command completes.
- ▶ **High Performance FICON for System z (zHPF) multitrack operations:** This new feature delivers a performance improvement in terms of higher throughput for the existing, chargeable zHPF feature.

4.4.2 Introduction to the DS8000 series features and functions

The IBM System Storage DS8000 series is a high-performance, high-capacity series of disk storage subsystems. It offers balanced performance and storage capacity that scales linearly up to hundreds of terabytes.

The IBM System Storage DS8700 highlights include:

- ▶ Robust, flexible, enterprise class, and cost-effective disk storage
- ▶ Exceptionally high system availability for continuous operations
- ▶ Centralized and simplified management
- ▶ IBM POWER5+™ processor technology for DS8300 and IBM POWER6 for DS8700
- ▶ Capacities from 1.1 TB to 460/1024 TB (FC/SATA)
- ▶ Point-in-time copy function with FlashCopy, FlashCopy SE, or Remote Pair FlashCopy, and Remote Mirror and Copy functions with Metro Mirror, Global Copy, Global Mirror, Metro/Global Mirror, z/OS Global Mirror, and z/OS Metro/Global Mirror with new Incremental Resync capability
- ▶ Support for a wide variety and intermix of OSs, including System i and System z

- Designed to increase storage efficiency and utilization, the DS8000 can be part of a Green Data Center

4.4.3 The DS8700: Latest member of the DS Family

The IBM System Storage DS8700 is designed to support the most demanding business applications with its exceptional all-around performance and data throughput. It is the next step of DS8000, which includes all functions from previous family members. This, combined with its world-class business resiliency and encryption features, provides a unique combination of high availability, performance, and security. Its tremendous scalability, broad server support, and virtualization capabilities can help simplify the storage environment by consolidating multiple storage systems onto a single DS8700.

The IBM System Storage DS8000 series encompasses the flagship disk enterprise storage products in the IBM System Storage portfolio. The DS8700 represents the latest in this series designed for high-performance, high-capacity, and resilient series of disk storage systems.

The IBM System Storage DS8700 provides new functional capabilities, allowing you to choose the combination of price and efficiency that is right for your application needs. New capabilities include:

- IBM POWER6 processor technology: The DS8700 features IBM's POWER6 server technology to help support high performance. Compared to the POWER5+ processor in previous models, the POWER6 processor can enable over a 50% performance improvement in I/O operations per second in transaction processing workload environments. Additionally, sequential workloads can receive as much as 150% bandwidth improvement. The DS8700 offers either a dual 2-way processor complex or a dual 4-way processor complex.
- PCI-e IO enclosures: To improve IOPS (I/O Operations Per Second) and sequential read/write throughput, the new IO enclosures are directly connected to the servers by point-to-point PCI-e cables. IO enclosures no longer share common "loops"; they connect directly to each internal server by separate cables and link cards.
- Four-port device adapters: Device adapter processor hardware has been upgraded to twice as fast processors for more IOPS performance to enable better utilization of SSD drives.
- Non-disruptive upgrade path for the DS8700 Model 941 and additional Model 94E expansion frames allows processor, cache, and storage enhancement to be performed concurrently without disrupting applications.
- Enhancements to disk encryption key management that can help address PCI-DSS (Payment Card Industry Data Security Standard) requirements:
 - Encryption deadlock recovery key: This supports the ability for IBM to restore access to a DS8700 when the encryption key for the storage is unavailable due to an encryption deadlock scenario.
 - Dual platform key server support: DS8000 requires an isolated key server in encryption configurations. The isolated key server currently defined is an IBM System x server. Dual platform key server support allows two separate server platforms to host the key manager with either platform operating in either *clear key* or *secure key* mode.
- Value based pricing/licensing: Operating Environment License is now priced based on the performance, capacity, speed, and other characteristics that provide value in customer environments.

Besides these new functions, the DS8700 inherits most of the features of its predecessors.

4.4.4 The DS8000 design keys

In this section we discuss key architectural basics of DS8000 and improvements of DS8700. From an architectural point of view, the DS8700 has not changed much with respect to the fundamental architecture of the predecessor DS8100 and DS8300 models. This ensures that the DS8700 can benefit from a very stable and well-proven operating environment, offering the optimum in availability. The hardware is optimized to provide higher performance, connectivity, and reliability.

Green Data Center

Faced with increasingly urgent warnings about the consequences of the projected rise in both energy demands and greenhouse gas emissions, governments and businesses alike are now focusing more attention than ever on the need to improve energy efficiency. Corporate data centers are well known as significant power users.

IBM is launching the Green Data Center environment in a very responsible *green* way, which is catching the attention of many companies. IBM products can provide huge increases in processing, with dramatically less power consumption and lower carbon output.

The DS8700 is one of these products and is reflected with a green colored stripe on the front covers of the machine (Base frame).

For more information about creating a green data center, call your IBM representative or visit following web page:

<http://www.ibm.com/cio>

IBM POWER processor technology

The DS8000 series exploits the IBM POWER5 technology, which is the foundation of the storage system LPARs. The DS8000 uses 64-bit POWER5 microprocessors in dual 2-way or dual 4-way processor complexes, with up to 256 GB of cache. The maximum host I/O operations per second of the DS8300 models.

The DS8700 exploits the IBM POWER6 technology. The SMP system features 2-way or 4-way, copper-based, SOI-based POWER6 microprocessors running at 4.7 GHz.

Compared to the POWER5+ processor in previous models, the POWER6 processor can enable over a 50% performance improvement in I/O operations per second in transaction processing workload environments. Additionally, sequential workloads can receive as much as 150% bandwidth improvement. The DS8700 offers either a dual 2-way processor complex or a dual 4-way processor complex.

Internal fabric

The DS8000 comes with a high bandwidth, fault tolerant internal interconnection, called RIO-G (Remote I/O) which is also used in the System p. The interconnection can operate at speeds up to 1 GHz and offers a 2 Gbps sustained bandwidth per link, with exceptional performance and reliability.

DS8700 uses direct point-to-point high speed PCIe connections to the I/O enclosures to communicate with the device and host adapters. Each single PCIe connection operates at a speed of 2GB/s in each direction. There are up to 16 PCIe connections from the processor complexes to the I/O enclosures.

Switched Fibre Channel Arbitrated Loop

The DS8000 uses switched Fibre Channel Arbitrated Loop (FC-AL) for its disk interconnection. This offers a point-to-point connection to each drive and device adapter, so that there are four paths available from the controllers to each disk drive.

Fibre Channel disk drives

The DS8000 offers a selection of industry standard Fibre Channel disk drives, including 146 GB (15K rpm), 300 GB (15K rpm), and 450 GB (15K rpm). The 450 GB 15K rpm Fibre Channel Disk Drive provides capacity, allowing a single system to scale up to 460 TB of Fibre Channel capacity. The DS8000 series also now allow customers to install 146 GB 15K rpm, 300 GB 15K rpm, and 450 GB 15K rpm Full Disk Encryption drive sets.

Serial ATA drives

With the 1 TB (7,200) Serial ATA (SATA) drives, the DS8000 capacity scales up to 1 PB (1024 TB = 1 petabyte = 1 PB). These SATA drives offer a cost-effective option for lower priority data.

Solid State Drives

With the Solid State Drives (SSDs), which are available in 73 GB and 146 GB, the DS8000 offers new opportunities for ultra high performance applications. The SSD drives are the best choice for I/O intensive workload. They provide up to 100 times the throughput and 10 times lower response time than 15K rpm spinning disks. Additionally, they also consume much less power.

For more information about SSDs, see *DS8000: Introducing Solid State Drives*, REDP-4522.

Host adapters

The DS8000 series offers host connectivity with four-port Fibre Channel/FICON Host Adapters. The DS8000 currently supports 4 Gbps Fibre Channel/FICON host adapters.

The 4 Gbps Fibre Channel/FICON Host Adapters, are offered in longwave and shortwave, auto-negotiate to either 4 Gbps, 2 Gbps, or 1 Gbps link speeds. Each port on the adapter can be individually configured to operate with Fibre Channel Protocol (FCP) (also used for mirroring) or FICON.

A DS8700 with the dual 4-way feature can support up to a maximum of 32 host adapters, which provide up to 128 Fibre Channel/FICON ports. Be aware that ESCON adapters are no longer supported.

IBM System Storage management console for the DS8000

The Hardware Management Console (HMC) is one of the focal points for configuration, Copy Services management, and maintenance activities. The management console is a dedicated workstation (Thinkpad) that is physically located (installed) inside the DS8000 and can proactively monitor the state of your system, notifying you and IBM when service is required. It can also be connected to your network to enable centralized management of your system using the IBM System Storage DS Command-Line Interface or storage management software utilizing the IBM System Storage DS Open API.

All base model ships with one Storage Hardware Management Console (HMC) and two Ethernet switches.

A mobile computer HMC (Lenovo ThinkPad), shown in Figure 4-16, will be shipped with a DS8700.

All communication of a user configuring or changing a DS8000 configuration by using the GUI or DSCLI is addressing the HMC.

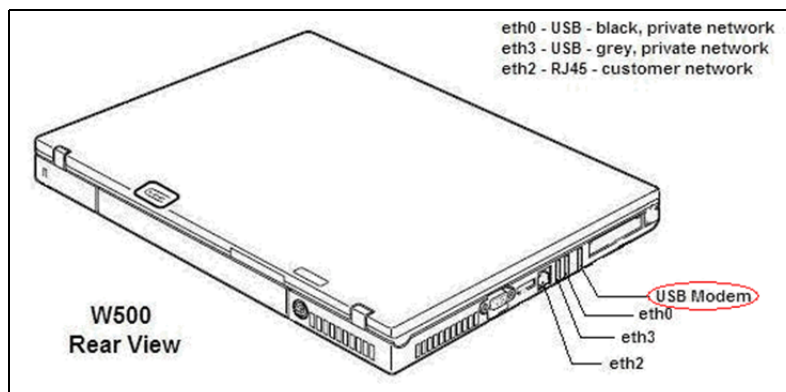


Figure 4-16 Mobile computer HMC

Internet Protocol: The DS8000 HMC supports IPv6, the next generation of the Internet Protocol. The HMC continues to support the IPv4 standard as well as mixed IPV4 and IPV6 environments.

Client and IBM services can also connect to the Hardware Management Console through a Web-base unit interface (WUI), which is connected to the client network. This connection brings improved response time for remote HMC operations. The connection uses a standard SSL based authentication with encryption capability and user ID and password protection.

An external management console is available as an optional feature and can be used as a redundant management console for environments with high availability requirements.

IBM System Storage Productivity Center management console

The DS8000 series utilizes the IBM System Storage Productivity Center (SSPC), an advanced management console that can provide a view of both IBM and non-IBM storage environments. SSPC can enable a greater degree of simplification for organizations grappling with the growing number of element managers in their environment. The SSPC is an external System x machine with pre-installed software, including IBM Tivoli Storage Productivity Center Basic Edition.

Utilizing IBM Tivoli Storage Productivity Center Basic Edition software, SSPC extends the capabilities available through the IBM DS Storage Manager. SSPC offers the unique capability to manage a variety of storage devices connected across the storage area network (SAN). The rich, user-friendly graphical user interface provides a comprehensive view of the storage topology, from which the administrator can explore the health of the environment at an aggregate or in-depth view. Moreover, the TPC Basic Edition, which is pre-installed on the SSPC, can be optionally licensed and used to enable more in-depth performance reporting, asset and capacity reporting, and automation for the DS8000, as well as to manage other resources, such as server file systems, tape drives and libraries.

Important: Any DS8700 shipped requires a minimum of one SSPC per data center to enable the launch of the DS8000 Storage Manager other than from the HMC.

TKLM Isolated Key server

The Tivoli Key Lifecycle Manager software performs key management tasks for IBM encryption enabled hardware, such as the IBM System Storage DS8000 Series Family and IBM encryption-enabled tape drives by providing, protecting, storing, and maintaining encryption keys that are used to encrypt information being written to, and decrypt information being read from, encryption enabled disks. TKLM operates on a variety of operating systems.

For DS8700 systems shipped with Full Disk Encryption (FDE) drives, two TKLM key servers are required. An isolated key server (IKS) with dedicated hardware and non-encrypted storage resources is required.

The isolated TKLM key server can be ordered from IBM. It is the same hardware as used for the SSPC. The following software is used on the isolated key server:

- ▶ Linux operating system
- ▶ Tivoli Key Lifecycle Manager V: Includes DB2 V9.1 FB4

No other hardware or software is allowed on the IKS.

For more information, see *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

Storage system logical partitions

The DS8000 series (with exception of DS8700) provides storage system logical partitions (LPARs) as a first in the industry. Logical partitioning introduces a new level of virtualization.

Storage LPAR technology (see Figure 4-17) enables the creation of two or potentially more completely separate storage subsystems running the same or various versions of microcode. With these separate resources, each storage system LPAR can be used for completely separate production, test, or other unique storage environments within a single physical DS8000 system. Each partition can be established to support specific performance requirements. You can run independent, virtual storage images with differing workloads, and with various operating system environments. The DS8000 series isolates and protects LPARs by hardware and the POWER5 Hypervisor firmware.

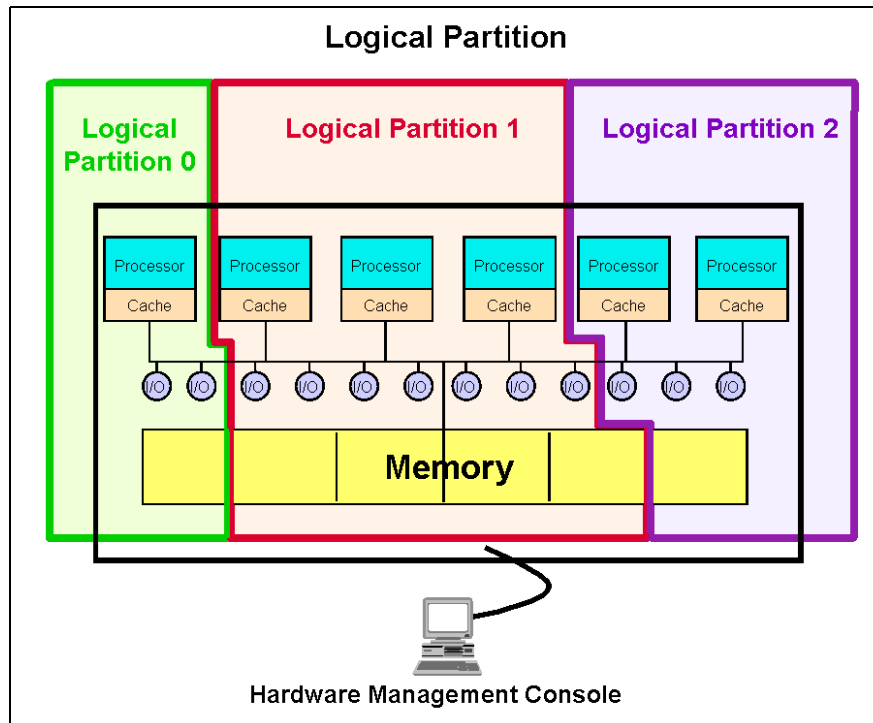


Figure 4-17 Logical partitions (LPARs)

IBM started offering an implementation of the server LPAR functionality in a disk storage system initially with the DS8300 Model 9A2, and with the Turbo Model 9B2; it provides two virtual storage systems within one physical storage system. Initially each storage system LPAR has access to:

- ▶ 50 percent of the processors
- ▶ 50 percent of the processor memory
- ▶ Up to 16 host adapters
- ▶ Up to 320 disk drives (up to 96 TB of capacity)

4.4.5 The DS8000 architecture

This section describes the architecture of the DS8000 systems and provides information about the following components:

- ▶ Processor complexes
- ▶ Disk subsystems (switched FC-AL loops)
- ▶ I/O enclosures (device and host adapters)
- ▶ Power and cooling
- ▶ Battery backup units
- ▶ Hardware Management Console (HMC)
- ▶ Network switches
- ▶ Frames and expansion frames

Figure 4-18 shows the components used to create the DS8000 Primary Frame.

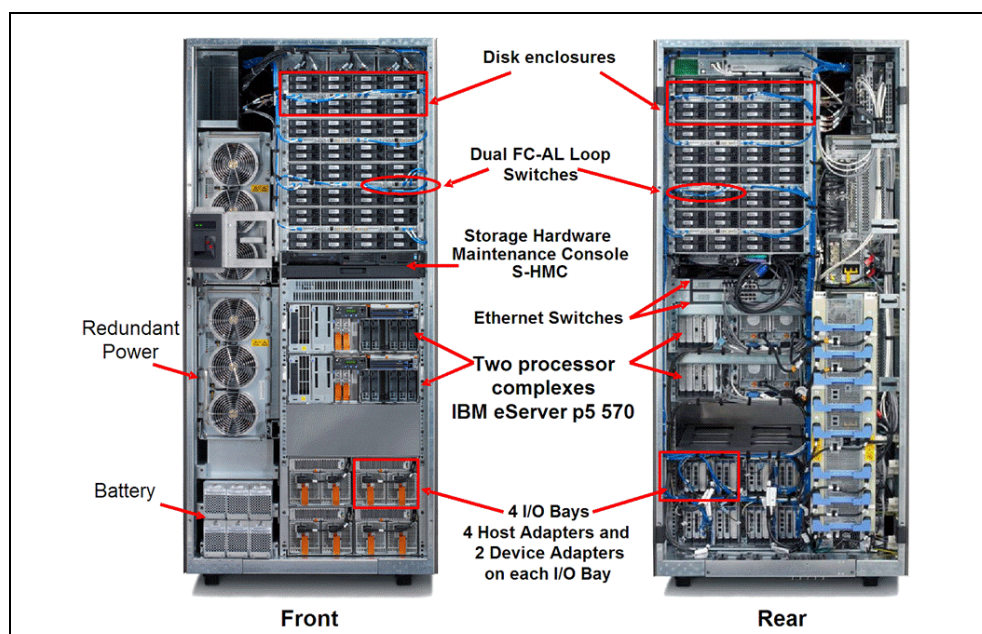


Figure 4-18 DS8000 - Primary Frame

The DS8000 Family technical specifications are shown in Table 4-2.

Table 4-2 DS8000 Family technical specifications

Models	DS8100 (931)	DS8300 (932, 9B2)	DS8700 (941, 94E)
Shared SMP processor configuration	POWER5+ dual 2-way	POWER5+ dual 4-way	POWER6 dual 2-way or 4-way
Other major processors	PowerPC, ASICs	PowerPC, ASICs	PowerPC, ASICs
Processor memory for cache and NVS (min/max)	16 GB/128 GB	32 GB/256 GB	32 GB/384 GB
Host adapter interfaces	4-port 4 Gbps or 2 Gbps Fibre Channel/FICON, 2-port ESCON	4-port 4 Gbps or 2 Gbps Fibre Channel/FICON, 2-port ESCON	4-port 4 Gbps Fibre Channel/ FICON
Host adapters (min/max)	2/16	2/32	2/32
Host ports (min/max)	4/64	4/128	4/128
Drive interface	FC-AL	FC-AL	FC-AL
Number of disk drives (min/max)	16/384	16/1024	16/1024
Device adapters	Up to eight 4-port FC-AL	Up to 16 4-port FC-AL	Up to 16 4-port FC-AL
Maximum physical storage capacity	384 TB	1024TB	1024TB

The processor complexes

The DS8000 base frame contains two IBM p5 570 processor complexes (till R5 code machines), which contain the processor and memory that drive all functions in the DS8000. The symmetric multiprocessor (SMP) p5 570 system features 2-way or 4-way (2-way means that each processor complex has 2 CPUs, while 4-way means that each processor complex has 4 CPUs.), copper-based, SOI-based POWER5 microprocessors running at 1.5 GHz or 1.9 GHz with 36 MB off-chip Level 3 cache configurations. The p5 570 also has I/O expansion capability using the RIO-G interconnect.

The SMP system features 2-way or 4-way, copper-based, SOI-based POWER5+ microprocessors running at 2.2 GHz with 36 MB off-chip Level 3 cache configurations. The system is based on a concept of system building blocks. The processor complexes are facilitated with the use of processor interconnect and system flex cables that enable as many as four 4-way System p processor complexes to be connected to achieve a true 16-way SMP combined system. How these features are implemented in the DS8000 might vary.

The DS8700 base frame contains also two processor complexes, based POWER6 microprocessors running at 4.7 GHz. The 941 model can have the 2-way processors feature or the 4-way processors feature. There can be two or four I/O enclosures. In this case, at least four I/O enclosures are required. More might be necessary depending on the number of disk drives.

The DS8700 uses DDR2 memory DIMMs. Using SMPs as the primary processing engine sets the DS8000 apart from other disk storage systems on the market. Additionally, the System p POWER6 processors used in the DS8700 series support the execution of two independent threads concurrently. This capability is referred to as *simultaneous multi-threading (SMT)*. The two threads running on the single processor share a common L1 cache. The SMP/SMT design minimizes the likelihood of idle or overworked processors, while a distributed processor design is more susceptible to an unbalanced relationship of tasks to processors.

The design decision to use SMP memory as I/O cache is a key element of IBM storage architecture. Although a separate I/O cache can provide fast access, it cannot match the access speed of the SMP main memory.

All memory installed on any processor complex is accessible to all processors in that complex. The addresses assigned to the memory are common across all processors in the same complex. Alternatively, using the main memory of the SMP as the cache leads to a partitioned cache. Each processor has access to the processor complex's main memory but not to that of the other complex. Keep this in mind with respect to load balancing between processor complexes.

The processor memory

The DS8100 Model 921 offers up to 128 GB of processor memory and the DS8300 Models 922 and 9A2 up to 256 GB. Half the memory is located in each processor complex. In addition, the Non-Volatile Storage (NVS) scales to the processor memory size selected, which can also help optimize performance.

The DS8700 offers up to 384 GB of processor memory. Half of this will be located in each processor complex. Also the DS8700 has the Non-volatile Storage (NVS).

Service processor and SPCN

The service processor (SP) performs predictive failure analysis based on any recoverable processor errors. The SP can monitor the operation of the firmware during the boot process, and it can monitor the operating system for loss of control. This enables the service processor to take appropriate action. The SPCN is the system power control network that is used to control the power of the attached I/O subsystem.

RIO-G

The RIO-G ports are used for inter processor communication only on a DS8000. RIO stands for remote I/O. The RIO-G is evolved from earlier versions of the RIO interconnect.

Each RIO-G port can operate at 1 GHz in bidirectional mode and is capable of passing data in each direction on each cycle of the port. It is designed as a high performance self-healing interconnect.

I/O enclosures

The DS8700 base frame contains I/O enclosures and adapters. There can be two or four I/O enclosures in a DS8700 base frame. The I/O enclosures hold the adapters and provide connectivity between the adapters and the processors. Device adapters and host adapters are installed in the I/O enclosure. Each I/O enclosure has 6 slots.

In the current implementation bridge cards are used to translate PCIe to PCI-X legacy adapters. Each PCIe slot supports PCI-X adapters with bridge cards running at 64-bit, 133 Mhz. Slots 3 and 6 are used for the device adapters. The remaining slots are available to install up to four host adapters per I/O enclosure.

Each I/O enclosure has the following attributes:

- ▶ 5U rack-mountable enclosure
- ▶ Six PCIe slots

Default redundant hot-plug power and cooling devices

The disk subsystems

The disk subsystem consists of three components:

- ▶ First, located in the I/O enclosures are the device adapters. These are RAID controllers that are used by the storage images to access the RAID arrays.
- ▶ Second, the device adapters connect to switched controller cards in the disk enclosures. This creates a switched Fibre Channel disk network.
- ▶ Third, there are the disks themselves. The disks are commonly referred to as disk drive modules (DDMs).

In the DS8000, a faster ASIC and a faster processor is used on the device adapter cards compared to adapters of other members of the DS8000 Family leading to higher throughput rates.

Ports

Each DS8000 device adapter (DA) card offers four FC-AL ports. These ports are used to connect the processor complexes to the disk enclosures. The adapter is responsible for managing, monitoring, and rebuilding the RAID arrays. The adapter provides remarkable performance thanks to a high function/high performance ASIC. To ensure maximum data integrity, it supports meta data creation and checking.

The device adapter design is shown in Figure 4-19.

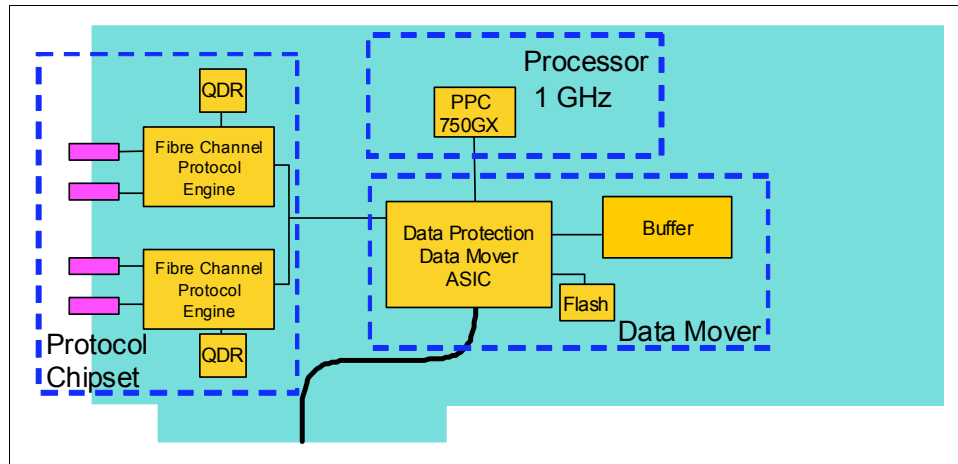


Figure 4-19 New DS8700 device adapter with 1Ghz processor

The DAs are installed in pairs because each storage partition requires its own adapter to connect to each disk enclosure for redundancy. This is why we call them *pairs*.

The DS8000 implements the concept of Array Across Loops (AAL) which splits each array site into two halves. Half of the site is located on the first disk loop of a device adapter (DA) pair and the other half is located on the second disk loop of that DA pair. It is implemented primarily to maximize performance; however it also provides higher level of redundancy with RAID-10 configurations, as shown in Figure 4-20.

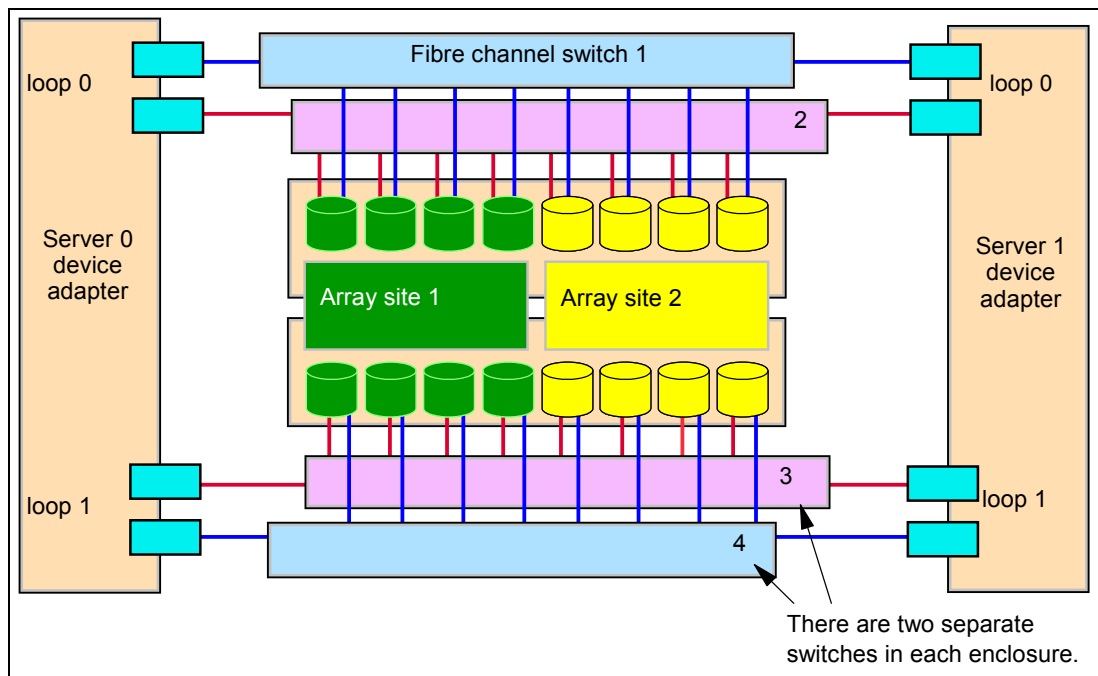


Figure 4-20 Array across loops

Disk drives

Each disk drive module (DDM) is hot pluggable and has two indicators. The green indicator shows disk activity, while the amber indicator is used with light path diagnostics to allow for easy identification and replacement of a failed DDM.

The DS8000 allows the choice of:

- ▶ Three unique Fibre Channel DDM types (available in both non-encrypted and encrypted versions):
 - 146 GB, 15K RPM drive
 - 300 GB, 15K RPM drive
 - 450 GB, 15K RPM drive

For more information about encrypted drives and inherent restrictions, see *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

- ▶ One Serial Advanced Technology Attachment (SATA) DDM drive type:
 - 1 TB, 7.2K RPM drive
- ▶ Two unique Solid State Drive (SSD) types:
 - 73 GB
 - 146 GB

For more information about Solid State Drives, see *DS8000: Introducing Solid State Drives*, REDP-4522.

The host adapters

The DS8000 offers enhanced connectivity with the availability of four-port Fibre Channel/FICON host adapters (Figure 4-21). Fibre Channel is a technology standard that allows data to be transferred from one node to another at high speeds and great distances (up to 10 km and beyond). The DS8000 uses Fibre Channel protocol to transmit SCSI traffic inside Fibre Channel frames. It also uses Fibre Channel to transmit FICON traffic, which uses Fibre Channel frames to carry System z I/O.

Ports

Each DS8000 Fibre Channel card offers four 4 Gbps Fibre Channel ports. The cable connector required to attach to this card is an LC type. Each 4 Gbps port independently auto-negotiates to either 4 or 2 or 1 Gbps link speed. Each of the 4 ports on one DS8000 adapter can also independently be either Fibre Channel protocol (FCP) or FICON. The personality of the port is changeable through the DS Storage Manager GUI or by using DSCLI commands. A port cannot be both FICON and FCP simultaneously, but it can be changed as required.

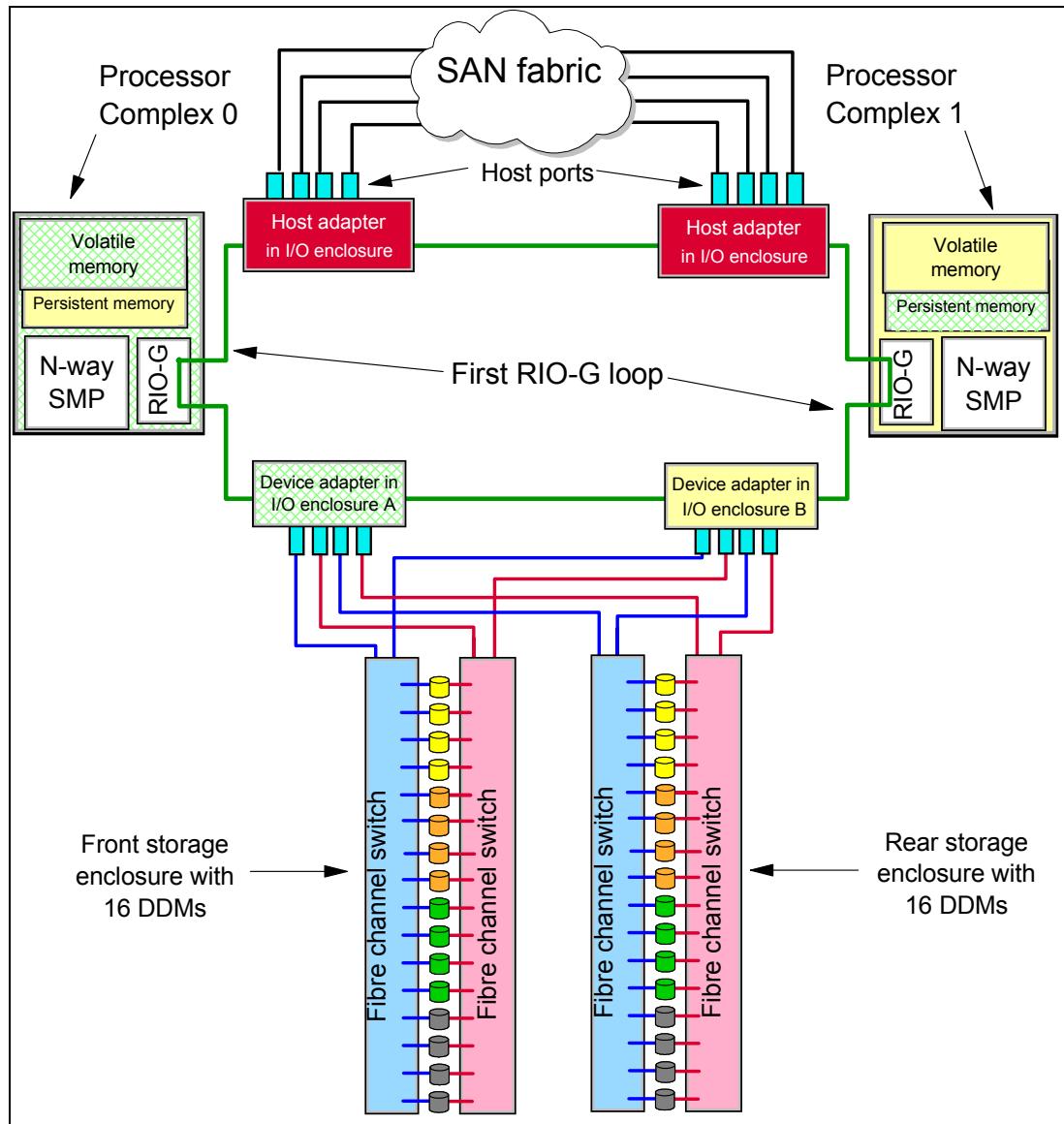


Figure 4-21 The DS8000 architecture

The attached hosts interact with software running on the complexes to access data on logical volumes. Each complex hosts at least one instance of this software (which is called a *server*), running in an LPAR. The servers manage all read and write requests to the logical volumes on the disk arrays. During write requests, the servers use fast-write where the data is written to volatile memory on one complex and persistent memory on the other complex. The server then reports the write as complete before it has been written to disk. This provides much faster write performance.

RIO-G/PCI-X

Both the device and host adapters operate on a high bandwidth fault-tolerant interconnect known as the RIO-G (shown in Figure 4-22).

The RIO-G ports are used for I/O expansion to external I/O drawers. The RIO-G is evolved from earlier versions of the RIO interconnect.

Each RIO-G port can operate at 1 GHz in bidirectional mode and can pass data in each direction on each cycle of the port. It is designed as a high performance self-healing interconnect. The p5 570 provides two external RIO-G ports, and an adapter card adds two more. Two ports on each processor complex form a loop.

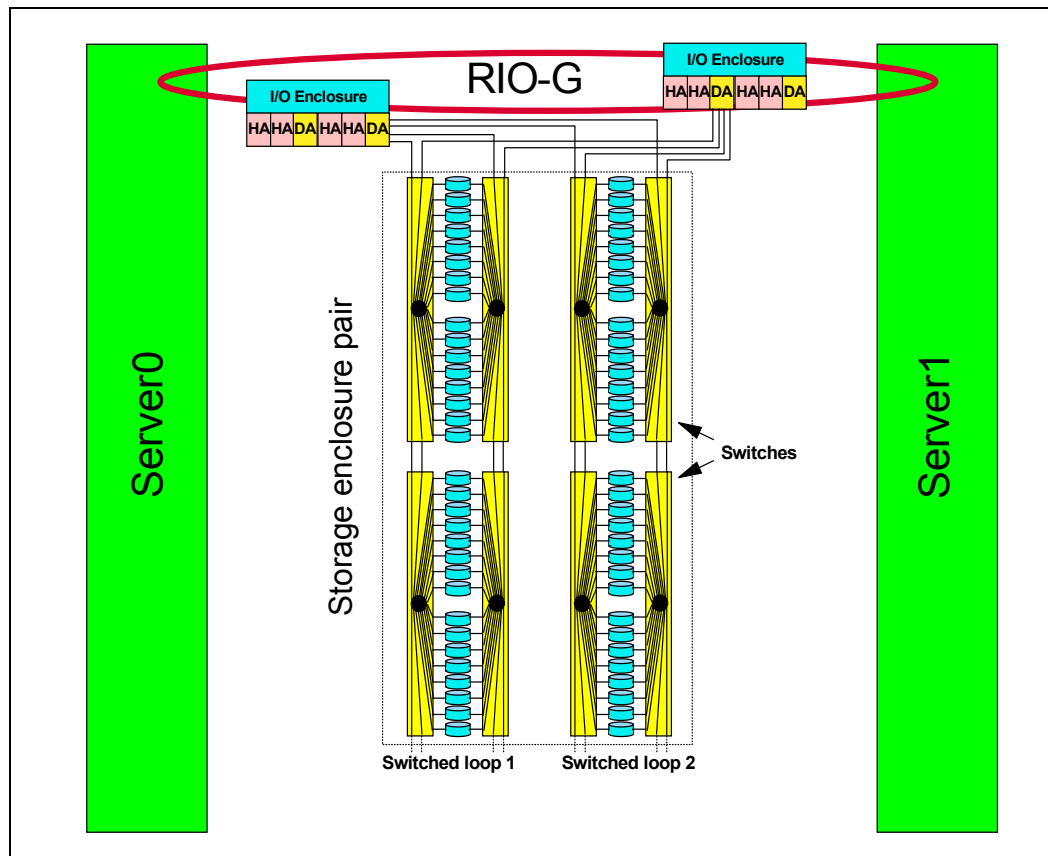


Figure 4-22 RIO-G design

The card itself is a PCI-X 64 Bit 133 MHz for DS8700 configuration. The card is driven by a new high function, high performance ASIC. To ensure maximum data integrity, it supports meta data creation and checking. Each Fibre Channel port supports a maximum of 509 host login IDs and 1,280 paths. This allows for the creation of very large storage area networks (SANs). The design of the card is depicted in Figure 4-23.

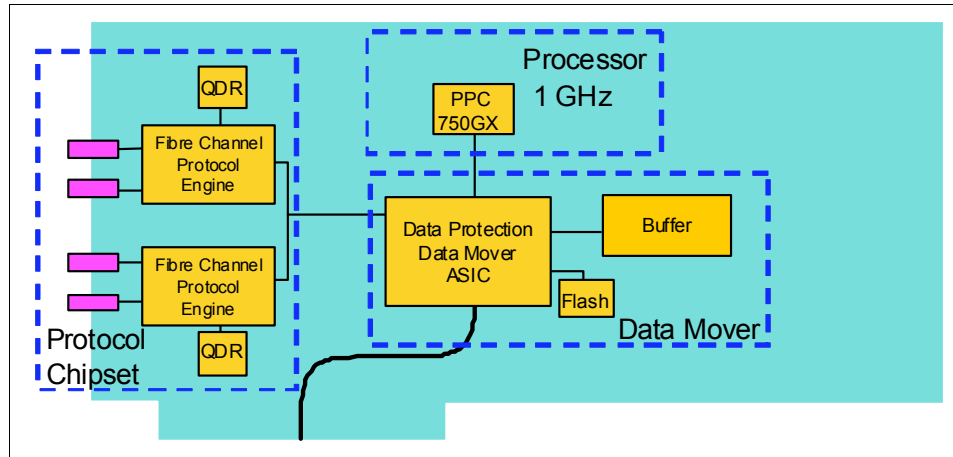


Figure 4-23 DS8000 Fibre Channel/FICON host adapter

The DS8700 storage facility consists of two POWER6 p 570 CECs. They form a processor complex that utilizes a RIO-G loop for processor communication and an PCIe infrastructure to communicate to the I/O subsystem (Figure 4-24).

Ethernet switches

Two 16-port gigabit Ethernet switches connect the disk enclosures and the processor complexes, which allow the creation of a fully redundant network. Each processor complex has multiple connections to each switch so that each server can access each switch. These switches are dedicated for use only by the DS8000.

Power and cooling

The DS8000 power and cooling system is highly redundant. There are two redundant primary power supplies (PPSs) in each frame of the DS8000, and each PPS can power the frame by itself. Furthermore, each processor and I/O enclosure has dual redundant power supplies.

The disk enclosures do not have separate power supplies because they draw power directly from the PPSs. They do however have cooling fans located in a plenum above the enclosures.

Battery backup assemblies

The backup battery assemblies help protect data in the event of a loss of external power. In the event of a complete loss of input AC power, the battery assemblies are used to allow the contents of NVS memory to be written to a number of DDMs internal to the processor complex, prior to power off.

The FC-AL DDMs are not protected from power loss unless the extended power line disturbance feature has been purchased.

DS8000 frames and enclosures

The base frame can contain up to eight disk enclosures, each can contain up to 16 disk drives. They are described as 16-packs because each enclosure can hold 16 disks. In a maximum configuration, the base frame can hold 128 disk drives. Each expansion frame can hold up to 16 disk enclosures which contain the disk drives. In a maximum configuration, an expansion frame can hold 256 disk drives (Figure 4-25).

The left side of the base frame (viewed from the front of the machine) is the frame power area. Only the base frame contains rack power control cards (RPC) to control power sequencing for the storage unit. It also contains a fan sense card to monitor the fans in that frame. The base frame contains two primary power supplies (PPSs) to convert input AC into DC power. The power area also contains two or three battery backup units (BBUs), depending on the model and configuration.

Between the disk enclosures and the processor complexes are two Ethernet switches and a Storage Hardware Management Console (HMC).

The base frame contains two processor complexes. These System p POWER5+/6 servers contain the processor and memory that drive all functions within the DS8000.

Finally, the base frame contains two or four I/O enclosures. These I/O enclosures provide connectivity between the adapters and the processors. The adapters contained in the I/O enclosures can be either device adapters (DAs) or host adapters (HAs), or both.

The communication path used for adapter to processor complex communication in DS8700 are now four lane (x4) PCIe Generation 2 connections providing a bandwidth of 2GB/s for each connection.

The inter processor complex communication still utilizes the RIO-G loop as in previous models of the DS8000 Family. However, this RIO-G loop no longer has to handle data traffic, which improves reliability to a great extent.

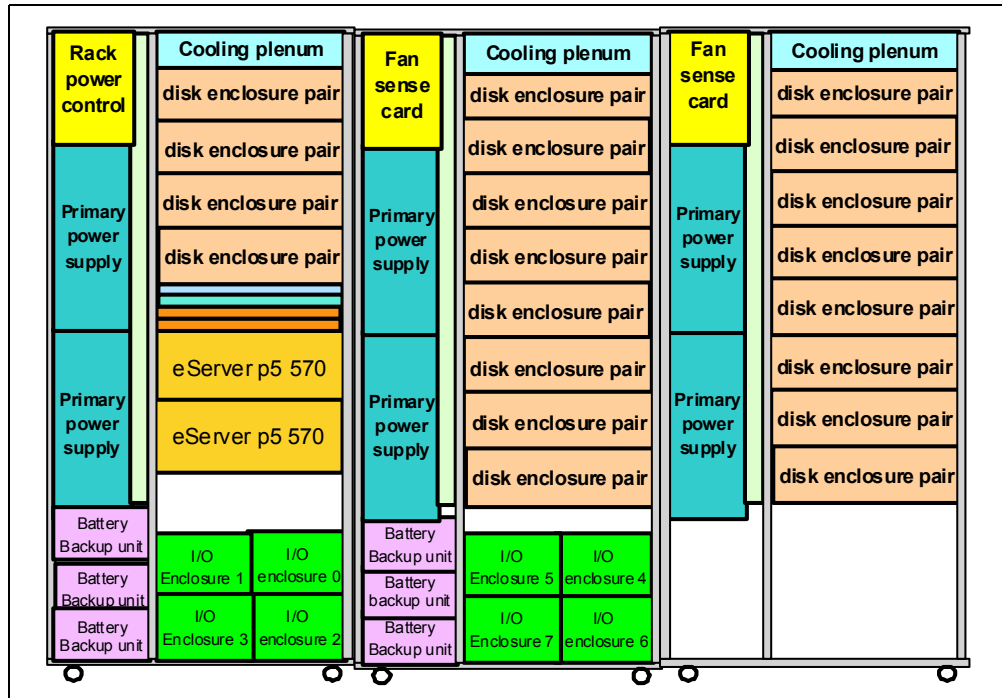


Figure 4-25 DS8000 frames

4.4.6 DS8000 models

One of the advantages of the DS8000 series is its linear scalability for capacity and performance. If your business grows rapidly, you might need much more storage capacity, faster storage performance, or both. The DS8000 series can meet these demands within a single storage unit.

DS8000 Turbo Models 931, 932, and 9B2

Before we get into the details about each model, we start with an explanation of the model naming conventions in the DS8000.

The DS8000 series currently has three Turbo models available: the DS8100 Turbo Model 931 and the DS8300 Turbo Models 932 and 9B2 (Figure 4-26). The difference in models is in the processors and in the capability of storage system LPARs. The predecessors of the DS8000 series Turbo models were the DS8000 series Models 921, 922, and 9A2.

The last position of the three characters 9xy refers to either a dual two-way processor complex model (9x1), a dual four-way processor complex model (9x2), or an Expansion Frame with no processors (9xE). The middle position of the three characters means LPAR or non-LPAR model: 93y indicates a non-LPAR model and 9By is an LPAR model. You can also order Expansion Frames with the base frame: 92E for non-LPAR models and 9AE for LPAR models.



Figure 4-26 Model 932 and 9B2

DS8100 Turbo Model 931 features

The DS8100 Turbo Model 931 has the following features:

- ▶ Two processor complexes, each with a System p5 POWER5+ 2.2 GHz two-way Computer Electronics Complex (CEC).
- ▶ Base frame with up to 128 DDMs for a maximum base frame disk storage physical capacity of 128 TB with SATA DDMs.
- ▶ Up to 128 GB of processor memory, also referred to as the cache. Note that the DS8000 supports concurrent cache upgrades.
- ▶ Up to 16 four-port Fibre Channel/FICON host adapters (HAs) of 2 or 4 Gbps. Each port can be independently configured as either:
 - FCP port to open systems hosts attachment
 - FCP port for Metro Mirror, Global Copy, Global Mirror, and Metro/Global Mirror connectivity
 - FICON port to connect to System z hosts
 - FICON port for z/OS Global Mirror connectivity

This totals up to 64 ports with any mix of FCP and FICON ports. ESCON host connection is also supported, but with ESCON, a host adapter contains only two ESCON ports and both must be ESCON ports. A DS8000 can have both ESCON adapters and Fibre Channel/FICON adapters at the same time.

The DS8100 Turbo Model 931 can connect to one Expansion Frame Model 92E. The base and Expansion Frame together allow for a maximum capacity of 384 DDMs: 128 DDMs in the base frame and 256 DDMs in the Expansion Frame. With all of these being 1 TB enterprise DDMs, this results in a maximum disk storage subsystem physical capacity of 384 TB.

DS8300 Turbo Models 932 and 9B2 features

The DS8300 Turbo Models 932 and 9B2 offer higher storage capacity than the DS8100, and they provide greater throughput by means of their dual four-way processor implementation.

The Model 9B2 also has the capability of providing two storage images through two storage system LPARs within the same physical storage unit. Both storage images split the processor, cache, and adapter resources in a 50:50 manner.

Both models provide the following features:

- ▶ Two processor complexes, each with a System p5 POWER5+ 2.2 GHz four-way CEC.
- ▶ Base frame with up to 128 DDMs for a maximum base frame disk storage physical capacity of 128 TB with FC DDMs. This is the same as the DS8100 base frame.
- ▶ Up to 256 GB of processor memory, also referred to as the cache. Note that the DS8000 supports concurrent cache upgrades.
- ▶ Up to 16 four-port Fibre Channel/FICON host adapters (HAs) of 2 or 4 Gbps in the base frame, when configured without an Expansion Frame. Each port can be independently configured as either:
 - FCP port to open systems hosts attachment
 - FCP port link for Metro Mirror, Global Copy, Global Mirror, and Metro/Global Mirror connectivity
 - FICON port to connect to System z hosts
 - FICON port for z/OS Global Mirror connectivity

This totals up to 64 ports for the base frame with any mix of FCP and FICON ports. ESCON host connection is also supported, but with ESCON, a host adapter contains only two ESCON ports and both must be ESCON ports. A DS8000 can have both ESCON adapters and Fibre Channel/FICON adapters at the same time.

The DS8300 models can support either one or two expansion frames. With expansion frame 92E or 9AE, you can expand the disk storage capacity and number of adapters as follows:

- ▶ With one expansion frame, you can support the following expanded capacity and number of adapters:
 - Up to 384 disk drives, for a maximum capacity of 384 TB with 1 TB SATA DDMs
 - Up to 32 host adapters (HAs); this can be an intermix of Fibre Channel/FICON (four-port) and ESCON (two-port) adapters
- ▶ With two Expansion Frames, the disk capacity of the DS8300 models expands to:
 - Up to 640 DDMs in total, for a maximum disk storage physical capacity of 640 TB when using 1 TB SATA DDMs
- ▶ With three Expansion Frames, the disk capacity of the DS8300 models expands to:
 - Up to 896 DDMs in total, for a maximum disk storage physical capacity of 896 TB when using 1 TB SATA DDMs
- ▶ With four Expansion Frames, the disk capacity of the DS8300 models expands to:
 - Up to 1024 DDMs in total, for a maximum disk storage physical capacity of 1 PB when using 1 TB SATA DDMs

Table 4-3 summarizes the model capabilities.

Table 4-3 Comparison of models for capacity

	2-way base frame only	2-way + expansion frame	4-way base frame only	4-way + expansion frame	4-way + two expansion frames
Disk drives: ▶ 73 GB 15k RPM ▶ 146 GB 10k RPM ▶ 300 GB 10k RPM ▶ 500 GB 7.5k RPM ▶ 1 TB SATA	16 to 128 (increment of 16)	16 to 384 (increment of 16)	16 to 128 (increment of 16)	16 to 384 (increment of 16)	16 to 640 (increment of 16)
Maximum capacity	128 TB	384 TB	128 TB	384 TB	640 TB

DS8700 Models 941 and 94E

The DS8700 series currently has one model available: the DS8700 model 941. You can also order Expansion Frames with the base frame: 94E.

In the following sections, we describe these models further:

▶ **DS8700 Model 941:**

This model features a dual two-way processor complex with only 64 DDMs and 8 FC Adapter cards, a dual two-way processor complex with 128 DDMs and 16 FC Adapter cards and a dual four-way processor complex with 128 DDMs and 16 FC Adapter cards.

▶ **DS8700 Model 94E:**

This model represents the Expansion Frame for the 941 model and includes additional FC Adapter cards to allow a maximum configuration of 32 FC Adapter cards.

Note: The 3th, 4th and 5th frame from an DS8300 (92E) can be reused in the DS8700.

The DS8700 Model 941 has the following features:

- ▶ Two processor complexes, each with a System p6 POWER6 4.7GHz two-way or four-way Central Electronics Complex (CEC).
- ▶ A two-way configuration requires two battery packs; a four-way configuration requires three battery packs.
- ▶ Base frame with up to 128 DDMs for a maximum base frame disk storage capacity of 128 TB with SATA DDMs.
- ▶ Up to 128 GB (two way) or 384 GB (four way) of processor memory, also referred to as the *cache*. Note that the DS8700 supports concurrent cache upgrades.
- ▶ Up to 16 four-port Fibre Channel/FICON host adapters (HAs) of 4 Gbps. Each port can be independently configured as either:
 - FCP port to open systems hosts attachment.
 - FCP port for Metro Mirror, Global Copy, Global Mirror, and Metro/Global Mirror connectivity.
 - FICON port to connect to System z hosts.
 - FICON port for z/OS Global Mirror connectivity.
 - This totals up to 64 ports with any mix of FCP and FICON ports.

- The DS8700 Model 941 can connect up to four Expansion Frame Model 94E/92E. Figure 4-15 on page 101 displays a front view of a DS8700 Model 941 and 94E with the covers off. The base and Expansion Frame together allow for a maximum capacity of 384 DDMs: 64 DDMs minimum or 128 DDMs maximum in the base frame and 256 DDMs in the Expansion Frame. With all of these being 1 TB enterprise DDMs, this results in a maximum disk storage subsystem capacity of 384 TB.

Figure 4-27 shows the maximum configuration of a DS8700 with the 941 base frame plus a 94E Expansion Frame and provides the front view of the basic structure and placement of the hardware components within both frames.

Note: A model 941 supports non-disruptive upgrades from a 64 DDM install to a full four expansion rack unit.

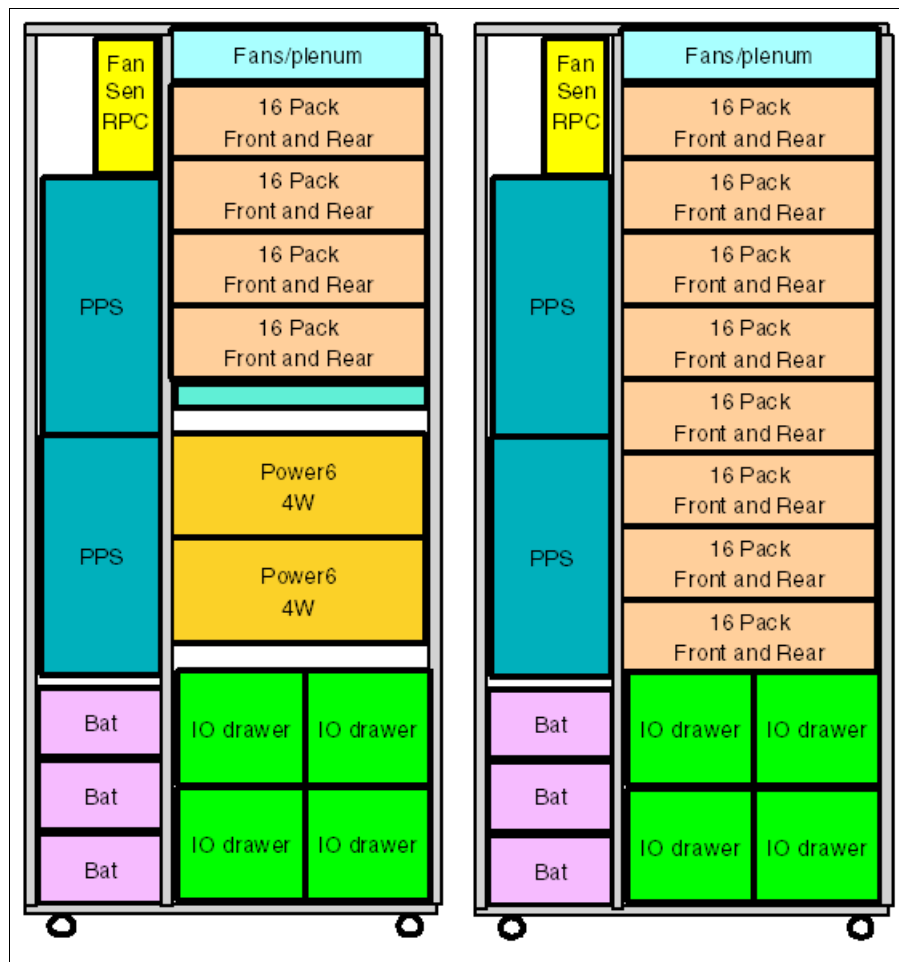


Figure 4-27 Maximum DS8700 configuration: 941 base unit and 94E expansion

Orders: New DS8700 systems can be ordered with Full Disk Encryption drives.

DS8700 models vary in processor type, maximum disk storage capacity, processor memory, and maximum host adapters. Table 4-4 provides a comparison of the DS8700 model 941 and its available combination of resources.

Table 4-4 DS8700 series model comparison 941 and additional resources

Base model	Images	Expansion model	Processor type	DDMs	Processor memory	Max. host adapters
941	1	None	2-way 4.7 GHz	<= 64	<=128 GB	<= 8
941	1	None	2-way 4.7 GHz	<= 128	<=128 GB	<= 16
941	1	None	4-way 4.7 GHz	<= 64	<= 384GB	<= 16
		1 x 94E		<= 384		<= 32
		2 x 94E		<= 640		
		3 x 94E		<= 896		
		4 x 94E		<= 1024		

Depending on the DDM sizes, which can vary within a 941 or 94E and the number of DDMs, the total capacity is calculated accordingly.

4.4.7 IBM Standby Capacity on Demand offering for the DS8000

Standby Capacity on Demand (Standby CoD) provides standby on demand storage for the DS8000 and allows you to access the extra storage capacity whenever the need arises. With Standby CoD, IBM installs up to 64 drives (in quantities of 16) in your DS8000. At any time, you can logically configure your Standby CoD capacity for use. It is a nondestructive activity that does not require intervention from IBM. Upon logical configuration, you will be charged for the capacity.

4.4.8 More information

For more information about the DS8000 series, see the following Redbooks publications:

- ▶ *The IBM TotalStorage DS8000 Series: Concepts and Architecture*, SG24-6452
- ▶ *IBM System Storage DS8000: Architecture and Implementation*, SG24-6786
- ▶ *DS8000 Copy Services for IBM System z*, SG24-6787
- ▶ *IBM System Storage DS8000: Copy Services in Open Environments*, SG24-6788
- ▶ *DS8000 Performance Monitoring and Tuning*, SG24-7146
- ▶ *IBM System Storage DS8700: Architecture and Implementation*, SG24-8786

4.5 Storage management

In this section we provide information about configuration and management capabilities of the IBM System Storage Enterprise Disk Systems.

4.5.1 DS Management Console

The IBM DS6000 and DS8000 series offer new management tools and interfaces that are applicable to both products.

IBM System Storage DS Storage Manager GUI

The DS Storage Manager is a web based GUI that is used to perform logical configurations and Copy Services management functions. It is accessed by a web browser.

You have the following options to use the DS Storage Manager:

- ▶ **Simulated (Offline) Configuration:**

With this option, you can create or modify logical configurations when disconnected from the network. After creating the configuration, you can save it and then apply it to the storage system later when connected to the network (Figure 4-28).



Figure 4-28 DS6000 Simulated Manager

- Real-time (Online) Configuration:

This option provides real-time management support for logical configuration and Copy Services features for a network attached storage system (Figure 4-29).

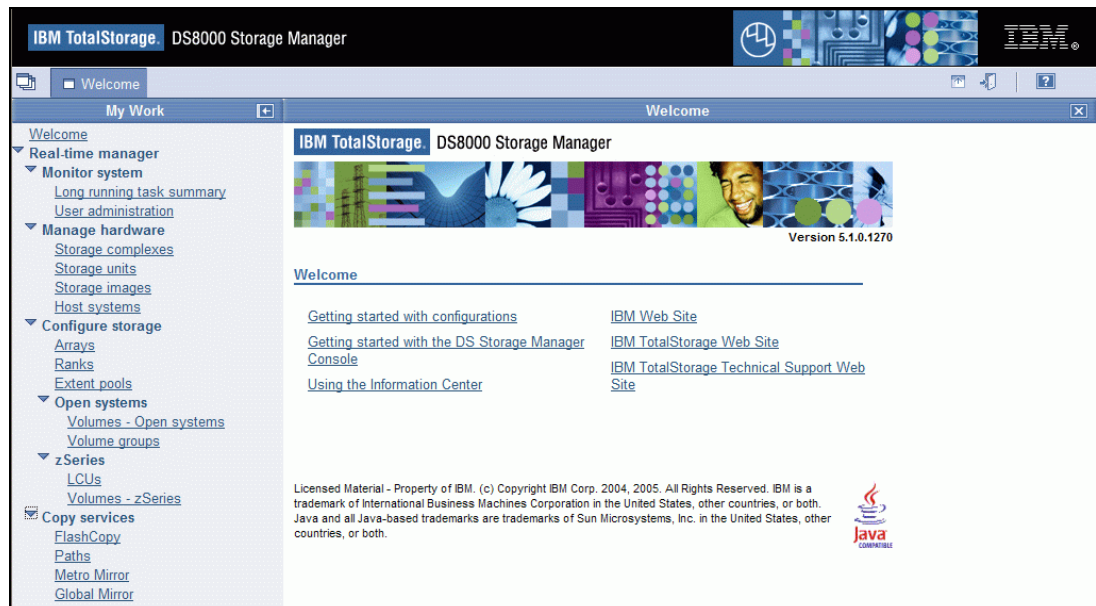


Figure 4-29 DS8000 Real-time Manager

The DS8000 Storage Manager contains more diagnostic features than the DS6000 GUI.

The HMC is the focal point for DS8000 management with multiple functions, including:

- DS8000 power control
- Storage provisioning
- Advanced Copy Services management
- Interface for on-site service personnel
- Call Home and problem management
- Remote support
- TKLM encryption functions

The HMC is the point where the DS8000 is connected to the customer network. It provides the services that the customer needs to configure and manage the storage, and it also provides the interface where service personnel will perform diagnostics and repair actions. The HMC is the contact point for remote support, both modem and VPN.

The DS Storage Manager provides a graphical user interface (GUI) to configure the IBM System Storage DS8700 series and manage DS8700 Copy Services. The DS Storage Manager GUI (DS GUI) is invoked from SSPC.

For DS GUI changes related to disk encryption, see *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

For DS GUI changes related to LDAP authentication, see *IBM System Storage DS8000: LDAP Authentication*, REDP-4505.

IBM System Storage DS Command-Line Interface (CLI)

The DS CLI is a single CLI which can perform a full set of commands for logical configuration and/or Copy Services activities. IDS CLI commands can also be saved in a script, which can enhance productivity because it eliminates the previous requirement (on ESS) to create and save a task using the GUI (Example 4-1).

The following list highlights a few of the functions that you can perform with the DS command-line interface:

- ▶ Create user IDs that can be used with the GUI and the DS CLI.
- ▶ Manage user ID passwords.
- ▶ Install activation keys for licensed features.
- ▶ Manage storage complexes and units.
- ▶ Configure and manage storage facility images.
- ▶ Create and delete RAID arrays, ranks, and extent pools.
- ▶ Create and delete logical volumes.
- ▶ Manage host access to volumes.
- ▶ Check the current Copy Services configuration that is used by the storage unit.
- ▶ Create, modify, or delete Copy Services configuration settings.
- ▶ Integration of LDAP policy usage and configuration.
- ▶ Implement encryption functionality.

Internet Protocol: For DS8000 with Licensed Machine Code (LMC) level 5.40xx.xx or later, you have the possibility to access the HMC over the newest Internet Protocol IPv6, IPv4, or both. Remember to have the DSCLI version corresponding to the LMC level installed on your system.

Example 4-1 DS CLI

```
dsccli> lssi
Date/Time: 27 June, 2005 9:47:23 IBM DSCLI Version: 5.1.0.276
Name ID          Storage Unit    Model WWNN          State ESSNet
=====
IBM.2107-75ABTV1 IBM.2107-75ABTV0 9A2    5005076303FFC663 Online Enabled
IBM.2107-75ABTV2 IBM.2107-75ABTV0 9A2    5005076303FFCE63 Online Enabled
dsccli> lsarray
Date/Time: 27 June 2006 9:48:27 IBM DSCLI Version: 5.1.0.276 DS: IBM.2107-7ABTV1
Array State Data RAIDtype  arsite Rank DA Pair DDMcap (10^9B)
=====
A0 Assigned Normal 5 (6+P+S) S1      R0    0          146.0
A1 Assigned Normal 5 (6+P+S) S2      R1    0          146.0
dsccli>
```

For more information, see the *IBM System Storage DS8000: Command-Line Interface User's Guide*, SC26-7916.

For DS CLI commands related to disk encryption, see the *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500.

4.5.2 IBM TotalStorage Productivity Center

The IBM TotalStorage Productivity Center is an open storage management solution that helps to reduce the effort of managing complex storage infrastructures, to increase storage capacity utilization and to improve administrative efficiency. It is designed to enable an agile storage infrastructure that can respond to on demand storage needs.

TPC is the integration point for storage and fabric management and replication. It provides a launchpad for the following IBM TotalStorage Open Software Family and Tivoli products:

- ▶ IBM TotalStorage Productivity Center for Disk
- ▶ IBM TotalStorage Productivity Center for Data
- ▶ IBM TotalStorage Productivity Center for Fabric
- ▶ IBM TotalStorage Productivity Center for Replication

For more information about the IBM TotalStorage Productivity Center, see Chapter 18, “IBM Tivoli Storage Productivity Center” on page 473.

4.6 Copy Services

Advanced Copy Services are enterprise-level, leading-edge functions that address an organization's needs for disaster recovery, data migration, and data duplication.

Copy Services in the DS Family includes the following optional licensed functions:

- ▶ IBM System Storage FlashCopy and IBM FlashCopy SE, which are point-in-time copy functions
- ▶ Remote mirror and copy functions, which include:
 - IBM System Storage Metro Mirror, previously known as synchronous PPRC
 - IBM System Storage Global Copy, previously known as PPRC Extended Distance
 - IBM System Storage Global Mirror, previously known as asynchronous PPRC
 - IBM System Storage Metro/Global Mirror, a 3-site solution to meet the most rigorous business resiliency needs
- ▶ Additionally for the System z users, the following options are available:
 - z/OS Global Mirror, previously known as Extended Remote Copy (XRC)
 - z/OS Metro/Global Mirror, a 3-site solution that combines z/OS Global Mirror and Metro Mirror

Many design characteristics of the DS systems and its data copy and mirror capabilities and features contribute to the protection of your data, 24 hours a day and seven days a week.

4.6.1 IBM System Storage FlashCopy: Point-in-time Copy feature

The IBM System Storage FlashCopy (see Figure 4-30) helps reduce or eliminate planned outages for critical applications. FlashCopy enables data to be copied in the background while making both source and copied data available to users almost immediately. The point-in-time copy created by FlashCopy is typically used when a copy of production data must be produced with minimal application downtime. It can be used for online backup, testing of new applications, or for copying a database for data mining purposes.

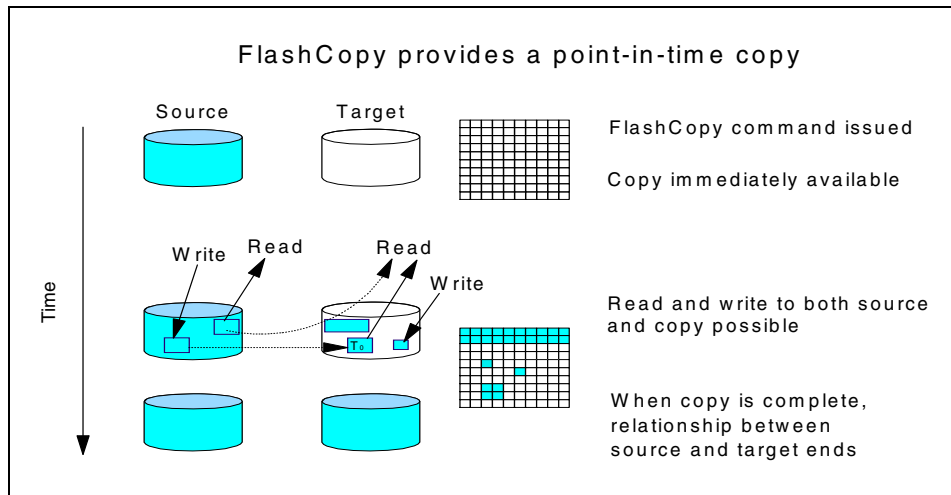


Figure 4-30 FlashCopy

FlashCopy is an additional charged feature. You have to order the Point-in-time Copy feature, which includes FlashCopy.

FlashCopy supports many advanced capabilities, including:

- ▶ **Data Set FlashCopy:**
Allows a FlashCopy of a data set in a System z environment.
- ▶ **Multiple Relationship FlashCopy:**
Allows a source volume to have multiple targets simultaneously.
- ▶ **Incremental FlashCopy:**
Allows update of a FlashCopy target without having to recopy the entire volume.
- ▶ **Consistency Group function:**
Used to help create a consistent point-in-time copy across multiple LUNs or volumes, and even across multiple DS6800 systems as well as across DS8000 series.
- ▶ **In-band commands over remote mirror link:**
In a remote mirror environment, commands to manage FlashCopy at the remote site can be issued from the local or intermediate site and transmitted over the remote mirror Fibre Channel links. This eliminates the need for a network connection to the remote site solely for the management of FlashCopy.

► **FlashCopy to a Remote Mirror Primary**

Allows a FlashCopy target volume to be used also as a remote mirror primary volume. This process allows you to create a point-in-time copy and then make a copy of that data at a remote site with Metro Mirror or Global Copy (Figure 4-31).

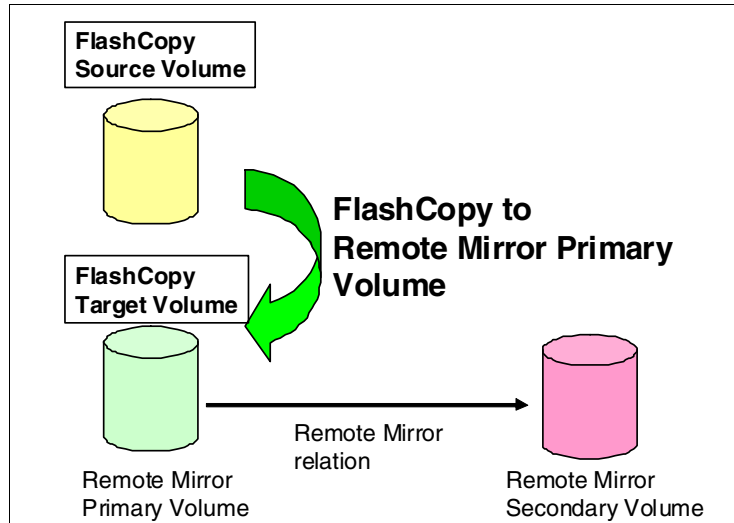


Figure 4-31 Establish FlashCopy on existing Remote Mirror and Copy Primary

4.6.2 Remote Mirror and Copy feature

Remote Mirror and Copy is another separately orderable priced feature, which includes Metro Mirror, Global Copy, and Global Mirror. The local and remote storage systems must have a Fibre Channel connection between them. The Fibre Channel ports used for Remote Mirror and Copy can be configured either as a dedicated remote mirror link or as a shared port between remote mirroring and Fibre Channel Protocol (FCP) data traffic.

The DS6000 and DS8000 series systems can participate in remote mirror and copy solutions with each other as well with other systems which support this service.

IBM supports the following remote mirror and copy solutions.

IBM System Storage Metro Mirror (Synchronous PPRC)

Metro Mirror is a remote-mirroring technique for all supported servers, including z/OS and open systems. It is designed to constantly maintain an up to date copy of the local application data at a remote site which is within the metropolitan area (typically up to 300 km away using DWDM). With synchronous mirroring techniques, data currency is maintained between sites, though the distance can have an impact on performance.

Metro Mirror is used primarily as part of a business continuance solution for protecting data against disk storage system loss or complete site failure (Figure 4-32).

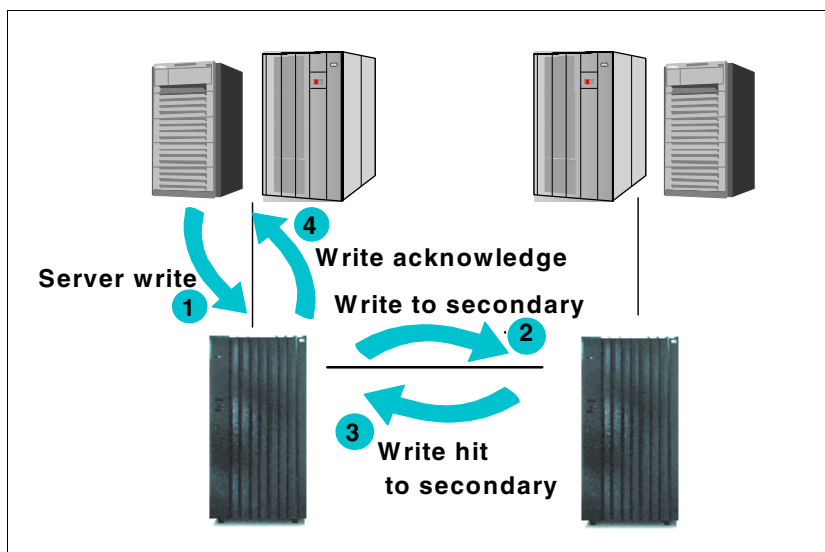


Figure 4-32 Metro Mirror

IBM System Storage Global Copy (Asynchronous PPRC-XD)

Global Copy (Figure 4-33) is an asynchronous remote copy function for z/OS and open systems for longer distances than are possible with Metro Mirror. With Global Copy, write operations complete on the primary storage system before they are received by the secondary system. This capability is designed to prevent the primary system's performance from being affected by wait-time from writes on the secondary system. Therefore, the primary and secondary copies can be separated by any distance. This function is appropriate for remote data migration, off-site backups, and transmission of inactive database logs at virtually unlimited distances.

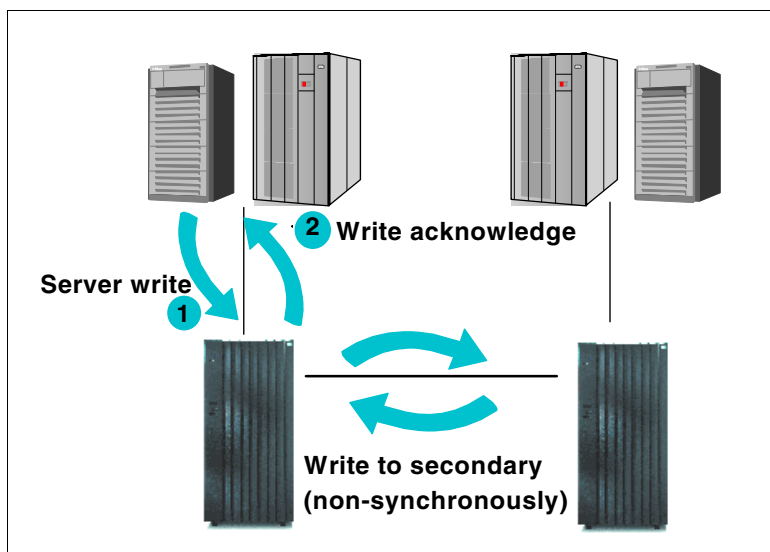


Figure 4-33 Global Copy

IBM System Storage Global Mirror (Asynchronous PPRC)

Global Mirror (Figure 4-34) copying provides a two-site extended distance remote mirroring function for z/OS and open systems servers. With Global Mirror, the data that the host writes to the storage unit at the local site is asynchronously shadowed to the storage unit at the remote site. A consistent copy of the data is then automatically maintained on the storage unit at the remote site. This two site data mirroring function is designed to provide a high-performance, cost-effective global distance data replication and disaster recovery solution.

Global Mirror provide the following benefits:

- ▶ Support for virtually unlimited distances between the local and remote sites. The distances are typically limited only by the capabilities of the network and channel extension products.
- ▶ A consistent and restartable copy of the data at the remote site, created with minimal impact to applications at the local site.
- ▶ Dynamic selection of the desired recovery point objectives
- ▶ Efficient synchronization of the local and remote sites with support for failover and failback modes.

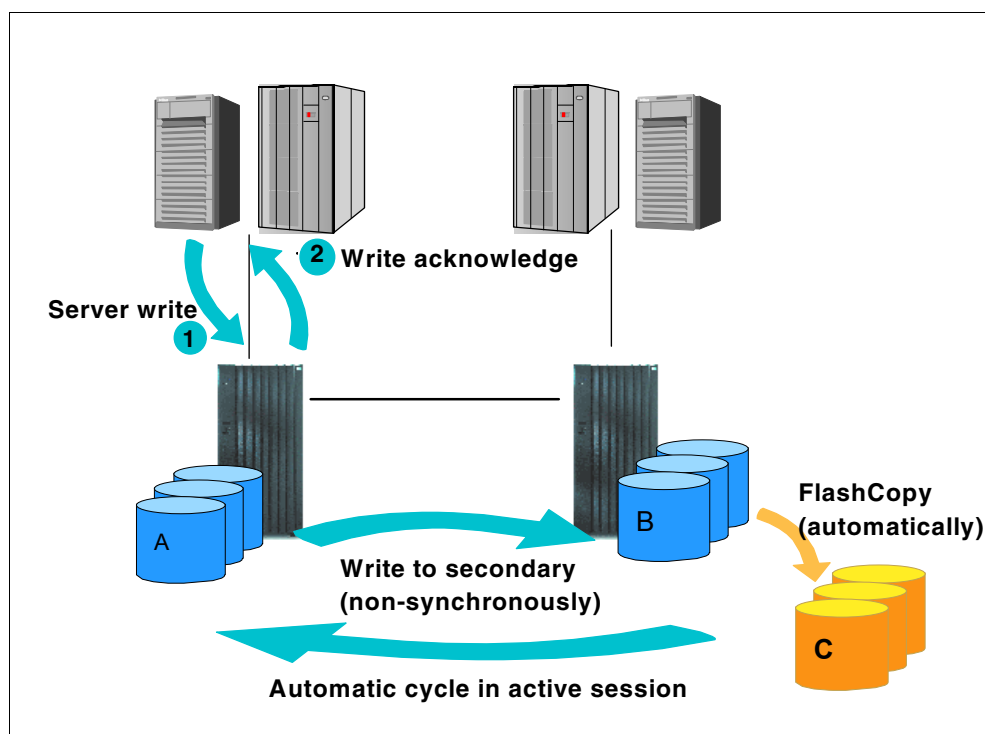


Figure 4-34 Global Mirror

IBM System Storage z/OS Global Mirror (Extended Remote Copy XRC)

z/OS Global Mirror (Figure 4-35) is a remote data mirroring function available for the z/OS and OS/390® operating systems. It maintains a copy of the data asynchronously at a remote location over unlimited distances. z/OS Global Mirror is well suited for large System z server workloads and can be used for business continuance solutions, workload movement, and data migration.

Attention: The DS6000 series systems can only be used as a target system in z/OS Global Mirror operations.

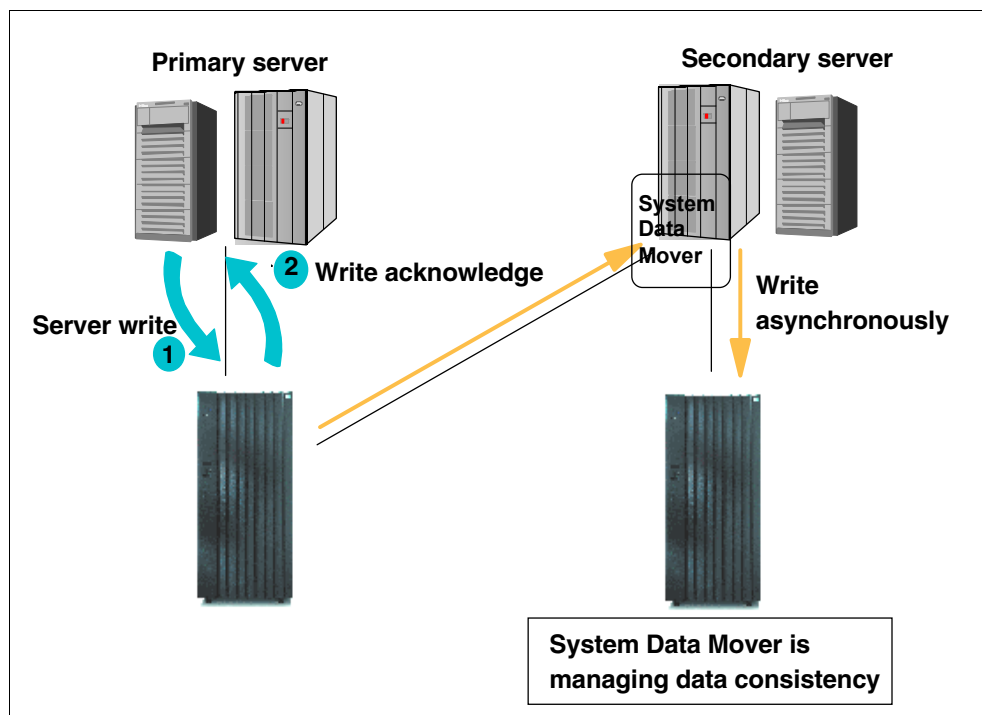


Figure 4-35 z/OS Global Mirror

IBM System Storage z/OS Metro/Global Mirror

This mirroring capability utilizes z/OS Global Mirror to mirror primary site data to a location that is a long distance away and also uses Metro Mirror to mirror primary site data to a location within the metropolitan area. This enables a 3-site high availability and disaster recovery z/OS solution for even greater protection from unplanned outages (Figure 4-36).

Attention: z/OS Metro/Global Mirror is not supported on DS6000 series.

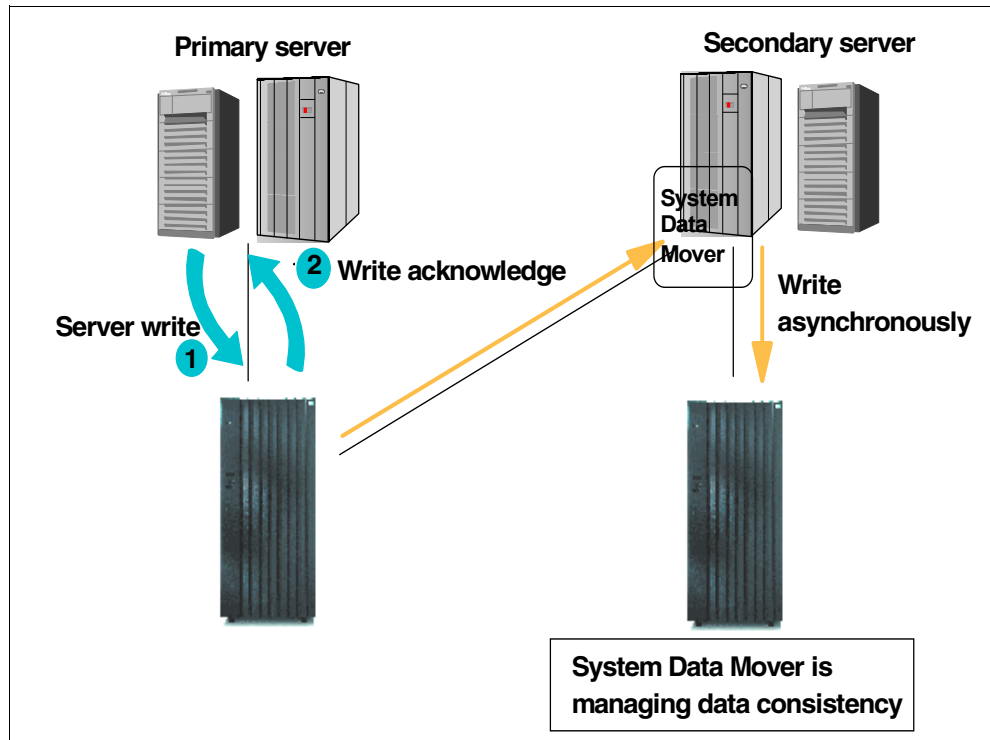


Figure 4-36 z/OS Metro/Global Mirror

For more up to date information, see *DS8000 Copy Services for IBM System z*, SG24-6787.

4.7 Additional information

In this section we provide complementary information about the support and requirements of the IBM System Storage Enterprise Family products.

Licensed Machine Code

The IBM System Storage Enterprise Systems are shipped with IBM Licensed Internal Code (LIC) that is licensed for use by a customer on a specific machine, designated by serial number, under the terms and conditions of the IBM Customer Agreement or the IBM Agreement for Licensed Internal Code. New releases of LIC offer improvements in both function and reliability. It is important that customers obtain the latest level of LIC.

All features and functions are available either as a standard or as an optional feature and include the following capabilities:

- ▶ Operating environment and assigned capacity
- ▶ Host attachment capabilities
- ▶ Parallel access volumes on System z
- ▶ LPAR and S-HMC capabilities on the DS8000 series
- ▶ Point-in-time Copy features
- ▶ Remote mirror and copy features

You can activate and manage licensed features from the System Storage DS Manager GUI on the DS6000 and DS8000 series (Figure 4-37).

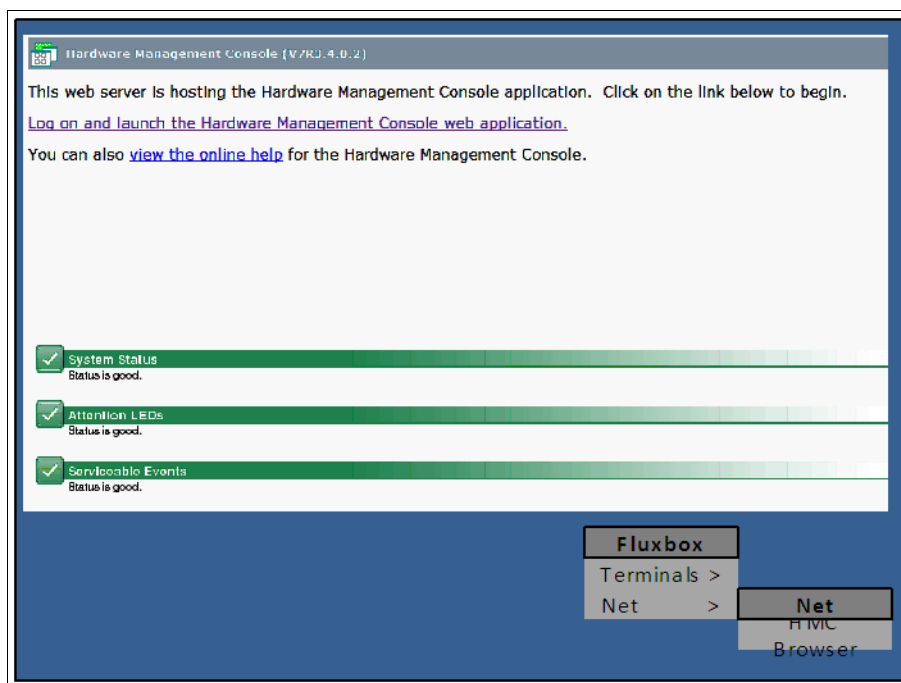


Figure 4-37 DS8700 console welcome screen

Supported environments

The IBM System Storage Enterprise Disk Systems can be connected across a broad range of server environments such as:

- ▶ System z
- ▶ System i
- ▶ System p
- ▶ System x
- ▶ Servers from SUN Microsystems
- ▶ Servers from Hewlett-Packard
- ▶ Other non-IBM based server platforms

New versions of operating systems, servers, file systems, host bus adapters, clustering products and SAN components are constantly announced in the market. Information about the supported environments changes frequently. Therefore you are strongly advised always to consult the online resources.

The Interoperability Matrix always provides the latest information about supported platforms, operating systems, host adapters and SAN infrastructure solutions. It contains detailed specifications about models and versions. It also lists special support items, such as boot support, and exceptions:

- ▶ The DS8000 Interoperability Matrix:
<http://www.ibm.com/servers/storage/disk/ds8000/pdf/ds8000-matrix.pdf>
- ▶ The DS6000 Interoperability Matrix:
<http://www.ibm.com/servers/storage/disk/ds6000/pdf/ds6000-interop.pdf>

4.8 IBM XIV Storage System

The IBM XIV Storage System is a fully scalable enterprise storage system that is based on a grid of standard, off-the-shelf hardware components. It has been designed with an easy to use and intuitive GUI that allows administrators to become productive in a very short time.

The XIV Storage System architecture is designed to deliver performance, scalability, and ease of management while harnessing the high capacity and cost benefits of Serial Advanced Technology Attachment (SATA) drives. The system employs off-the-shelf products as opposed to traditional offerings that use proprietary designs thus requiring more expensive components.

Here are some key properties of this ground breaking solution:

- ▶ Superior architecture boasting massive parallelism and sophisticated distributed algorithms that yield superior power and value
- ▶ Superior performance through maximized utilization of all disks, true distributed cache implementation coupled with more effective cache bandwidth, and practically zero overhead incurred by snapshots
- ▶ Superior reliability through distributed architecture, redundant components, self-monitoring and auto-recovery processes; ability to sustain failure of a complete disk module and three more disks with minimal performance degradation
- ▶ Market leading auto-restoration of system redundancy in case of disk failure (less than 30 minutes for 1 TB drives – compared to several hours for the closest competitor)
- ▶ Superior administration and zero management overhead: no need to ever tune the system in order to attain high-end performance, even after adding or removing disks; no need to make compromises when utilizing snapshots; non-restricted creation and management of unlimited snapshots with practical zero overhead

4.8.1 System models and components

The IBM XIV Storage System Family consists of two machine types, the XIV Storage System (Machine type 2812) Model A14 and the XIV Storage System (Machine type 2810) Model A14. The 2812 Model A14 supports a 3-year warranty to complement the 1-year warranty offered by the existing and functionally equivalent, 2810 Model A14.

The majority of hardware and software features are the same for both machine types. The major differences are listed in Table 4-5.

Table 4-5 Machine type comparisons

Machine type	2810-A14	2812-A14
Warranty	1 year	3 years
CPUs per Interface Module	1 or 2	2
CPUs per Data Module	1	1

New orders for both machine types feature a new low voltage CPU (dual CPU in interface modules), for less power consumption.

Both machine types are available in the following configurations:

- ▶ 6 modules (including 3 Interface Modules)
- ▶ 9 -15 modules (including 6 Interface Modules)

Both machine types include the following components (Figure 4-38):

- ▶ 3-6 Interface Modules, each with 12 SATA disk drives
- ▶ 3-9 Data Modules, each with 12 SATA disk drives
- ▶ An Uninterruptible Power Supply (UPS) module complex comprising three redundant UPS units
- ▶ Two Ethernet switches and an Ethernet Switch Redundant Power Supply (RPS)
- ▶ A Maintenance Module
- ▶ An Automatic Transfer Switch (ATS) for external power supply redundancy
- ▶ A modem, connected to the Maintenance Module for externally servicing the system (note that the modem, feature number 910, is not available in all countries).

All of the modules in the system are linked through an internal redundant Gigabit Ethernet network, which enables maximum bandwidth utilization and is resilient to at least any single component failure.

The system and all of its components come pre-assembled and wired in a lockable rack.

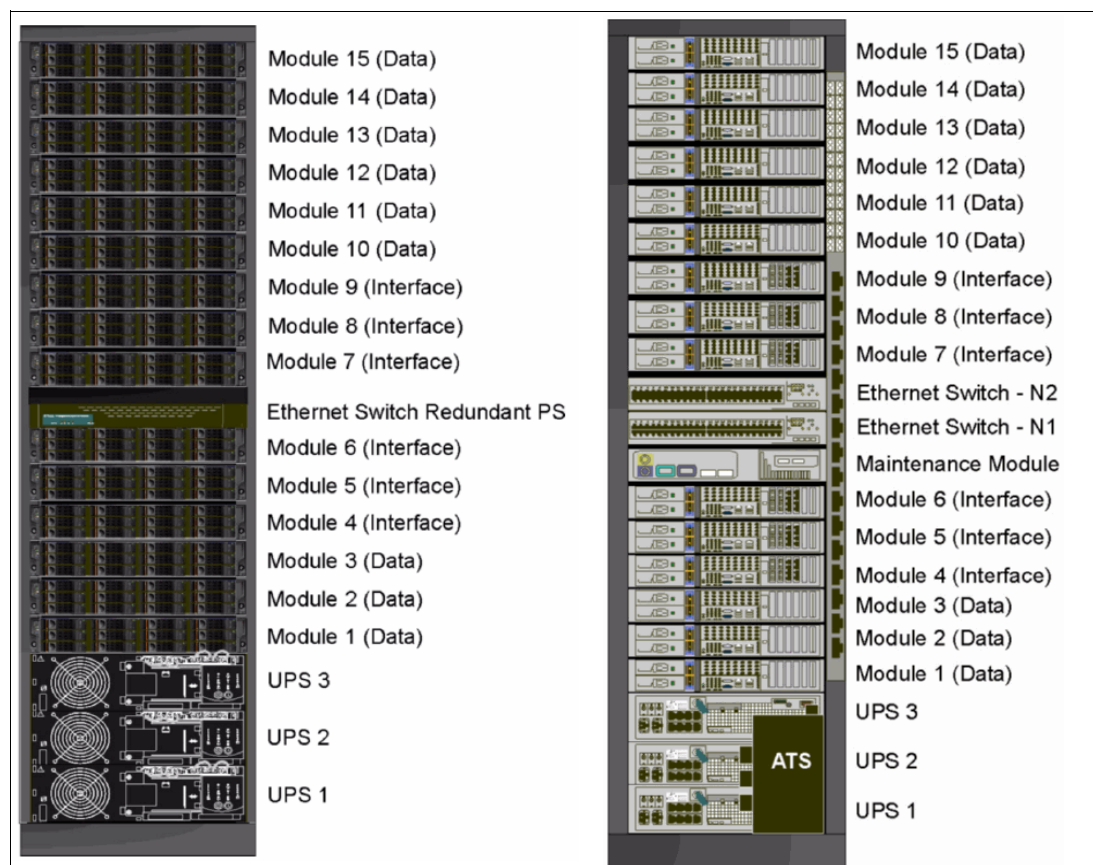


Figure 4-38 Components of XIV system

The XIV Storage System architecture incorporates a variety of features designed to uniformly distribute data across internal resources. This unique data distribution method fundamentally differentiates the XIV Storage System from conventional storage subsystems, thereby offering numerous availability, performance, and management benefits across both physical and logical elements of the system.

Hardware elements

In order to convey the conceptual principles that comprise the XIV Storage System architecture, it is useful to first provide a glimpse of the physical infrastructure.

The primary components of the XIV Storage System are known as *modules*. Modules provide processing, cache, and host interfaces and are based on “off the shelf” Intel based systems. They are redundantly connected to one another through an internal switched Ethernet network, as shown in Figure 4-39. All of the modules work together concurrently as elements of a grid architecture, and therefore, the system harnesses the powerful parallelism inherent in such a distributed computing environment.

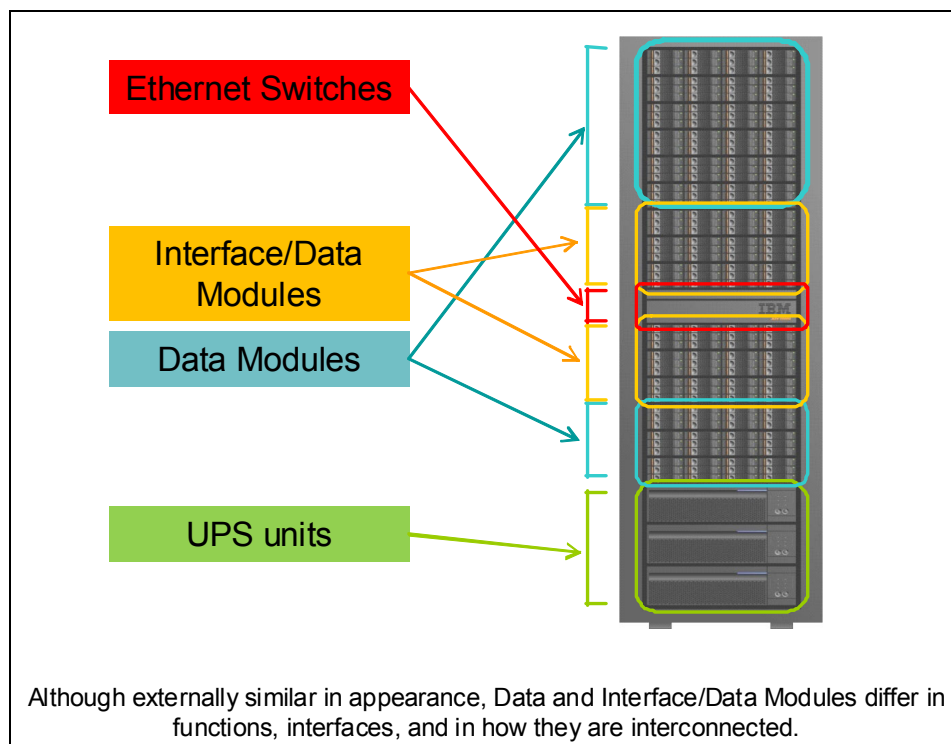


Figure 4-39 IBM XIV Storage System major hardware elements

Figure 4-40 on page 140 summarizes the main hardware characteristics of the IBM XIV Storage System 2810-A14 and 2812-A14.

All XIV hardware components come pre-installed in a standard 19” data center class rack. At the bottom of the rack, an Uninterruptible Power Supply (UPS) module complex, which is made up of three redundant UPS units, is installed and provides power to the various system components.

Fully populated configurations

A fully populated rack contains 9 Data Modules and 6 Interface Modules for a total of 15 modules. Each module is equipped with the following connectivity adapters:

- ▶ USB ports
- ▶ Serial ports
- ▶ Ethernet adapters
- ▶ Fibre Channel adapters (Interface Modules only)

Each module also contains twelve 1 TB Serial Advanced Technology Attachment (SATA) disk drives. This design translates into a total usable capacity of 79 TB (180 TB raw) for the complete system. For information about *usable capacity*, see Figure 4-40.

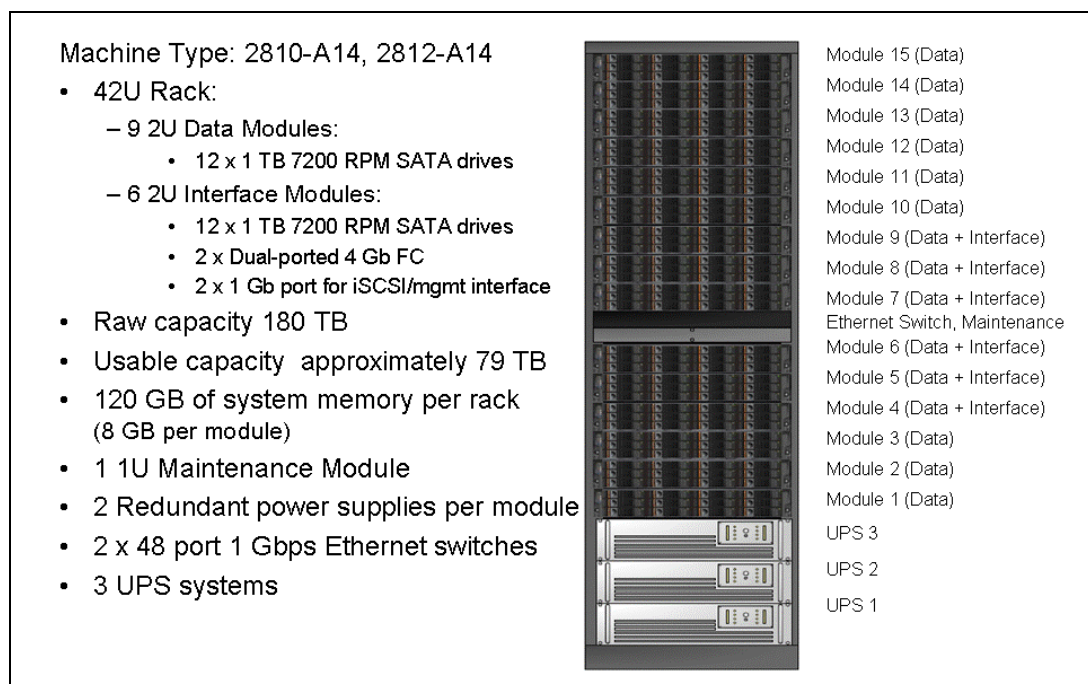


Figure 4-40 Hardware overview - Machine type 2810_2812 - model A14

Data Modules

At a conceptual level, the Data Modules function as the elementary “building blocks” of the system, providing storage capacity, processing power, and caching, in addition to advanced system-managed services. The Data Module’s ability to share and manage system software and services are key elements of the physical architecture, as depicted in Figure 4-41.

Interface Modules

Interface Modules are equivalent to Data Modules in all aspects, with the following exceptions:

- ▶ In addition to disk, cache, and processing resources, Interface Modules are designed to include both Fibre Channel and iSCSI interfaces for host system connectivity, Remote Mirroring, and Data Migration activities. Figure 4-41 on page 141 conceptually illustrates the placement of Interface Modules within the topology of the XIV IBM Storage System architecture.
- ▶ The system services and software functionality associated with managing external I/O reside exclusively on the Interface Modules.

Ethernet switches

The XIV Storage System contains a redundant switched Ethernet network that transmits both data and metadata traffic between the modules. Traffic can flow in any of the following ways:

- ▶ Between two Interface Modules
- ▶ Between two Data Modules
- ▶ Between an Interface Module and a Data Module

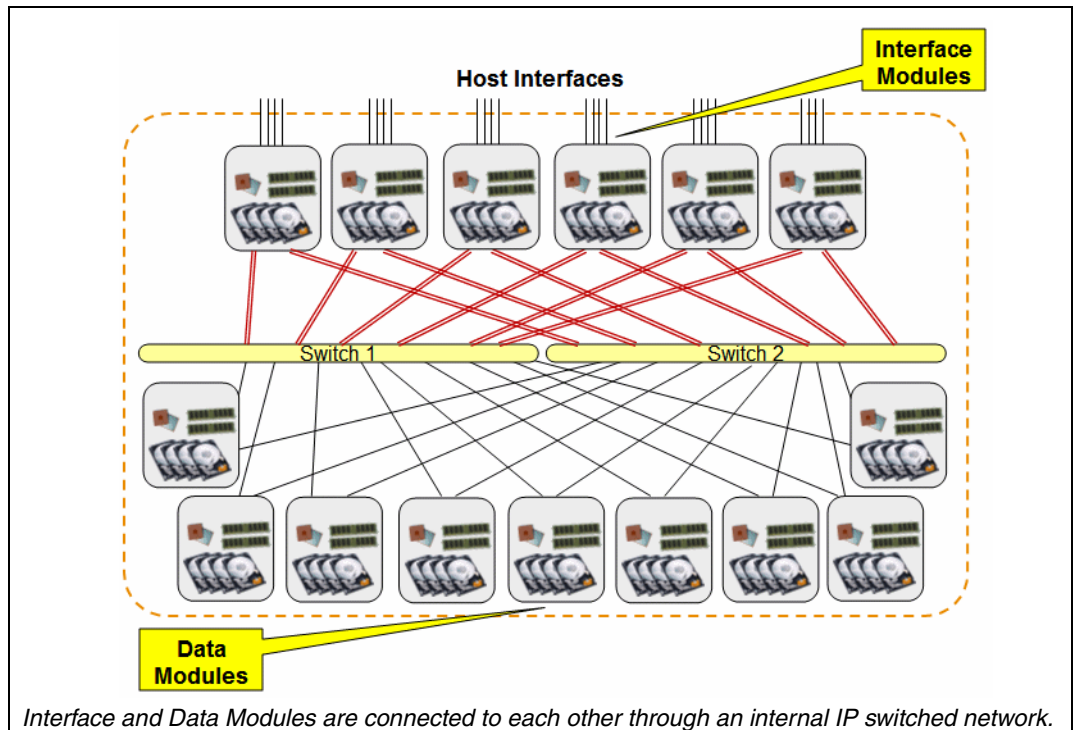


Figure 4-41 Architectural overview

Note: Figure 4-41 depicts the conceptual architecture only. Do not misinterpret the number of connections and such as a precise hardware layout.

4.8.2 Key design features

This section describe the key design features of the XIV Storage System architecture.

Massive parallelism

The system architecture ensures full exploitation of all system components. Any I/O activity involving a specific logical volume in the system is always inherently handled by all spindles. The system harnesses all storage capacity and all internal bandwidth, and it takes advantage of all available processing power, which is as true for host-initiated I/O activity as it is for system-initiated activity, such as rebuild processes and snapshot generation. All disks, CPUs, switches, and other components of the system contribute to the performance of the system at all times.

Workload balancing

The workload is evenly distributed over all hardware components at all times. All disks and modules are utilized equally, regardless of access patterns. Despite the fact that applications might access certain volumes more frequently than other volumes, or access certain parts of a volume more frequently than other parts, the load on the disks and modules will be balanced perfectly.

Pseudo-random distribution ensures consistent load-balancing even after adding, deleting, or resizing volumes as well as adding or removing hardware. This balancing of all data on all system components eliminates the possibility of a hot-spot being created.

Self-healing

Protection against double disk failure is provided by an efficient rebuild process that brings the system back to full redundancy in minutes. In addition, the XIV Storage System extends the self-healing concept, resuming redundancy even after failures in components other than disks.

True virtualization

Unlike other system architectures, storage virtualization is inherent to the basic principles of the XIV Storage System design. Physical drives and their locations are completely hidden from the user, which dramatically simplifies storage configuration, letting the system lay out the user's volume in the optimal way. The automatic layout maximizes the system's performance by utilizing system resources for each volume, regardless of the user's access patterns.

Thin provisioning

The system supports thin provisioning, which is the capability to allocate actual storage to applications on a just-in-time and as needed basis, allowing the most efficient use of available space and as a result, significant cost savings compared to traditional provisioning techniques. This is achieved by defining a logical capacity that is larger than the physical capacity and utilizing space based on what is consumed rather than what is allocated.

Processing power

The IBM XIV Storage System open architecture benefits from the latest processor technologies and is more scalable than solutions that are based on a closed architecture. The IBM XIV Storage System avoids sacrificing the performance of one volume over another, and therefore requires little to no tuning.

4.8.3 IBM XIV Storage System software

The IBM XIV system software 10.1 (or later) provides the functions of the system, which include:

- ▶ **Bundled advanced features:**

All the features of the XIV including advanced features such as migration and mirroring are included free of charge and apply to the entire storage capacity.

- ▶ **Non-Disruptive Code Load (NDCL):**

System software code can be upgraded without requiring downtime. This enables 'non-stop' production environments to remain running while new code is upgraded.

The code upgrade is run on all modules in parallel and the process is fast enough to minimize impact on hosts applications.

No data migration or rebuild process is allowed during the upgrade. Mirroring, if any, will be suspended during the upgrade and automatically reactivated upon completion.

Storage management operations are also not allowed during the upgrade, although the status of the system and upgrade progress can be queried. It is also possible to cancel the upgrade process up to a point of no return.

Note that the NDCL does not apply to specific components firmware upgrades (for instance, module BIOS and HBA firmware). Those require a phase in / phase out process of the impacted modules.

► Support for 16000 snapshots:

The snapshot capabilities within the XIV Storage System Software utilize a metadata, redirect-on-write design that allows snapshots to occur in a sub-second time frame with little performance overhead. Up to 16000 full or differential copies can be taken. Any of the snapshots can be made writable, and then snapshots can be taken of the newly writable snapshots. Volumes can even be restored from these writable snapshots.

► Remote Mirroring:

The IBM XIV Storage System enables mirroring for business continuity, as well as to distribute workloads across the organization's various data centers, improving efficiency by utilizing existing assets for more diverse uses, such as backup, decision support, testing, and application development.

The system offers mirroring with broad support:

- Replication on volume-based granularity, with up to 16 target XIV systems supported
- All mirroring based on differential snapshot technology; allows moving of minimal amounts of data, avoiding the high cost of communications
- Synchronous and asynchronous mirroring
- Mirroring of volumes and/or consistency groups
- Ability to specify Recovery Point Objectives (RPOs) per mirror
- Flexible replication schedules
- Concurrent and bidirectional mirroring
- FC and iSCSI-based mirroring, with the mirroring protocol agnostic to the host protocol (mirroring by iSCSI volumes served by FC and vice versa)
- Flexible RPO per volume from 30 seconds to 24 hours, enabling flexible planning and priority setting based on business needs

The XIV system's remote mirroring capability provides a flexible, cost effective, simple to implement, and simple to test solution for a disaster recovery site.

► Support for thin provisioning:

Thin provisioning allows administrators to over-provision storage within storage pools; this is done by defining logical volume sizes that are larger than the physical capacity of the pool. Unlike other approaches, the physical capacity only needs to be larger than the actual written data, not larger than the logical volumes. Physical capacity of the pool needs to be increased only when actual written data increases.

► Support for in-band data migration of heterogeneous storage:

The XIV Storage System is also capable of acting as a host, gaining access to volumes on an existing legacy storage system. The XIV is then configured as a proxy to respond to requests between the current hosts and the legacy storage while migrating all existing data in the background. In addition, XIV supports thick-to-thin data migration, which allows the XIV Storage System to reclaim any allocated space that is not occupied by actual data.

► Authentication using Lightweight Directory Access Protocol (LDAP):

LDAP can be used to provide user logon authentication allowing the XIV Storage System to integrate with Microsoft Active Directory (AD) or Sun Java Systems Directory Server (formerly Sun ONE Directory). Multiple directory servers can be configured to provide redundancy if one becomes unavailable.

► Robust user auditing with access control lists:

The XIV Storage System Software offers the capability for robust user auditing with Access Control Lists (ACLs) in order to provide more control and historical information.

- Support for Tivoli Storage Productivity Center:

Tivoli Storage Productivity Center (TPC) can now discover XIV Storage Systems and all internal components, manage capacity for storage pools including allocated, unallocated, and available capacity with historical trending on utilization. It can also receive events and define policy-based alerts based on user-defined triggers and thresholds.

IBM XIV Storage Manager GUI

The XIV Storage Manager GUI acts as the management console for the XIV Storage System. A simple and intuitive GUI enables storage administrators to manage and monitor all system aspects easily, with almost no learning curve. Figure 4-42 shows one of the top level configuration panels.



Figure 4-42 The IBM XIV Storage Manager GUI

The GUI is also supported on the following platforms:

- Microsoft Windows 2000, Windows ME, Windows XP, Windows Server 2003, and Windows Vista
- Linux (Red Hat 5.x or equivalent)
- AIX 5.3, AIX 6.1
- Solaris v9, Solaris v10
- HPUX 11i v2, HPUX 11i v3

The GUI can be downloaded at the following site:

<ftp://ftp.software.ibm.com/storage/XIV/GUI/>

The XIV system's virtualization and simplicity, together with its powerful GUI management, significantly reduce the time and effort required for managing storage resources. Note that GUI and XCLI (command line interface) are packaged together.

GUI enhancements:

- ▶ You can use the organization's LDAP (Lightweight Directory Access Protocol) server to manage users, passwords and access lists.
- ▶ Cluster private mapping: You can define private LUNs for hosts even if they are members of a cluster.
- ▶ Additional statistics filters:
 - System statistics per host ports
 - System capacity usage over time
- ▶ Up to 15 systems can be managed with a single GUI (up from 9 systems).
- ▶ Demo mode provides the ability to run the GUI without connecting to any live system, using predefined sample data and configuration. To access Demo mode, login with the user name "P10DemoMode".
- ▶ Snapshots Tree and Snapshot Groups Tree views can now be used to actively manage snapshots (for example, create, delete, map, and so on.)
- ▶ You can execute tasks for multiple objects at once, for example, delete volumes, activate mirroring, and so on.
- ▶ Capacity usage thresholds: You can adjust the thresholds that determine the alert level for capacity usage by pool volumes and snapshots.
- ▶ Gradual events retrieval: Events are retrieved from the system gradually, to allow better performance as well as the ability to retrieve any number of events that match the used filter (up to a maximum of 10,000).

Advanced support features:

- ▶ Support logs: You can automatically collect all the system logs required by IBM-XIV support for timely analysis of the system status and health.
- ▶ Customer contact information forms: Use these forms to facilitate communication.
- ▶ VPN setup: VPN allows secure connectivity to IBM-XIV Remote Support Center.

IBM XIV Storage System XCLI

The XIV Storage System offers a comprehensive set of Extended Command Line Interface (XCLI) commands to configure and monitor the system. All the functions available in the GUI are also available in the XCLI. The XCLI can be used in a shell environment to interactively configure the system or as part of a script to perform lengthy and/or complex tasks. Example 4-2 shows a command being run in the XCLI interactive mode (XCLI session).

Example 4-2 The XCLI interactive mode

```
>> config_get
Name                               Value
dns_primary
dns_secondary
email_reply_to_address
email_sender_address
email_subject_format               {severity}: {description}
iscsi_name                         iqn.2005-10.com.xivstorage:000019
machine_model                      A14
machine_serial_number              MN00019
machine_type                       2810
ntp_server
snmp_community                     XIV
```

snmp_contact	Unknown
snmp_location	Unknown
snmp_trap_community	XIV
support_center_port_type	Management
system_id	19
system_name	XIV MN00019

The XCLI is supported on the following platforms:

- ▶ Microsoft Windows 2000, Windows ME, Windows XP, Windows Server 2003, and Windows Vista
- ▶ Linux (Red Hat 5.x or equivalent)
- ▶ AIX 5.3, AIX 6.1
- ▶ Solaris v9, Solaris v10
- ▶ HP-UX 11i v2, HP-UX 11i v3

The XCLI can be downloaded at the following site:

<ftp://ftp.software.ibm.com/storage/XIV/GUI/>

4.8.4 Parallelism

Here we provide an overview of the concept of *parallelism*. Complete information about the XIV System can be found in the IBM Redbooks publication, *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

The concept of parallelism pervades all aspects of the XIV Storage System architecture by means of a balanced, redundant data distribution scheme in conjunction with a pool of distributed (or grid) computing resources. In order to explain the principle of parallelism further, it is helpful to consider the ramifications of both the hardware and software implementations independently. We subsequently examine virtualization principles in 4.8.5, “Full storage virtualization” on page 148.

Important: The XIV Storage System exploits parallelism at *both* the hardware and software levels.

Hardware parallelism and grid architecture

The XIV grid design in Figure 4-43 involves the following characteristics:

- ▶ Both Interface Modules and Data Modules work together in a distributed computing sense. However, the Interface Modules also have additional functions and features associated with host system connectivity.
- ▶ The modules communicate with each other through the internal, redundant Ethernet network.
- ▶ The software services and distributed computing algorithms running within the modules collectively manage all aspects of the operating environment.

Design principles

The XIV Storage System grid architecture, by virtue of its distributed topology and “off the shelf” Intel components, ensures that the following design principles are possible:

- ▶ Performance:
 - The relative effect of the loss of a module is minimized.
 - All modules are able to participate equally in handling the total workload.

- ▶ **Compatibility:**
 - Modules consist of standard “off the shelf” components.
 - Because components are not specifically engineered for the system, the resources and time required for the development of newer hardware technologies are minimized.
- ▶ **Scalability:**
 - Computing resources can be dynamically changed.
 - “Scaled out” by adding new modules to accommodate both new capacity and new performance demands.
 - “Scaled up” by upgrading modules.

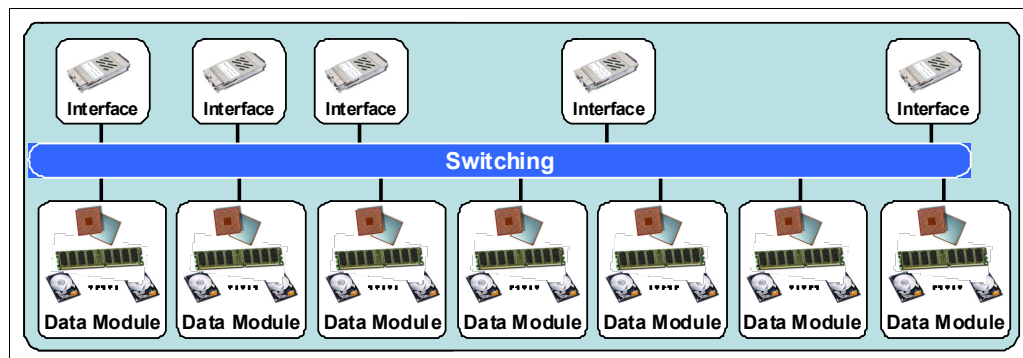


Figure 4-43 IBM XIV Storage System scalable conceptual grid architecture

Proportional scalability

Within the XIV Storage System, each module contains all of the pertinent hardware elements that are necessary for a grid topology (processing, caching, and storage). All modules are connected through a scalable network. This aspect of the grid infrastructure enables the relative proportions of cache, processor, disk, and interconnect bandwidth to remain optimal even in the event that modules are added or removed:

- ▶ **Linear cache growth:** The total system cache size and cache bandwidth increase linearly with disk capacity, because every module is a self-contained computing resource that houses its own cache (in terms of both host-to-cache and cache-to-disk).
- ▶ **Proportional interface growth:** Interface Modules house Ethernet and Fibre Channel host interfaces and are able to access not only the local resources within the module, but the entire system. With every Interface Module added, the system proportionally scales both the number of host interfaces and the bandwidth to the internal resources.
- ▶ **Constant switching capacity:** The internal switching capacity is designed to scale proportionally as the system grows, preventing bottlenecks regardless of the number of modules.
- ▶ **Embedded processing power:** Because each module incorporates its own processing power in conjunction with cache and disk components, the ability of the system to perform processor-intensive tasks, such as aggressive prefetch caching, sophisticated cache updates, snapshot management, and data distribution, is always maintained regardless of the system capacity.

Software parallelism

In addition to the hardware parallelism, the XIV Storage System also employs sophisticated algorithms to achieve optimal parallelism.

Modular software design

The XIV Storage System internal operating environment consists of a set of software functions that are loosely coupled with the hardware modules. These software functions reside on one or more modules and can be redistributed among modules as required, thus ensuring resiliency under changing hardware conditions.

An example of this modular design resides specifically in the interface modules. All six interface modules actively manage system services and software functionality associated with managing external I/O. Also, three of the interface modules deliver the system's management interface service for use with the XIV Storage System.

Data distribution algorithms

Data is distributed across all drives in a *pseudo-random* fashion. The patented algorithms provide a uniform yet random spreading of data across all available disks to maintain data resilience and redundancy. Figure 4-44 on page 150 provides a conceptual representation of the pseudo-random data distribution within the XIV Storage System.

4.8.5 Full storage virtualization

In this part we give the overview for concept of *virtualization*. See *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659 for complete information about the XIV System.

The data distribution algorithms employed by the XIV Storage System are innovative in that they are deeply integrated into the system architecture itself, instead of at the host or storage area network level. The XIV Storage System is unique in that it is based on an innovative implementation of full storage virtualization within the system itself.

IBM XIV Storage System virtualization design

The implementation of full storage virtualization employed by the XIV Storage System eliminates many of the potential operational drawbacks that can be present with conventional storage subsystems, while maximizing the overall usefulness of the subsystem.

The XIV Storage System virtualization offers the following benefits:

- ▶ Easier volume management:
 - Logical volume placement is driven by the distribution algorithms, which manage all of the data in the system collectively without deference to specific logical volume definitions.
 - Any interaction, whether host or system driven, with a specific logical volume in the system is inherently handled by all resources; it harnesses all storage capacity, all internal bandwidth, and all processing power currently available in the system.
 - Logical volumes are not exclusively associated with a subset of physical resources (can be dynamically resized and thinly provisioned).
- ▶ Consistent performance and scalability:
 - Hardware resources are always utilized equally, because all logical volumes always span all physical resources and are therefore able to reap the performance potential of the full system and maintain data integrity.
 - There are no “pockets” of capacity, “orphaned” disk space, or resources that are inaccessible due to array mapping constraints or data placement.

► Flexible snapshots:

- Full storage virtualization incorporates snapshots that are differential in nature; only updated data consumes physical capacity:
 - Many concurrent snapshots (Up to 16000 volumes and snapshots can be defined.)
Multiple concurrent snapshots are possible because a snapshot uses physical space only after a change has occurred on the source.
 - Multiple snapshots of a single master volume can exist independently of each other.
 - Snapshots can be cascaded, in effect, creating snapshots of snapshots.
- Creation and deletion of snapshots do not require data to be copied and hence occur immediately.
- When updates occur to master volumes, the system's virtualized logical structure enables it to preserve the original point-in-time data associated with any and all dependent snapshots by redirecting the update to a new physical location on disk. This process, which is referred to as *redirect on write*, occurs transparently from the host perspective and uses the virtualized remapping of the updated data to minimize any performance impact associated with preserving snapshots, regardless of the number of snapshots defined for a given master volume.

Note: The XIV snapshot process uses “redirect on write,” which is more efficient than the “copy on write” that is used by many other storage subsystems.

► Data migration efficiency:

- XIV supports thin provisioning. When migrating from a system that only supports regular (or thick) provisioning, XIV allows thick-to-thin provisioning of capacity.
- Due to the XIV pseudo-random distribution of data, the performance impact of data migration on production activity is minimized, because the load is spread evenly over all resources.

Logical system concepts

In this section, we elaborate on the logical system concepts, which form the basis for the system full storage virtualization.

Logical constructs

The XIV Storage System logical architecture incorporates constructs that underlie the storage virtualization and distribution of data, which are integral to its design. The logical structure of the system ensures that there is optimum granularity in the mapping of logical elements to both modules and individual physical disks, thereby guaranteeing an equal distribution of data across all physical resources.

Partitions

The fundamental building block of logical volumes is known as a *partition*. Partitions have the following characteristics on the XIV Storage System:

- All partitions are 1 MB (1024 KB) in size.
- A partition contains either a primary copy or secondary copy of data:
 - Each partition is mapped to a single physical disk:
 - This mapping is dynamically managed by the system through innovative data distribution algorithms in order to preserve data redundancy and equilibrium.

- The storage administrator has no control or knowledge of the specific mapping of partitions to drives.
- Secondary copy partitions are always placed in another *Module* than the one containing the primary copy partition.

Important: In the context of the XIV Storage System logical architecture, a partition consists of 1 MB (1024 KB) of data. Do not confuse this definition with other definitions of the term “partition.”

The diagram in Figure 4-44 illustrates that data is uniformly, yet randomly distributed over all disks. Each 1 MB of data is duplicated in a primary and secondary partition. For the same data, the system ensures that the primary partition and its corresponding secondary are not located within the same module.

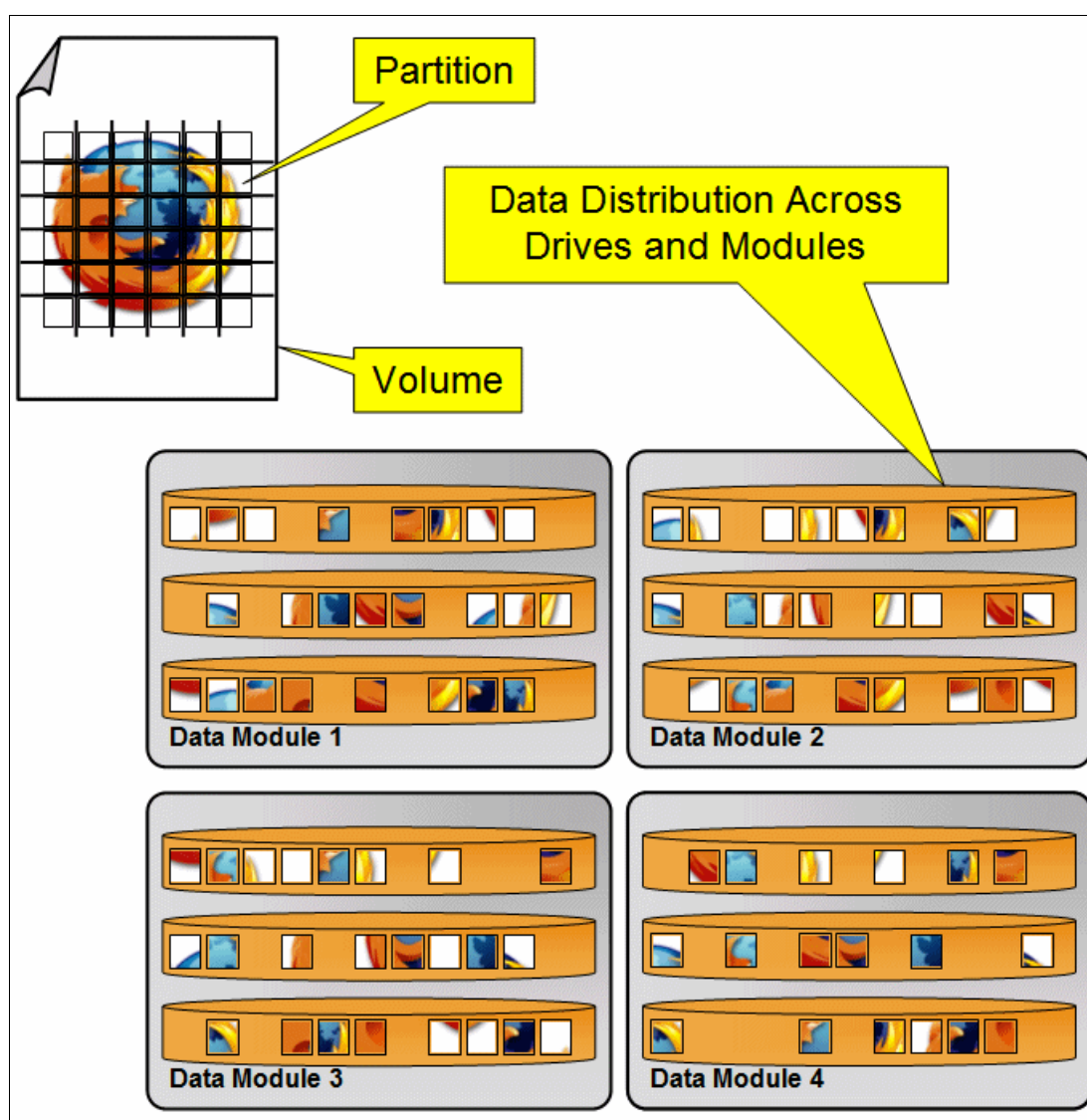


Figure 4-44 Pseudo-random data distribution¹

¹ Copyright 2005-2008 Mozilla. All Rights Reserved. All rights in the names, trademarks, and logos of the Mozilla Foundation, including without limitation, Mozilla, Firefox, as well as the Firefox logo, are owned exclusively by the Mozilla Foundation.

Logical volumes

The XIV Storage System presents logical volumes to hosts in the same manner as conventional subsystems; however, both the granularity of logical volumes and the mapping of logical volumes to physical disks differ:

- ▶ As discussed previously, every logical volume is comprised of 1 MB (1024 KB) constructs of data known as partitions.
- ▶ The physical capacity associated with a logical volume is *always a multiple of 17 GB (decimal)*.

Therefore, while it is possible to present a block-designated logical volume to a host that is *not* a multiple of 17 GB, the *actual physical space* that is allocated for the volume will always be the sum of the minimum number of 17 GB increments needed to meet the block-designated capacity.

Physical capacity: The initial physical capacity actually allocated by the system upon volume creation can be less than this amount.

- ▶ The maximum number of volumes that can be concurrently defined on the system is limited by:
 - The logical address space limit:
 - The logical address range of the system permits up to 16 377 volumes, although this constraint is purely logical, and therefore, is not normally a practical consideration.
 - Note that the same address space is used for both volumes and snapshots.
 - The limit imposed by the logical and physical topology of the system for the minimum volume size.

The physical capacity of the system, based on 180 drives with 1 TB of capacity per drive and assuming the minimum volume size of 17 GB, limits the maximum volume count to 4 605 volumes. Again, a system with active snapshots can have more than 4 605 addresses assigned collectively to both volumes and snapshots, because volumes and snapshots share the same address space.

Important: The logical address limit is ordinarily not a practical consideration during planning, because under most conditions, this limit will not be reached; it is intended to exceed the adequate number of volumes for all conceivable circumstances.

Storage Pools

Storage Pools are administrative boundaries that enable storage administrators to manage relationships between volumes and snapshots and to define separate capacity provisioning and snapshot requirements for such uses as separate applications or departments. Storage Pools are not tied in any way to physical resources, nor are they part of the data distribution scheme.

Snapshots

A *snapshot* represents a point-in-time copy of a volume. Snapshots are like volumes except snapshots incorporate dependent relationships with their source volumes, which can be either logical volumes or other snapshots. Because they are not independent entities, a given snapshot does not necessarily wholly consist of partitions that are unique to that snapshot. Conversely, a snapshot image will not share all of its partitions with its source volume if updates to the source occur after the snapshot was created.

Logical volume layout on physical disks

The XIV Storage System manages the distribution of logical volumes over physical disks and modules by means of a dynamic relationship between primary data partitions, secondary data partitions, and physical disks. This virtualization of resources in the XIV Storage System is governed by the data distribution algorithms.

Distribution table

The Distribution table is created at system startup, and contains a mapping of every primary and secondary partition, as well as the Module and physical disk they reside on. When hardware changes occur, a new Distribution table is created and delivered to every module. Each module retains redundant copies of the Distribution table.

Volume layout

At a conceptual level, the data distribution scheme can be thought of as an mixture of mirroring and striping. While it is tempting to think of this scheme in the context of RAID 1+0 (10) or 0+1, the low-level virtualization implementation precludes the usage of traditional RAID algorithms in the architecture.

As discussed previously, the XIV Storage System architecture divides logical volumes into 1 MB partitions. This granularity and the mapping strategy are integral elements of the logical design that enable the system to realize the following features and benefits:

- ▶ Partitions that make up a volume are distributed on all disks using what is defined as a *pseudo-random distribution function*, as introduced in “Software parallelism” on page 147:
 - The distribution algorithms seek to preserve the equality of access among all physical disks under all conceivable conditions and volume access patterns. Essentially, while not truly random in nature, the distribution algorithms in combination with the system architecture preclude the occurrence of “*hot-spots*”:
 - A fully configured XIV Storage System contains 180 disks, and each volume is allocated across at least 17 GB (decimal) of capacity that is distributed evenly across all disks.
 - Each logically adjacent partition on a volume is distributed across another disk; partitions are *not* combined into groups before they are spread across the disks.
 - The pseudo-random distribution ensures that logically adjacent partitions are never striped sequentially across physically adjacent disks. See “Software parallelism” on page 147 for a further overview of the partition mapping topology.
 - Each disk has its data mirrored across all other disks, *excluding the disks in the same module*.
 - Each disk holds approximately one percent of any other disk in other modules.
 - Disks have an equal probability of being accessed regardless of aggregate workload access patterns.

Goal distribution: When the number of disks or modules changes, the system defines a new data layout that preserves redundancy and equilibrium. This target data distribution is called the *goal distribution*.

- ▶ As discussed previously in “IBM XIV Storage System virtualization design” on page 148:
 - The storage system administrator does *not* plan the layout of volumes on the modules.
 - Provided that there is space available, volumes can always be added or resized instantly with negligible impact on performance.
 - There are no unusable pockets of capacity known as “orphaned spaces.”

- ▶ When the system is scaled out through the addition of modules, a new *goal distribution* is created whereby just a minimum number of partitions are moved to the newly allocated capacity to arrive at the new distribution table.

The new capacity is fully utilized within several hours and with no need for any administrative intervention. Thus, the system automatically returns to a state of equilibrium among all resources.

- ▶ Upon the failure or phase-out of a drive or a module, a new goal distribution is created whereby data in non-redundant partitions is copied and redistributed across the remaining modules and drives.

The system rapidly returns to a state in which all partitions are again redundant, because all disks and modules participate in achieving the new goal distribution.

System usable capacity

The XIV Storage System reserves physical disk capacity for:

- ▶ Global spare capacity
- ▶ Metadata, including statistics and traces
- ▶ Mirrored copies of data

Global spare capacity

The dynamically balanced distribution of data across all physical resources by definition obviates the inclusion of dedicated spare drives that are necessary with conventional RAID technologies. Instead, the XIV Storage System reserves capacity on each disk in order to provide adequate space for the redistribution or rebuilding of redundant data in the event of a hardware failure.

This global spare capacity approach offers advantages over dedicated hot spare drives, which are used only upon failure and are not used otherwise, therefore reducing the number of spindles that the system can utilize for better performance. Also, those non-operating disks are typically not subject to background scrubbing processes, whereas in XIV, all disks are operating and subject to examination, which helps detect potential reliability issues with drives.

The global reserved space includes sufficient capacity to withstand the failure of a full module and a further three disks, and will still allow the system to execute a new *goal distribution*, and to return to full redundancy.

Important: The system will tolerate multiple hardware failures, including up to an entire module in addition to three subsequent drive failures *outside of the failed module*, provided that a new goal distribution is fully executed before a subsequent failure occurs. If the system is less than 100% full, it can sustain more subsequent failures based on the amount of unused disk space that will be allocated at the event of failure as a spare capacity.

Note: The XIV Storage System does not manage a global reserved space for snapshots. We explore this topic in the next section.

Metadata and system reserve

The system reserves roughly 4% of the physical capacity for statistics and traces, as well as the distribution table.

Net usable capacity

The calculation of the *net usable capacity* of the system consists of the total disk count, less disk space reserved for sparing (which is the equivalent of one module plus three more disks), multiplied by the amount of capacity on each disk that is dedicated to data (that is 96% because of metadata and system reserve), and finally reduced by a factor of 50% to account for data mirroring achieved by the secondary copy of data.

Storage Pool concepts

While the hardware resources within the XIV Storage System are virtualized in a global sense, the available capacity in the system can be administratively portioned into separate and independent Storage Pools. The concept of *Storage Pools* is purely administrative. Essentially, Storage Pools function as a means to effectively manage a related group of similarly provisioned logical volumes and their snapshots.

Improved management of storage space

Storage Pools form the basis for controlling the usage of storage space by imposing a capacity quota on specific applications, a group of applications, or departments, enabling isolated management of relationships within the associated group of logical volumes and snapshots.

A *logical volume* is defined within the context of one and only one Storage Pool. As Storage Pools are logical constructs, a volume and any snapshots associated with it can be moved to any other Storage Pool, as long as there is sufficient space.

As a benefit of the system virtualization, there are no limitations on the size of Storage Pools or on the associations between logical volumes and Storage Pools. In fact, manipulation of Storage Pools consists exclusively of metadata transactions and does not trigger any copying of data. Therefore, changes are completed instantly and without any system overhead or performance degradation.

Consistency Groups

A *Consistency Group* is a group of volumes of which a snapshot can be made at the same point in time, thus ensuring a consistent image of all volumes within the group at that time. The concept of a Consistency Group is common among storage subsystems in which it is necessary to perform concurrent operations collectively across a set of volumes, so that the result of the operation preserves the consistency among volumes. For example, effective storage management activities for applications that span multiple volumes, or for creating point-in-time backups, is not possible without first employing Consistency Groups.

This consistency between the volumes in the group is paramount to maintaining data integrity from the application perspective. By first grouping the application volumes into a Consistency Group, it is possible to later capture a consistent state of all volumes within that group at a given point-in-time using a special snapshot command for Consistency Groups.

Issuing this type of a command results in the following process:

1. Complete and destage writes across the constituent volumes.
2. Instantaneously suspend I/O activity simultaneously across all volumes in the Consistency Group.
3. Create the snapshots.
4. Finally, resume normal I/O activity across all volumes.

The XIV Storage System manages these suspend and resume activities for all volumes within the Consistency Group.

Note: Additional mechanisms or techniques, such as those provided by the Microsoft Volume Shadow copy Services (VSS) framework, might still be required to maintain full application consistency.

Storage Pool relationships

Storage Pools facilitate the administration of relationships among logical volumes, snapshots, and Consistency Groups.

The following principles govern the relationships between logical entities within the Storage Pool:

- ▶ A logical volume can have multiple independent snapshots. This logical volume is also known as a *master volume*.
- ▶ A *master volume* and all of its associated snapshots are always a part of only one Storage Pool.
- ▶ A volume can only be part of a single Consistency Group.
- ▶ All volumes of a Consistency Group must belong to the same Storage Pool.

Storage Pools have the following characteristics:

- ▶ The size of a Storage Pool can range from 17 GB (the minimum size that can be assigned to a logical volume) to the capacity of the entire system.
- ▶ *Snapshot reserve capacity* is defined within each Storage Pool and is effectively *maintained separately* from logical, or master, volume capacity. The same principles apply for thinly provisioned Storage Pools, with the exception that space is not *guaranteed* to be available for snapshots due to the potential for hard space depletion, which is discussed in *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659:
 - Snapshots are structured in the same manner as logical, or master, volumes.

Attention: The snapshot reserve needs to be a minimum of 34 GB. The system preemptively deletes snapshots if the snapshots fully consume the allocated available space.

- As mentioned before, snapshots will only be automatically deleted when there is inadequate physical capacity available within the context of each Storage Pool. This process is managed by a snapshot deletion priority scheme. Therefore, when the capacity of a Storage Pool is exhausted, only the snapshots that reside in the affected Storage Pool are deleted in order of the deletion priority.
- ▶ The space allocated for a Storage Pool can be dynamically changed by the storage administrator:
 - The Storage Pool can be increased in size. It is limited only by the unallocated space on the system.
 - The Storage Pool can be decreased in size. It is limited only by the space that is consumed by the volumes and snapshots that are defined within that Storage Pool.
- ▶ The designation of a Storage Pool as a regular pool or a thinly provisioned pool can be dynamically changed even for existing Storage Pools. Thin provisioning is discussed in-depth in “Capacity allocation and thin provisioning” on page 156.

- ▶ The storage administrator can relocate logical volumes between Storage Pools without any limitations, provided there is sufficient free space in the target Storage Pool:
 - If necessary, the target Storage Pool capacity can be dynamically increased prior to volume relocation, assuming there is sufficient unallocated capacity available in the system.
 - When a logical volume is relocated to a target Storage Pool, sufficient space must be available for all of its snapshots to reside in the target Storage Pool as well.

Notes:

- ▶ When moving a volume into a Storage Pool, the size of the Storage Pool is not automatically increased by the size of the volume. Likewise, when removing a volume from a Storage Pool, the size of the Storage Pool does not decrease by the size of the volume.
- ▶ The system defines capacity using decimal metrics. Using decimal metrics, 1 GB is 1 000 000 000 bytes. Using binary metrics, 1 GB is 1 073 741 824 bytes.

Capacity allocation and thin provisioning

Thin provisioning is a central theme of the virtualized design of the XIV system, because it uncouples the virtual, or apparent, allocation of a resource from the underlying hardware allocation.

Capacity allocation

The following benefits emerge from the XIV Storage System's implementation of thin provisioning:

- ▶ Capacity associated with specific applications or departments can be dynamically increased or decreased per the demand imposed at a given point in time, without necessitating an accurate prediction of future needs. Physical capacity is only committed to the logical volume when the associated applications execute writes, as opposed to when the logical volume is initially allocated.
- ▶ Because the total system capacity is architected as a globally available pool, thinly provisioned resources share the same “buffer” of free space, which results in highly efficient aggregate capacity utilization without pockets of inaccessible unused space.

With the static, inflexible relationship between logical and physical resources commonly imposed by traditional storage subsystems, each application's capacity must be managed and allocated independently. This situation often results in a large percentage of the total system capacity remaining unused, because the capacity is confined within each volume at a highly granular level.

- ▶ Capacity acquisition and deployment can be more effectively deferred until actual application and business needs demand additional space, in effect facilitating an on demand infrastructure.

Thin provisioning

The idea behind thin provisioning is simple. Users define volumes with any logical size, and acquire and install only the physical capacity needed for data that is actually written.

This simple idea is similar to others used widely in other IT areas, for example:

- ▶ Virtual memory in a computer gives each process the perception of a huge memory space, while physical memory is much smaller.
- ▶ File systems allocate user quotas that can total more than the available space.
- ▶ Communication infrastructure paradigms have switched from dedicating capacity per user or application to statistical multiplexing of multiple users or applications.

With the XIV system, thin provisioning functions similarly to the aforementioned concepts. Thin provisioning principles are as follows:

- ▶ For a system, two types of capacities are defined:
 - Hard Capacity. The physical disk capacity available to applications (net capacity, after taking into account redundancy and spares)
 - Soft Capacity. The logical capacity of the entire system. Volumes are defined out of this capacity.
- ▶ When defining a volume, its size is only limited by the remaining soft capacity of the system.
- ▶ At the time of provisioning, the new volume is defined as formatted, meaning all zeros, and does not consume any hard capacity. As applications start to write data, the written areas consume hard capacity.

Soft capacity: Any system must potentially be expanded to its soft capacity. When defining the soft capacity, the customer must pre-plan how storage capacity will be added, reserve the required floor space, and ensure that the future power and cooling requirements can be met.

Thin provisioning conceptual examples

It is helpful to examine the following basic examples, because they incorporate all of the concepts inherent to the XIV Storage System's implementation of thin provisioning.

System-level thin provisioning conceptual example

Figure 4-45 on page 158 depicts the incremental allocation of capacity to both a regular Storage Pool and a thinly provisioned Storage Pool within the context of the global system soft and hard sizes. This example assumes that the soft system size has been defined to exceed its hard size. The unallocated capacity shown within the system's soft and hard space is represented by a discontinuity in order to convey the full scope of both the logical and physical view of the system's capacity. Each increment in the diagram represents 17 GB of soft or hard capacity.

When a regular Storage Pool is defined, only one capacity is specified, and this amount is allocated to the Storage Pool from *both the hard and soft global capacity within the system*.

When a thinly provisioned Storage Pool is defined, both the soft and hard capacity limits for the Storage Pool must be specified, and these amounts are deducted from the system's global available soft and hard capacity, respectively.

In the next example we focus on the regular Storage Pool introduced in Figure 4-45.

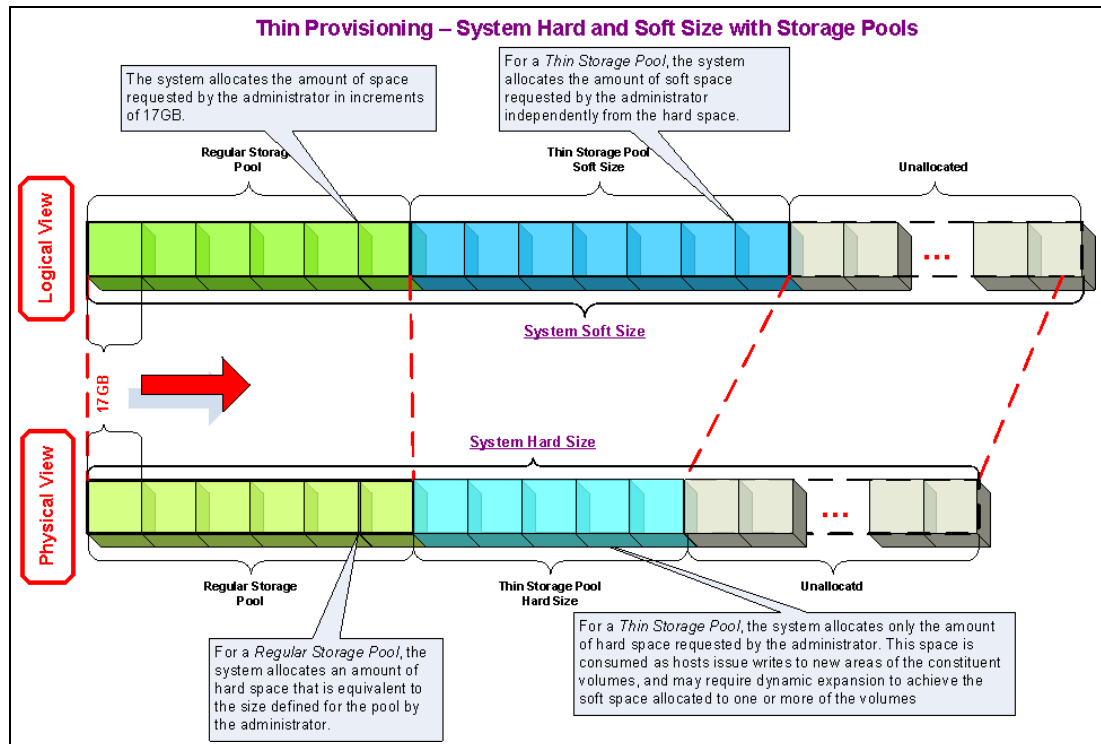


Figure 4-45 Thin provisioning at the system level

Regular Storage Pool conceptual example

Figure 4-46 represents a focused view of the regular Storage Pool that is shown here in Figure 4-45 and depicts the division of both soft and hard capacity among volumes within the pool. Note that the regular pool is the same size (102 GB) in both diagrams.

First, consider Volume 1. Although Volume 1 is defined as 19,737,900 blocks (10 GB), the soft capacity allocated will nevertheless be comprised of the minimum number of 17 GB increments needed to meet or exceed the requested size in blocks, which is in this case only a single 17 GB increment of capacity. The host will, however, see exactly 19,737,900 blocks. When Volume 1 is created, the *system does not initially allocate any hard capacity*. At the moment that a host writes to Volume 1, even if it is just to initialize the volume, the system will allocate 17 GB of hard capacity. The hard capacity allocation of 17 GB for Volume 1 is illustrated in Figure 4-46 on page 159, although clearly this allocation will never be fully utilized as long as the host-defined capacity remains only 10 GB.

Unlike Volume 1, Volume 2 has been defined in terms of gigabytes and has a soft capacity allocation of 34 GB, which is the amount that is reported to any hosts that are mapped to the volume. In addition, the hard capacity consumed by host writes has not yet exceeded the 17 GB threshold, and hence, the system has thus far *only allocated one increment of 17 GB hard capacity*. However, because the hard capacity and the soft capacity allocated to a regular Storage Pool are equal by definition, the remaining 17 GB of soft capacity assigned to Volume 2 is effectively preserved and will remain available within the pool's hard space until it is needed by Volume 2. In other words, because the pool's soft capacity does not exceed its hard capacity, there is no way to allocate soft capacity to effectively "overcommit" the available hard capacity.

The final reserved space within the regular Storage Pool shown in Figure 4-46 is dedicated for the snapshot usage. The diagram illustrates that the specified snapshot reserve capacity of 34 GB is effectively deducted from both the hard and soft space defined for the regular

Storage Pool, thus guaranteeing that this space will be available for consumption collectively by the snapshots associated with the pool. Although snapshots consume space granularly at the partition level, as discussed in “Storage Pool relationships” on page 155, the snapshot reserve capacity is still defined in increments of 17 GB.

The remaining 17 GB within the regular Storage Pool have not been allocated to either volumes or snapshots. Note that all soft capacity remaining in the pool is “backed” by hard capacity; the remaining unused soft capacity will always be less than or equal to the remaining unused hard capacity.

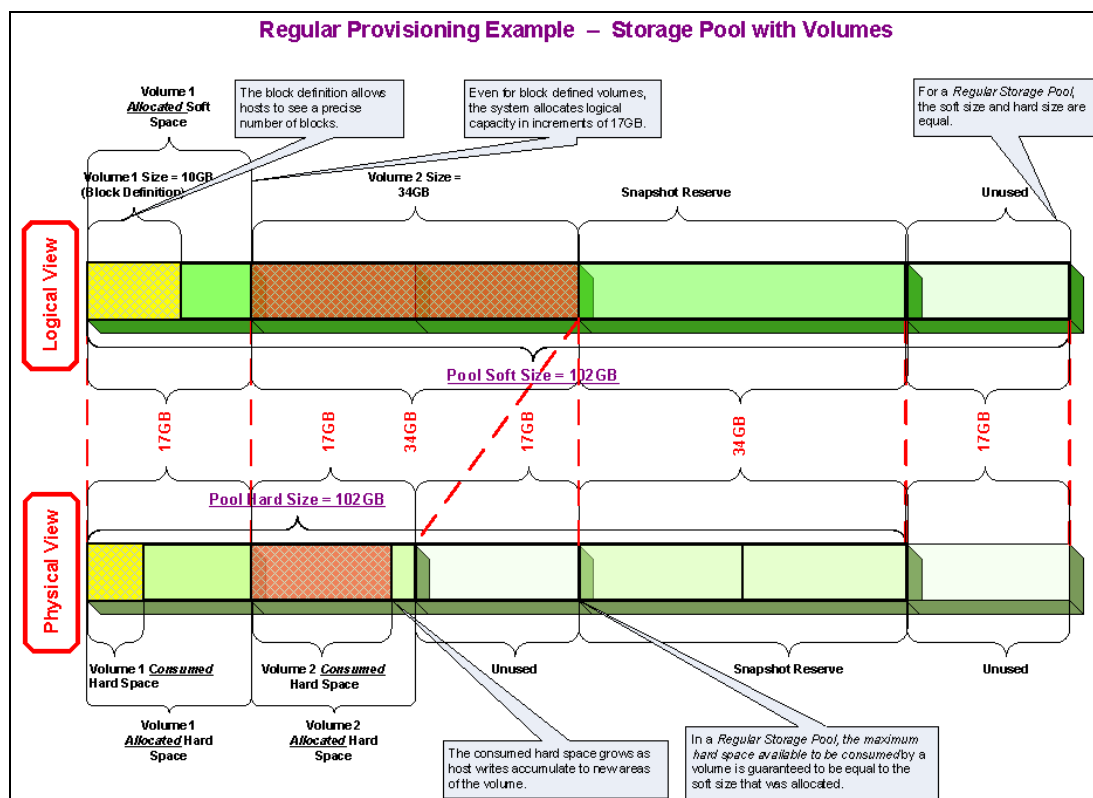


Figure 4-46 Volumes and snapshot reserve space within a regular Storage Pool

Thinly provisioned Storage Pool conceptual example

The thinly provisioned Storage Pool that was introduced in Figure 4-45 on page 158 is explored in detail in Figure 4-47 on page 160. Note that the hard capacity and the soft capacity allocated to this pool are the same in both diagrams: 136 GB of soft capacity and 85 GB of hard capacity are allocated. Because the available soft capacity exceeds the available hard capacity by 51 GB, it is possible to thinly provision the volumes collectively by up to 66.7%, assuming that the snapshots are preserved and the remaining capacity within the pool is allocated to volumes.

Consider Volume 3 in Example 4-47. The size of the volume is defined as 34 GB; however, less than 17 GB has been consumed by host writes, so only 17 GB of hard capacity have been allocated by the system. In comparison, Volume 4 is defined as 51 GB, but Volume 4 has consumed between 17 GB and 34 GB of hard capacity and therefore has been allocated 34 GB of hard space by the system. It is possible for either of these two volumes to require up to an additional 17 GB of hard capacity to become fully provisioned, and therefore, at least 34 GB of additional hard capacity must be allocated to this pool in anticipation of this requirement.

Finally, consider the 34 GB of snapshot reserve space depicted in Figure 4-47. If a new volume is defined in the unused 17 GB of soft space in the pool, or if either Volume 3 or Volume 4 requires additional capacity, the system will sacrifice the snapshot reserve space in order to give priority to the volume requirements. Normally, this scenario does not occur, because additional hard space must be allocated to the Storage Pool as the hard capacity utilization crosses certain thresholds.

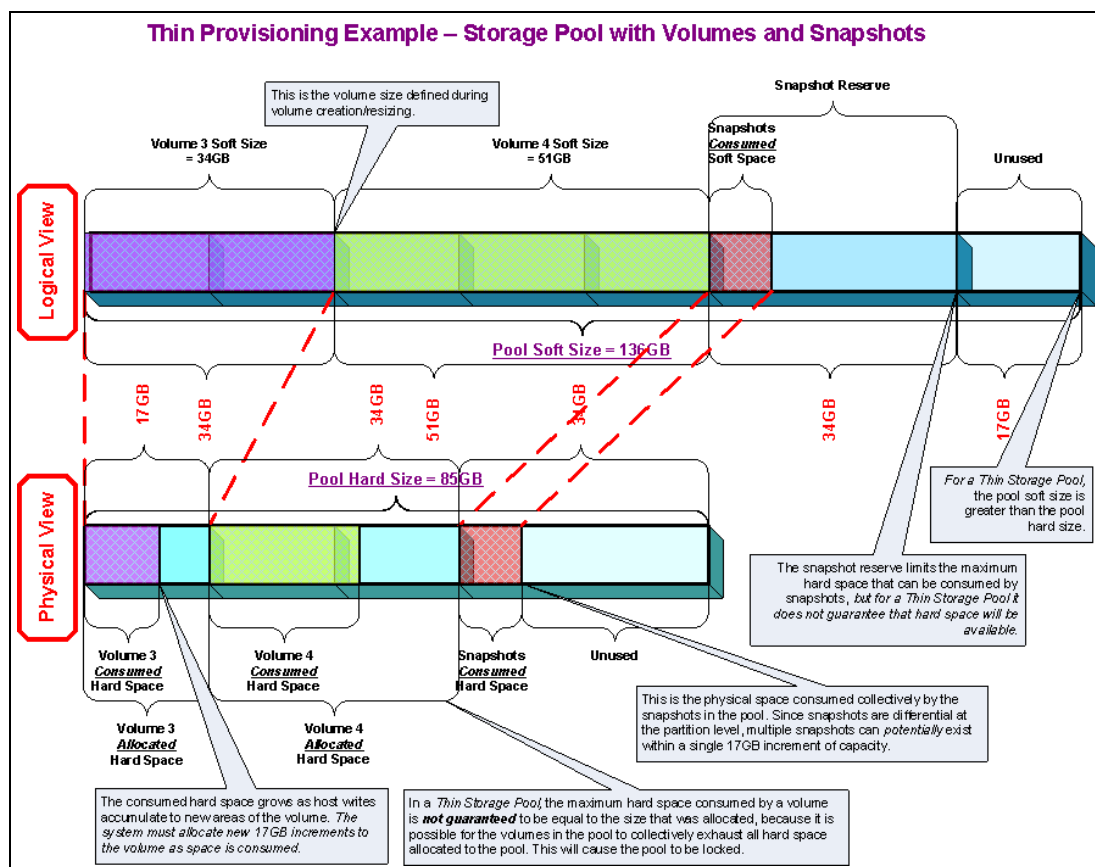


Figure 4-47 Volumes and snapshot reserve space within a thinly provisioned Storage Pool

Depletion of hard capacity

Using thin provisioning creates the inherent danger of exhausting the available physical capacity. If the soft system size exceeds the hard system size, the potential exists for applications to fully deplete the available physical capacity.

Important: Upgrading the system beyond the full 15 modules in a single frame is currently not supported.

► Snapshot deletion:

As mentioned previously, snapshots in regular Storage Pools can be automatically deleted by the system in order to provide space for newer snapshots, or in the case of thinly provisioned pools, to permit more physical space for volumes.

► Volume locking:

If more hard capacity is still required after all the snapshots in a thinly provisioned Storage Pool have been deleted, *all the volumes in the Storage Pool are locked, thereby preventing any additional consumption of hard capacity.* There are two possible behaviors for a locked volume: read only (the default behavior) or no I/O at all.

Important: Volume locking prevents writes to all volumes in the Storage Pool.

It is very important to note that thin provisioning implementation in the XIV Storage System manages space allocation within each Storage Pool, so that hard capacity depletion in one Storage Pool will never affect the hard capacity available to another Storage Pool. There are both advantages and disadvantages:

- ▶ Because Storage Pools are independent, thin provisioning volume locking on one Storage Pool never cascades into another Storage Pool.
- ▶ Hard capacity cannot be reused across Storage Pools, even if a certain Storage Pool has free hard capacity available, which can lead to a situation where volumes are locked due to the depletion of hard capacity in one Storage Pool, while there is available capacity in another Storage Pool. Of course, it is still possible for the storage administrator to intervene in order to redistribute hard capacity.

4.8.6 Reliability, availability, and serviceability

The XIV Storage System's unique modular design and logical topology fundamentally differentiate it from traditional monolithic systems, and this architectural divergence extends to the exceptional reliability, availability, and serviceability aspects of the system. In addition, the XIV Storage System incorporates autonomic, proactive monitoring and self-healing features that are capable of not only transparently and automatically restoring the system to full redundancy within minutes of a hardware failure, but also taking preventive measures to preserve data redundancy even before a component malfunction actually occurs.

For further reading about the XIV Storage System's parallel modular architecture, see 4.8.4, "Parallelism" on page 146.

Resilient architecture

As with any enterprise class system, redundancy pervades every aspect of the XIV Storage System, including the hardware, internal operating environment, and the data itself. However, the design elements, including the distribution of volumes across the whole of the system, in combination with the loosely coupled relationship between the underlying hardware and software elements, empower the XIV Storage System to realize unprecedented resiliency.

The resiliency of the architecture encompasses not only high availability, but also excellent maintainability, serviceability, and performance under non-ideal conditions resulting from planned or unplanned changes to the internal hardware infrastructure, such as the loss of a module.

Availability

The XIV Storage System maximizes operational availability and minimizes the degradation of performance associated with non-disruptive planned and unplanned events, while providing for the capability to preserve the data to the fullest extent possible in the event of a disaster.

Write path redundancy

Data arriving from the hosts is temporarily placed in two separate caches before it is permanently written to disk drives located in separate modules. This design guarantees that the data is always protected against possible failure of individual modules, even before the data has been written to the disk drives.

Figure 4-48 illustrates the path taken by a write request as it travels through the system. The diagram is intended to be viewed as a conceptual topology, so do not interpret the specific numbers of connections and so forth as literal depictions. Also, for purposes of this discussion, the Interface Modules are depicted on a separate level from the Data Modules. However, in reality the Interface Modules also function as Data Modules. The following numbered steps correspond to the numbers in Figure 4-48:

1. A host sends a write request to the system. Any of the Interface Modules that are connected to the host can service the request, because the modules work in an active-active capacity. Note that the XIV Storage System does not load balance the requests itself. Load balancing must be implemented by storage administrators to equally distribute the host requests among all Interface Modules.
2. The Interface Module uses the system configuration information to determine the location of the primary module that houses the referenced data, which can be either an Interface Module, including the Interface Module that received the write request, or a Data Module. The data is written only to the local cache of the primary module.
3. The primary module uses the system configuration information to determine the location of the secondary module that houses the copy of the referenced data. Again, this module can be either an Interface Module or a Data Module, but it will not be the same as the primary module. The data is redundantly written to the local cache of the secondary module.

After the data is written to cache in both the primary and secondary modules, the host receives an acknowledgement that the I/O is complete, which occurs independently of copies of either cached, or dirty, data being destaged to physical disk.

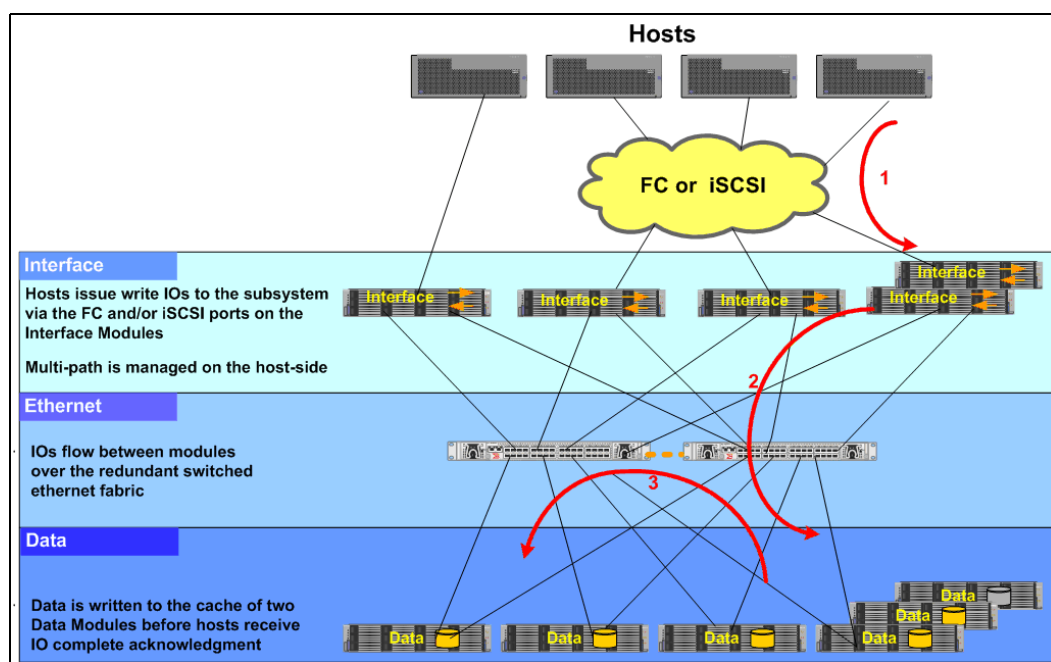


Figure 4-48 Write path

System quiesce and graceful shutdown

When an event occurs that compromises both sources of power to the XIV Storage System's redundant uninterruptible power supplies, the system executes the graceful shutdown sequence. Full battery power is guaranteed during this event, because the system monitors available battery charge at all times and takes proactive measures to prevent the possibility of conducting write operations when battery conditions are non-optimal.

Due to the XIV Storage System's grid topology, a system quiesce event essentially entails the graceful shutdown of all modules within the system. Each module can be thought of as an independent entity that is responsible for managing the destaging of "dirty" data, that is, written data that has not yet been destaged to physical disk. The dirty data within each module consists of equal parts primary and secondary copies of data, *but will never contain both primary and secondary copies of the same data*.

► Write cache protection:

Each module in the XIV Storage System contains a local, independent space reserved for caching operations within its system memory.

Each module contains 8 GB of high speed volatile *memory* (a total of 120 GB), from which 5.5 GB (and 82.5 GB overall) is dedicated for caching data.

Destage: The system does *not* contain non-volatile memory space reserved for write operations. However, the close proximity of the cache and the drives, in conjunction with the enforcement of an upper limit for dirty, or non-destaged, data on a per-drive basis, ensures that the full destage will occur while operating under battery power.

► Graceful shutdown sequence:

The system executes the graceful shutdown sequence under either of these conditions:

- The battery charge remaining in two or more universal power supplies is below a certain threshold, which is conservatively predetermined in order to provide adequate time for the system to fully destage all dirty data from cache.
- The system detects the loss of external power for more than 30 seconds.

► Power on sequence:

Upon startup, the system will verify that the battery charge levels in all uninterruptible power supplies exceed the threshold necessary to guarantee that a graceful shutdown can occur. If the charge level is inadequate, the system will halt the startup process until the charge level has exceeded the minimum required threshold.

I/O activity: If the battery charge is inadequate, the system will remain fully locked until the battery charge has exceeded the necessary threshold to safely resume I/O activity.

Rebuild and redistribution

As discussed in "Data distribution algorithms" on page 148, the XIV Storage System dynamically maintains the pseudo-random distribution of data across all modules and disks while ensuring that two copies of data exist at all times *when the system reports Full Redundancy*. Obviously, when there is a change to the hardware infrastructure as a result of a failed component, data must be restored to redundancy and distributed, or when a component is added, or *phased-in*, a new data distribution must accommodate the change.

Goal distribution

The process of achieving a new goal distribution while simultaneously restoring data redundancy due to the loss of a disk or module is known as a *rebuild*. Because a rebuild occurs as a result of a component failure that compromises full data redundancy, there is a period during which the *non-redundant data* is both restored to full redundancy and homogeneously redistributed over the remaining disks.

The process of achieving a new goal distribution (only occurring when redundancy exists) is known as a *redistribution*, during which all data in the system (including both primary and secondary copies) is redistributed, when it is a result of the following events:

- ▶ The replacement of a failed disk or module following a rebuild, also known as a “phase-in”
- ▶ When one or more modules are added to the system, known as a “scale out” upgrade

Following the completion of goal distribution resulting from a rebuild or phase-out, a subsequent redistribution must occur when the system hardware is fully restored through a phase-in.

Goal distribution: The goal distribution is transparent to storage administrators and cannot be changed. In addition, the goal distribution has many determinants depending on the precise state of the system.

Important: Never perform a phase-in to replace a failed disk or module until after the rebuild process has completed. These operations must be performed by the IBM XIV technician anyway.

Preserving data redundancy

Whereas conventional storage systems maintain a static relationship between RAID arrays and logical volumes by preserving data redundancy only across a subset of disks that are defined in the context of a particular RAID array, the XIV Storage System dynamically and fluidly restores redundancy and equilibrium across all disks and modules in the system during the rebuild and phase-out operations.

See “Logical volume layout on physical disks” on page 152 for a detailed discussion of the low-level virtualization of logical volumes within the XIV Storage System. The proactive phase-out of non-optimal hardware through autonomic monitoring and the modules’ cognizance of the virtualization between the logical volumes and physical disks yield unprecedented efficiency, transparency, and reliability of data preservation actions, encompassing both rebuilds and phase-outs.

Rebuild examples

When the full redundancy of data is compromised due to a module failure, as depicted in Figure 4-49, the system immediately identifies the non-redundant partitions and begins the rebuild process. Because none of the disks within a given module contain the secondary copies of data residing on any of the disks in the module, the secondary copies are read from the remaining modules in the system. Therefore, during a rebuild resulting from a module failure, there will be concurrently 168 disks (180 disks in the system minus 12 disks in a module) reading, and 168 disks writing, as is conceptually illustrated in Figure 4-49.

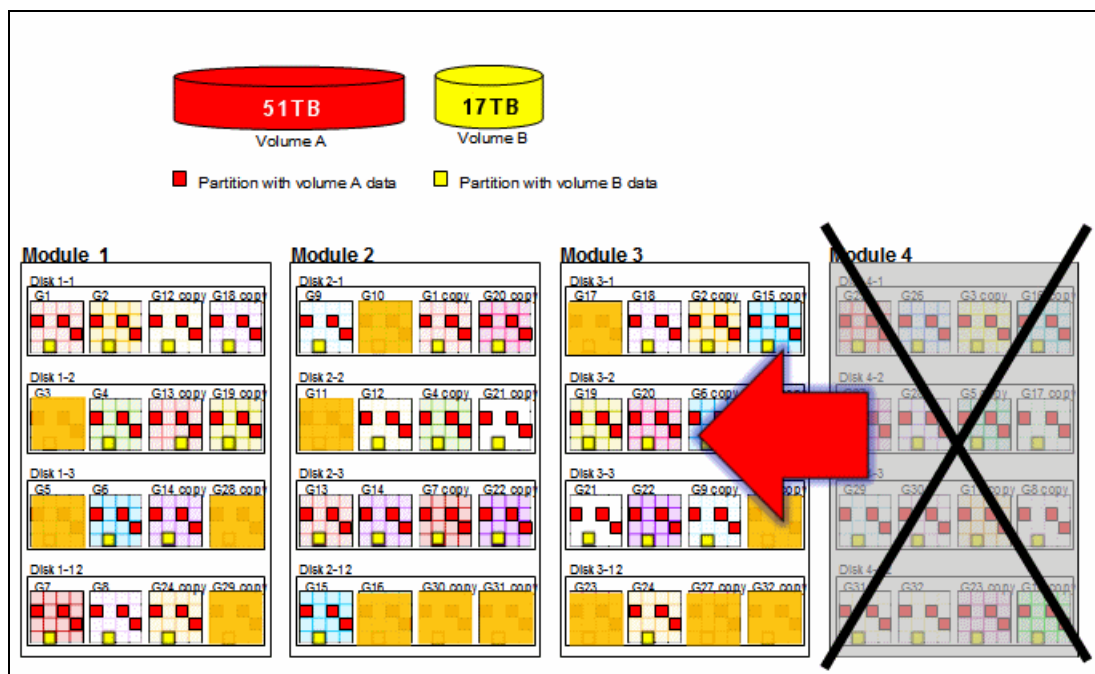


Figure 4-49 Non-redundant group of partitions following module failure

Figure 4-50 depicts a denser population of redundant partitions for both volumes A and B, thus representing the completion of a new goal distribution, as compared to Figure 4-49, which contains the same number of redundant partitions for both volumes distributed less densely over the original number of modules and drives.

Finally, consider the case of a single disk failure occurring in an otherwise healthy system (no existing phased-out or failed hardware). During the subsequent rebuild, there will be only 168 disks reading, because there is no non-redundant data residing on the other disks within the same module as the failed disk. Concurrently, there will be 179 disks writing in order to preserve full data distribution.

Attention: Figure 4-49 and Figure 4-50 conceptually illustrate the rebuild process resulting from a failed module. *The diagrams are not intended to depict in any way the specific placement of partitions within a real system, nor do they literally depict the number of modules in a real system.*

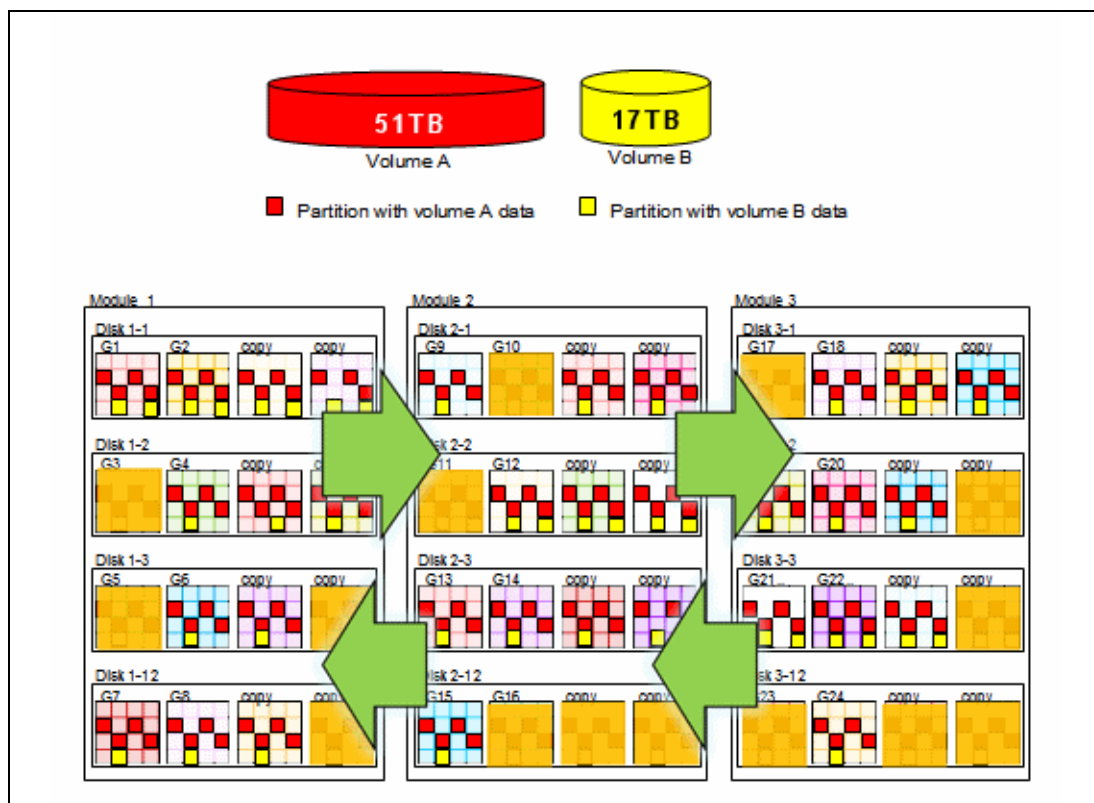


Figure 4-50 Performing a new goal distribution following module failure

Transient soft and hard system size

The capacity allocation that is consumed for purposes of either restoring non-redundant data during a rebuild, or creating a tertiary copy during a phase-out, will be sourced based on availability.

In the event that sufficient unallocated hard capacity is available, the system will withhold allocating reserve spare space to complete the rebuild or phase-out process in order to provide additional protection. As a result, it is possible for the system to report a maximum soft size that is temporarily less than the allocated soft capacity. The soft and hard system sizes will not revert to the original values until a replacement disk or module is phased-in, and the resultant redistribution completes.

Important: While it is possible to resize or create volumes, snapshots, or Storage Pools while a rebuild is underway, we strongly discourage these activities until the system has completed the rebuild process and restored full data redundancy.

Redistribution

The XIV Storage System homogeneously redistributes all data across all disks whenever new disks or modules are introduced or phased in to the system. This redistribution process is not equivalent to the “striping volumes on all disks” employed in traditional systems. When new capacity is added and new volumes are allocated, ordinary RAID striping algorithms do not intelligently redistribute data to preserve equilibrium for all volumes through the pseudo-random distribution of data. Thus, the XIV Storage System employs dynamic volume-level virtualization, obviating the need for ongoing manual volume layout planning.

The redistribution process is triggered by the “phase-in” of a new drive or module and differs from a rebuild or phase-out in the following respects:

- ▶ The system does not need to create secondary copies of data to reinstate or preserve full data redundancy.
- ▶ The distribution density, or the concentration of data on each physical disk, decreases instead of increasing.
- ▶ The redistribution of data performs differently, because the concentration of write activity on the new hardware resource is the bottleneck.

Minimized exposure

In this section we describe other features that contribute to the XIV Storage System reliability and availability.

Disaster recovery

All high availability SAN implementations must account for the contingency of data recovery and business continuance following a disaster, as defined by the organization’s recovery point and recovery time objectives. The provision within the XIV Storage System to efficiently and flexibly create nearly unlimited snapshots, coupled with the ability to define Consistency Groups of logical volumes, constitutes integral elements of the data preservation strategy. In addition, the XIV Storage System’s synchronous data mirroring functionality facilitates excellent potential recovery point and recovery time objectives as a central element of the full disaster recovery plan.

Proactive phase-out and self-healing mechanisms

The XIV Storage System can seamlessly restore data redundancy with minimal data migration and overhead. A further enhancement to the level of reliability standards attained by the XIV Storage System entails self-diagnosis and early detection mechanisms that autonomically phase out components before the probability of a failure increases beyond a certain point. In real systems, the failure rate is not constant with time, but rather increases with service life and duty cycle. By actively gathering component statistics to monitor this trend, the system ensures that components will not operate under conditions beyond an acceptable threshold of reliability and performance. Thus, the XIV Storage System’s self-healing mechanisms dramatically increase the already exceptional level of availability of the system, because they virtually preclude the possibility of data redundancy from ever being compromised along with the associated danger, however unlikely, of subsequent failures during the rebuild process.

The autonomic attributes of the XIV Storage System cumulatively impart an enormous benefit to not only the reliability of the system, but also the overall availability, by augmenting the maintainability and serviceability aspects of the system. Both the monetary and time demands associated with maintenance activities, or in other words, the total cost of ownership (TCO), are effectively minimized by reducing reactive service actions and enhancing the potential scope of proactive maintenance policies.

▶ Disk scrubbing:

The XIV Storage System maintains a series of scrubbing algorithms that run as background processes concurrently and independently scanning multiple media locations within the system in order to maintain the integrity of the redundantly stored data. This continuous checking enables the early detection of possible data corruption, alerting the system to take corrective action to restore the data integrity before errors can manifest themselves from the host perspective. Thus, redundancy is not only implemented as part of the basic architecture of the system, but it is also continually monitored and restored as required.

In summary, the data scrubbing process has the following attributes:

- Verifies the integrity and redundancy of stored data
- Enables early detection of errors and early recovery of redundancy
- Runs as a set of background processes on all disks in parallel
- Checks whether data can be read from partitions and verifies data integrity by employing checksums
- Examines a partition approximately every two seconds

► Enhanced monitoring and disk diagnostics:

The XIV Storage System continuously monitors the performance level and reliability standards of each disk drive within the system, using an enhanced implementation of Self-Monitoring, Analysis and Reporting Technology (SMART) tools. As typically implemented in the storage industry, SMART tools simply indicate whether certain thresholds have been exceeded, thereby alerting that a disk is at risk for failure and thus needs to be replaced.

However, as implemented in XIV Storage System, the SMART diagnostic tools, coupled with intelligent analysis and low tolerance thresholds, provide an even greater level of refinement of the disk behavior diagnostics and the performance and reliability driven reaction. For instance, the XIV Storage System measures the specific values of parameters including, but not limited to:

- Reallocated sector count: If the disk encounters a read or write verification error, it designates the affected sector as “reallocated” and relocates the data to a reserved area of spare space on the disk. Note that this spare space is a parameter of the drive itself and is *not* related in any way to the system reserve spare capacity that is described in “Global spare capacity” on page 153. The XIV Storage System initiates phase-out at a much lower count than the manufacturer says.
- Disk temperature: The disk temperature is a critical factor that contributes to premature drive failure and is constantly monitored by the system.
- Raw read error: The raw read error count provides an indication of the condition of the magnetic surface of the disk platters and is carefully monitored by the system to ensure the integrity of the magnetic media itself.
- Spin-up time: The spin-up time is a measure of the average time that is required for a spindle to accelerate from zero to 7,200. The XIV Storage System recognizes abnormal spin-up time as a potential indicator of an impending mechanical failure.

Likewise, for additional early warning signs, the XIV Storage System continually monitors other aspects of disk-initiated behavior, such as spontaneous reset or unusually long latencies. The system intelligently analyzes this information in order to reach crucial decisions concerning disk deactivation and phase-out. The parameters involved in these decisions allow for a very sensitive analysis of the disk health and performance.

► Redundancy-supported reaction:

The XIV Storage System incorporates *redundancy-supported reaction*, which is the provision to exploit the distributed redundant data scheme by intelligently redirecting reads to the secondary copies of data, thereby extending the system’s tolerance of above average disk service time when accessing primary data locations. The system will reinstate reads from the primary data copy when the transient degradation of the disk service time has subsided. Of course, a redundancy-supported reaction itself might be triggered by an underlying potential disk error that will ultimately be managed autonomically by the system according to the severity of the exposure, as determined by ongoing disk monitoring.

- Flexible handling of dirty data:

In a similar manner to the redundancy-supported reaction for read activity, the XIV Storage System can also make convenient use of its redundant architecture in order to consistently maintain write performance. Because intensive write activity directed to any given volume is distributed across all modules and drives in the system, and the cache is independently managed within each module, the system is able to tolerate sustained write activity to an under-performing drive by effectively maintaining a considerable amount of “dirty,” or unwritten, data in cache, thus potentially circumventing any performance degradation resulting from the transient, anomalous service time of a given disk drive.

Non-disruptive code load

Non-disruptive code load (NDCL) enables upgrades to the IBM XIV Storage System software from a current version (starting with Version 10.1) to a later version without disrupting the application service.

The code upgrade is run on all modules in parallel and the process is fast enough to minimize impact on the host applications.

No data migration or rebuild process is allowed during the upgrade. Mirroring, if any, will be suspended during the upgrade and automatically reactivated upon completion.

Storage management operations are also not allowed during the upgrade, although the status of the system and upgrade progress can be queried. It is also possible to cancel the upgrade process up to a point of no return.

Note that the NDCL does not apply to specific components firmware upgrades (for instance, module BIOS, HBA firmware). Those require a phase in / phase out process of the impacted modules.

4.9 Comparison of products

Here we compare the DS6000, the DS8000, and XIV to each other and to other systems.

DS6000 series compared to DS4000

Those familiar with the DS4000 Family will find substantial difference between the DS4000 series and DS6000 series of products. Both product families have about the same footprint and capacity, but their functions differ. DS6000 offers higher performance compared to the DS4000.

The DS6000 offers enterprise capabilities not found in midrange offerings, because it originates from the ESS roots and provides support for System z and System i servers as well as open systems.

The implementation of FlashCopy on the DS4000 varies compared to the DS6000 series. For example, the target system on a DS4000 cannot be used for production, while it can on the DS6000.

While the DS4000 also offers remote copy solutions, these functions are not compatible with the DS6000.

DS6000 series compared to DS8000 series

You can think of the DS6000 as the smaller sibling of the DS8000. The DS8000 and the DS6000 can work with each other perfectly because of their common switched FC-AL architecture.

All Copy Services (with the exception of z/OS Global Mirror) are available on both systems. You can do Metro Mirror, Global Mirror, and Global Copy between the two series. The CLI commands and the DS Storage Manager GUI look the same for both systems.

So it is very easy to have a mixed environment with DS8000 and DS6000 systems to optimize the cost-effectiveness of your storage solution, while providing the cost efficiencies of common skills and management functions.

The DS8000 can deliver a higher throughput and a dramatic leap in performance. It provides high I/O bandwidth and self-healing interconnect with RIO-G architecture which is only available on the DS8000. At the same time, the DS8000 offers extensive virtualization capabilities with the industry's first implementation of storage logical partitioning. Storage system LPARs are not available on the DS6000.

XIV compared to midrange solutions

XIV offers superior reliability. Midrange systems are hampered by extremely long rebuild times. Restoring redundancy after failure of a 1 TB drive can take over 6 hours, during which the system is susceptible to a double disk failure. Moreover, non-uniform data distribution schemes featured by midrange systems subject a subset of system drives to exceedingly increased stress levels during normal operation and rebuild processes that increase the likelihood of drive failure. XIV, by comparison, has the best rebuild time in the industry (<40 minutes for 1 TB), which dramatically reduces the likelihood of double disk failure, and features a patented uniform data distribution that distributes load among drives equally, thus eliminating spike loads altogether.

XIV offers superior performance. Midrange systems feature inconsistent performance, limited cache size, and limited active passive load balancing. Importantly, they require constant tuning to address performance degradation due to snapshot usage, configuration changes and component failures. Midrange systems might also resort to cache write through mode upon cache node failure which results in very low performance. XIV never resorts to writethrough mode, features huge cache (>80 GB per rack), active load balancing, data distribution that offers high tier1 performance without requiring any tuning, and consistent performance at all times (even after configuration changes, or component failure). Furthermore, XIV's performance is not restricted by snapshot usage.

XIV offers superior ease of management. Midrange systems are plagued by cumbersome provisioning processes and extensive configuration and planning (for example, assuring proper layout of disk groups). XIV's system virtualization obviates any knowledge of internal system design, features streamlined and quick provisioning of space, and is never operating in need of tuning.

XIV offers a superior feature set. Midrange systems are typically missing important capabilities such as thin provisioning, and offer key functionality at extra cost, including support for snapshots, mirroring, data migration, and reporting. XIV offers all product features at no extra cost.

XIV offers superior TCO less power and higher density per GB. Midrange systems rely on expensive FC drives to attain reasonable performance, which in turn contribute to high power requirements. XIV's utilization of SATA drives coupled with its unique architecture, can keep power costs down, while maintaining world class performance that out-performs midrange systems.

4.10 More information

See the following publications for more information about the topics covered in this chapter:

- ▶ *IBM System Storage DS8000: Architecture and Implementation*, SG24-6786
- ▶ *IBM System Storage DS8700: Architecture and Implementation*, SG24-8786
- ▶ *DS8000 Performance Monitoring and Tuning*, SG24-7146
- ▶ *Migrating to IBM System Storage DS8000*, SG24-7432
- ▶ *IBM TotalStorage Productivity Center for Replication Using DS8000*, SG24-7596
- ▶ *IBM System Storage DS8000: Copy Services in Open Environments*, SG24-6788
- ▶ *DS8000 Copy Services for IBM System z*, SG24-6787
- ▶ *IBM System Storage DS8000 Series: IBM FlashCopy SE*, REDP-4368
- ▶ *IBM DS8000 and z/OS Basic HyperSwap*, REDP-4441
- ▶ *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500
- ▶ *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, REDP-4504
- ▶ *IBM System Storage DS8000: LDAP Authentication*, REDP-4505
- ▶ *DS8000: Introducing Solid State Drives*, REDP-4522
- ▶ *DS8000 Thin Provisioning*, REDP-4554
- ▶ *IBM System Storage DS Command-line Interface User's Guide* GC53-1127-03
- ▶ *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659
- ▶ *White Paper IBM XIV Storage System: Thin Provisioning Reinvented*:
ftp://service.boulder.ibm.com/storage/disk/xiv/ibm_xiv_thin_provisioning_paper.pdf
- ▶ IBM XIV website:
<http://www.ibm.com/systems/storage/disk/xiv/xivsystem.org>



IBM System Storage N series

In this chapter we introduce the IBM System Storage N series and describe its hardware and software. The IBM System Storage N series provides a range of reliable, scalable storage solutions for a variety of storage requirements. These capabilities are achieved by using network access protocols such as Network File System (NFS), Common Internet File System (CIFS), HTTP, iSCSI and FCoE as well as storage area network technologies such as Fibre Channel (FC). Utilizing built-in Redundant Array of Inexpensive Disks (RAID) technologies.

All data is well protected with options to enhance protection through mirroring, replication, Snapshots, and backup. The N series systems are also characterized by simple management interfaces that make installation, administration, and troubleshooting straightforward. The IBM System Storage N series is designed from the ground up as a stand-alone storage system.

Advantages of using this type of flexible storage solution include the following capabilities:

- ▶ Tune the storage environment to a specific application while maintaining flexibility to increase, decrease, or change access methods with minimal disruption, and keeping its availability.
- ▶ React easily and quickly to changing storage requirements. If additional storage is required you can expand it quickly and non-disruptively. When existing storage exists but is deployed incorrectly, you have the capability to reallocate available storage from one application to another quickly and simply.
- ▶ Maintain availability and productivity during upgrades. If outages are necessary, you can keep them to the shortest time possible.
- ▶ Create effortless backup/recovery solutions that operate commonly across all data access methods.
- ▶ Simplify your infrastructure with file and block-level services in a single system.
- ▶ Change the deployment of storage resources non-disruptively, easily, and quickly. Online storage resource redeployment is possible.
- ▶ Easily and quickly implement the upgrade process. Non-disruptive upgrade is possible.
- ▶ Strong data protection solutions and support for online backup/recovery are provided.

5.1 IBM N series storage systems functionality

The IBM System Storage N series storage systems offer multiprotocol connectivity using internal storage or storage provided by expansion units (Figure 5-1). IBM System Storage N series systems are designed to provide integrated block-level and file-level data access, allowing concurrent operation in IP SAN (iSCSI), FC SAN, NFS, and CIFS environments. Other storage vendors might require the operation of multiple systems to provide this functionality, thus this makes the N series a perfect storage consolidation system. IBM System Storage N series storage systems are designed to avoid costly downtime, both planned and unplanned, and improve your access to mission-critical data, thereby helping you gain a competitive advantage.

Figure 5-1 shows the ability to connect multiple data access methods.

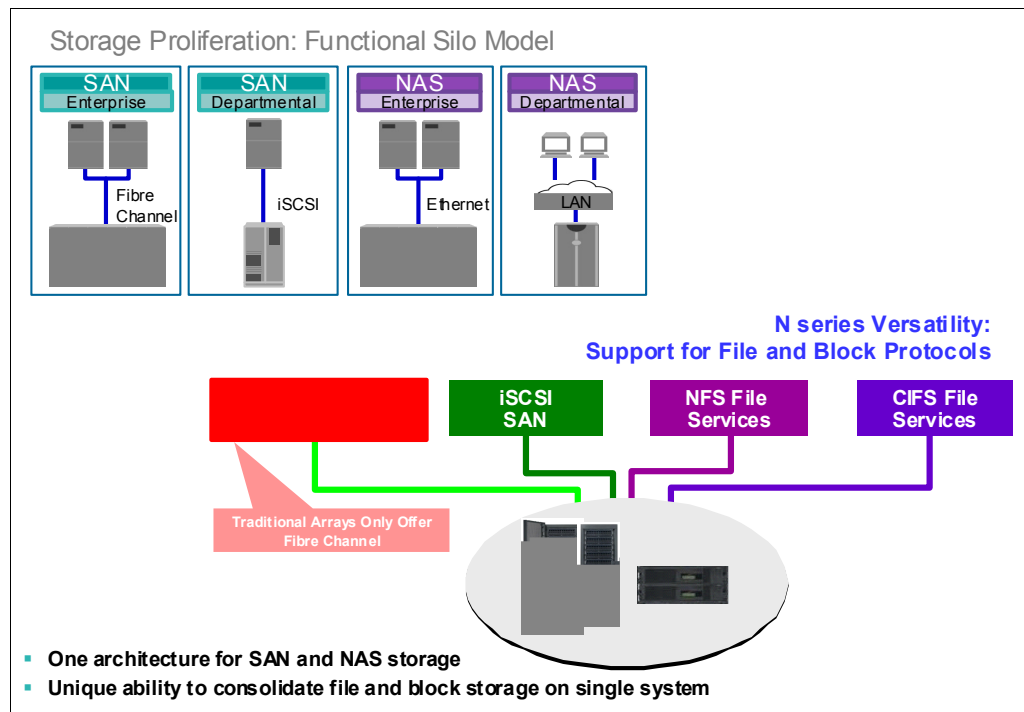


Figure 5-1 IBM System Storage N series A models

The N series is a specialized, *thin server* storage system with a customized operating system, similar to a stripped-down UNIX kernel, hereafter referred to as Data ONTAP®. With this customized operating system, many of the server operating system functions that you are familiar with are not supported. Data ONTAP improves performance and reduces costs by eliminating unnecessary functions that do not pertain to a storage system.

The N series systems come with preconfigured software and hardware, and with no monitor or keyboard for user access. This is commonly called a *headless* system. A storage administrator accesses the systems and manages the disk resources from a remote console using a web browser or command line.

One of the typical characteristics of an N series storage systems product is its ability to be installed rapidly, using minimal time and effort to configure the system. N series is integrated seamlessly into the network, making it especially attractive when limited resources of time and skills are of consideration in the decision process.

5.1.1 Drive flexibility

IBM System Storage N series products are designed to provide network-attached storage for environments where customers have a need to utilize their storage investment in a multifaceted environment. IBM System Storage N series storage systems provide customers with a tremendous amount of versatility by allowing this solution to be populated with both Fibre Channel disk drives and SATA disk drives. An N series populated with Fibre Channel disk drives might be suitable for mission-critical high-performance data transaction environments, whereas an N series populated with SATA disk drives might be attractive to customers who wish to use the platform for disk-to-disk backup scenarios, disaster recovery scenarios, archive data, or data such as home directories, which do not have the demands of high-performance transactional environments. See Table 5-1.

Table 5-1 Drive positioning

Requirement	Fibre Channel drives	SAS drives	SATA drives
Online, high-performance, mission-critical production data repository	X	X	
Near-line storage used for tiered storage or infrequently accessed data	X	X	X
Data retention to help meet the needs of customers required to store data in non-erasable and non-rewritable (WORM) formats		X	X

5.1.2 Near-line storage solution

IBM System Storage N series with SATA drives offers near-line storage. Figure 5-2 shows an example of traditional disk-based backup and recovery. The left side shows the primary storage characterized by a higher cost and very fast performing system. On the far right, the archive targets are represented traditionally by tape or optical jukeboxes.

The N series secondary with SATA drives reduces access times to read and write data. A few years ago the concept of near-line storage in the middle for disk staging was introduced and enabled organizations to do daily backups to disk. Additionally, backups to tape can be done weekly or bi-weekly, thereby reducing the amount of data that needed to be written to tape. Moreover, data that is online is available for faster recovery. The other advantage that this provides is that you can profit from your existing investment in primary storage, your backup application, and tape libraries.

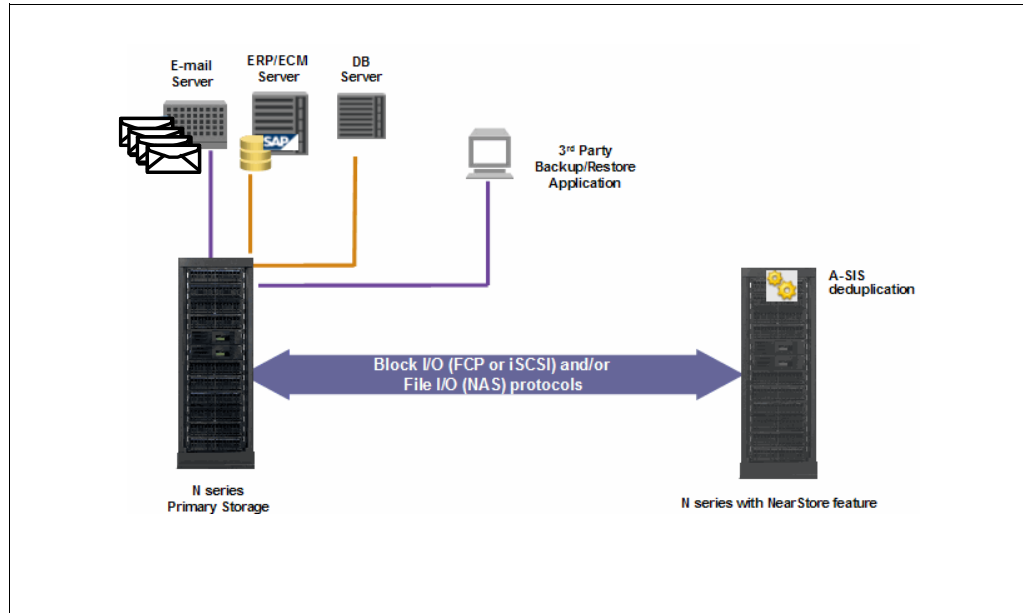


Figure 5-2 Near-line storage solution

5.2 IBM System Storage N series hardware

In the following sections we discuss the N series models available today:

- ▶ N3000 series
- ▶ N6000 series
- ▶ N7000 series

Table 5-2 provides details about these models.

Table 5-2 IBM System Storage N series storage systems and their disk capacity

A1X and A2X models	Maximum number of drives	Maximum raw capacity in TB
N3300	68	68
N3600	104	104
N6040 ^a	420	420
N6060 ¹	672	672
N6070 ¹	840	840
N7700	840	840
N7900	1176	1176

a. The N6000 models series also act as a gateway with Data ONTAP 7.3 and later.

5.2.1 Maximum number of network interfaces

Beginning with Data ONTAP 7.3, storage systems (except for the smallest models) are no longer limited to 128 interfaces per storage system. Larger storage systems can accommodate from 256 to 1,024 network interfaces per system depending on the system's memory and whether they are in a clustered, active/active configuration.

Table 5-3 shows the new maximum number of network interfaces allowed for each system. The total number of interfaces can include physical, Virtual Interface (VIF), virtual local area network (VLAN), vh, and loopback interfaces.

The number of physical interfaces depends on the storage system model. Up to 16 VIFs can be supported by each storage system. The maximum number of VLANs that can be supported equals the maximum number of network interfaces shown in Table 5-3 minus the total number of physical interfaces, VIFs, vh, and loopback interfaces supported by the storage system.

Table 5-3 Storage system memory and maximum number of interfaces

Storage system memory	Maximum number of interfaces
2 GB or less	128
2 GB or less in an active/active configuration	256
6 GB or less	256
6 GB or less in an active/active configuration	512
More than 6 GB	512
More than 6 GB in an active/active configuration	1024

5.2.2 IBM System Storage N series A models hardware reference

Table 5-4 provides a quick reference for A model hardware.

Table 5-4 A model's hardware quick reference

Function	N3300	N3600	N6040	N6060	N6070	N7700	N7900
Maximum raw capacity in TB A1X models	68	104	420	672	840	840	1176
Maximum raw capacity in TB A2X models	68	104	420	672	840	840	1176
Fibre Channel disk drives	10 K RPM - 300 GB 15 K RPM - 144 GB, 300 GB, 450 GB						
Serial Advanced Technology Attachment (SATA) disk drives	7.2 K RPM - 500 GB, 750 GB, 1 TB						
Maximum number of disks	68	104	420	672	840	840	1176
Expansion units supported	EXN1000 (SATA), EXN2000 (FC 2 Gbps), EXN4000 (FC 4 Gbps)						

Table 5-5 provides a quick reference for N series Gateway (G) models.

Table 5-5 N series G models quick reference

Function	N7700	N7900
Maximum raw capacity in TB G10/G20 models	840	1176
Max. number of logical units (LUNs) on back-end disk storage array	840	1176
Max LUN size in GB	1000	1000
Maximum volume size in TB	16	16

5.2.3 IBM System Storage N series A and G models hardware quick reference

Table 5-6 provides a quick reference for IBM System Storage N series A and G model hardware.

Table 5-6 Storage system reference

Function	N3300	N3600	N6040	N6060	N6070	N7700	N7900
Network protocol support	NFS V2/V3/V4 over UDP or TCP, PCNFSD V1/V2 for (PC) NFS client authentication, Microsoft CIFS, iSCSI, FCP, FCoE, VLD, HTTP 1.0, HTTP1.1 virtual host						
Other protocol support	SNMP, NDMP, LDAP, NIS, DNS						
Onboard I/O ports per node	4 x GbE 4 x FC	4 x GbE 4 x cal FC	4 x GbE 8 x FC ^a		4 x Gbe 8 x FC ^a	6 x GbE 8 x FC	6 x GbE 8 x FC
PCI expansion slots per node	N/A	1 x PCI-E	4 x PCI-E	4 x PCI-E	4 x PCI-E	3 x PCI-x5 10 x PCI-E	3 x PCI-E 10 x PCI-E
Non-volatile RAM (NVRAM) in MB per node	128	256	512	1024	2048	2048	4096
Memory in GB per node	1	2	4	8	16	16	32
Rack space per node	2U for two nodes	4U for two nodes	6U for two nodes	6U for two nodes	6U for two nodes	6U	6U
Processors (A1X)	One 2.2 GHz Mobile Celeron®	One 2.2 GHz Mobile Celeron	2.4 GHz AMD Dual-core 64-bit Opteron	Two 2.6 GHz AMD Dual-core 64-bit Opteron	Two 2.6 GHz AMD Dual-core 64-bit Opteron	Two 2.6 GHz AMD dual-core Opteron	Four 2.6 GHz AMD dual-core Opteron
Processors (A2X)	Two 2.2 GHz Mobile Celeron	Two 2.2 GHz Mobile Celeron	Two 2.4 GHz AMD Dual-core 64-bit Opteron	Four 2.6 GHz AMD Dual-core 64-bit Opteron	Four 2.6 GHz AMD Dual-core 64-bit Opteron	Four 2.6 GHz AMD dual-core Opteron	Eight 2.6 GHz AMD dual-core Opteron

a. 4 Gbps

5.2.4 IBM System Storage N3000 introduction

The N3000 systems are designed to provide primary and secondary storage for midsize enterprises. IT administrators can consolidate fragmented application-based storage and unstructured data into one unified, easily managed, expandable platform. N3000 systems offer integrated block-level and file-level data access, intelligent management software, and data protection capabilities (similar to those offered in higher-end N series systems) in a cost-effective package. N series innovations include Serial-Attached SCSI (SAS) disk drive support, expandable I/O connectivity, and onboard remote management.

N3000 systems

The N3000 systems are designed as the entry point to the entire N series family. These systems can provide the following key advantages:

- ▶ High availability: Utilizes proven features including a high-performing and scalable operating system, data management software, and redundancy features
- ▶ Backup and recovery features: Designed to support disk-based backup, with file or application-level recovery with Snapshot and SnapRestore® software features
- ▶ Simple replication and disaster recovery: Designed to provide an easy-to-deploy mirroring solution that is highly tolerant of wide area network (WAN) interruptions
- ▶ Management simplicity: Self-diagnosing systems designed to enable on-the-fly provisioning
- ▶ Versatile, single, integrated architecture: Designed to support concurrent block I/O and file serving over Ethernet and Fibre Channel SAN infrastructures

The N3000 series is compatible with the entire family of N series unified storage systems, which feature a comprehensive line-up from top-to-bottom of hardware and software designed to address a variety of possible deployment environments:

- ▶ N3300:
 - 2859-A10 Single Node
 - 2859-A20 Clustered
- ▶ N3600:
 - 2862-A10 Single Node
 - 2862-A20 Clustered

The N3000 series supports Ethernet and Fibre Channel environments, enabling economical NAS, FC, and iSCSI deployments. The N3000 system functions as a *unification engine*, which is designed to allow you to simultaneously serve both file-level and block-level data across a single or multiple networks demanding procedures that, for certain solutions, require multiple separately managed systems.

N3000 storage systems can offer significant advantages for distributed enterprises with remote and branch office sites. These organizations and others can utilize the SnapVault® and SnapMirror software functions to implement a cost-effective data protection strategy by mirroring data to a central corporate data center.

There are no additional PCI adapter slots in the N3300 storage system. On the other hand, the N3600 storage system has one available PCIe adapter slot per node. For an A20 model, identical adapters must be added in pairs, one to each node, so that both nodes are populated with identical types of PCIe adapters.

N3300 and N3600

The N3300 and 3600 systems (Figure 5-3 and Figure 5-4) provide multiple I/O connectivity options, a small footprint to hold high-density Serial-Attached SCSI (SAS) drives, and external expansion using either low-cost SATA drives or Fibre Channel disks for production applications while using Data ONTAP's Snapshot technology. SAS is the next generation of SCSI that combines the advantages of parallel SCSI and serial FC.

For further systems administration time and cost advantages, the systems come standard with Remote Onboard Management capabilities to simplify remote system monitoring, cycle power, execute firmware upgrades, enter console commands, and run diagnostics to maintain the reliability of the system as well as business-critical data.



Figure 5-3 N3300



Figure 5-4 N3600

Figure 5-5 and Figure 5-6 show the rear views of N3300 and N3600, respectively. The single node models (A10) use a single control unit, while the dual-node clustered models (A20) use two control units.



Figure 5-5 N3300 rear view



Figure 5-6 N3600 rear view

N3300 is a 2U-high device with capacity for 12 internal SAS drive bays and supports up to two external disk expansion units. Each controller has dual gigabit Ethernet ports, dual 4 Gbps Fibre Channel ports (Figure 5-7), one console port, and one remote management port.

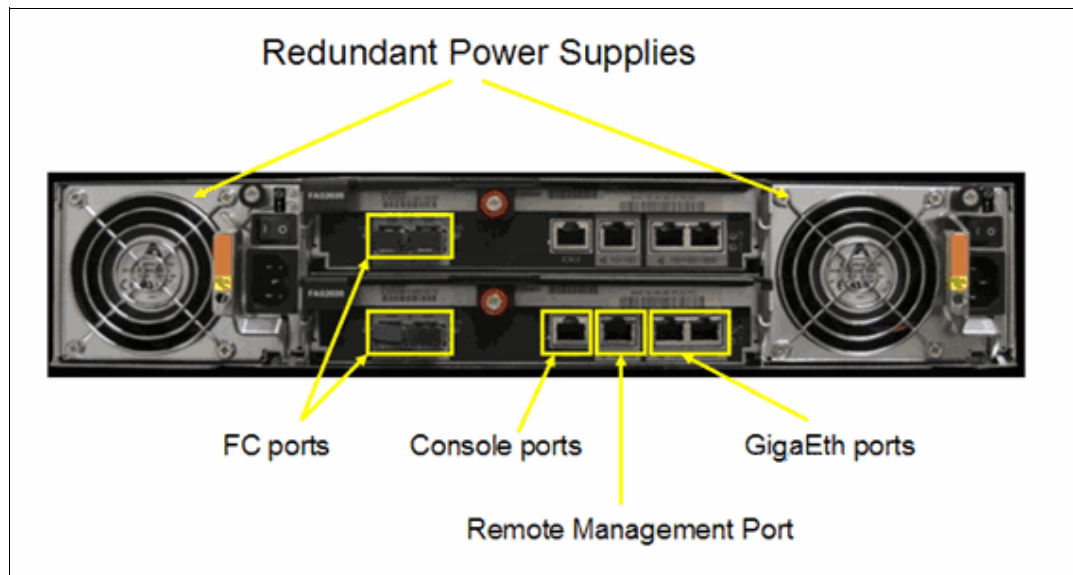


Figure 5-7 External ports on N3300

Tips: The N3300 series supports SAS, FC, and SATA disk technologies. The N3300 controller chassis supports from 6 to 12 SAS or SATA disk drives. All disks must be of the same type. The N3300 controller can be configured with 0 disk drives and storage in disk expansion units like EXN1000/SATA or EXN4000/FC.

The N3600 has redundant power supplies (Figure 5-8).

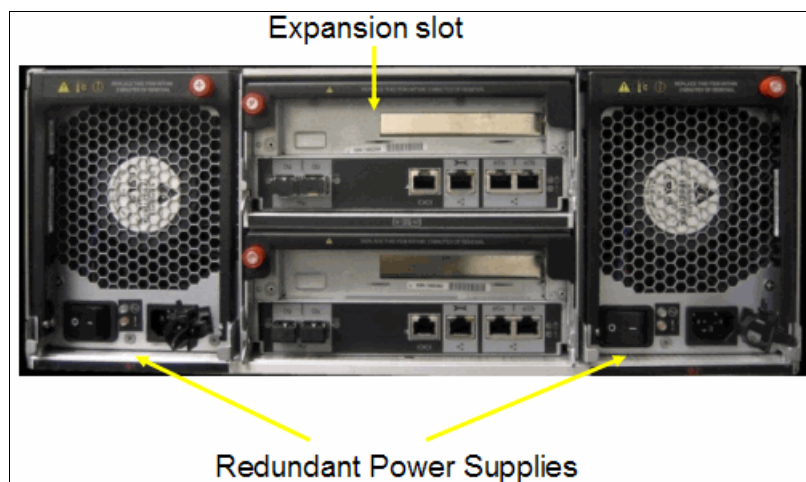


Figure 5-8 N3600 power supplies and expansion slot

N3600 is a 4U-high device with capacity for 20 internal SAS drive Bays. N3600 can support up to six external disk expansion units. Each controller has dual gigabit Ethernet ports and dual 4 Gbps Fibre Channel ports (Figure 5-9). It also has one console port and one remote management port, as well as one PCIe slot on each controller for an expansion card.

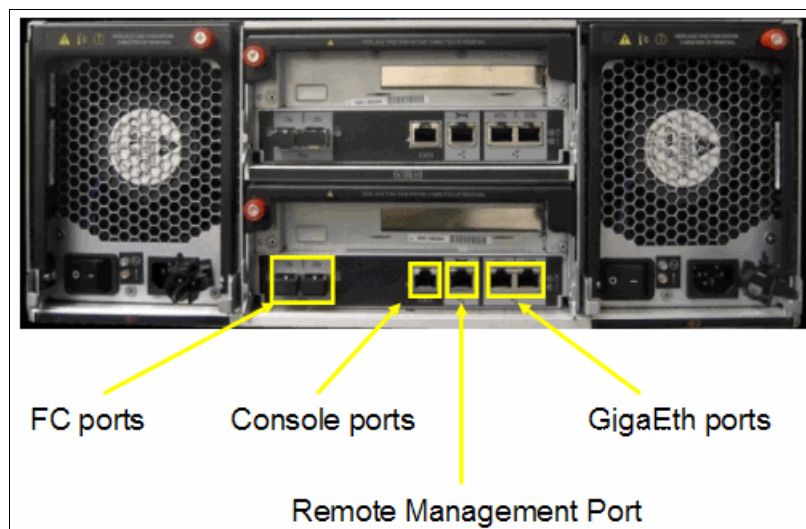


Figure 5-9 External ports on N3600

Tips: The N3600 series supports SAS, FC, and SATA disk technologies. The controller chassis supports either 20 SAS or 20 SATA disk drives. The N3600 requires a minimum of six SAS drives in the controller chassis.

The N3000 series is a small form-factor appliance that conserves scarce and valuable space in data centers or remote office locations and it is engineered for small to medium-sized enterprises.

Note: For N3000 quick reference of model specifications, see Table 5-4 on page 177 through Table 5-6 on page 178.

For detailed specifications of N3000 series models, see the website:

<http://www-03.ibm.com/systems/storage/network/n3000/appliance/specification.html>

5.2.5 IBM System Storage N6000 introduction

IBM N6000 series systems offer a versatile storage platform for handling the large amounts of diverse data moving through your business. With an N6000 series system, you can consolidate these varied data sets onto a unified storage platform supporting simultaneous block and file services for business and technical applications. The N6000 systems enable you to connect your heterogeneous server environment (including Windows, UNIX, and Linux servers) and clients to one storage system by using standard storage protocols and interfaces.

The IBM N6000 A series comes in the following models:

- ▶ N6040:
 - 2858-A10 Single Node
 - 2858-A20 Clustered
- ▶ N6060:
 - 2858-A12 Single Node
 - 2858-A22 Clustered
- ▶ N6070:
 - 2858-A11 Single Node
 - 2858-A21 Clustered

Figure 5-10 shows the front view of the N6000 series.



Figure 5-10 N6000 front view

Figure 5-11 shows the rear view of the N6000 series.

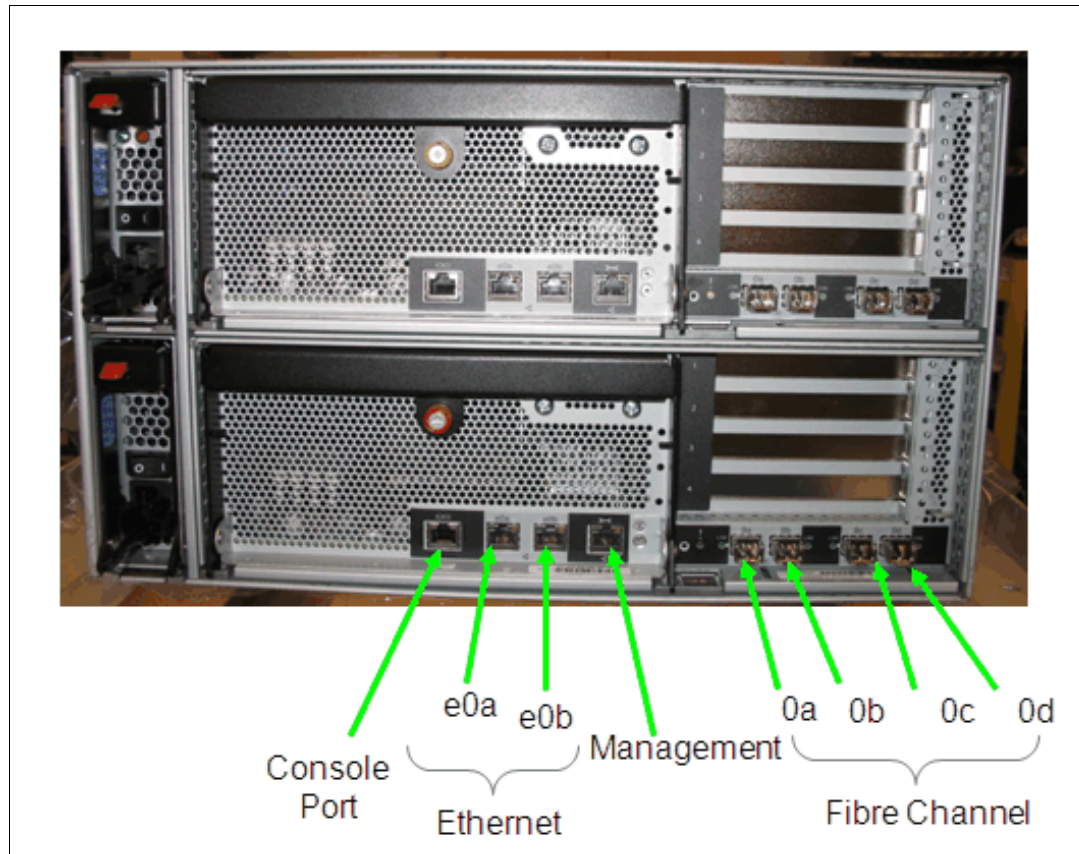


Figure 5-11 N6000 rear view

Note: For N6000 quick reference of model specifications, see Table 5-4 on page 177 through Table 5-6 on page 178.

For detailed specifications of N6000 series models, see the website:

<http://www.ibm.com/systems/storage/network/n6000/appliance/specifications.html>

5.2.6 IBM System Storage N7000 introduction

The IBM System Storage N7000 series offers additional choices to organizations facing the challenges of enterprise data management. The IBM System Storage N7000 series is designed to deliver high-end enterprise storage and data management value with midrange affordability. Built-in enterprise serviceability and manageability features help support your efforts to increase reliability, simplify and unify storage infrastructure and maintenance, and deliver exceptional economy.

- ▶ The IBM N7000 A series comes in two models:
 - N7700:
 - 2866-A11
 - 2866-A21 Clustered
 - N7900:
 - 2867-A11
 - 2867-A21 Clustered
- ▶ FC or SATA (both can be used behind a single controller but not in the same drawer)

Like its N5000 predecessor, the front of the N7000 series unit has the LCD display and the three standard LEDs indicating system activity, status, and power (Figure 5-12).



Figure 5-12 N7000 front view

From the rear of the N7000, you can see the redundant power supplies marked, the NVRAM card, the gigabit Ethernet interfaces, and the Fibre Channel interfaces. The console port and RLM port are also located on the rear (Figure 5-13).

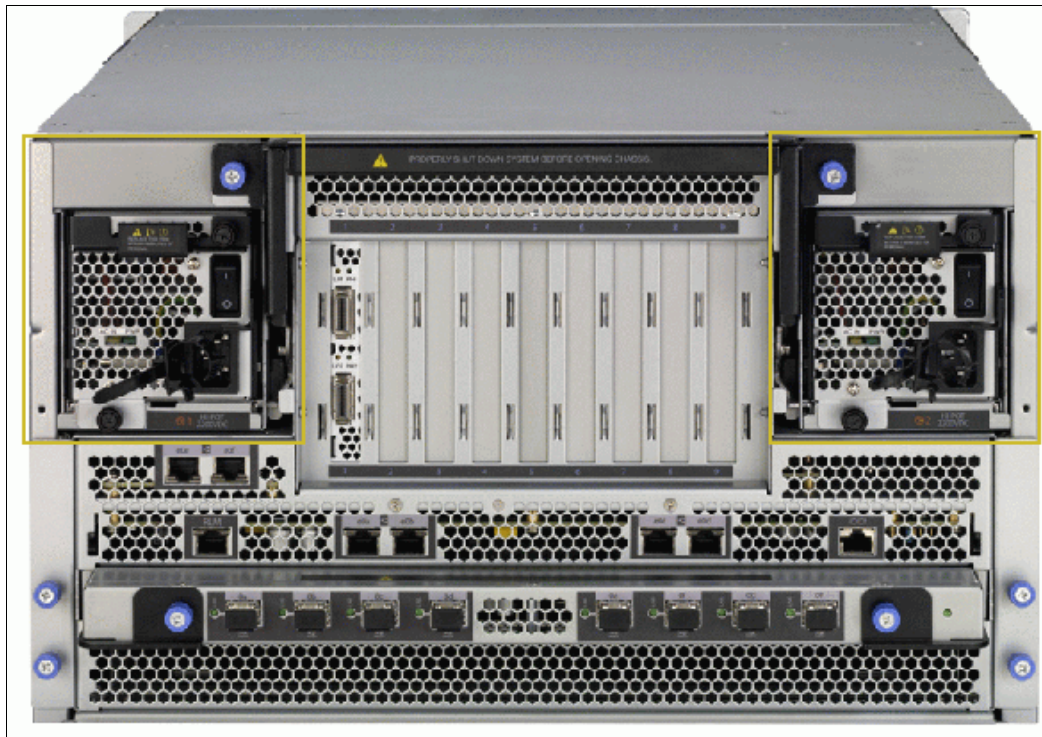


Figure 5-13 N7000 rear view

Each N7000 node requires 6U of rack space. Each expansion unit requires 3U of rack space. Each N7000 node requires at least one expansion unit (Figure 5-14).



Figure 5-14 N7000 racked

A dual-node N7900 supports a maximum of 84 storage expansion units. Each rack holds a maximum of 12 expansion units. The N7000 products are installed by IBM Service or by a qualified IBM Business Partner. These systems are not for customer setup.

Figure 5-15 shows the N7000 with multiple expansion units.



Figure 5-15 Clustered N7000 with multiple expansion units

Note: For N7000 quick reference of model specifications, see Table 5-4 on page 177 through Table 5-6 on page 178.

For detailed specifications of N7000 series models, see the website:

<http://www.ibm.com/systems/storage/network/n7000/appliance/specification.html>

5.3 IBM System Storage N series Gateways

The IBM System Storage N series Gateways, an evolution of the N5000 series product line, is a network-based virtualization solution that virtualizes tiered, heterogeneous storage arrays, allowing customers to utilize the dynamic virtualization capabilities available in Data ONTAP across multiple tiers of IBM and vendor-acquired storage (Figure 5-16).

Like all IBM System Storage N series storage systems, the IBM System Storage N series Gateway family is based on the industry-hardened Data ONTAP microkernel operating system, which unifies block and file storage networking paradigms under a common architecture and brings a complete suite of IBM System Storage N series advanced data management capabilities for consolidating, protecting, and recovering mission-critical data for enterprise applications and users.

Figure 5-16 shows the N series gateway in a heterogeneous SAN environment.

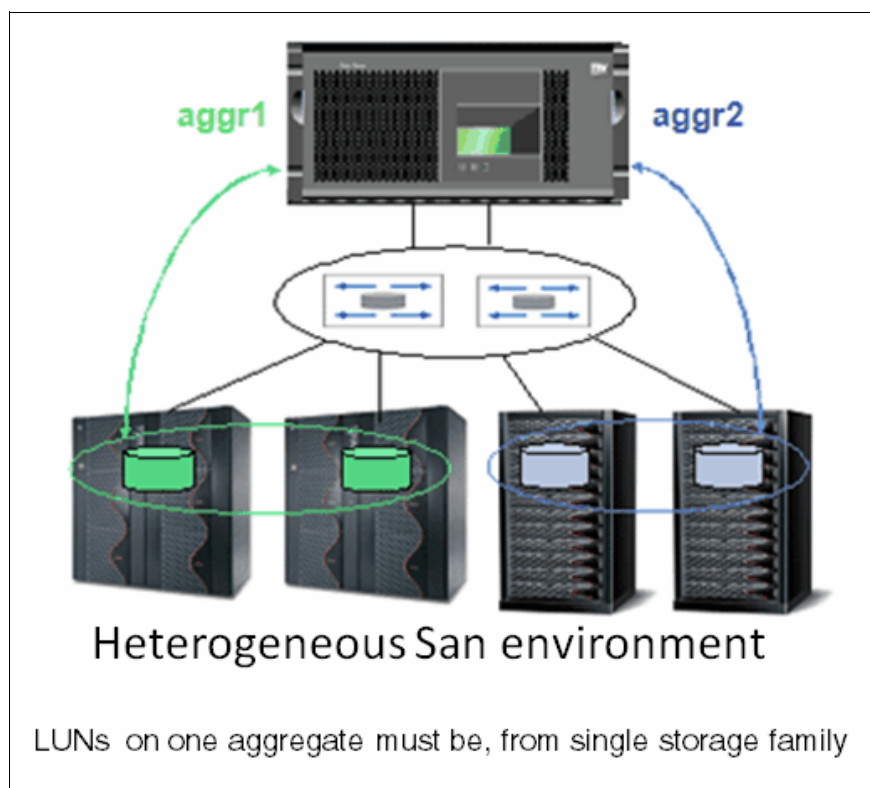


Figure 5-16 Heterogeneous storage

The industry's most comprehensive virtualization solution, the N series Gateways provides proven and innovative data management capabilities for sharing, consolidating, protecting, and recovering mission-critical data for enterprise applications and users and seamlessly integrates into mission-critical enterprise-class SAN infrastructures. These innovative data management capabilities, when deployed with disparate storage systems, simplify heterogeneous storage management (Figure 5-16).

The IBM System Storage N series Gateway presents shares, exports, or LUNs that are built on flexible volumes that reside on aggregates. The N series Gateway is also a host on the storage array SAN. N series Gateways can take storage array LUNs (which are treated as disks) and virtualize them through Data ONTAP, presenting a unified management interface.

The IBM System Storage N series Gateway offers customers new levels of performance and scalability and a robust portfolio of proven data management software for sharing, consolidating, protecting, and recovering mission-critical data.

IBM System Storage N series storage systems seamlessly integrate into mission-critical SAN environments and provide a simple, elegant data management solution by decreasing management complexity, improving asset utilization, and streamlining operations to increase business agility and reduce total cost of ownership for organizations that are looking for ways to utilize SAN-attached storage to create a consolidated storage environment for the various classes of applications and storage needs throughout their enterprise. These customers are looking for ways to increase utilization, simplify management, improve consolidation, enhance data protection, enable rapid recovery, increase business agility, deploy heterogeneous storage services, and broaden centralized storage usage by provisioning SAN capacity for business solutions requiring NAS, SAN, or IP SAN data access.

Figure 5-17 shows the Gateway topology for N series gateways.

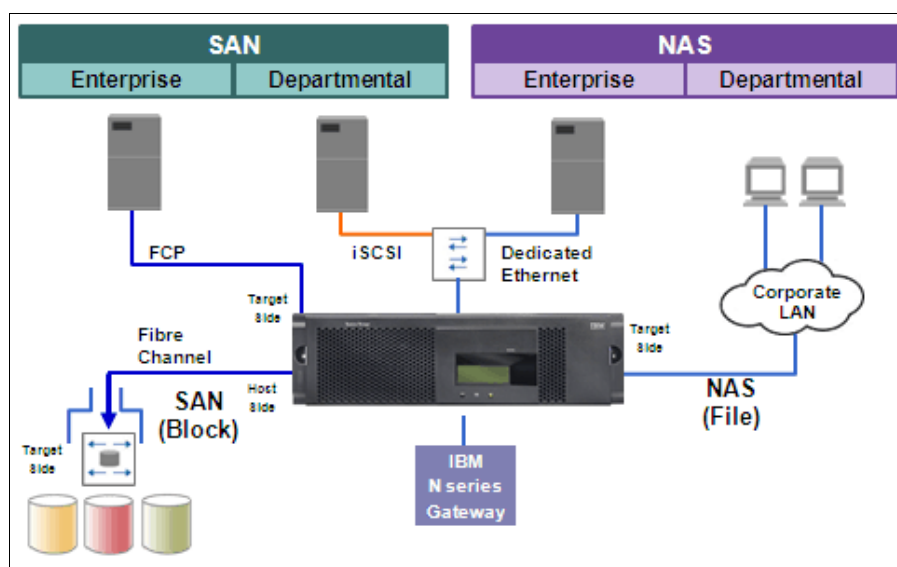


Figure 5-17 Gateway topology

These customers have:

- ▶ Significant investments or a desire to invest in a SAN architecture
- ▶ Excess capacity or an attractive storage cost for SAN capacity expansion
- ▶ Increasing requirements for both block access (FCP, iSCSI) and file access (NFS, CIFS, and so on)
- ▶ Increasing local or remote shared file services and file access workloads

They are seeking solutions to:

- ▶ Cost-effectively increase utilization
- ▶ Consolidate distributed storage, Direct Access Storage, and file services to SAN storage
- ▶ Simplify storage management
- ▶ Improve storage management business practices

With Data ONTAP, the N series Gateway now supports attachment of heterogeneous storage systems as well as IBM expansion units of the type used with N series storage systems.

There are two halves to set up: An N series Gateway implementation can be thought of as a front-end implementation and a back-end implementation:

- ▶ The front-end setup includes configuring the N series Gateway for all protocols (NAS or FCP), implementing any snap features (Snapshot, SnapMirror, SnapVault, and so on), and setting up backup including NDMP dumps to tapes.
- ▶ The back-end implementation includes all tasks required to set up the N series Gateway system up to the point where it is ready for Data ONTAP installation. These include array LUN formatting, port assignment, cabling, switch zoning, assigning LUNs to the N series Gateway system, creating aggregates, and loading Data ONTAP.

5.3.1 IBM System Storage N series Gateway highlights

IBM System Storage N series Gateway provides a number of key features that enhance the value and reduce the management costs of utilizing a storage area network (Figure 5-18). An N series Gateway:

- ▶ Simplifies storage provisioning and management
- ▶ Lowers storage management and operating costs
- ▶ Increases storage utilization
- ▶ Provides comprehensive, simple-to-use data protection solutions
- ▶ Improves business practices and operational efficiency
- ▶ Transforms conventional storage systems into a better managed storage pool

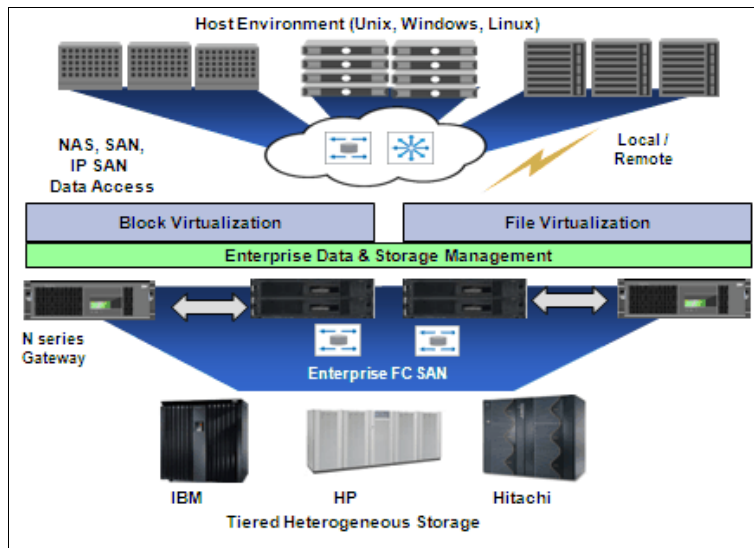


Figure 5-18 Tiered heterogeneous storage

5.3.2 IBM System Storage N7700 and N7900 Gateway models

IBM System Storage N7000 series Gateway models offers additional choices to organizations facing the challenges of enterprise data management. The IBM System Storage N7000 series is designed to deliver high-end enterprise storage and data management value with midrange affordability. Built-in enterprise serviceability and manageability features help support your efforts to increase reliability, simplify and unify storage infrastructure and maintenance, and deliver exceptional economy.

The IBM System Storage N series Gateway models N7700 and N7900 deliver all the feature function that the N5000 series did but with increased processing, memory, NVRAM, and total storage capacity. The N7700 and N7900 are designed with the high-end of the enterprise environments. The N7000 series Gateway hardware is identical to the A10 and A20 models, the difference being in the enabled features and disk attachment by Data ONTAP.

The IBM System Storage N7000 G series comes in two models:

- ▶ N7700:
 - 2866-G11
 - 2866-G21 Clustered
- ▶ N7900:
 - 2867-G11
 - 2867-G21 Clustered

LUN sizing

Gateway support for LUN sizes is as follows:

- ▶ Maximum LUN size: 1024 GB
- ▶ Minimum LUN size: 100 MB

Data ONTAP definition: One GB is equal to 1000 x 1024 x 1024 bytes.

Therefore, the maximum LUN size that Data ONTAP supports means $1024 * 1000 * 1024 * 1024 = 1,048,576,000,000$ bytes.

LUN mapping

Storage subsystem LUNs are converted to disks for the IBM System Storage N series Gateway (Figure 5-19). When external storage is used the logical disk count equals the physical LUN count associated with the N series. The diagram shows an example of an array LUN mapped to a gateway disk.

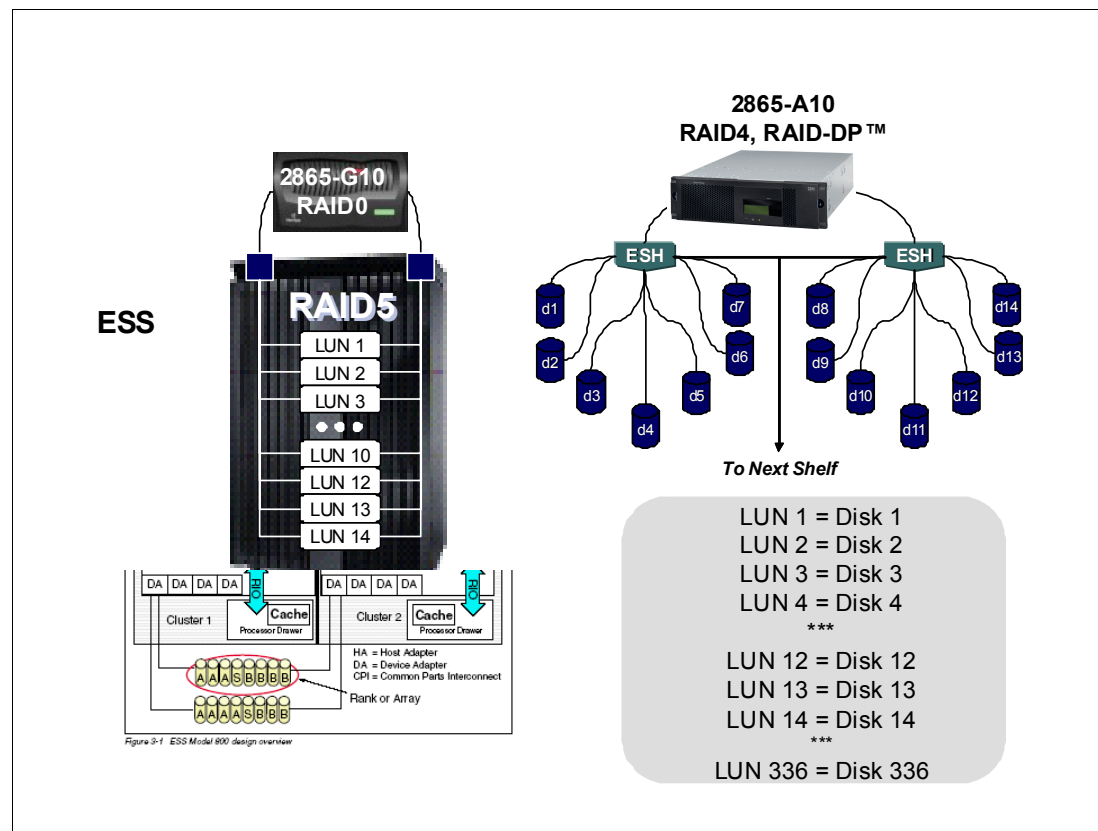


Figure 5-19 LUN to N series Gateway disk relationship

LUNs are added to the Gateway through the same volume wizard that we use on the N series A models (Figure 5-20).

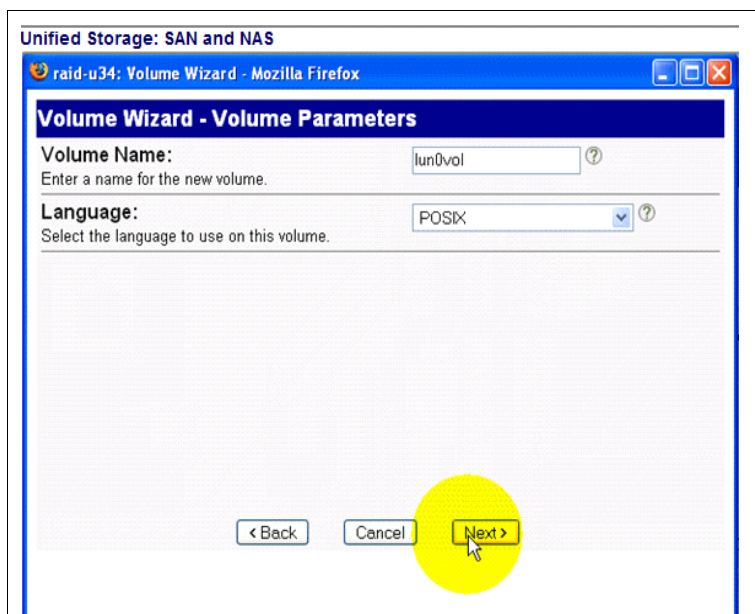


Figure 5-20 Volume wizard

Attention: Do not map LUN 0 to gateway systems, even if LUN 0 is a storage LUN.

For all N7000 model specifications, see Table 5-4 on page 177 through Table 5-6 on page 178.

5.4 N series expansion units

There are currently two disk storage expansion units specifically designed for the IBM System Storage N series storage systems:

- ▶ IBM EXN4000 Fibre Channel disk storage expansion unit with 4 Gbps interfaces
- ▶ IBM EXN1000 Serial Advanced Technology Attachment storage expansion unit

Note: EXN expansion units can be used for attachment to a gateway with Data ONTAP 7.3 and later.

Multiple EXN1000s, each having unique SATA disk drive feature codes, can be attached to the same N series storage system on the same Fibre Channel loop. Multiple EXN4000s, each having unique Fibre Channel disk drive feature codes, can be attached to the same N series storage system on the same Fibre Channel loop.

For the latest storage expansion unit support information, see the following website:

<http://www.ibm.com/storage/support/nas/>

5.4.1 Intermixing EXN units with N series A models

EXN4000s are Fibre Channel disk storage expansion units. Intermixing Fibre Channel and SATA disk drives in a supported N series storage system configuration is supported as follows:

- ▶ Intermixing Fibre Channel disk expansion units with SATA disk expansion units on the same loop is not supported.
- ▶ EXN4000s (Fibre Channel disk drives) and EXN1000s (SATA disk drives) can be attached to the same N series storage system only if the Fibre Channel disk expansion units (EXN4000s) are on separate loops from the SATA disk expansion units (EXN1000s).

Figure 5-21 shows the Speed switches on the EXN4000.

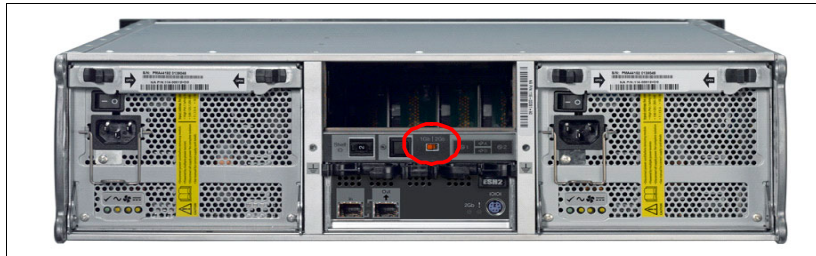


Figure 5-21 Speed switches

5.4.2 EXN1000

The EXN1000 uses the same shelf and hardware as the EXN4000 so it has the same dimensions. It also supports the same number of disks per shelf (14). The main differences are as follows:

- ▶ Drive type supported: SATA versus Fibre Channel
- ▶ Interface module: AT-FCX versus ESH2

Figure 5-22 shows the EXN1000 front view.



Figure 5-22 EXN1000 expansion unit: front view

Data ONTAP supports up to 400 RAID groups per storage system or cluster. When configuring your aggregates, keep in mind that each aggregate requires at least one RAID group and that the total of all RAID groups in a storage system cannot exceed 400.

5.4.3 EXN4000

The EXN4000 uses the same shelf and hardware as the EXN1000, so it has the same dimensions. EXN4000 also supports the same number of disks per shelf (14). EXN4000 uses ESH4 as its controller module. ESH4 refers to the third-generation, multiloop speed ESH module. ESH4 can function at 1 Gb, 2 Gb, or 4 Gb loop speed when it works with EXN4000. The ESH4 has LEDs that indicate whether the module is functioning normally, whether there are any problems with the hardware, and the loop speed operation of the EXN4000. The main differences are:

- ▶ The following drive type is supported: Fibre Channel versus SATAI.
- ▶ The EXN4000 FC storage expansion unit will run at 2 Gbps FC when attached to systems that do not have 4 Gbps capability.
- ▶ ENX4000 can also be added to loops with existing EXN2000 loops.

Figure 5-23 shows the EXN4000 front view.



Figure 5-23 EXN 4000 front view

Figure 5-24 shows the EXN4000 rear view.



Figure 5-24 EXN4000 rear view

5.5 IBM System Storage N series standard software features

Table 5-7 lists no-charge features that are available with the IBM System Storage N series.

Table 5-7 Standard software features

Data ONTAP	Operating system software that optimizes data serving and allows multiple protocol data access.
FTP	File Transfer Protocol (FTP), a standard Internet protocol, is a simple way to exchange files between computers on the Internet
Telnet	The Telnet protocol provides a general, bi-directional communications facility. It provides user-oriented command-line login sessions between hosts.
Snapshot	This enables online backups, providing near instantaneous access to previous versions of data without requiring complete, separate copies.
FlexVol®	FlexVol creates multiple flexible volume on a large pool of disks; dynamic, nondisruptive (thin) storage provisioning; and space and time efficiency. These flexible volumes can span multiple physical volumes without regard to size.
FlexShare™	FlexShare gives administrators the ability to utilize existing infrastructure and increase processing utilization without sacrificing the performance of critical business needs. With the use of FlexShare, administrators can confidently consolidate various applications and data sets on a single storage system. FlexShare gives administrators the control to prioritize applications based on how critical they are to the business.
Disk sanitization	Disk sanitization is the process of physically obliterating data by overwriting disks with specified byte patterns or random data in a manner that prevents recovery of current data by any known recovery methods. This feature enables you to carry out disk sanitization by using three successive byte overwrite patterns per cycle. By default, six cycles are performed.
FilerView®	This is a Web-based administration tool that allows IT administrators to fully manage N series storage systems from remote locations. It provides simple and intuitive Web-based single-appliance administration.
SnapMover®	This migrates data among N series clusters with no impact on data availability and no disruption to users.
AutoSupport	AutoSupport is a sophisticated, event-driven logging agent featured in the Data ONTAP operating software and inside each N series system. It continuously monitors the health of your system and issues alerts if a problem is detected. These alerts can also be in the form of email.
SecureAdmin™	SecureAdmin is a Data ONTAP module that enables authenticated, command-based administrative sessions between an administrative user and Data ONTAP over an intranet or the Internet.
Domain Name System (DNS)	The N series supports using a host naming file or a specified DNS server and domain.
Network Information Service (NIS)	The N series provides NIS client support and can participate in NIS domain authentication.
Integrated automatic Redundant Array of Independent Disks (RAID) manager	The IBM System Storage N series and Data ONTAP provide integrated RAID management with RAID-Double Parity (default) and RAID 4.

iSCSI Host Attach Kit for AIX, Windows, Linux	A host support kit includes support software and documentation for connecting a supported host to an iSCSI network. The support software includes programs that display information about storage and programs to collect information needed by customer support to diagnose problems.
Systems Manager	The basic idea behind the System Manager (SM) is to provide comprehensive management and capability of one of more arrays by way of a simple, easy-to-use, intuitive GUI.
SyncMirror®	SyncMirror is the synchronous mirror of a volume. It maintains a strict physical separation between the two copies of your mirrored data. In case of an error in one copy, the data is still accessible without any manual intervention.

5.6 Optional software

Table 5-8 provides information about optional chargeable features.

Table 5-8 Optional software

Common Internet File System (CIFS)	This provides file system access for Microsoft Windows environments.
NFS	This provides file system access for UNIX and Linux environments.
Hypertext Transfer Protocol (HTTP)	Hypertext Transfer Protocol allows a user to transfer displayable web pages and related files.
FlexCache™ for NFS	This provides efficient caching of files/volumes in a local N series storage system when the source volume resides in a remote location N series storage system, thus avoiding inefficient use of bandwidth resources.
FlexClone®	This is designed to provide instant replication of data volumes/sets without requiring additional storage space at the time of creation.
FlexScale	This is used exclusively to improve performance by managing the additional cache provided by Performance Accelerator Modules. These modules can be added as an option to an N series storage system.
Multistore	This permits an enterprise to consolidate a large number of Windows, Linux, or UNIX file servers onto a single storage system. Many virtual N series storage systems on one physical appliance ease migration and multi-domain failover scenarios.
SnapLock	This provides non-erasable and non-rewritable data protection that helps enable compliance with government and industry records retention regulations. Snaplock is not available starting in DATA ONTAP 7.3.0. It is available in Data ONTAP 7.3.1.
SnapMirror	This is remote mirroring software that provides automatic block-level incremental file system replication between sites. It is available in synchronous, asynchronous, and semi-synchronous modes of operation.
SnapRestore	This allows rapid restoration of the file system to an earlier point in time, typically within a few seconds.
SnapVault	This provides disk-based backup for N series systems by periodically backing up a Snapshot copy to another system

SnapDrive®	SnapDrive enables Windows and UNIX applications to access storage resources on N series storage systems, which are presented to the Windows 2000 or later operation system as locally attached disks. For UNIX it allows you to create storage on N series storage systems in the form of logical unit numbers (LUNs), file systems, logical volumes, or disk groups.
SnapManager	Host software for managing Snapshots for backup and restore operations. There are various versions of SnapManager that integrate easily with critical applications, in particular: ▶ SnapManager for MS Exchange ▶ SnapManager for SQLServer ▶ SnapManager for MS SharePoint ▶ SnapManager for Oracle ▶ SnapManager for SAP ▶ SnapManager for Virtual Infrastructures, which automates and simplifies backup and recovery of primary storage used by VMWare Virtual Infrastructure
SnapValidator®	For Oracle deployments, SnapValidator can be used to provide an additional layer of integrity checking between the application and N series storage. SnapValidator allows Oracle to create checksums on data transmitted to N series storage for writes to disk and include the checksum as part of the transmission.
Single Mailbox Recovery for Exchange (SMBR)	SMBR is a software option from SnapManager that is designed to take near-instantaneous online backups of Exchange databases, verify that the backups are consistent, and rapidly recover Exchange within levels (storage group, database, folder, single mailbox, or single message). The potential results are improved service to internal clients, reduced infrastructure expenses, and significant time savings for Exchange administrators.
Operations Manager SRM	The File Storage Resource Manager (FSRM) feature of Operations Manager provides monitoring and management of storage resources, including applications, files, file systems, and networks.
Operations Manager Core	Operations Manager provides remote, centralized management of IBM System Storage N series data storage infrastructure, including global enterprise, storage network, and so on.
MetroCluster	MetroCluster software provides an enterprise solution for high availability over wide-area networks between two clustered nodes of a single N series storage system.
NearStore® option	This provides enhanced performance in a disk-based, secondary storage device used for enterprise applications.
Advanced Single Instance Storage (ASIS)	This is designed to significantly improve physical storage efficiency and network efficiency by enabling the sharing of duplicate data blocks. ASIS provides a data-deduplication solution native to N series.
Virtual File Manager (VFM®)	IBM System Storage N series Virtual File Manager (VFM) software is a comprehensive solution for managing unstructured file data. It is designed to provide data management functionality for server and storage consolidation, migration, remote office data management, and disaster recovery features while avoiding disruption to users. It provides all of this functionality through automated policy-based data management utilizing a global namespace.
Cluster fail over (CFO)	This ensures high data availability for business-critical requirements by eliminating a single point of failure. It must be ordered for A2X clustered configurations or upgrades from A1X to A2X. Its active/active pairing delivers even more <i>nines to the right of the decimal point</i> .

5.7 Performance Accelerator Module

The performance of a disk system is heavily dependent on disk count as well as RPM of each disk. While a greater disk count and higher RPM provide better performance, this also increases the power consumption and space usage in the data center. The Performance Accelerator Module (PAM) has been introduced in the N6000, and N7000 series as a means of improving performance by adding modules with additional cache memory.

Each PAM currently consists of 16 GB of memory, and it is possible to add more than one PAM to an N series storage system. PAMs can be configured to operate in three modes:

- ▶ Mode 1: Behaves more like main memory
- ▶ Mode 2: Preferentially caches meta data
- ▶ Mode 3: Caches what otherwise might have been flushed

Figure 5-25 shows a photograph of the PAM.



Figure 5-25 IBM System Storage N series Performance Accelerator Module (PAM)

Here are advantages offered by the PAM:

- ▶ Alternative to adding more disk drives as a means of reducing latency and increasing response time and throughput
- ▶ PAM device functions as an intelligent read cache
- ▶ Automatically puts active data where access can be fast
- ▶ Greater I/O throughput without adding high-performance disk drives to a disk-bound storage system
- ▶ Effective for random read intensive workloads such as file services, OLTP databases, messaging, and virtual infrastructure (Exchange and virtual environments)
- ▶ Results can be predicted for an existing storage system

Using tools based on predictive cache statistics, it is possible to determine the optimal settings and the number of modules to deploy.

PAMs are available as optional adapters in N6000, and N7000 N series systems. When including PAMs, it is necessary to add the FlexScale licensed software.

5.8 References

To learn more about NAS:

- See the Redbooks publication:

IBM System Storage N series, SG24-7129.

- Browse the website:

<http://www.ibm.com/systems/storage/network/>



IBM Scale Out Network Attached Storage

The IBM Scale Out Network Attached Storage (SONAS) is a highly scalable system designed to provide a clustered NAS system with a single name space for CIFS, NFS, and FTP services. The system consists of two to thirty Interface nodes (2851-SI1), one to thirty storage pods consisting of the Storage node (2851-SS1), Storage controller (2851-DR1) and attached disks, and one to two Management nodes (2851-SM1). The maximum configuration provides up to 14.4 PB of storage capacity in a single clustered, highly-redundant system.

The storage used in the SONAS system can be high performance 15K RPM SAS hard disk drives or high capacity 7.2K RPM SATA hard disk drives, allowing users to tune the configuration to suit their needs.

There three rack configurations (2851-RXA, -RXB, and -RXC) that be populated with the systems and switches needed to configure according to the needs of the environment.

6.1 Overview

SONAS is designed to address the new storage challenges posed by the continuing explosion of data. Utilizing mature technology from the IBM High Performance Computing experience, and based upon the IBM General Parallel File System (GPFS™), SONAS is an easy-to-install, turnkey, modular, scale out NAS solution that provides the performance, clustered scalability, high availability and functionality that are essential to meeting strategic Petabyte Age and cloud storage requirements.

The high-density, high-performance SONAS can help organizations consolidate and manage data affordably, reduce crowded floor space, and reduce management expense associated with administering an excessive number of disparate storage systems. With its advanced architecture, SONAS virtualizes and consolidates multiple filers into a single, enterprise-wide file system, which can translate into reduced total cost of ownership, reduced capital expenditure, and enhanced operational efficiency.

SONAS provides a global namespace that enables your storage infrastructure to scale to extreme amounts of data, from terabytes to petabytes. Within the solution, centralized management, provisioning, control, and automated information life-cycle management (ILM) are integrated as standard features to provide the foundation for a truly cloud storage-enabled solution.

SONAS provides extreme scale out capability, a globally clustered NAS file system built upon IBM GPFS. The global namespace is maintained across the entire global cluster of multiple storage pods and multiple interface nodes. All inter-face nodes and all storage nodes share equally in the cluster to balance workloads dynamically and provide parallel performance to all users and storage, while also assuring high availability and automated failover.

SONAS will eventually support a full complement of standards-based network protocols, including Common Internet File System (CIFS), Network File System (NFS), Secure Copy Protocol (SCP), Hypertext Transfer Protocol (HTTP) and File Transfer Protocol (FTP).

6.2 How SONAS works

SONAS is a scale out NAS offering designed to manage vast repositories of information in enterprise environments requiring very large capacities, high levels of performance, and high availability. The system supports multiple petabytes of storage and up to a billion files in a single file system and supports up to 256 file systems per system. Data access to the system from client computers is provided using industry standard network file protocols NFS v2/v3, CIFS, and FTP (including z/OS® NFS client support).

In addition, the system supports traditional NAS features and functions, including:

- ▶ Snapshots
- ▶ Quotas (user, group, and fileset level)
- ▶ Integration with user directory servers such as Microsoft® Active Directory (AD) and Lightweight Directory Access Protocol (LDAP)
- ▶ Command line interface (CLI) and browser-based graphical user interface (GUI)
- ▶ Call home and remote support capabilities

In addition to these traditional features, the system supports advanced features, including integrated Tivoli® Storage Manager backup client.

The advanced architecture of SONAS virtualizes and consolidates the file space into a single, enterprise-wide file system as shown in Figure 6-1.

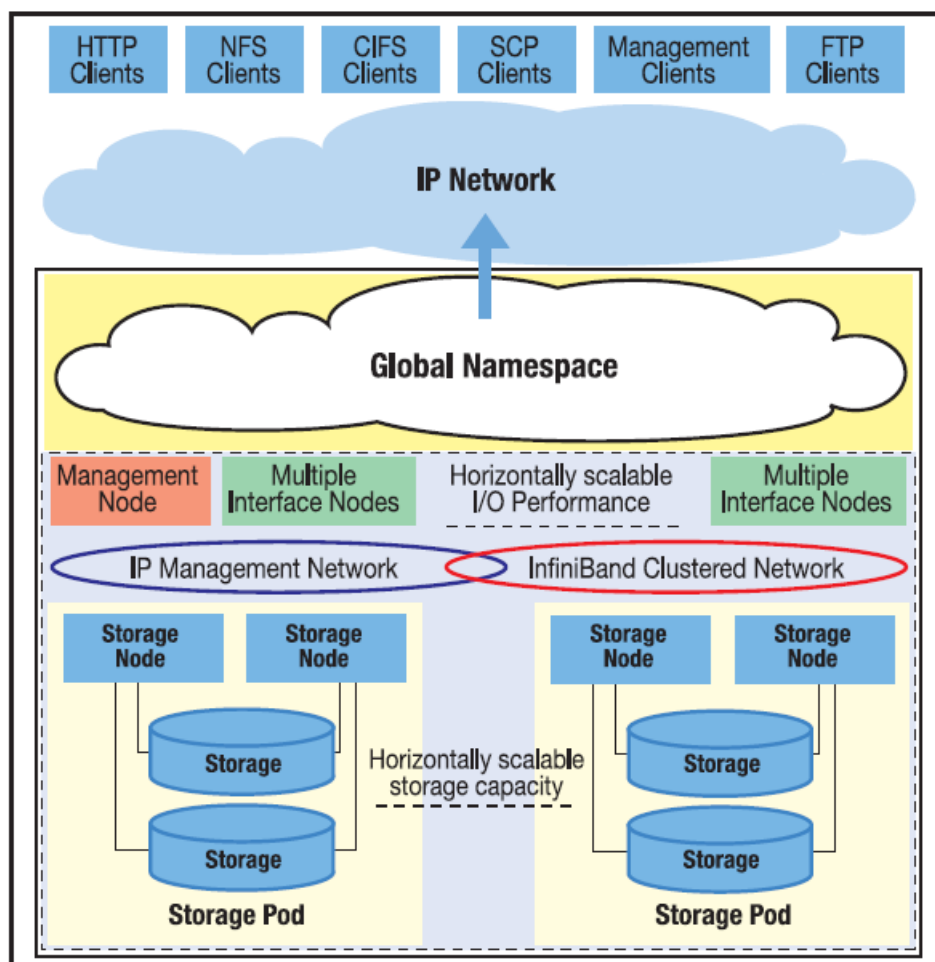


Figure 6-1 SONAS Architecture

The SONAS system consists of:

- ▶ Management node
- ▶ Interface nodes
- ▶ Storage nodes
- ▶ Storage controllers
- ▶ Disk storage expansion units
- ▶ Ethernet and InfiniBand switches and racks

The Management node provides a central point for the system administrator to configure, monitor, and manage the system. The Management node supports both a browser-based graphical user interface (GUI) and a command line interface (CLI). It also provides a System Health Center for monitoring the overall health of the system.

The Interface node provides the network port connection and the file services for the IBM SONAS system. These nodes can operate at up to 10 Gb speeds with optional adapters to providing extremely fast access to data. They are connected to the rest of the SONAS system by a redundant high speed InfiniBand data network.

The system has the concept of a storage pod. Each storage pod consists of a pair of storage nodes and one or two storage controllers. Each storage controller contains 60 disks, and can have one storage expansion attached for 60 additional disks per controller. Each storage pod supports a maximum of 240 hard disk drives. A maximum of 30 storage pods are supported for a maximum of 7200 hard disk drives in a single system.

These systems are configured within three pre-defined, customizable rack configurations. The Base Rack (2851-RXA) contains the management node, the InfiniBand switches, and from two to six interface nodes with up to one storage pod. Each Storage Expansion Rack (2851-RXB) contains up to two storage pods, for a maximum of 480 disks. The Interface Expansion Rack (2851-RXC) contains up to twenty additional interface nodes.

6.3 SONAS features

SONAS is designed to deliver cloud storage at petabyte scale, enabling cloud applications and services to be uncoupled from the underlying storage infrastructure, which allows the business to adjust quickly to change. As a result, SONAS can easily integrate with your organization's strategies to develop a more dynamic enterprise. Next we discuss important features of SONAS.

6.3.1 Centrally managed, centrally deployed

SONAS enables you to cluster and centrally deploy and manage a large scale out NAS storage environment, in order to gain a unified centralized management capability for security, provisioning, and configuration.

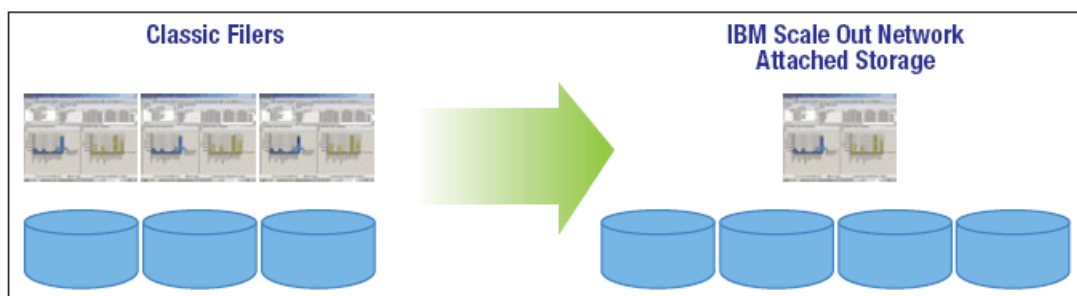


Figure 6-2 SONAS unified management view

The global cluster provides intelligent load balancing, dynamic thin provisioning, geographical distribution, performance optimization, and advanced replication, all globally applied and centrally managed.

6.3.2 Automated, integrated tiered storage and information life-cycle management

SONAS supports end users across the organization with a comprehensive suite of storage capabilities, including tiered storage and ILM.

When file systems grow from terabytes to petabytes and to millions and billions of files, the task of managing and migrating this scale of data can become an immense challenge. Backup and archive becomes complex, expensive, and slow because of the large amount of data that must be scanned to identify what must be backed up. SONAS is designed to solve these problems by providing extreme scale out capability in a single global namespace, with automated, integrated policy-driven ILM and hierarchical storage management built in.

The integrated ILM capability allows you to manage a petabyte-level global namespace in logical storage pools instead of individual disks. The architecture enables easy dynamic growth and transparent data migration, supporting the ability to make changes to your storage infrastructure without having to plan weeks in advance. You can centrally provision storage nearly instantaneously, on a quota basis, by user, group or location. Extensive global reporting and global access control capabilities are standard features.

SONAS provides a powerful high-performance scan engine that makes ILM possible at the level of petabytes of data with millions and billions of files. The high-performance scan engine enables fast searches, with ILM data movement, back-ups, and restores done in parallel by all nodes in the system.

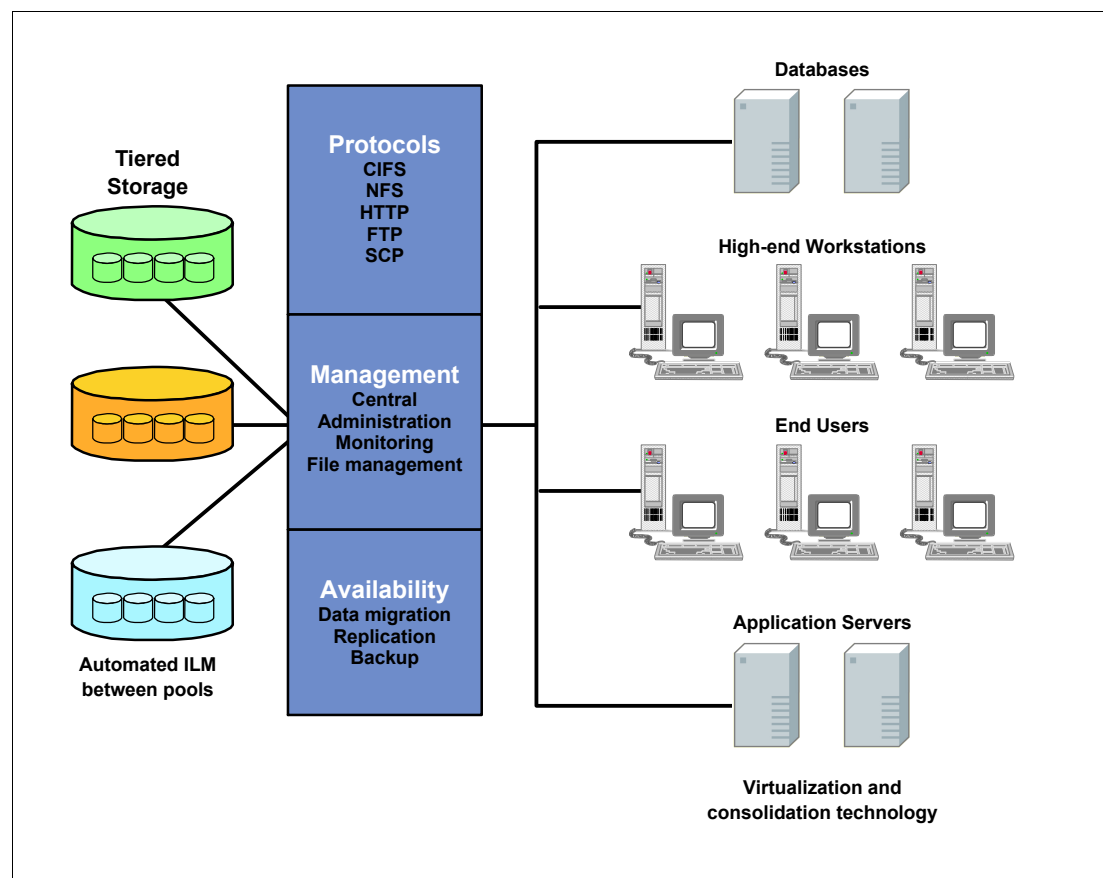


Figure 6-3 SONAS Tiered storage and ILM

The clustered architecture means all nodes share in performing the high-performance scan, and all nodes can participate in the transparent movement of ILM data or backup data to other storage pools or external removable media. ILM is performed at the individual file level, thus greatly optimizing storage usage while enabling maximum business flexibility. Physical data movement between pools, or migration and recall to and from tape, is transparent to the global namespace and to the logical user paths, directory locations, and applications.

6.3.3 Protecting data and replicating between sites

SONAS offers data protection through space-efficient snapshots that protect against accidental deletion or modification of files and that enable you to restore at the file level. The solution also provides for synchronous replication as well as bandwidth-optimized asynchronous replication to multiple sites.

6.3.4 Utilizing unique capabilities and modular hardware

In addition to extreme scale out capability and integrated ILM, SONAS enables your organization to utilize an array of uniquely scalable, globally clustered hardware capabilities. For instance, it has wide area network (WAN) capabilities that enable local caching of remote files for fast service.

The solution is self-tuning, with all data automatically striped across all disks in a logical storage pool for high performance. All nodes share equally in the reading and writing of files in parallel from all disks, thus providing intelligent load balancing, eliminating of hot spots, and ensuring very high performance. Storage capacity is only allocated when written, providing dynamic thin provisioning.

The solution is also self-healing, providing automatic failover, failback, and fully dynamic load balancing. The modular configuration of interface nodes and storage node building blocks can be upgraded or downgraded dynamically, giving you the ability to scale in real time.

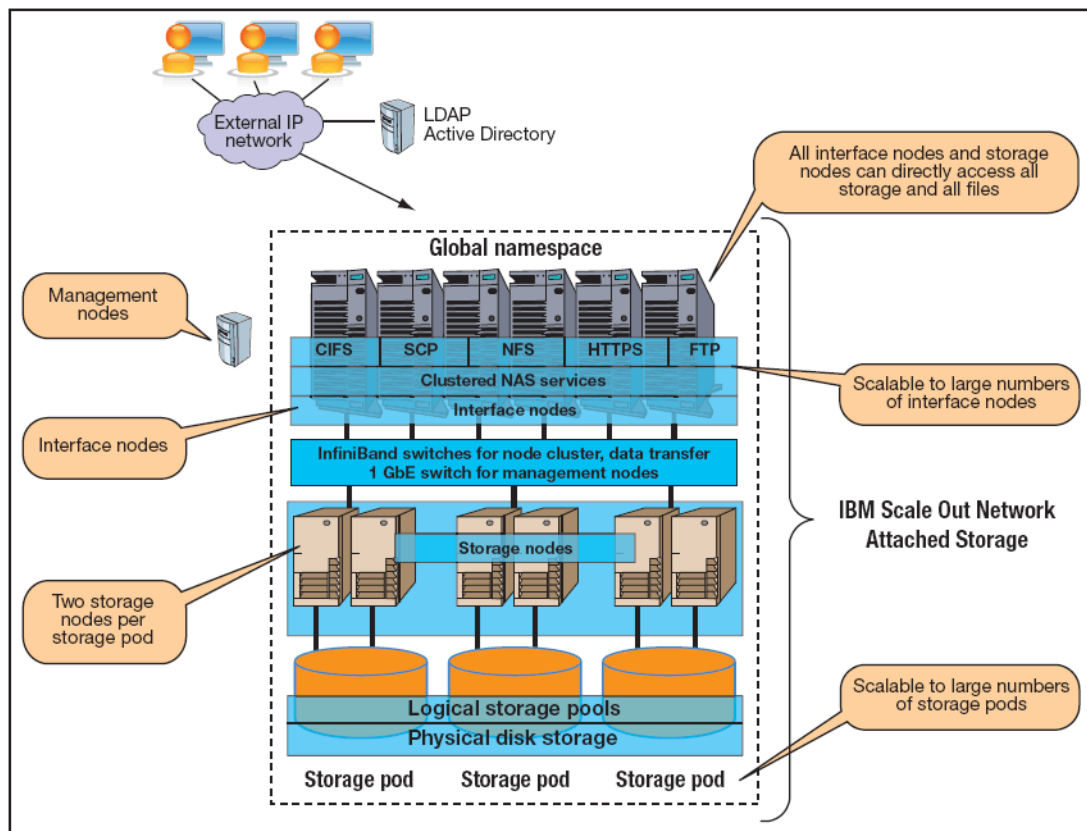


Figure 6-4 SONAS - Storage performance, scalability, and flexibility

Organizations can take advantage of numerous flexible, modular, hardware configuration possibilities. The initial versions of SONAS will support any combination of up to 30 interface nodes and 30 storage pods, flexibly configured to suit your requirements.

Storage pods range from 4U to 16U in size. Each pod can provide reliable storage from 27 to 480 terabytes of data. Storage pods and interface nodes can be combined into one globally clustered scale out NAS system, storing up to 14.4 petabytes of data. This provides extreme flexibility in configuring, adding and subtracting interface nodes and storage nodes in any way your workload requires.

For example, a system for archival might be configured with 2 interface nodes and 20 storage pods, whereas a system for email might be configured with 20 interface nodes and 2 storage pods. SONAS provides extreme high density storage, and exploits leading-edge technology such as InfiniBand.

6.3.5 Driving a dynamic infrastructure

SONAS supports the most dynamic infrastructures, including remote office and disaster recovery infrastructures. Organizations in every industry can gain significant competitive advantages by implementing dynamic infrastructures that enable them to improve service, reduce cost and manage risk. IBM Scale Out Network Attached Storage (SONAS) is designed to serve as a highly dynamic, flexible, elastic storage element of your infrastructure, driving greater levels of responsiveness through consolidation, resiliency, and scalability.

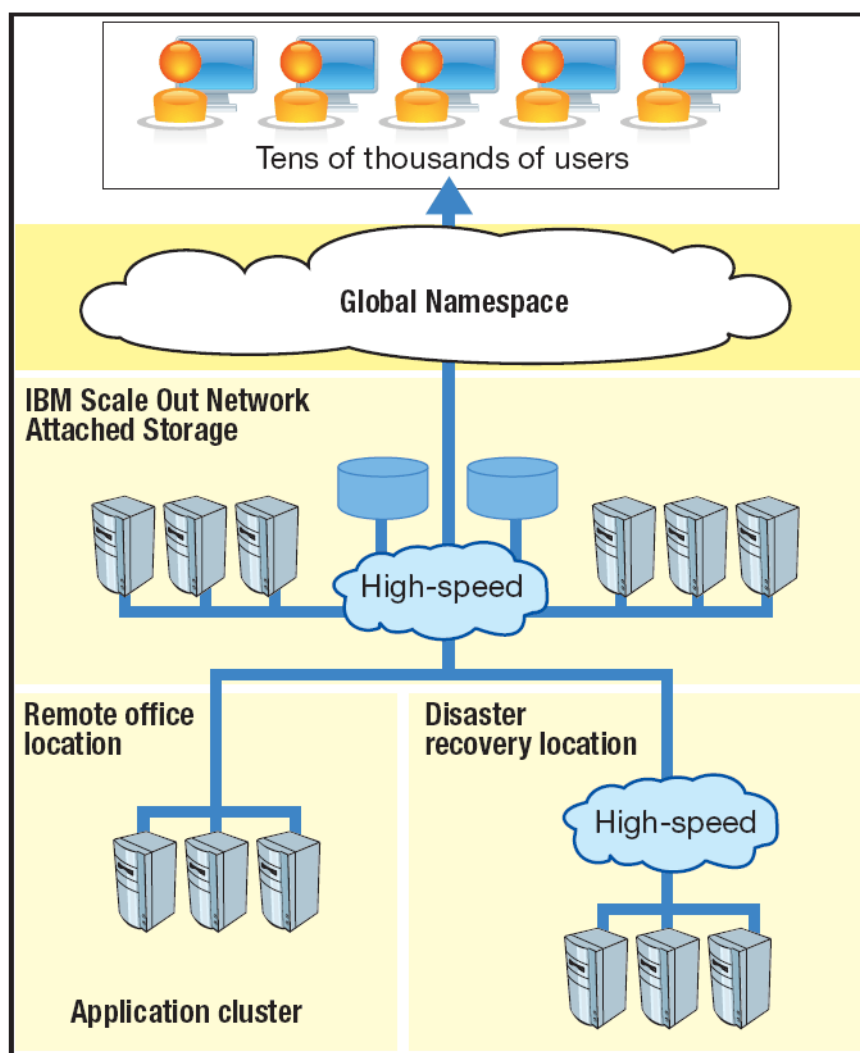


Figure 6-5 SONAS - Dynamic Infrastructure

6.4 Components within a SONAS solution

The main components of the IBM Scale Out Network Attached Storage (SONAS) system are the Interface node (2851-SI1), Management node (2851-SM1), Storage node (2851-SS1), Storage controller (2851-DR1), and Disk Storage expansion unit (2851-DE1). These components are assembled into one of three rack configurations with Ethernet and InfiniBand (IB) switches.

The minimum SONAS system configuration consists of:

- ▶ Two Interface nodes
- ▶ Two 36-port 4X DDR InfiniBand switches
- ▶ One Management node
- ▶ Two Storage nodes with one Storage controller

The maximum size of a SONAS system, in terms of Interface nodes and Storage pods, depends on whether the 36-port or 96-port InfiniBand switch is selected. Each of the following major components requires one IB port on each IB switch:

- ▶ Management node
- ▶ Interface node
- ▶ Storage node

Three InfiniBand ports of each IB switch are reserved for the following components:

- ▶ Two ports are reserved for future use.
- ▶ One port is for the required Management node.
- ▶ The remaining IB ports are available for Interface nodes and Storage nodes.

6.4.1 Interface node

The Interface node provides the connections to the customer's IP network for attaching to the SONAS system for network file serving capabilities (CIFS, NFS, FTP). A SONAS system contains a minimum of two Interface nodes and a maximum of 30.

The Interface node (2851-SI1) is a 2U server containing the following components:

- ▶ Two Intel® Xeon® E5530 (Nehalem EP) quad-core processors
- ▶ 32 GB of DDR3 memory standard, with feature to add an additional 32 GB of memory
- ▶ Four onboard 10/100/1000 Ethernet ports (two available for customer use)
- ▶ Two 300 GB 2.5-inch SFF 10K RPM SAS Slim-HS hard disk drives with mirroring between the two HDDs
- ▶ Four PCIe Gen 2.0 x8 adapter slots (two available for customer use)
- ▶ Integrated Baseboard Management controller (IBMC)
- ▶ Two redundant hot-swappable power supplies
- ▶ Six redundant hot-swappable cooling fans

The Interface node contains two redundant hot-swappable 300 GB 2.5-inch 10K RPM SAS HDDs with mirroring between them for high-availability. The HDDs contain the SONAS System Software product which hosts the operating system and all other software needed for an operational SONAS system.

Two of the PCIe adapter slots are already populated with two single-port 4X Double Data Rate (DDR) InfiniBand Host Channel Adapters (HCA). The two HCAs attach to two independent InfiniBand switches in the SONAS system and interconnect the Interface nodes to the Management node(s) and the Storage nodes.

Two of the PCIe adapter slots are available for customer use to add more adapters for host IP interface connectivity.

Each Interface node comes standard with two 10/100/1000 Ethernet ports available to connect to the customer's IP network. Additional Ethernet connectivity can be achieved by adding one of the following features:

- ▶ 1100 (Quad-port 1GbE NIC)
- ▶ 1101 (Dual-port 10 Gb Converged Network Adapter)

Each Interface node is connected to the two InfiniBand switches by one InfiniBand cable to each switch, a total of two cables. The InfiniBand Host Channel Adapters (HCA) in the Interface nodes have an X4 port. The 36-port InfiniBand switches have QSFP connectors and the 96-port IB switches have X4 connects.

For an Interface node (2851-SI1) in the base rack (2851-RXA) no InfiniBand cables need to be ordered. Copper InfiniBand cables are automatically provided for all Interface nodes in the base rack. The length of the copper InfiniBand cables provided is based on the position of the Interface node in the rack.

6.4.2 Management node

The Management node provides the user interface for configuring, administering, and monitoring the SONAS system. A single Management node is required. The Management node contains two redundant hot-swappable 300 GB 2.5-inch 10K RPM SAS HDDs with mirroring between them for high-availability. The hard disk drives contain the SONAS System Software product which hosts the operating system and all other software needed for an operational SONAS system.

The Management node (2851-SM1) is a 2U server containing the following components:

- ▶ Two Intel Xeon E5530 (Nehalem EP) quad-core processors
- ▶ 32 GB of DDR3 memory
- ▶ Four onboard 10/100/1000 Ethernet ports (two for connecting to customer network for management access)
- ▶ Two 300 GB 2.5-inch SFF 10K RPM SAS Slim-HS hard disk drives with mirroring between the two HDDs
- ▶ One non-mirrored 300 GB 2.5-inch SFF 10K RPM SAS Slim-HS hard disk drive for centralized log/trace file collection
- ▶ Four PCIe Gen 2.0 x8 adapter slots
- ▶ Integrated Baseboard Management Controller (IBMC)
- ▶ Two redundant hot-swappable power supplies
- ▶ Six redundant hot-swappable cooling fans

Two of the PCIe x8 adapter slots are already populated with two single-port 4X Double Data Rate (DDR) InfiniBand Host Channel Adapters (HCA). The two HCAs attach to two independent InfiniBand switches in the SONAS system and interconnect the Management node to the other components of the SONAS system.

The Management node comes with two InfiniBand copper cables to connect it to the two InfiniBand switches in the base rack. Both the primary Management node and the optional redundant Management node are assumed to be in the SONAS base rack with the two InfiniBand switches.

6.4.3 Storage node

The Storage node provides the InfiniBand connection to the InfiniBand cluster interconnect and direct Fibre Channel attachment to the SONAS Storage controller. Storage nodes must be configured in high-availability (HA) pairs. The two Storage nodes in the HA pair are connected to one or two SONAS Storage controllers.

The Storage node (2851-SS1) is a 2U server containing the following components:

- ▶ Two Intel Xeon E5530 (Nehalem EP) quad-core processors
- ▶ 8 GB of DDR3 memory
- ▶ Four onboard 10/100/1000 Ethernet ports
- ▶ Two 300 GB 2.5-inch SFF 10K RPM SAS Slim-HS hard disk drives with mirroring between the two HDDs
- ▶ Four PCIe Gen 2.0 x8 adapter slots
- ▶ Integrated Baseboard Management Controller (IBMC)
- ▶ Two redundant hot-swappable power supplies
- ▶ Six redundant hot-swappable cooling fans

Two of the onboard Ethernet ports connect the storage node to the internal private management network, and two for a NULL Ethernet connection to the DASD Disk Storage controller.

The Storage node contains two redundant hot-swappable 300 GB 2.5-inch 10K RPM SAS HDDs with mirroring between them for high-availability. The hard disk drives contain the SONAS System Software product which hosts the operating system and all other software needed for an operational SONAS system.

All of the PCIe x8 adapter slots in the storage node are already populated with adapters. Two of the PCIe adapter slots are populated with two single-port 4X DDR InfiniBand HCAs for attaching to the two InfiniBand switches in the SONAS system. The other two PCIe x8 adapter slots are populated with two dual-port 8 Gbps Fibre Channel Host Bus Adapters (HBAs) for attaching to the SONAS Storage controller.

The SONAS Storage controller (2851-DR1) is a high-performance 4U Storage controller with:

- ▶ Dual redundant hot-swappable RAID controllers
- ▶ Dual redundant hot-swappable power supply and cooling modules
- ▶ Support for 60 3.5-inch SAS/SATA HDDs

Each RAID controller contains:

- ▶ Cache of 4 GB
- ▶ Two 8 Gbps FC host ports
- ▶ One drive-side SAS expansion port

The Storage controller supports RAID 1, 5 and 6. One high-density disk Storage expansion unit can be attached to the Storage controller. The Storage controller comes with four FC optical cables for attaching the Storage controller to the two Storage nodes.

All sixty hard disk drives in the Storage controller must be the same, therefore a quantity of six of feature #1300, #1301, or #1310 must be ordered.

Table 6-1 shows the total raw storage capacity of the Storage controller when populated with the various hard disk drive features.

Table 6-1 Raw storage capacity of the storage controller

Quantity	Feature	Description	Total raw storage capacity
6	1300	10-pack of 7.2K RPM 1 TB SATA	60 TB
6	1301	10-pack of 7.2K RPM 2 TB SATA	120 TB
6	1310	10-pack of 15K RPM 450 GB SAS	27 TB

Features #9011, #9012, #9013, and #9014 are plant-only feature and are used to indicate into which Storage controller location (1-4) this particular Storage controller is to be placed during manufacturing. The rack must be populated from the bottom of the rack (EIA position 1) to the top of the rack (EIA position 42) for stability reasons. Therefore, the following rules apply to the placement of Storage controllers within the rack:

- ▶ You cannot have a Storage controller in location #4 (EIA positions 29-32) unless you have a Storage controller in location #3 (EIA positions 21-24).
- ▶ You cannot have a Storage controller in location #3 (EIA positions 21-24) unless you have a Storage controller in location #2 (EIA positions 9-12).
- ▶ You cannot have a Storage controller in location #2 (EIA positions 9-12) unless you have a Storage controller in location #1 (EIA positions 1-4).

Also, features #9013 and #9014 are not valid when plant-merging the storage controller into the Base Rack (2851-RXA), because the Base Rack only holds two Storage controllers.

The SONAS disk Storage expansion unit (2851-DE1) is a 4U high-density disk enclosure supporting 60 3.5 inch SATA or SAS hard disk drives. One disk Storage expansion unit might be attached to a single SONAS Storage controller. Each disk Storage expansion unit contains the following features:

- ▶ Two SAS switch modules which provide SAS connections to the storage controller
- ▶ Two redundant hot-swappable power supply and cooling units
- ▶ Support for 60 3.5-inch SAS or SATA HDDs

The Disk Storage expansion unit must be physically located in the same rack and in adjacent EIA positions as the SONAS Storage controller to which it is attaching. All sixty hard disk drives in the disk Storage expansion unit must be the same, therefore a quantity of six of feature #1300, #1301, or #1310 must be ordered.

Table 6-2 shows the total raw storage capacity of the disk Storage expansion unit when populated with the various types of hard disk drive features.

Table 6-2 Raw storage capacity of the disk Storage expansion unit

Quantity	Feature	Description	Total raw storage capacity
6	1300	10-pack of 7.2K RPM 1 TB SATA	60 TB
6	1301	10-pack of 7.2K RPM 2 TB SATA	120 TB
6	1310	10-pack of 15K RPM 450 GB SAS	27 TB

Features #9021, #9022, #9023, and #9024 are plant-only features and are used to indicate into which disk Storage expansion unit location (1-4) this particular disk Storage expansion unit is to be placed during manufacturing. A disk Storage expansion unit must be attached to a Storage controller. The Storage controller attached to a disk Storage expansion unit must be in the EIA position immediately adjacent and below the disk Storage expansion unit.

Therefore, the following rules apply to the placement of Storage controllers within the rack:

- ▶ You cannot have a disk Storage expansion unit in disk Storage expansion unit location #1 (EIA positions 5-8) unless you have a Storage controller in Storage controller location #1 (EIA positions 1-4).
- ▶ You cannot have a disk Storage expansion unit in disk Storage expansion unit location #2 (EIA positions 13-16) unless you have a Storage controller in Storage controller location #2 (EIA positions 9-12)
- ▶ You cannot have a disk Storage expansion unit in disk Storage expansion unit location #3 (EIA positions 25-28) unless you have a Storage controller in storage controller location #3 (EIA positions 21-24)
- ▶ You cannot have a disk Storage expansion unit in disk Storage expansion unit location #4 (EIA positions 33-36) unless you have a Storage controller in Storage controller location #4 (EIA positions 29-32)

Additionally, features #9023 and #9024 are not valid when plant-merging the Storage controller into the Base Rack (2851-RXA), because the Base Rack only holds two disk Storage expansion units.

6.4.4 Infiniband switches

The major components of a SONAS system (Interface node, Storage node, and Management node) are interconnected by a high-performance low-latency InfiniBand 4X Double Data Rate (DDR) fabric. Two redundant InfiniBand switches are incorporated inside each SONAS system.

For small and medium configuration, a 1U 36-port 4X DDR InfiniBand switch is available. For larger configurations a 7U 96-port 4X DDR InfiniBand switch is available. Two identical IB switches must be ordered for a SONAS system, either two 36-port IB switches or two 96-port IB switches. These IB switches are described in the following sections.

36-port IB switch

The SONAS 36-port IB switch (2851-I36) is a 4X DDR InfiniBand switch providing 36 QSFP ports, with each port operating at 20 Gbps. This IB switch provides a maximum backplane bandwidth of 1.44 Tbps and contains an embedded IB fabric manager. The switch provides two redundant hot-swap power supplies.

96-port IB switch

The 96-port IB switch (2851-I96) provides a 96-port 4X DDR InfiniBand switch providing up to 96 4X DDR X4 switch ports. The 96-port IB switch is intended for large SONAS system configurations. The IB switch provides a maximum backplane bandwidth of 3.84 Tbps.

The 96-port IB switch comes standard with the following features:

- ▶ Two 96-port InfiniBand Switch Fabric boards
- ▶ One 24-port 4X DDR Line Board
- ▶ One Hi-Memory Management board containing an embedded IB fabric manager
- ▶ Two Power Supply Units (PSU)
- ▶ All fan assemblies

The 96-port switch comes standard with one 24-port 4X DDR line board providing 24 4X DDR (20 Gbps) IB ports. Up to three additional SLB-2024 24-port 4X DDR line boards can be added for a total of 96 ports.

The 96-port IB switch comes standard with two SLB-2004 Switch Fabric boards. Up to two additional SLB-2004 Switch Fabric boards can be added to provide additional backplane bandwidth.

The 96-port IB switch comes standard with two power supplies. Up to two additional PSUs can be added for redundancy. The two standard power supplies are capable of powering a fully configured 96-port IB switch with the following boards:

- ▶ Four SFV-2004 Switch Fabric boards
- ▶ Four SLB-2024 24-port 4X DDR line boards
- ▶ Two SMB-HM Hi-Memory Management boards

The following options can be added to the 96-port InfiniBand switch:

- ▶ FC 1500 96-port InfiniBand Switch Fabric board
- ▶ FC 1501 24-port 4X DDR InfiniBand Line Board
- ▶ FC 1502 Hi-Memory Management Board
- ▶ FC 1503 Power Supply

6.5 Summary of features and benefits

Here is a summary of the features and benefits of SONAS.

Massive scalability:

- ▶ Supports multiple petabytes of storage for organizations that need billions of files in a single file system.
- ▶ Supports the ability to rapidly add massive amounts of storage capacity. Supports up to 256 file systems per system

Flexibility:

- ▶ Access your data in a single global namespace allowing all users a single, logical view of files through a single drive letter such as a Z drive
- ▶ Tiered storage for highly available primary data and secondary storage for less accessed data.
- ▶ Industry standard protocols: CIFS, NFS, FTP

Consolidation:

- ▶ Your organization can consolidate and manage data to avoid problems associated with administering an array of disparate storage systems.
- ▶ It is highly scalable, helps conserve floor space, and can help to reduce your capital expenditure and enhance your operational efficiency.
- ▶ Its advanced architecture virtualizes and consolidates your file space into a single, enterprise-wide file system, which can translate into reduced total cost of ownership.

GPFS:

- ▶ GPFS integrates into organizational environments by bringing together mixed server and storage components to provide a common view to enterprise file data.
- ▶ It also provides online storage management, scalable access and information lifecycle tools to manage large volumes of data.

Data protection:

- ▶ Snapshots, up to 256 per file system
- ▶ Supports RAID 5, 6

Cloud storage:

- ▶ In a cloud environment, applications and services are not tethered to specific hardware components. Instead, processing is handled across a distributed, globally accessible network of resources, which are dispensed on demand.
- ▶ Self-managing, autonomic system that enables capacity, provisioning and other IT service management decisions to be made dynamically, without human intervention or increased administrative costs.
- ▶ Seamless elasticity to scale computing resources up or down, as required, to fulfil changing organizational needs without service interruption.
- ▶ Highly resilient and secure applications, and an underlying infrastructure capable of meeting expected levels of availability, reliability and integrity.
- ▶ A highly standardized environment that facilitates simultaneous service deployment and upgrades for all users, no matter where they reside.
- ▶ Lowers the cost of service access by economies of scale.
- ▶ Clouds can provide rapid access to computing capacity at a lower cost of ownership enabling companies to perform operations that might have previously been unaffordable or impractical.

Interface:

- ▶ Supports both Command line interface (CLI) and browser-based graphical user interface (GUI).

Advanced features:

- ▶ Integrated Tivoli® Storage Manager backup/archive client.

Drive support:

- ▶ Supports both SAS and SATA drives.

6.6 More information

For further information about SONAS, see the following websites:

<http://www-03.ibm.com/systems/storage/network/sonas/index.html>

<http://publib.boulder.ibm.com/infocenter/sonasic/sonas1ic/index.jsp>



IBM Information Archive

In this chapter we discuss the IBM Information Archive, which represents the next generation information retention solution. It is designed as an archiving repository for all types of content (structured or unstructured) to help organizations of any size address complete information retention needs; business, legal, or regulatory.

To help you meet the growing challenge of efficiently managing and protecting business information, IBM offers IBM Information Archive, an integrated storage repository solution for archiving multiple types of data from multiple sources. Designed to help you store, retrieve, manage, share, and secure regulated and non-regulated data, the Information Archive is designed to be highly secure and offer extraordinary cost efficiency.

As you address your information retention needs: whether keeping valuable content for long periods of time, meeting industry retention regulations or addressing corporate governance, you need an archiving solution that is secure, scalable, but also cost-effective.

For complete information about the IBM Information Archive, see the product page at the following website:

<http://www.ibm.com/systems/storage/disk/archive/>

Note: Although IBM Information Archive is sometimes shortened to “IA”, in this chapter we are using the full name to avoid confusion with other possible products thusly abbreviated.

7.1 Data retention and compliance

There is a rapidly growing class of data that is best described by the way in which it is managed rather than the arrangement of its bits. The most important attribute of this kind of data is its retention period, hence it is called retention managed data, and it is typically kept in an archive or a repository. In the past it has been variously known as archive data, fixed content data, reference data, unstructured data, or other terms implying its read-only nature. It is often measured in terabytes and is kept for long periods of time, sometimes forever.

Not only are there numerous state and governmental regulations that must be met for data storage, but there are also industry-specific and company-specific ones as well. And of course these regulations are constantly being updated and amended. Organizations need to develop a strategy that must ensure that the correct records are kept for the correct period of time. Needless to say these records must be readily accessible when they need to be retrieved at the request of regulators or auditors. It is easy to envisage the exponential growth in data storage that will result from these regulations and the concomitant requirement for a means of managing this data. Overall, the management and control of retention managed data is a significant challenge for the IT industry when taking into account factors such as cost, latency, bandwidth, integration, security, and privacy.

7.1.1 Characteristics of retention managed data

When considering the safekeeping of retention managed data, companies also need to consider storage and data characteristics that differentiate it from transactional data.

Storage characteristics of retention managed data are:

- ▶ Variable data retention period: This is usually a minimum of a few months, up to forever.
- ▶ Variable data volume: Many customers start with 5 TB to 10 TB of storage in an enterprise. It also usually consists of a large number of small files.
- ▶ Data access frequency: Write Once Read Rarely or even Write Once Read Never.
- ▶ Data read/write performance: Write handles volume; Read varies by industry and application.
- ▶ Data protection: Requirement for non-erasability, non-rewritability, and deletion when the retention policy expires

Data characteristics of retention managed data are:

- ▶ Data lifecycle: Typically, the data is frequently accessed near its creation, then the access frequency diminishes exponentially, near zero. Certain industries have peaks that require access, such as check images in tax season.
- ▶ Data rendering: Ability to view or use data in a very old data store. There is a risk that the data format used to store the data might no longer be supported in 25 years.
- ▶ Data mining: With all this data being saved, there is intrinsic value in the content of the archive that can be exploited using data mining. But this implies indexing and categorization of the data when it is initially written.

For additional information about managing data, see the following Redbooks publications:

- ▶ *ILM Library: Information Lifecycle Management Best Practices Guide*, SG24-7251
- ▶ *ILM Library: Techniques with Tivoli Storage and IBM TotalStorage Products*, SG24-7030
- ▶ *IBM Enterprise Content Management and DR550 for E-mail Archiving and Records Management Overview*, REDP-4284

7.1.2 IBM strategy and positioning

Regulations and other business imperatives, as briefly outlined before, stress the need for an Information Lifecycle Management process and tools to be in place. The unique experiences that IBM has had with the broad range of ILM technologies, and its broad portfolio of offerings and solutions, can help businesses address this particular need and provide them with the best solutions to manage their information throughout its lifecycle.

IBM provides a comprehensive, open, set of solutions to help. IBM has products that provide content management, data retention management, and sophisticated storage management, along with the storage systems to house the data. To specifically help companies with their risk and compliance efforts, the IBM Risk and Compliance framework is another tool designed to illustrate the infrastructure capabilities needed to help address the myriad of compliance requirements. Using the framework, organizations can standardize on the use of common technologies to design and deploy a compliance architecture that can help them deal more effectively with compliance initiatives

7.2 IBM Information Archive offering

IBM Information Archive (as seen in Figure 7-1), is a consolidated hardware and software offering built with the latest IBM technologies enabling the archive of data with scalability to multiple petabytes.

IBM Information Archive offers' policy-based, data retention capabilities designed to support non-erasable, nonrewritable (NENR)/write Once Read Many (WORM) data storage protection and help address the needs of regulated and non-regulated industries with long-term data retention and protection requirements.

In addition to data encryption capabilities, each collection within the IBM Information maintains a set of tamper-proof audit logs, which provide an immutable and retention-protected provenance record for documents in the collection. Audit logs track document ownership and system lifecycle events including document creation and deletion, changes to retention policies, and system software upgrades.



Figure 7-1 IBM Information Archive

IBM Information Archive is the next generation information retention solution that brings together off-the-shelf IBM hardware and software products along with customized applications for ease of management and use. It is a universal, scalable, and secure storage repository for structured and unstructured information as a compliant or non-compliant integrated archive appliance.

IBM Information Archive is an integrated IBM software and hardware solution that retains and protects data by enforcing the industry's most stringent data retention laws and addresses the complete information retention needs of mid-size and enterprise clients.

7.2.1 IBM Information Archive

IBM Information Archive is:

- ▶ The next generation information retention solution
- ▶ A universal storage repository for all types of content of structured (database) and unstructured (files, email, images, and docs)
- ▶ A robust, scalable, secure information retention, fully integrated hardware, software, and services solution
- ▶ A solution that addresses the complete business, legal, or regulatory information retention needs of mid-size and enterprise clients

IBM Information Archive is easy to implement, configure, and manage, allowing you to start archiving the next day. It looks like a file server; just simply drag and drop files to IBM Information Archive. Built to industry standards, IBM Information Archive works with any archiving application that supports standard NAS interface. There is no additional integration work required.

IBM Information Archive allows you to optimize your storage infrastructure immediately. You can do this by offloading your production files to IBM Information Archive right away and gain storage efficiencies, utilizing tape, and automatically migrating to tape for cost and energy savings, and storing a single instance of the same files creating storage space savings.

IBM Information Archive allows you to quickly archive, locate and retrieve information and offers index / search on data and metadata. IBM Information Archive automates “file” lifecycle (from ingestion to deletion) by policy driven administration and enforces retention policies set forth by the application, or allows administrators set retention policies (time or event based).

Retain your data with confidence and protect, secure, and tamper-proof your information with customizable retention policies to meet your organization's needs.

IBM Information Archive scales up to 304 TB terabytes (raw) with disk and petabytes with attached external tape allowing you to consolidate your data and be ready for growth.

Because the IBM Information Archive solution allows use of multiple collections per IBM Information Archive, you can handle multiple archive applications at once under one management point and with it being a cloud ready smart business system, you have a unified and simplified smart business system infrastructure which will adapt as business needs evolve.

As you can see in Figure 7-2, IBM Information Archive allows input by NFS and supported applications that can write to the Tivoli Storage Manager API client. After the data is stored in the appropriate secure collection on internal disk, you have the capability to send to external tape or other supported device allowing for greater scalability.

IBM Information Archive also provides a single, task oriented GUI to manage the entire solution. The IBM Information Archive Administrative Console can be used as follows:

Archive Management:

- ▶ Document collection configuration (creation, modification, deletion)
- ▶ Reporting (documents and capacity)

System Management:

- ▶ System configuration
- ▶ System monitoring and alerting (SNMP and email)

Service:

- ▶ Retrieve logs from all servers and SW components
- ▶ Launch Director and RSM GUIs

Auditing:

- ▶ Download audit logs per document collection
- ▶ Configure access to audit logs for feeding into TCIM, and so on

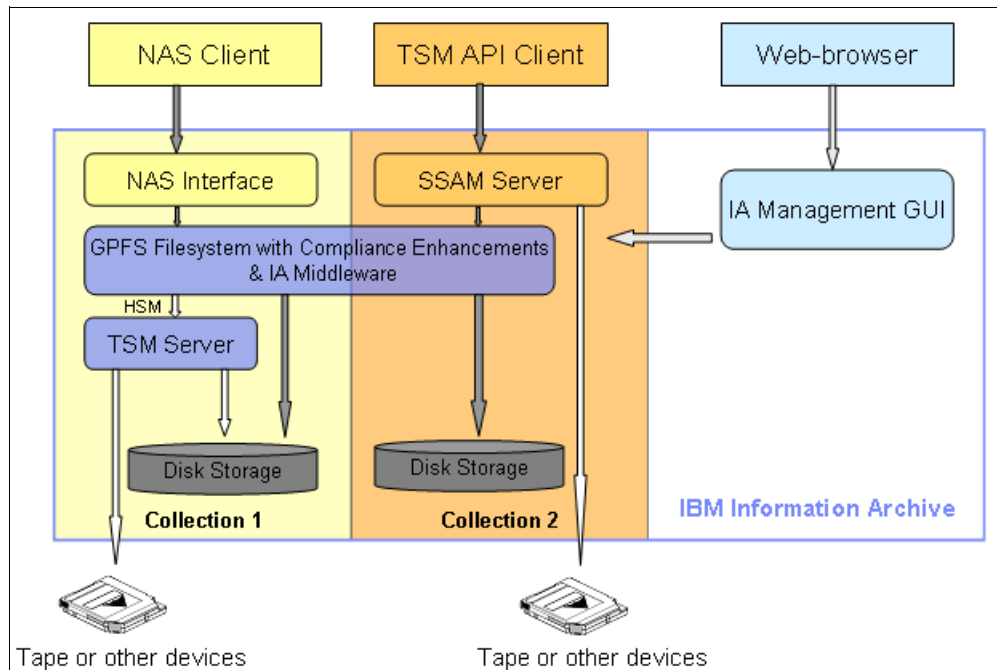


Figure 7-2 Information Archive High Level Architecture

7.2.2 IBM Information Archive components

IBM Information Archive includes the following hardware and software components:

Hardware

Integrates IBM hardware:

- ▶ Numerous System x servers
- ▶ SATA Storage
- ▶ IBM Tape Solutions

Software

Integrates IBM software:

- ▶ System Storage Archive Manager
- ▶ Tivoli Storage Manager - HSM
- ▶ Scale Out File System - GPFS
- ▶ Management Console SW
- ▶ Content Manager
- ▶ DB2 + Net Search Extender
- ▶ CommonStore
- ▶ eDiscovery Manager

Specifications

Here we summarize all the specifications for the IBM Information Archive:

General Ingest - General Parallel File System (GPFS) Node:

- ▶ 2U System x server
- ▶ Two quad-core Processors
- ▶ 2U System x server
- ▶ 24 GB memory

- ▶ 2 x Dual Ported Fibre Channel (FC) Host Bus Adapter (HBAs)
- ▶ 4 1 GB copper Ethernet ports
- ▶ Redundant 146 GB 15K rpm Serial Attached SCSI (SAS)
- ▶ Optional: Dual ported 1 Gb fibre Ethernet card (for Fibre customer infrastructure)
- ▶ Up to 3 nodes per cluster

Management node:

- ▶ 1U System x server
- ▶ 2 x Dual Ported Fibre Channel (FC) Host Bus Adapter (HBAs)
- ▶ One quad-core processor
- ▶ 4 GB memory
- ▶ Ethernet daughter card (4 ports total on planar and daughter)
- ▶ Optional: Dual ported 1 Gb fiber Ethernet card (for fibre customer infrastructure)

Private FC SAN Data Network:

- ▶ All disk controllers accessible to all Ingest/GPFS nodes
- ▶ 4 Gb FC Storage Area Network (SAN) by 8 Gb capable FC switches
- ▶ Dual-redundant 24-port FC switches
- ▶ Private dual/redundant Ethernet network for management traffic (and later for indexing I/O traffic) with dual 24 port Ethernet switches

Redundant Array of Independent Disk (RAID) Controllers:

- ▶ System Storage Mid Range Disk Controller with 16 HDD/3u disk controllers with hardware remote mirroring support
- ▶ Dual-redundant active/active RAID controllers in each Disk Controller
- ▶ RAID 6 data protection
- ▶ 2 GB cache per Disk Controller
- ▶ 2 x 4 Gb FC ports per controller (for server access). 2x 4 Gb FC ports per controller for remote mirroring
- ▶ Serial Advanced Technology Attachment (SATA) 1 TB HDDs
- ▶ 112 TB of raw capacity in rack 1, 2x96 TB raw in rack 2

Disk Expansion drawers:

- ▶ System Storage Expansion Unit with up to 16 HDD/3u expansion, up to six per Disk Controller
- ▶ SATA 1 TB HDDs

Service console:

- ▶ Integrated service console with Keyboard- Video-Mouse (KVM)
- ▶ All inside locked racks
- ▶ Call Home, Remote Support by Remote Support Manager (RSM) server running on System x (1u) server with 1 processor and 2 GB of RAM

Rack:

- ▶ Standard 36u 19" rack with lock
- ▶ One IBM Information Archive (main) rack and up to one (storage expansion) rack

Tape Systems:

- ▶ Tape attachment is fully supported

IBM Information Archive

The IBM Information Archive has the ability to archive information from multiple sources as seen in Figure 7-3.

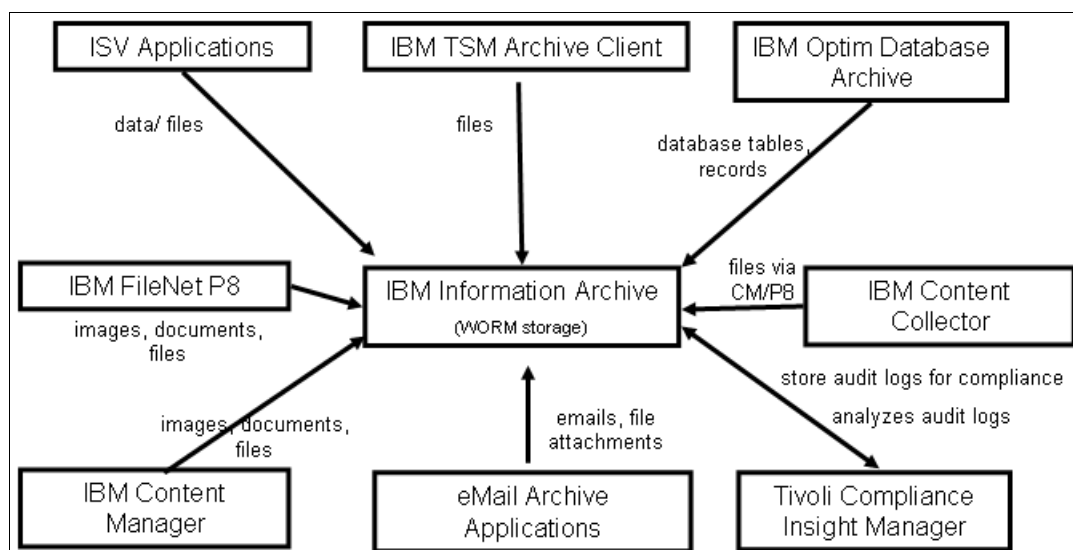


Figure 7-3 IBM Information Archive Logical Input View

Introducing collections

A single Information Archive can be partitioned into three archive collections, and each collection is assigned to a storage node. The IBM Information Archive can have up to three storage nodes thus, up to three archive collections and each collection can be customized to support various protection levels. Each collection is accessible by Network for File System (NFS) or System Storage Archive Manager (SSAM) protocols. IBM Information Archive supports multiple ingest and input models including custom applications, as shown in Figure 7-4.

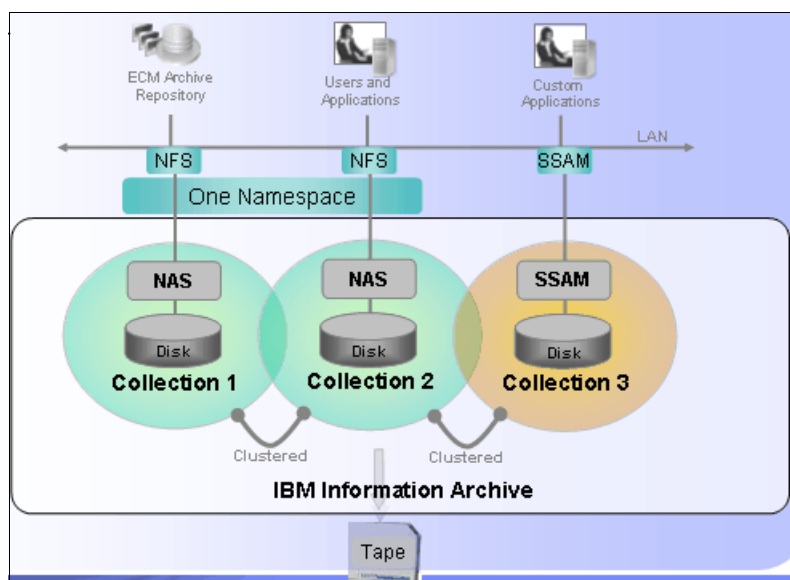


Figure 7-4 IBM Information Archive Advanced 3x3 Collection Architecture

Features and benefits

In order to meet the widest range of retention demands, the IBM Information Archive allows the IT Administrator to customize retention policies based on unique and dynamically changing business requirements.

It provides three information protection levels with up to three archive collections within a single system. This 3x3 combination provides you the greatest flexibility of protection levels to help your organization meet unique data retention requirements. With various collections, you have the flexibility to assign customized levels of protection, and retention policies for various kinds of data based on business rules.

These customizable retention policies include time-based and event-based information retention policies to protect information from intentional or accidental deletions or modifications. They also include a retention hold and deletion hold feature to set indefinite period of retention on a file during circumstances such as legal discovery.

There are many additional features and benefits of the IBM Information Archive:

- ▶ Designed to offer quick time-to-value; can install and configure quickly and easily
- ▶ Ease of use:
 - Ability to manage the archive from a single point
 - Ability to manage data with fewer resources
 - Designed to help reduce cost
- ▶ Information Protection Levels and Retention policies that address unique data retention needs with flexible information protection levels:
 - Ability to store data in up to three collections per system and assign various levels of protection, and retention policies for various kinds of data
 - Allows compliant and non-compliant storage within the same appliance foot-print
 - Offers Patent-Pending Enhanced Tamper protection
 - Accepts retention policies from applications or creates them
 - Accepts time-based and event-based information retention policies to protect information from intentional or accidental deletions or modifications
 - Accepts retention hold and deletion hold policies to set an indefinite period of retention for a file, such as during legal discovery
- ▶ Grows with your needs; can add storage to increase capacity and nodes for performance
- ▶ Industry standard interfaces:
 - Eliminates the need for customized application programming interface (API) or add on feature requirements
 - Helps manage operational costs
- ▶ Data-deduplication and compression:
 - Helps lower total cost of ownership (TCO) by providing embedded deduplication and compression capabilities
 - Designed to optimize storage capacity and to improve productivity
- ▶ Tiered storage management:
 - Designed to allow a cost-effective mix on nearline and offline storage (disk and/or tape)
 - Contains costs while utilizing storage technology for optimum usage
- ▶ Universal storage repository for all types of content, structured and unstructured, compliant or non compliant

- ▶ Numerous other capabilities:
 - Provides up to three information protection levels offering maximum flexibility
 - Stores information by multiple access methods
 - Scales up to 304 TB (raw capacity)
 - Maintains data integrity until deletion is permitted by retention policy
 - Enhanced security and protection with data encryption option
 - Helps optimize storage consumption with data deduplication and compression features
 - Lowers Total Cost of Ownership (TCO) by allowing use of mixed media (disk and tape)
 - Increases data security through patent-pending Enhanced Tamper Protection feature

Additional information

For additional information about the following topics, see these websites:

- ▶ IBM Information Archive:
<http://www.ibm.com/systems/storage/disk/archive>
- ▶ *IBM Information Archive V1.1 - Introduction and Planning Guide*, SC27-2324-01:
<http://www.ibm.com/support/docview.wss?rs=1328&uid=ssg1S7003077>
- ▶ *IBM Information Archive, V1.1 - Installing and Configuring IBM Information Archive*, GC27-2326-01:
<http://www.ibm.com/support/docview.wss?rs=1328&uid=ssg1S7003104>
- ▶ *IBM Information Archive V1.1 - User's Guide*, SC27-2325-0:
<http://www.ibm.com/support/docview.wss?rs=1328&uid=ssg1S7003103>
- ▶ *IBM Information Archive V1.1 - Service Guide*, SC27-2327-01:
<http://www.ibm.com/support/docview.wss?rs=1328&uid=ssg1S7003105>



Part 2

Tape systems

In Part 2 we review IBM TotalStorage and System Storage Tape Drives, IBM Tape Automation products, and IBM Tape Virtualization products.



IBM TotalStorage and System Storage Tape Drives

In this chapter we provide information about IBM tape drives starting from the entry level TS2230 Tape Drive Express model all the way to the enterprise capable IBM System Storage TS1130 Tape Drive. We briefly discuss Linear Tape-Open (LTO) technology and related Ultrium specifications, as well as TS1120/TS1130 tape technologies.

We also briefly consider IBM midrange and enterprise storage media that is used in LTO type and TS1120/1130 type drives.

8.1 LTO technology

Linear Tape-Open (LTO) is an open format technology. This means that users can have multiple sources of products and media. The LTO technology establishes a new open format specification for high-capacity, high-performance storage products and addresses a growing customer need for improved data interchange across platforms.

LTO technology was developed jointly by IBM, Hewlett Packard (HP) and Seagate in 1997 to provide a clear and viable choice in an increasing complex array of tape storage options. The consortium created two specifications, Accelis and Ultrium.

The current technology provider companies are IBM, HP, and Certance LLC (owned by Quantum). The Accelis technology has not been pursued by manufacturers as it is apparent that Ultrium meets the market needs.

Here we describe technical standards defined for LTO technology:

- ▶ Cartridge dimension (approximate) is 4.1 x 4.0 x 0.8 inches (105 x 102 x 21 mm).
- ▶ The single-hub design allows for the cartridge to be optimally packed with media. High capacity is further enhanced by the use of an LTO technology data compression algorithm with two control modes to maximize compression efficiency.
- ▶ There are four currently available Ultrium formats:
 - Ultrium format generation 1: 200 GB capacity (2:1 compression) per cartridge. Ultrium 1 drives are no longer being actively sold, although the media can still be read by Ultrium 3 drives, and read/written by Ultrium 2 drives.
 - Ultrium format generation 2: 400 GB capacity (2:1 compression) per cartridge
 - Ultrium format generation 3: 800 GB capacity (2:1 compression) per cartridge with 80 MBps
 - Ultrium format generation 4: 1.6 TB capacity (2:1 compression) per cartridge with 120 MBps and AME Encryption (SAS and FC4)

LTO technology delivers the following benefits:

- ▶ Easily integrates into current operating environments
- ▶ Roadmap protects investment today and in the future as each new generation reads previous generations formats
- ▶ Migration paths focused on increasing maximum transfer rates to reduce backup window. In addition, capacity will double in each successive generation
- ▶ Simplified product planning means faster cycle time for new features
- ▶ Compliance testing ensures that LTO Ultrium drives and media cartridges conform to the specification to deliver data interchange among multiple vendors' products
- ▶ Instant access to data; this is made possible by LTO CM (Cartridge Memory), which is a passive, contactless silicon storage device that is physically a part of the cartridge
- ▶ Dedicated Dual Servo; the servo bands are pre-written on the tape during the tape cartridge manufacture process. If one servo element becomes defective, the head will continue to track as a result of the second "redundant" servo system

For details about the LTO or Ultrium, see:

<http://www.lto-technology.com>
<http://www.ultrium.com>

8.1.1 Interleaved recording

The LTO drive uses an interleaved, serpentine, longitudinal recording format. The first set of 8 or 16 data tracks is written from near the physical beginning of the tape to near the physical end of the tape. The head then repositions to the next set of tracks for the return. This process continues until all tracks are written and the tape is full.

The format of the recording of the data and servo tracks is defined as part of the LTO specification in order to meet the requirement for interchange between various manufacturers' implementations.

8.1.2 Servo tracks

Servo tracks (also called servo bands) enable accurate positioning of the tape drive head over the data track, ensuring that the head does not stray onto an adjacent track. They are necessary to support high-data densities on the tape where the tracks are very close together. The servo bands are written at time of cartridge manufacture, before the cartridge is usable for data storage and retrieval. If the servo bands are erased, the tape becomes unusable.

Servo tracks are like lane markings on a multi-lane highway. Imagine the difficulty of driving on the highway without any lane markings. Lane markings help by positioning you on the lane, just as servo tracks support the drive recording head to position on the data tracks (Figure 8-1).

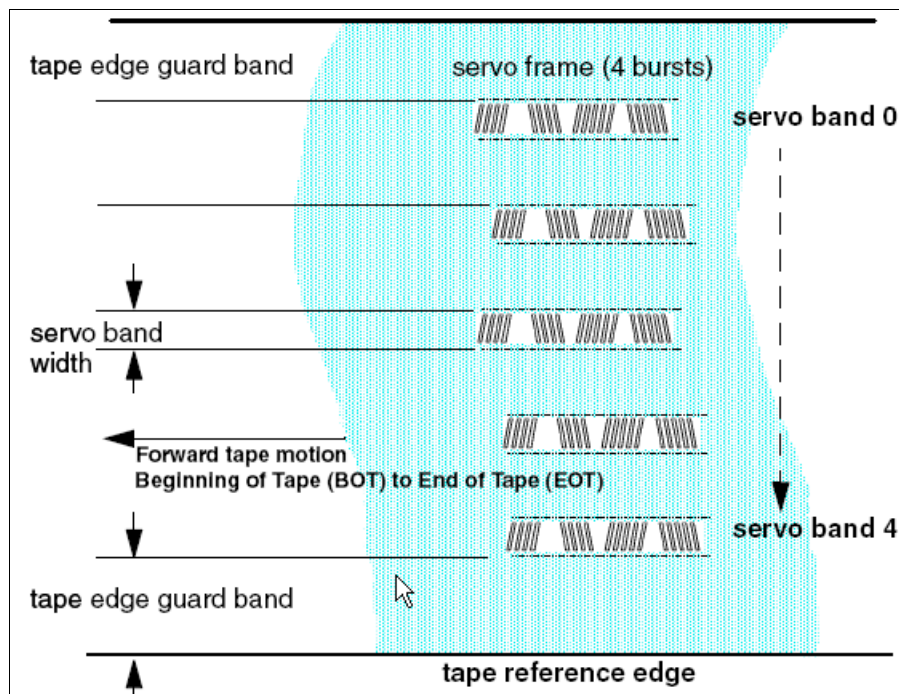


Figure 8-1 Servo band position and nomenclature

8.1.3 Track following

Each pair of servo bursts is at an angle to each other, and the servo heads move such that they keep a constant value for the distance between the bursts. In this way the servo is able to follow a straight line within the servo band; any small deviation away from the correct path causes a variation (plus or minus) in the gap between the bursts (Figure 8-2). Provided that the servo head element follows a straight line along the servo band, then the distance “x” shown in the figure remains constant. IBM LTO drives use two servo bands simultaneously during write operations to provide two sources of servo information, and therefore increased accuracy.

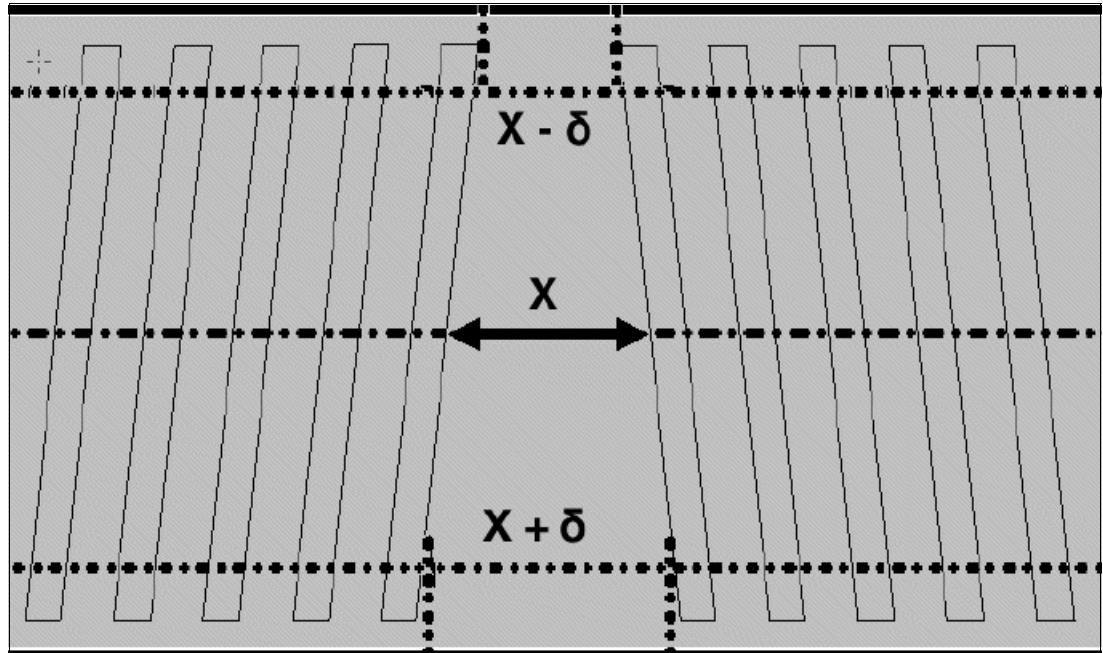


Figure 8-2 Magnified servo band showing a pair of servo bursts

Longitudinal positioning

The LTO servo band is designed not only for track following, but also for recording the longitudinal position (LPOS). The absolute location down the length of the tape and the manufacturer data are recorded in LPOS *words* approximately every quarter-inch (.7 cm) along the tape. The LPOS word consists of symbols constructed from bit sequences (ones and zeros); these bits are encoded within the servo frames. Each servo frame encodes one bit using the first pair of servo bursts. When servo stripes 2 and 3 (out of the five) are shifted inward (Figure 8-3), this encodes a zero; when servo stripes 2 and 3 are shifted outward, this encodes a one.

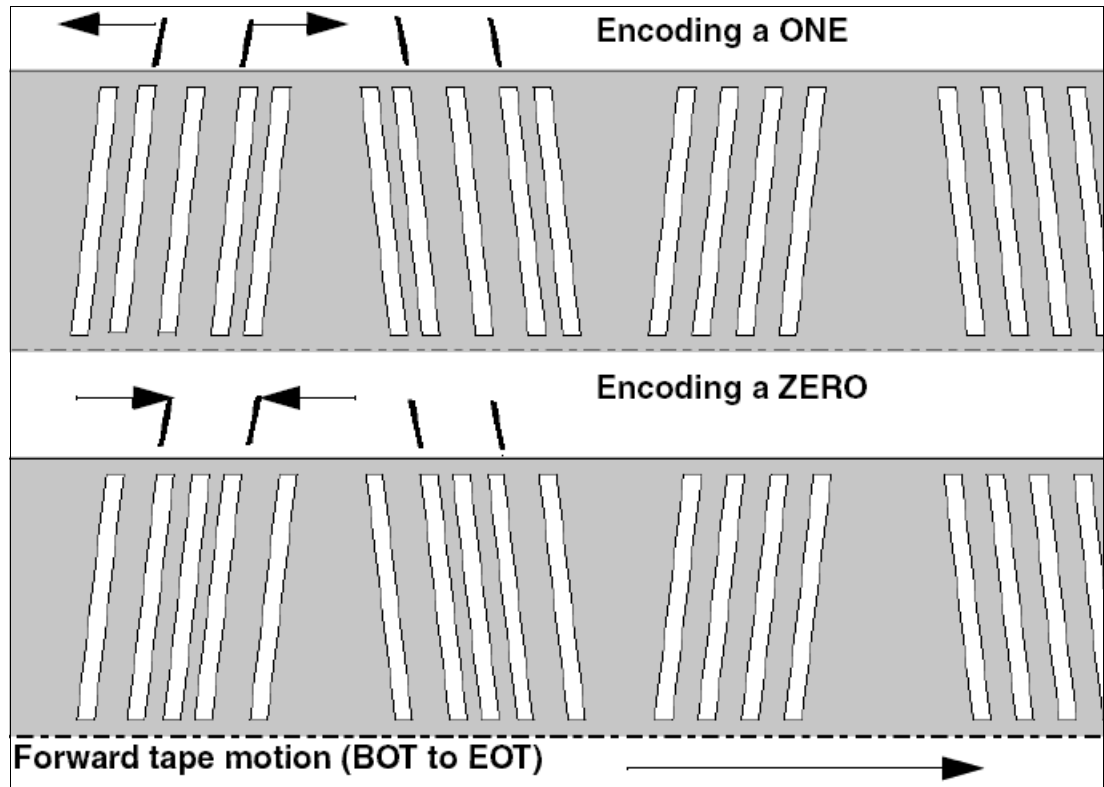


Figure 8-3 Encoding bits using the servo stripes within the servo bursts

The LPOS word contains 36 bits and therefore has a length of 36 servo frames. Each of the 5 servo bands on the tape can be uniquely identified by the relative positions of the frames down the tape, in adjacent servo bands. The offset of the frames between servo band n and servo band $n+1$ are specific to each servo band (0 and 1, 1 and 2, 2 and 3, or 3 and 4). Therefore, the drive can move the head directly from the physical beginning of the tape to a specific logical position for reading or writing.

Data tracks

The area between adjacent servo bands is a data band. There are four data bands numbered 2, 0, 1, and 3, where data band number 2 is nearest the reference edge of the tape and data band 3 is farthest away, as in Figure 8-4. The data bands are written in sequence beginning with 0 (in the center of the tape) and ending with 3.

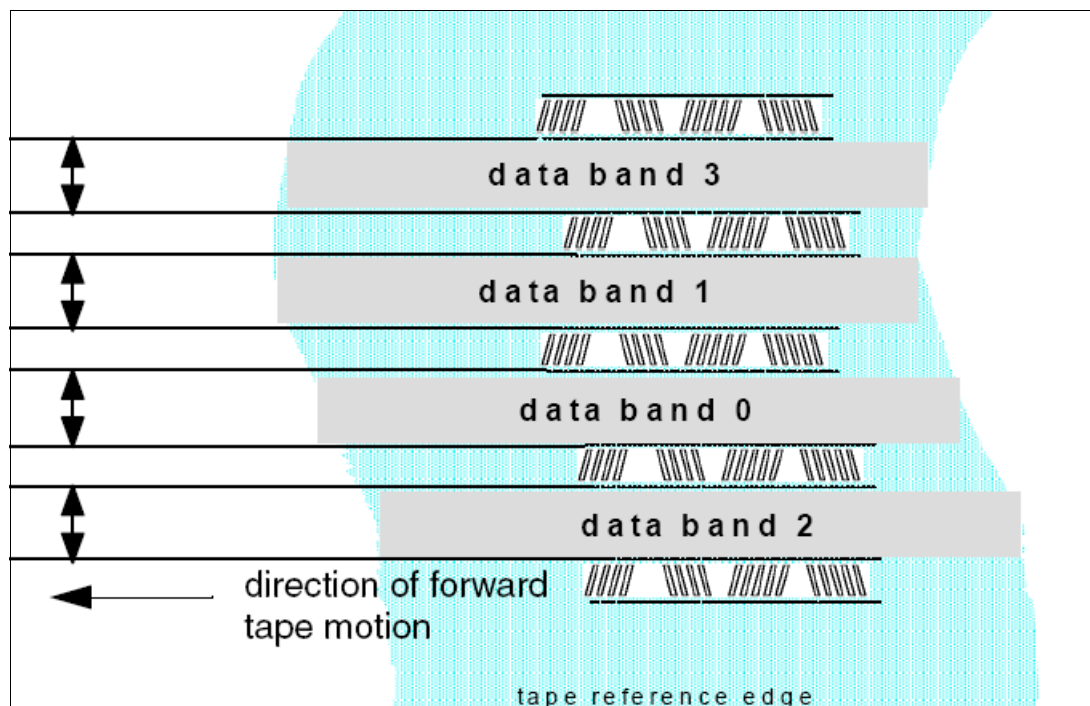


Figure 8-4 Four data bands written between the servo tracks

Each data band consists of numbers of tracks that are recorded eight tracks at a time (for Ultrium 1 and Ultrium 2) from one end of the tape to the other end. For Ultrium 3 and Ultrium 4, the tracks are recorded 16 tracks at a time. The process is:

1. The head is positioned over data band 0, and the first set of eight (16) tracks is written from the physical beginning of the tape (BOT) to the physical end of the tape (EOT).
2. The head physically repositions (using a separate servo position within the same servo bands) and switches electronically³ to a second set of eight (16) write elements in order to write eight (16) tracks in the reverse direction back to the physical beginning of the tape.
3. The head physically repositions again, and, switching back to the first set of write elements, writes another set of tracks to the physical end of the tape.
4. The head continues to switch and index in this manner until all the tracks are written, and the head is back at the physical beginning of the tape (Ultrium 1 and 2) or at the end of the tape for Ultrium 3.
5. The head moves to data band 1 to continue writing the data.

For Ultrium 1, 96 data tracks (8 tracks times 6 forward and backward writes) coexist in one data band. For Ultrium 2, there are 128 data tracks in one data band (8 tracks times 8 forward and backward writes), and for Ultrium 3 there are 176 tracks in one data band (16 tracks times 6+5 forward or backward writes). For Ultrium 4, 224 data tracks (16 tracks times 14 forward and backward writes) are in one data band.

A group of tracks recorded concurrently in the physical forward or the physical backward direction is called a wrap. Wraps recorded while the tape is moving from BOT to EOT are forward wraps; wraps recorded while the tape is moving from EOT to BOT are reverse wraps. The wraps are recorded in a serpentine fashion, as described: a forward wrap, then a reverse wrap. They are numbered sequentially in the order that they are processed, starting with wrap 0. Therefore, for Ultrium 1 six forward wraps and six reverse wraps make up a data band. For Ultrium 2, eight forward and eight reverse wraps make up a data band. The individual tracks within a wrap are interleaved with tracks from other wraps; in other words, adjacent tracks are not part of the same wrap (Figure 8-5).

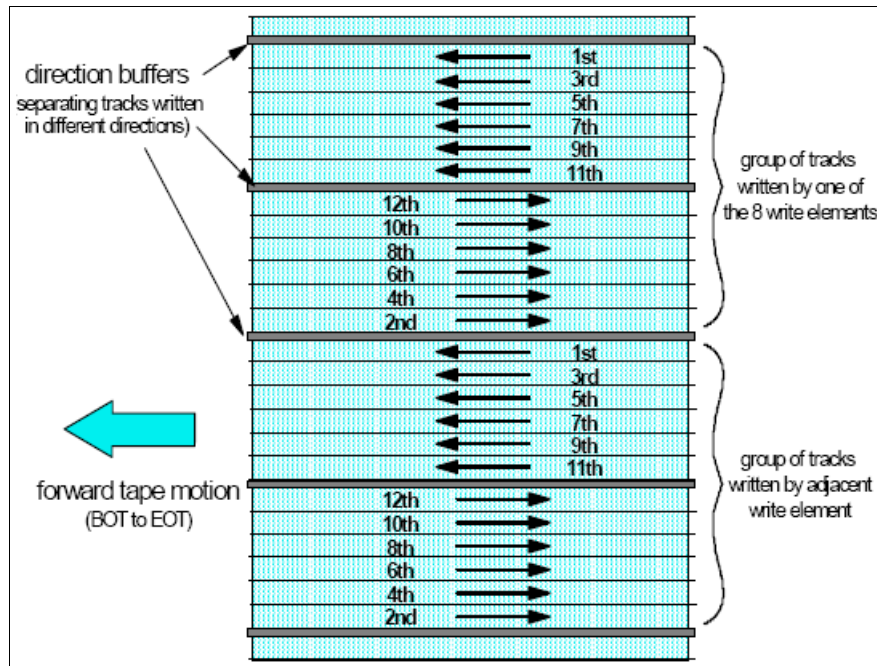


Figure 8-5 Portion of data band showing Ultrium 1 track-writing sequence

Backwards compatibility

The LTO standard specifies a backwards compatibility of writing one generation backwards and reading two generations backwards. To make this possible, LTO uses a technique called *shingling*. When using shingling, a write track can overlap the bottom of a previously written track. LTO Generation 2 uses shingling when writing data to an LTO Generation 2 Cartridge. The very first two passes write to the tape in the normal way. The following passes can partially overwrite previously written data tracks. The IBM LTO Generation 2 write head width is that of the LTO Generation 1. Therefore the LTO Generation 2 drives can write an LTO 1 cartridge in full track width, and when writing to an LTO 2 cartridge it uses the shingling write function. To read the residual LTO generation 2 data tracks, the read head must of course be narrower than the LTO generation 1 read head (Figure 8-6).

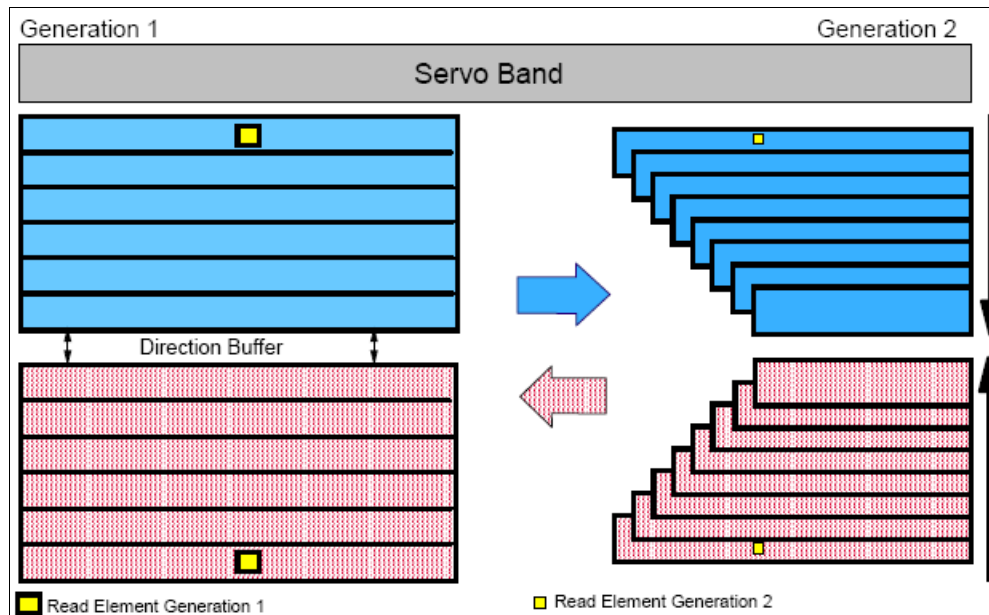


Figure 8-6 Shingling - writing to tape generation 1 versus generation 2

Again, LTO Generation 3 is conceptually the same, but technically different. The Ultrium 3 data is written 16 tracks at a time utilizing the shingling technique. To achieve compatibility, the write heads are equal to the residual track width of the Generation 2 format, and the spacing between alternate (every second) Generation 3 write heads is the same as between the Generation 2 write heads. Thus Generation 3 data is written using the shingling method, but Generation 2 data is written using only every second write head and writing the full Generation 2 data width. Generation 1 data can similarly be read by every other read head.

8.1.4 Linear density

The linear density for LTO 1 is 4880 bits per mm. The linear density was improved for LTO 2 to 7398 bits per mm, and for LTO 3 to 9638 bits per mm. Both LTO 1 and LTO 2 cartridges are of the same length, 610 m (2000 feet). To achieve the required doubling of capacity, the LTO 3 tape is slightly longer, 680 m (2231 feet).

8.2 IBM Ultrium 1, 2, 3, and 4 highlights

In this section, we show highlights of the IBM Ultrium 1, 2, 3, and 4 Tape Drives, such as drive and cartridge compatibility, LTO drive performance, drive operations, reliability, and cleaning the drive. We also illustrate the IBM LTO Ultrium family of tape drives and libraries and an overview of the multipath architecture.

Table 8-1 shows the compatibility among the four generations of LTO data cartridges. The following rules describe the compatibility among the various LTO cartridge generations:

- ▶ Data cartridges one generation prior are read/write compatible.
- ▶ Data cartridges two generations prior are read only.
- ▶ Data cartridges three generations prior are not supported.

However, the Ultrium 4 drive only allows an Ultrium 3 cartridge to be written at the LTO Generation 3 operating point (400 GB). More specifically, the Ultrium 4 drive does not allow an Ultrium 3 cartridge (400 GB) to be reformatted to the Ultrium 4 format (800 GB), which is generally true when a data cartridge is used in a higher generation Ultrium tape drive

Table 8-1 IBM Ultrium Tape Compatibility

IBM Ultrium Tape Drive	IBM TotalStorage LTO Ultrium Data Cartridge			
	800 GB (Ultrium 4)	400 GB (Ultrium 3)	200 GB (Ultrium 2)	100 GB (Ultrium 1)
Ultrium 4	Read/Write	Read/Write	Read only	N/A
Ultrium 3	N/A	Read/Write	Read/Write	Read only
Ultrium 3	N/A	N/A	Read/Write	Read/Write
Ultrium 1	N/A	N/A	N/A	Read/Write

8.2.1 WORM tape format

The IBM Ultrium Write Once Read Many (WORM) cartridges (machine type 3589) were designed for applications, such as archiving and data retention, as well as those applications requiring an audit trail. The IBM Ultrium 3 and Ultrium 4 tape drives provide WORM support for these types of applications. The IBM Ultrium WORM cartridges work with the IBM System Storage LTO Ultrium Tape Drive to help prevent the alteration or deletion of user data. In addition, IBM has taken several steps to reduce the possibility of tampering with the information:

- ▶ The bottom of the cartridge is molded in a color other than that used for rewritable cartridges
- ▶ The special cartridge memory helps protect the WORM nature of the media
- ▶ A unique format is factory-written on each WORM cartridge.

The Ultrium 400 GB WORM format, based on LTO specifications, provides a tape cartridge capacity of up to 400 GB native physical capacity (800 GB with 2:1 compression). You can use these cartridges in IBM TotalStorage Ultrium 3 Tape Drives, the third generation LTO Ultrium Tape Drive in the IBM TotalStorage LTO Ultrium family of products. The WORM 800 data cartridge provides a capacity of up to 800 GB native physical (1600 GB with 2:1 compression). The 800 GB WORM can only be used in the IBM Ultrium 4 tape drives.

8.2.2 Reliability

The IBM LTO Ultrium tape format differs from earlier IBM products and the reliability and availability features include:

- ▶ Data integrity: The drive performs a read after write, for verification.
- ▶ Power loss: No recorded data is lost as a result of normal or abnormal power loss while the drive is reading or writing data.
- ▶ Error correction: Two levels of error correction that can provide recovery from longitudinal media scratches.
- ▶ Integrated head cleaner: During the load process, a brush integrated into the drive mechanism cleans the head before it is used with the tape.

- ▶ Surface control guiding: Guides the tape along the tape path using the surface of the tape rather than the edges to control tape motion.
- ▶ Flat lap head: Improves contact between the read and write recording elements and the tape, giving higher quality recording and readback of data.
- ▶ Statistical Analysis and Reporting System (SARS): Only IBM LTO drives provide this level of preventive diagnostic reporting to assist in isolating failures between media and hardware.

For more information, see the website:

<http://www.ibm.com/servers/storage/tape/lto/>

8.3 IBM LTO Tape Drives

The IBM LTO Ultrium family of tape drives is comprised of four product offerings:

- ▶ IBM System Storage TS2230 Tape Drive Express Model
- ▶ IBM System Storage TS2240 Tape Drive Express Model
- ▶ IBM System Storage TS2340 Tape Drive Express Model
- ▶ IBM TotalStorage 3580 Tape Drive

8.3.1 IBM System Storage TS2230 Tape Drive Express Model

The IBM System Storage TS2230 Tape Drive Express LTO3 HH Mode is the entry-level IBM System Storage tape product family offering. By utilizing advanced Linear Tape-Open technology and the half-high format, the TS2230 Tape Drive is suited for handling the backup, save and restore, and archival data storage needs of a wide range of small systems. See Figure 8-7 on page 237.

The TS2230 has a storage capacity of up to 800 GB (with 2:1 compression) in conjunction with the IBM TotalStorage LTO Ultrium 400 GB data cartridge, which is double the capacity of the Ultrium 2 technology. Along with its higher capacity, the performance of the TS2230 Tape Drive has more than doubled over the previous generation of half-high LTO drives in the market for a native data transfer rate of up to 60 MBps. The TS2230 Tape Drive provides an excellent alternative to slower and smaller capacity 1/4-inch, 4 mm and 8 mm DLT/SDLT tape drives.

The TS2230 Tape Drive houses the Half-High Ultrium 3 Tape Drive. The TS2230 came available at the end of 2006 and might be an ideal tape drive for small clients who want a reliable tape drive with LTO technology. The TS2230 Tape Drive provides an excellent migration path from digital linear tape (DLT or SDLT), 1/4-inch, 4mm, or 8mm tape drives. The TS2230 Tape Drive Express Model incorporates third-generation IBM LTO Ultrium technology. The maximum tape drive throughput data rate performance is up to 60 MB/s native data transfer rate. IBM Ultrium 3 Tape Drives can read and write LTO Ultrium 2 Data Cartridges and can read LTO Ultrium 1 Data Cartridges.

The TS2230 can be attached to IBM System i, IBM System p, IBM System x, Microsoft Windows, Hewlett-Packard UNIX (HP-UX), Sun Solaris, UNIX, Linux, and PC servers. To determine the latest update of the supported servers, see the website:

<http://www-03.ibm.com/systems/storage/tape/library.html>

TS2230 highlights: The IBM System Storage TS2230 Tape Drive Express Model has the following features (Figure 8-7):

- ▶ Half-high LTO 3
- ▶ Lower list price than full-high LTO 3 tape drives
- ▶ LVD Ultra160 SCSI attach
- ▶ 3 Gbps SAS attach
- ▶ WORM capable
- ▶ Up to 800 GB compressed; 400 GB native capacity per cartridge based on 2:1 compression
- ▶ Up to 60 MBps native sustained data transfer rate based on 2:1 compression



Figure 8-7 Front view of IBM System Storage TS2230 Tape Drive Express Model

8.3.2 IBM System Storage TS2240 Tape Drive Express Model

The System Storage TS2240 Tape Drive, as seen in Figure 8-8 on page 238, is an excellent tape storage solution for businesses requiring backup or low-cost, real-time archival storage of their data. The TS2240, with a half-high form factor, offers the same high capacity of full-high LTO 4 tape drives. The TS2240 has a physical storage capacity of up to 1.6 TB (with 2:1 compression) in conjunction with the new IBM System Storage LTO Ultrium 800 GB data cartridge, which provides up to double the capacity of Ultrium 3 cartridges.

The native data transfer performance of the TS2240 Tape Drive has increased over the previous LTO half-high generation to up to 120 MBps. The TS2240 Tape Drive continues to provide an excellent alternative to slower and smaller capacity 1/4-inch, 4 mm and 8 mm DLT/SDLT tape drives.

The IBM System Storage TS2240 Tape Drive (3580 Model H4S) is an external stand-alone or rack-mountable shelf unit, designed for the family of IBM Ultrium Tape products. The TS2240 Tape Drive is a high-capacity data storage device that offers high performance and is designed for backup and restore by midrange Open Systems applications. The TS2240 Tape Drive incorporates the Linear Tape-Open (LTO) IBM System Storage Ultrium 4 Half-High Tape Drive.

The TS2240 Tape Drive Model H4S is encryption-capable and supports Application-Managed Encryption (AME). The TS2240 Tape Drive can use the T10 encryption method. Encryption is only supported with the LTO Ultrium 4 Data Cartridge. The TS2240 is a client-replaceable unit (CRU). In case the TS2240 has a failure, IBM provides you a replacement rather than repairing the tape drive.

TS2240 highlights: The IBM System Storage TS2240 Tape Drive Express Model has the following features (Figure 8-8):

- ▶ Features the new IBM Ultrium 4 Tape Drive in a half-high form factor
- ▶ IBM Ultrium 4 technology is designed to support encryption of data and continues to support write once, read many (WORM) operation
- ▶ Hardware encryption available

- ▶ 3 Gbps SAS attach
- ▶ Up to 1.6 TB compressed; 800 GB native capacity per cartridge based on 2:1 compression
- ▶ Up to 120 MBps native sustained data transfer rate based on 2:1 compression
- ▶ Large internal data transfer buffer
- ▶ Backwards compatibility with previous generations of LTO Ultrium cartridges



Figure 8-8 Front View of IBM System Storage TS2240 Tape Drive Express Model

8.3.3 IBM System Storage TS2340 Tape Drive Express Model

The IBM System Storage TS2340 Tape Drive features SCSI Ultra160 Low Voltage Differential (LVD) as well as 3 Gbps SAS interface for connection to a wide spectrum of open system servers. The TS2340 is designed to exceed your most demanding backup and restore needs. The IBM Ultrium 4 technology is designed to support encryption of data. The hardware encryption and decryption core and control core resides in the IBM Ultrium 4 tape drive (available to the TS2340 with the 3 Gbps SAS interface).

A larger internal data buffer helps improve data access rates and reduce cartridge fill and rewind times along with dynamic channel calibration to help increase data throughput. In addition to reading and writing to LTO Ultrium 3 tape cartridges, the TS2340 can read LTO Ultrium 2 cartridges as seen in Figure 8-9 on page 239

The TS2340 Tape Drive is available in two models, which are determined by the attachment interfaces. The TS2340 Tape Drive Model L43 uses a Small Computer Systems Interface (SCSI) Ultra160 Low Voltage Differential (LVD) attachment, and Model S43 uses a 3 Gbps Serial-Attached SCSI (SAS) interface for connecting to Open Systems servers.

The TS2340 Model S43 has two SFF-8088 interfaces for connecting to Open Systems servers. Write Once Read Many (WORM) cartridges are supported and recognized when loaded.

The TS2340 Tape Drive Model S43 is encryption-capable and supports Application-Managed Encryption (AME). The TS2340 Tape Drive uses the T10 encryption method. Encryption is only supported with the LTO Ultrium 4 Data Cartridge. The TS2340 is a client-replaceable unit (CRU). In case the TS2340 has a failure, IBM provides you a replacement.

For error codes and messages, there is a Single Character Display (SCD) at the front of the TS2340 Tape Drive. The TS2340 Tape Drive can be attached to IBM System i, IBM System p, IBM System x, Microsoft Windows, Hewlett-Packard UNIX (HP-UX), Sun Solaris, UNIX, Linux, and PC servers.

To determine the latest supported servers, see the website:

<http://www.ibm.com/servers/storage/tape/compatibility/pdf/>

TS2340 highlights: The IBM System Storage TS2340 Tape Drive Express Model has the following features (Figure 8-9):

- ▶ Model L43 – One IBM Ultrium 4 Tape Drive, LVD Ultra160 SCSI attach
- ▶ Model S43 – One IBM Ultrium 4 Tape Drive, 3 Gbps SAS attach
- ▶ Next generation of LTO technology
- ▶ Hardware encryption available
- ▶ WORM capable
- ▶ Up to 1.6 TB compressed; 800 GB native capacity per cartridge based on 2:1 compression
- ▶ Up to 120 MBps native sustained data transfer rate based on 2:1 compression
- ▶ Large internal data transfer buffer
- ▶ Backwards compatibility with previous generations of LTO Ultrium cartridges



Figure 8-9 Front View of IBM System Storage TS2340 Tape Drive Express Model

8.3.4 IBM TotalStorage 3580 Tape Drive

The IBM TotalStorage 3580 Tape Drive is the smallest in the family of IBM Ultrium tape solutions. It is an external, stand-alone, Small Computer System Interface (SCSI)-attached tape drive that attaches to System i, System p, System x, RS/6000® SP, and other UNIX and PC servers supporting OS/400, IBM AIX, Sun Solaris, HP-UX, Microsoft Windows NT®, Microsoft Windows 2000 Server, Microsoft Windows Server 2003, and Red Hat and SUSE LINUX using a supported SCSI adapter. See Figure 8-10 on page 240.

The IBM TotalStorage 3580 model L33 Tape Drive is an external drive incorporating the 3rd and latest generation of IBM LTO technology. This is an external stand-alone or rack-mountable unit, similar to previous models of the 3580 and is the entry point for the family of IBM Ultrium tape products as seen in Figure 8-10. The 3580 Tape Drive provides an excellent migration path from digital linear tape (DLT or SDLT), 1/4-in., 4mm, or 8mm tape drives. The 3580 model L33 can read and write LTO Ultrium 2 Data Cartridges and read LTO Ultrium 1 Data Cartridges.

For additional information regarding the IBM TotalStorage 3580 Tape Drive, see the website:
<http://www.ibm.com/systems/storage/tape/3580/index.html>

3580 product details: The IBM TotalStorage 3580 Tape Drive has the following features (Figure 8-10):

- ▶ Integrates into the following storage environments: Server (non-IBM & IBM), Automated library, and SAN-attached
- ▶ Adheres to the widely supported LTO specification which promotes standardization and allows for multiple media and drive providers
- ▶ Provides superior performance and capacity attributes for unattended backup within midrange and enterprise server environments
- ▶ Compatible with major operating systems and ISV applications
- ▶ Uses advanced technologies that optimize throughput, increase cartridge capacity and provide superior data protection



Figure 8-10 3580 Model L33/L3H

8.3.5 IBM 3592 Tape Drive

The IBM TotalStorage 3592 Enterprise Tape Drive has been replaced by the IBM System Storage TS1120 and TS1130 Tape Drive.

8.4 IBM System Storage TS1100 Tape Drive Family

The IBM System Storage TS1120 and TS1130 tape drives (machine types 3592-E05 and 3592-E06/EU6) offer a design focused on high capacity, performance, and high reliability for storing mission critical data. With the September 2003 introduction of the first generation of the new family of tape drives, IBM advanced its high-end half-inch cartridge tape technology to a new level. The second generation of the 3592 family enhanced the capacity and performance characteristics, which have now become increased by the IBM System Storage TS1130 Model E06/EU6 Tape Drive, the third generation of the 3592 family, which provides the unprecedented capacity of 1 TB of uncompressed data.

The following key features are included:

- ▶ Virtual backhitch, which is the optimum adaptive format and algorithm designed for improved start/stop write synchronize performance

- ▶ High performance and robust dual microprocessor architecture. One microprocessor operates the host attachment interface (running what is essentially proven 3590 host attach microcode), while the other microprocessor is allowed to focus strictly on writing data and reading data from tape. Each microprocessor is designed to reset the other microprocessor to act as a fail-safe
- ▶ S/390® and System z attachment through ESCON and FICON by means of the existing J70 controller, as well as the TS1120 Tape Controller
- ▶ Statistical Analysis Recording System (SARS) algorithm with extended mount count
- ▶ Fast random access performance when operating on any of the Short Length Cartridge (SLC) types
- ▶ Support of an enhanced capacity scaling and segmentation format when operating on the full length Read/Write (R/W) cartridge type JA and JB, enabling very fast locate and read times
- ▶ Streaming Lossless Data Compression (SLDC) algorithm (enhancement of the Lempel-Ziv class 1 (LZ-1) data compression algorithm)
- ▶ Cartridge memory of 4 K designed for the 3592 to support advanced features
- ▶ An advanced interleaved bidirectional serpentine recording technique that writes 16 data tracks at a time on a 3592 cartridge
- ▶ The timing-based servo technique guarantees the most accurate positioning of the Read/Write heads
- ▶ Improved reliability and availability are achieved by:
 - Single Field Replaceable Unit (FRU)
 - Redundant, hot-pluggable power supplies
 - Retention of Fibre Channel Worldwide Name ID during service action
- ▶ Advanced technology by:
 - Robust loader mechanism
 - Elimination of drive pneumatics and mechanical adjustments
 - Straighter and shorter tape path for better tape tracking
 - Speed matching to reduce backhitching
 - Channel calibration to optimize performance and data integrity
- ▶ The drive has a large data buffer with read-ahead buffer management
- ▶ Speed matching for medium data rates when operating from a host that cannot sustain the maximum 3592 data rate
- ▶ Cartridge memory chip, which holds information about that specific cartridge, the media in the cartridge, and the data on the media
- ▶ High-resolution tape directory for improved nominal and average access times for locate operations
- ▶ Performance/Capacity scaling to allow clients to trade off capacity for improved access times

8.4.1 The 1 TB background

On 5 April 2002, IBM achieved an unprecedented feat—1 TB of data was recorded, without compression, to a tape containable in a half-inch format tape cartridge. This was a technological accomplishment that set the foundation for the 3592 Tape Drive family. By using the evolutionary progression of technology building blocks that IBM had set in place over the preceding five years, an Enterprise Tape Drive roadmap was laid out that can ultimately reach and exceed 1 TB in native cartridge capacity over several tape drive generations.

The 3592 Model J1A became the first tape drive generation of the Enterprise Tape family. It enabled storage of 300 GB of data to a cartridge even if it is incompressible (900 GB with 3:1 compressible data). An even more innovative achievement was that these same cartridges are designed to be reused by the second generation of 3592 tape drives, the TS1120 Model E05 to store significantly more data; using the JA media, we can store 500 GB and with the high capacity cartridge, 700 GB without compression. Now, with the third generation of IBM 3592 Tape Drive, IBM kept its promise documented in the roadmap for 3592 tape drives. The TS1130 Tape Drive is not just a new drive but the IBM commitment to further tape technology development.

8.4.2 Recording format

The IBM 3592 Tape Drive uses an advanced interleaved bidirectional serpentine recording technique that writes eight or 16 (depending on the drive) data tracks at a time on a 3592 cartridge. The 3592 cartridge is a half-inch, advanced metal particle, dual layer tape, and the tape layout consists of five servo bands (which are prerecorded on the tape) and four data bands where the data is written (Figure 8-11). The servo bands provide location information to control the positioning of the head as it writes and reads data within the data band. This design is explained in detail in “Servo tracks” on page 96.

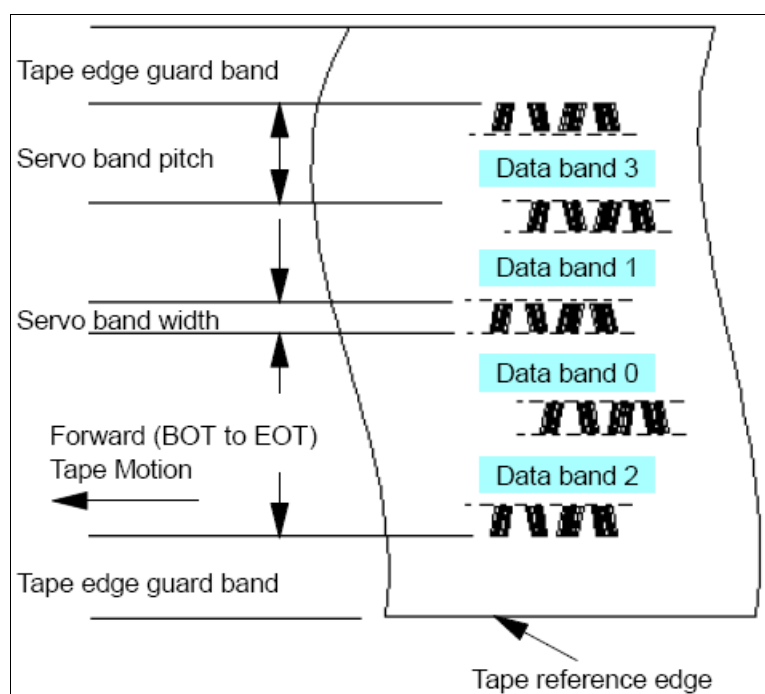


Figure 8-11 The layout of the servo and data bands on the 3592 media

As shown in Figure 8-11, the area between adjacent servo bands is a data band. There are four data bands, each with a number of data tracks (128 - 288, which is unique for each model), on the 3592 media. The data bands are numbered 2, 0, 1, and 3, with data band 2 nearest the Tape Reference Edge and data band 3 farthest from the Tape Reference Edge.

As shown in Figure 8-12 (here, for example, is a J1A written cartridge), each data band is actually composed of eight data sub-bands, one for each of the eight write heads. Each sub-band is written by a given write-head position in a technique called a linear serpentine, which means that the tape moves back and forth longitudinally while the head is indexed up or down laterally at each pass. This technique makes it possible to write multiple distinct tracks in a given data sub-band.

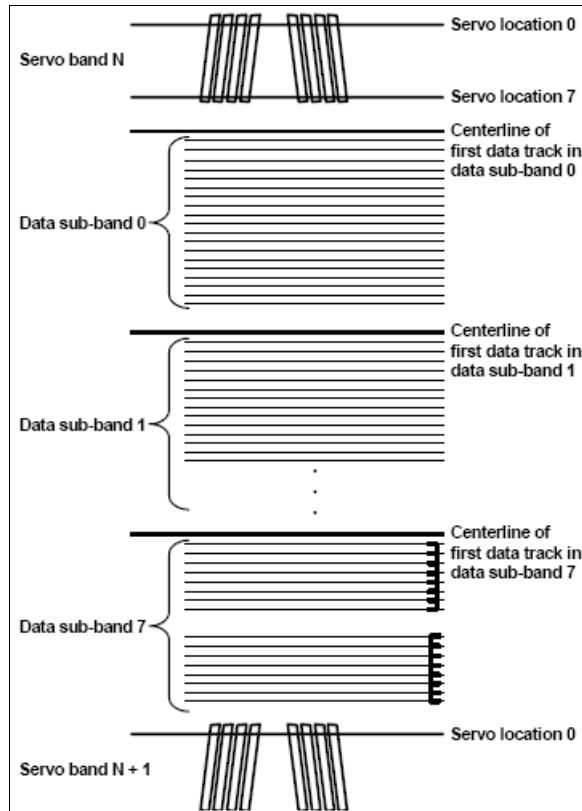


Figure 8-12 Section of tape showing one data band and its surrounding servo bands

8.5 IBM System Storage TS1120 Tape Drive

The IBM System Storage TS1120 Tape Drive (machine type 3592-E05) is the second generation 3592-type drive. It is supported in IBM tape libraries, IBM frames that support stand-alone installation, and in an IBM TotalStorage 3592 Model C20 (3592 C20 frame) attached to a Sun StorageTek 9310 library.

The TS1120 tape drive uses the existing 3592 media, which is available in rewritable or WORM format to store 100 GB, 500 GB, and 700GB native capacity depending on cartridge media. The 3592 JA/JW media helps reduce resources to lower total costs, whereas the 3592 JJ/JR media is designed to support applications that require rapid access to data.

TS1120 tape drives can be shared among supported open system hosts on a SAN, or between FICON and ESCON mainframe hosts when attached to an IBM System Storage TS1120 Tape Controller Model C06 (TS1120 tape controller). Sharing drives optimizes drive utilization and helps reduce infrastructure requirements. Figure 8-13 shows the TS1120 tape drive. The TS1120 features dual-port 4 Gbps Fibre Channel interfaces.



Figure 8-13 IBM TS1120 tape drive

8.5.1 Highlights

The TS1120 tape drive offers the following features:

- ▶ Supports IBM Systems and selected open system platforms
- ▶ Supported on existing IBM and Sun StorageTek automation
- ▶ Offers native data transfer rate of up to 104 MBps - over 2.5 times as fast as the IBM TotalStorage 3592 Tape Drive and over seven times as fast as the 3590
- ▶ Data transfer rate up to 260 MBps with 3:1 compression
- ▶ Up to 700 GB native cartridge capacity, a 1.6 times increase over the IBM TotalStorage 3592 Tape Drive and a five times as much as the 3590 H Models
- ▶ Supports up to 1.5 TB (uncompressed in a System z9® environment) on a 3592 JA/JW cartridge
- ▶ Supports fast access with a 3592 JJ/JR cartridge
- ▶ Helps support regulatory compliance requirements for records retention with WORM media
- ▶ Up to 2.1 TB on a 3592 JB/JX cartridge
- ▶ Supports encryption and key management:
 - Supports a single encryption key management approach, which can help reduce audit and compliance costs
 - Helps to avoid the need for host-based encryption of data or the use of specialized encryption appliances

8.5.2 Media support

The TS1120 tape drive can initialize short length JJ cartridges to 60 GB or 100 GB and initialize (or re-initialize) standard length JA cartridges to 60, 100, 300 or 500 GB to support fast access to data or to help address data growth and facilitate interchange. At typical compression ratios, the 3592 JA cartridge can provide usable capacity of up to 1 TB in an open system environment, and up to 1.5 TB in an IBM System z9 environment when used with a TS1120 tape drive.

8.5.3 More information

For additional product details, see:

<http://www.ibm.com/systems/storage/tape/drives>

8.5.4 IBM System Storage TS1120 Tape Controller Model C06

The IBM System Storage TS1120 Tape Controller Model C06 (model type 3592-C06) provides performance and reliability for IBM System z customers. The IBM System Storage TS1120 Tape Controller Model C06, has up to four 4 Gbps FICON attachments. The TS1120 Tape Controller also has up to eight ESCON attachments, or an intermix of ESCON and FICON attachments. Up to 16 IBM System Storage TS1120 Tape Drives or IBM TotalStorage 3592 Tape Drives can be attached to a single TS1120 Tape Controller.

The controller can be installed in an IBM System Storage 3952 Tape Frame Model F05 , an IBM 3953 Tape Frame Model F05, or in a stand-alone rack, supporting 3592 Tape Drives installed in IBM 3494 frames, IBM 3584 frames, IBM 3592 Model C20 frames, and stand-alone racks.

TS1120 tape drives can be shared among supported open system hosts on a Storage Area Network (SAN), or between FICON and ESCON mainframe hosts when attached to an IBM System Storage TS1120 Tape Controller (TS1120 tape controller). Sharing drives optimizes drive utilization and helps reduce infrastructure requirements.

The TS1120 Tape Controller provides up to 1.7 times the throughput of the IBM TotalStorage 3592 Tape Controller Model J70, with 4 Gbps FICON attachment using the 3592 Model J1A Tape Drive.

The TS1120 Tape Controller has many of the reliability and availability characteristics that the 3592 Model J70 offered, such as redundant power supplies with automatic failover and hot swap capabilities, redundant cooling, and support for the IBM TS3000 System Console (TSSC) attachment, previously known as the IBM TotalStorage Master Console (TSMC).

The TS1120 Tape Controller also includes application performance and capacity enhancements that are available with the 3592 Model J70, such as capacity scaling commands. The effect of capacity scaling is to potentially reduce the average locate time to a random record (from load point) to as little as 30% of the normal locate time. Figure 8-14 shows the IBM TS1120 Tape Controller.



Figure 8-14 IBM TS1120 Tape Controller

Highlights: IBM System Storage TS1120 Tape Controller:

- ▶ Enables System z attachment for the IBM System Storage TS1120 Tape Drive
- ▶ Supports the IBM System Storage TS3400, TS3500 and TotalStorage 3494 Enterprise Tape Libraries, StorageTek 9310 Powderhorn Tape Library; and stand-alone rack and frame environments.
- ▶ Designed to fully exploit the high performance, capacity, and reliability of the TS1120 tape drive
- ▶ Designed to support the 700GB extended capacity, 500GB standard capacity and 100 GB fast-access formats of the IBM 3592 tape cartridges
- ▶ Enables tape drive and tape cartridge sharing between ESCON and FICON enabled System z servers
- ▶ Supports TS1120 tape drive encryption

8.6 IBM System Storage TS1130 Tape Drive

The IBM System Storage TS1130 Tape Drive (also referred to as the 3592 Model E06 and EU6) is the third tape drive generation of the IBM 3592 tape family (Figure 8-15). This generation provides higher capacity and performance compared to the predecessor 3592 Model E05.

The TS1130 records in two recording formats supporting both encryption and non-encryption. Enterprise format 3 (EFMT3) is used to represent the non-encrypted recording format, and enterprise encrypted format 3 (EEFMT3) is used to denote the encrypted recording format. With these recording formats, the non-compressed capacity of the extended length MEDIA9 and MEDIA10 cartridges is increased from 700 GB to 1 TB.

The 3592 Model E06 is downward read compatible (n-2) to the 3592 Model J1A format (EFMT1) and is downward write compatible (n-1) to the 3592 Model E05 formats (EFMT2/EEFMT2). All current media types are supported.



Figure 8-15 Front View of TS1130 Tape

The 3592-E06 supports all of the same subsystems and automation environments as the previous models. The 3592-E06 and 3592 cartridges support multiple automation libraries and can be easily transported between environments (consider the compatibility requirements for silo-attached 3592 tape drives).

The TS1120 tape controller offers ESCON and FICON attachment of TS1130 tape drives in a TS3400, TS3500 or 3494 tape library, 3952 C20 frame or rack to support drives in a 3494 tape library the tape controller must reside in an IBM 3952 Tape Frame Model F05.

To help optimize drive utilization and reduce infrastructure requirements, the TS1130 tape drives can be shared among supported open system hosts on a Storage Area Network (SAN) or between IBM FICON and ESCON mainframe hosts when attached to an IBM System Storage TS1120 Tape Controller.

To further protect mission-critical data, the TS1130 tape drive provides advanced technology, such as error correction code and factory-written servo tracks on the tape cartridge, to provide precise head positioning. In addition, unique functions such as virtual backhitch and a high-resolution directory help improve small file write performance and improve recall access to data, respectively.

The IBM 3592 cartridges used in the TS1130 tape drive are available in short, standard and extended lengths and in rewritable and WORM formats. Cartridges can be ordered in packs of 20 and can be labeled and initialized, initialized only, or unlabeled and uninitialized. An RFID label option is also available.

Highlights: The following features are included:

- ▶ Provides information security with support for encryption and key management
 - Supports a single encryption key management approach, which can help reduce audit and compliance costs
 - Helps to avoid the need for host-based encryption of data or the use of specialized encryption appliances
- ▶ Optimizes information retention with support for existing IBM tape automation
- ▶ Supports Write Once Read Many (WORM) cartridges to help satisfy compliance requirements
- ▶ Offers high performance and high capacity for storage consolidation
- ▶ Up to 3 TB on a 3592 JB/JX cartridge
 - High capacity helps reduce the need for additional tape drives and cartridges
 - Use fewer drives to reduce complexity of your tape infrastructure
- ▶ 320 MBps in open system environments where data typically compresses at 2:1
- ▶ 350 MBps in a mainframe environment where data typically compresses at 3:1
- ▶ Virtual backhitch and high resolution directory which will improve access to data and reduce wear and tear on media
- ▶ IBM Power Systems, System i, System p, System z and System x support

8.7 IBM Midrange and Enterprise Storage Media

There is separate tape media for each tape drive type - LTO, 3592, and 3590. Each drive type only supports its own media - e.g. you cannot use an LTO cartridge in a 3592 tape drive. When an IBM Tape Library is ordered, it is supplied with one Ultrium data cartridge or one 3592 data cartridge and one cleaning cartridge at no charge. Each member of the IBM 35xx and TS3xxx Tape Library family has unique media features and rules that apply when placing the order. The following information will assist you in ordering additional media.

If you want to order media for your Ultrium, 3590, or 3592 tape drives, go to this website:

<http://www.ibm.com/servers/storage/media/distributors/index.html>

Select your geographic location and your country, and a list of authorized distributors is presented.

For a list of tape media compatibility, see the website:

http://www.ibm.com/systems/storage/media/tape_compatibility.html

8.7.1 LTO cartridge

IBM is delivering high-capacity media that provides clients the ability to meet the demanding growth in data storage. The increase in cartridge capacity reduces the amount of equipment, space and human intervention required for daily tape operations. In addition, it reduces the number of cartridges needed for backup and restore operations, and helps to lower operational costs throughout the enterprise.

For additional information about IBM LTO media, see the website:

<http://www.ibm.com/systems/storage/media/>

The Ultrium cartridge is a single-reel cartridge. This means that the whole tape is wrapped around a single reel when the cartridge is not loaded in a drive. During the loading process, the threader of the drive catches the leader pin of the tape and threads it through the drive and the machine reel. During the read/write process the tape is stored on the machine reel and the cartridge.

A top and rear view of the tape cartridge is shown in Figure 8-16.

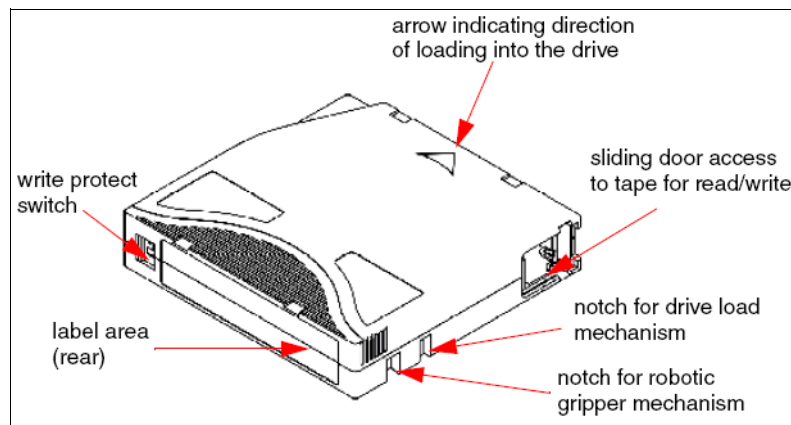


Figure 8-16 Ultrium cartridge view from top and rear

A top and front view of the tape cartridge is shown in Figure 8-17.

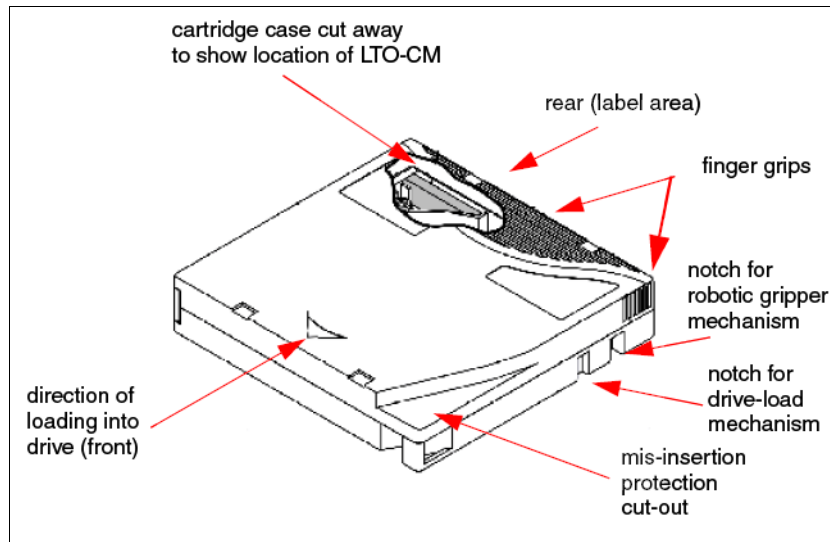


Figure 8-17 Ultrium cartridge view from top and front

IBM TotalStorage LTO Ultrium 100 GB Data Cartridge

This data cartridge offers the following features:

- ▶ Provides the capacity, performance and reliability for midrange and network tape storage environments
- ▶ Contains half-inch metal particle tape that features a native data capacity of 100GB and a compressed data capacity of 200GB
- ▶ Includes a Linear Tape-Open Cartridge Memory (LTO - CM) chip that contains cartridge specific information to communicate with Ultrium tape drives during loading/unloading by a radio frequency component
- ▶ Adheres to widely supported Linear Tape-Open (LTO) design specifications

IBM TotalStorage LTO Ultrium 200 GB Data Cartridge

This data cartridge offers the following features:

- ▶ Provides capacity, performance and reliability for midrange and network tape automated storage environments
- ▶ Contains half-inch metal particle tape; cartridges feature a native data capacity of 200GB and a compressed data capacity of 400GB (2:1 compression)
- ▶ Includes a Linear Tape-Open Cartridge Memory (LTO - CM) chip that contains cartridge specific information to communicate with Ultrium tape drives during loading/unloading by a radio frequency component
- ▶ Adheres to widely supported Linear Tape-Open (LTO) Ultrium 2 design specifications

IBM TotalStorage LTO Ultrium 400 GB Data Cartridge

This data cartridge offers the following features:

- ▶ High-capacity, high-reliability tape cartridge for LTO Generation 3 storage environments
- ▶ Provides 400 GB storage physical capacity (800 GB with 2:1 compression)
- ▶ Available in rewritable and Write Once Read Many (WORM) cartridge models

- ▶ Designed to help improve read/write reliability by storing the IBM statistical media read/write analysis on every use of the cartridge
- ▶ Designed to adhere to widely supported Linear Tape-Open (LTO) Ultrium 3 design specifications
- ▶ Gray standard cartridge color distinguishes the cartridges from previous generations of IBM LTO Ultrium cartridges
- ▶ Two-tone WORM cartridge color distinguishes WORM cartridges from standard cartridges

IBM TotalStorage LTO Ultrium 800 GB Data Cartridge

This data cartridge offers the following features:

- ▶ High-capacity, high-reliability tape cartridge for LTO Ultrium Generation 4 storage environments
- ▶ Provides 800 GB storage physical capacity (1.6 TB with 2:1 compression)
- ▶ Available in rewritable and write once, read many cartridge (WORM) models
- ▶ Data that is encrypted by LTO Ultrium 4 drives can be stored on both rewritable and WORM media to securely maintain data integrity
- ▶ Designed to help improve read/write reliability by storing the IBM statistical media read/write analysis on every use of the cartridge when using IBM LTO tape drives
- ▶ Green standard cartridge color which distinguishes the cartridges from previous generations of IBM LTO Ultrium cartridges
- ▶ Two-tone WORM cartridge color distinguishes WORM cartridges from standard cartridges

8.7.2 IBM 3590 and 3590E Cartridge

IBM 3590 Extended High-Performance Cartridge Tape continues the IBM focus on tape technology by doubling the length and providing twice the capacity of IBM 3590 High-Performance Cartridge Tape. These tape products increase the data storage capacity of half-inch cartridge tape. In addition, both types of cartridges can reside in the same physical space as IBM 3490 cartridges; they can also coexist in an IBM TotalStorage Enterprise Tape Library 3494.

- ▶ Designed to provide outstanding capacity, performance and reliability in 128-, 256- or 384-track formats
- ▶ Helps protect existing investments by residing in the IBM TotalStorage Enterprise Tape Library 3494 with IBM 3490 cartridges
- ▶ Supports flexibility with a choice of native media capacities: 10GB/20GB, 20GB/40GB or 30GB/60GB

For more information about IBM 3590 media, see the website:

<http://www.ibm.com/systems/storage/media/3590/index.html>

8.7.3 IBM 3592 Cartridge

IBM 3592 Tape Cartridges are designed to work with the first-generation IBM TotalStorage Enterprise Tape Drive 3592 Model J1A (3592 J1A tape drive) and the second-generation IBM System Storage TS1120 Tape Drive (TS1120 Tape Drive). 3592 tape media can be seen in Figure 8-18.

The 3592-J1A uses four media cartridge types: JA, JJ, JW, and JR. The 3592-E05 and the 3592-E06 use six media cartridge types: JA, JJ, JW, JR, JB, and JX. All six cartridge types contain the same dual-coat, advanced-particle media. Capacity on these media types depends on whether the cartridge is used by Model 3592-J1A, 3592-E05, or 3592-E06.

Cartridges are available in three lengths and in either re-writeable or Write Once, Read Many (WORM) formats. The short length 3592 JJ/JR cartridges provide rapid access to data and the standard length 3592 JA/JW cartridges provide high capacity. The 3592 Extended Data JB and Extended WORM JX 700GB high capacity cartridges support TS1120 tape drives only.

The first-generation 3592 J1A tape drive can initialize short length JJ cartridges to 60GB and initialize (or re-initialize) standard JA length cartridges to either 60GB (to support fast time to data) or 300GB (to support high capacity).

The second-generation TS1120 tape drive can initialize short length JJ cartridges to 60 or 100GB and initialize (or re-initialize) standard length JA cartridges to 60, 100, 300 or 500GB to support fast access to data or to help address data growth and facilitate interchange. 3592 Extended Data JB cartridges can be initialized to 140GB for fast access to data, can be segmented to 609GB or used at full 700GB capacity.

Highlights: The following features are included (Figure 8-17):

- ▶ Designed for the first generation IBM TotalStorage Enterprise Tape Drive 3592 or the second generation IBM System Storage TS1120 Tape Drive
- ▶ Available in rewritable and write once, read many (WORM) cartridge models
- ▶ The JJ and JR cartridges provide fast access to application information
- ▶ The JA/JW and JB/JX cartridges support high capacity requirements
- ▶ Memory chip tracks cartridge and tape drive usage

For more information about IBM 3592 media, see the website:

<http://www.ibm.com/systems/storage/media/3592/index.html>



Figure 8-18 IBM 3592 Tape Media

Additional information about how the 3592 can be used in your environment can be found in the following material:

IBM TotalStorage 3494 Tape Library: A Practical Guide to Tape Drives and Tape Automation, SG24-4632



IBM Tape Automation Products

This chapter discusses IBM Enterprise Tape Products, which provide the highest levels of performance and reliability of any IBM tape subsystem.

This chapter covers the following products:

- ▶ IBM System Storage:
 - IBM System Storage TS2900 Tape Autoloader Express
 - IBM System Storage TS3100 Tape Library
 - IBM System Storage TS3200 Tape Library
 - IBM System Storage TS3310 Tape Library
 - IBM System Storage TS3400 Tape Library
 - IBM System Storage TS3500 Tape Library
- ▶ IBM TotalStorage 3593 Tape System:
 - The TS7740 includes the functionality of the previously required IBM 3593 Tape System consisting of one or more frames and the Library Manager.
 - The 3953 Tape System is still required for attachment of TS1100 tape drives installed in the TS3500 Tape Library and connected to a System z host through tape controllers.
- ▶ Other models:
 - IBM TotalStorage 3581 Tape Autoloader: The IBM TotalStorage 3581 Tape Autoloader has been replaced by the IBM System Storage TS3100 Tape Library Express Model.
 - IBM TotalStorage 3582 Tape Library: The IBM TotalStorage 3582 Tape Library has been replaced by the IBM System Storage TS3200 Tape Library Express Model.
 - IBM TotalStorage 3583 Tape Library: The IBM TotalStorage 3583 Tape Library has been replaced by the IBM System Storage TS3310 Tape Library.
 - The IBM TotalStorage 3494 Tape Library: The IBM TotalStorage 3494 Tape Library has been replaced by the TS3500 Tape Library.

9.1 Overview

IBM System Storage provides a complete spectrum of tape libraries highlighting high performance and capacity for entry, midrange and enterprise systems. The IBM libraries handle backups, save and restore, and archival data storage needs.

9.1.1 IBM System Storage TS2900 Autoloader

The IBM System Storage TS2900 Tape Autoloader is a single-drive, low profile automated tape solution and the first entry automation solution for small- to mid-market tape environments. The TS2900 utilizes the technology of IBM half-high LTO tape drives to help create a high-capacity tape storage solution suited for handling backup and archival data storage for the Windows, Linux and other open system environments. The TS2900 is available with IBM HH LTO 3 or HH LTO 4 tape technology for a reliable, low entry-priced solution.

9.1.2 IBM System Storage TS3100 Tape Library

The IBM System Storage TS3100 Tape Library is well-suited for handling backup, save and restore, and archival data-storage needs for small to medium-size environments. It has a single Ultrium 3 tape drive and 22 tape-cartridge capacity.

9.1.3 IBM System Storage TS3200 Tape Library

The IBM System Storage TS3200 Tape Library is designed for backup, save and restore, and archival data-storage needs for small to medium-size environments. The TS3200 is an external 4U stand-alone or rack-mountable unit that incorporates up to two Ultrium 3 tape drives and 44 tape cartridges, as well as 3 media mail slots and 1 dedicated cleaning cartridge slot.

9.1.4 IBM System Storage TS3310 Tape Library

IBM System Storage TS3310 Tape Library offers simple, rapid expansion as processing needs grow. Its entry level configuration is a single 5 EIA rack unit high library. Over time, as the need for tape backup expands, you can add an additional 9U expansion module, with space for additional cartridges, tape drives and a redundant power supply.

9.1.5 IBM System Storage TS3400 Tape Library

The IBM System Storage TS3400 Tape Library offers the high capacity and performance advantage of the IBM System Storage TS1130 or TS1120 Tape Drives in a smaller automation footprint for IBM Power Systems, System x, System z and other open systems environments.

9.1.6 IBM System Storage TS3500 Tape Library

The IBM System Storage TS3500 Tape Library combines IBM automation and drive technology to provide a highly scalable, automated tape library for System z and open systems backup and archive in midrange to enterprise environments.

9.1.7 IBM TotalStorage 3494 Tape Library

The IBM TotalStorage 3494 Tape Library is an excellent solution for today's large storage requirements. Modular, flexible, and reliable, the 3494 uses 3590 tape drives, is cost-effective, and is backed by the service expertise of the IBM service organization. The 3494 offers a variety of models and features to fit all your needs and will grow as your business grows.

9.2 IBM System Storage TS2900 Tape Autoloader Express

The IBM System Storage TS2900 Tape Autoloader (Machine Type 3572) provides compact high capacity, low-cost solutions for simple unattended data backup. The library has a compact 1U form factor with easy access to tape cartridges by a removable magazine. The IBM System Storage TS2900 Tape Autoloader is an external stand-alone or rack-mountable unit that incorporates an IBM Ultrium 3 or Ultrium 4 Half-High Tape Drive. It is equipped with a Serial Attached SCSI (SAS) host adapter attachment that has a data transfer rate of up to 3 Gbps. See Figure 9-1.



Figure 9-1 IBM System Storage TS2900 Tape Autoloader Express

The IBM System Storage TS2900 Tape Autoloader has a removable cartridge magazine providing nine data cartridge slots, including a configurable two slot I/O station. IBM System Storage TS2900 Tape Autoloader is an entry point for IBM Linear Tape-Open (LTO) tape automation. This autoloader uses the IBM patented high density (HD) slot technology.

The cartridges supported in the IBM System Storage TS2900 Tape Autoloader include:

- ▶ IBM LTO Tape Cartridge Ultrium 4 (800 GB native physical capacity)
- ▶ IBM LTO Tape Cartridge Ultrium 3 (400 GB native physical capacity)
- ▶ IBM LTO Tape Cartridge Ultrium 2 (200 GB native physical capacity)
- ▶ IBM LTO Tape Cartridge Ultrium 1 (100 GB native physical capacity)
- ▶ WORM cartridges

Note that cartridge support is dependent on the tape drive that is installed in the IBM System Storage TS2900 Tape Autoloader. Only the cartridge types that are supported by the LTO3 or LTO4 drive that is installed are supported in the TS2900.

The library media capacity can be further increased by using hardware compression (2:1 compression factor). The library supports the movement of any media generation in any slot or drive. It is set in the host application to limit any operations that might result in unsupported media issues.

The TS2900 has a 3 Gbps single-port SFF-8088 SAS connector. This is the same connector for all of the other IBM tape products that support the SAS interface. The drive is integrated into the library. There is no drive field-replaceable unit (FRU) or client-replaceable unit (CRU), because the entire library is a CRU.

Designed for tape automation, The TS2900 supports 3 Gbps attachment to IBM Power Systems, IBM System x, Intel and other open systems server platforms. To determine the latest supported servers, see the website:

<http://www.ibm.com/systems/storage/tape/>

The IBM Ultrium 4 LTO Tape Drive with a Fibre Channel or a SAS interface is encryption-capable. The IBM LTO4 Tape Drive supports Application-Managed Encryption (AME), System-Managed Encryption (SME), and Library-Managed Encryption (LME).

9.2.1 TS2900 features

These are the features for the IBM System Storage TS2900 Tape Autoloader Express Model:

- ▶ IBM Half-high LTO 4 supports the encryption of data for increased security with archived data
- ▶ Supports HH LTO 3 or HH LTO4 tape technology from IBM with 3 Gbps SAS interface
- ▶ A single I/O station to help support continuous library operation
- ▶ Removable magazine to facilitate the offsite relocation of media and archival data
- ▶ Features IBM half-high LTO technology designed for reliable performance in small-to-medium open system environments
- ▶ HH LTO 3 tape drives can read and write to LTO 2 Media and read LTO 1 media; HH LTO 4 tape drives can read and write to LTO 3 media and read LTO 2 media
- ▶ Supports WORM media
- ▶ Bar code reader allows for the option of synchronous or random access of media cartridges, facilitating operation
- ▶ Adheres to LTO specifications

9.2.2 TS2900 specifications

Table 9-1 summarizes the capabilities of the IBM System Storage TS2900 Tape Library.

Table 9-1 IBM System Storage TS2900 Tape Autoloader Express

Model	IBM System Storage TS2900 Tape Autoloader Express
Drive options	Ultrium 3 Half high: 3 Gbps SAS Ultrium 4 Half high: 3 Gbps SAS
Optional Feature codes	N/A
Tape drive type	IBM LTO Ultrium 3 Half-Height and Full Height IBM LTO Ultrium 4 Half-Height and Full Height
Number of drives	1
Number of tape cartridges	9
Number of mail slots	1
Physical capacity	LTO 3: Up to 3.6 TB (7.2 TB with 2:1 compression) LTO 4: Up to 7.2 TB (14.4 TB with 2:1 compression)
Data transfer rate	LTO3: Up to 60 MBps LTO4: Up to 120 MBps
Attachment and systems support	The TS2900 supports 3 Gbps attachment to IBM Power Systems, IBM System x, Intel and other open systems server platforms.

For more information about the IBM System Storage TS2900 Tape Autoloader Express, see the Redbooks publication, *IBM System Storage Tape Library Guide for Open Systems*, SG24-5946.

For specific details, setup, and operations help, see the website:

<http://www.ibm.com/support/docview.wss?uid=publgc27221200>

9.3 IBM System Storage TS3100 Tape Library Express Model

IBM System Storage TS3100 Tape Library provides compact, high-capacity, low-cost solutions for simple, unattended data backup. It is designed to support the new IBM LTO Ultrium 4 Half-High (HH) Tape Drive and the IBM LTO Ultrium 3 HH Tape Drive, to help increase capacity and performance with Low Voltage Differential (LVD) SCSI and 3 Gb Serial Attached SCSI (SAS) attachments.

TS3100 media capacity is up to 8.8 TB (17.6 TB with 2:1 compression) data storage per unit. Figure 9-2 shows the IBM System Storage TS3100 Tape Library.



Figure 9-2 IBM System Storage TS3100 Tape Library

The IBM System Storage TS3100 Tape Library Express Model comes as two separate products. They are very similar with the basic difference being the amount of drives that can be installed. The two express model products are as follows:

- ▶ The IBM System Storage TS3100 Tape Library Express Model comes with a single Ultrium 3 or Ultrium 4 tape drive with 24 cartridge capacity.
- ▶ The IBM System Storage TS3100 Tape Library Express Model featuring half-high Ultrium technology allows up to 2 Ultrium Half-High tape drive(s) and 24 tape cartridge capacity.

The IBM System Storage TS3100 Tape Library is available as follows:

- ▶ 3573 2UL - IBM System Storage TS3100 Tape Library Model L2U Driveless:
 - Ultrium 4 Full Height: LVD SCSI (95P5002), 3 Gbps SAS (95P5006); Fibre Channel (95P5004)
 - Ultrium 3 Full Height: LVD SCSI (23R7260); 4 Gbps Fibre Channel (23R7261)
 - Ultrium 4 Half-Height: 3 Gbps SAS (45E2243)
 - Ultrium 3 Half-Height: LVD SCSI (95P4998); 3 Gbps SAS (95P5000)

9.3.1 TS3100 features

These are the features for the IBM System Storage TS3100 Tape Library Express Model:

- ▶ IBM LTO Ultrium 3 or IBM LTO Ultrium 4 in either full or half height technologies
- ▶ Supports 4 Gbps Fibre Channel and LVD SCSI attachment
- ▶ Ultrium 4 native data rate of 120 MBps
- ▶ Ultrium 3 Full Height native data rate up to 80 MBps
- ▶ Ultrium 3 Half Height native data rate up to 60 MBps
- ▶ Ultrium 4 native data capacities up to 19.2 TB (38.4 with 2:1 compression)
- ▶ Ultrium 3 native data physical capacities up to 9.6 TB (up to 19.2 TB using 2:1 compression)
- ▶ 2U form factor with 24 data-cartridge slots. The cartridges are housed in two removable magazines, left and right.
- ▶ A single mail slot on the left cartridge magazine provides import/export facilities for cartridges, without interrupting library operation
- ▶ Sequential or random access mode with a standard bar-code reader
- ▶ Operator front control panel contains power button, front panel LEDs, control keys, and the operator control panel display.
- ▶ Remote library management through a web interface

9.3.2 TS3100 specifications

Table 9-2 summarizes the capabilities of the IBM System Storage TS3100 Tape Library.

Table 9-2 IBM System Storage TS3100 Tape Library Express specifications

Model	TS3100 Tape Library Express
Drive options	Ultrium 4 Full Height: LVD SCSI (95P5002), 3 Gbps SAS (95P5006); Fibre Channel (95P5004) Ultrium 3 Full Height: LVD SCSI (23R7260); 4 Gbps Fibre Channel (23R7261) Ultrium 4 Half-Height: 3 Gbps SAS (45E2243) Ultrium 3 Half-Height: LVD SCSI (95P4998); 3 Gbps SAS (95P5000)
Optional feature codes	Transparent LTO Encryption 45E3081 Rack mount 23R6998 Right-side magazine set 23R6999 Left-side magazine 45E2237 Ultrium Tape Cartridges Ultrium 4 media: 95P4278 Ultrium 3 media: 95P2020 Ultrium Cleaning Cartridge 23R7008
Tape drive type	IBM LTO Ultrium 3 Half-Height and Full Height IBM LTO Ultrium 4 Half-Height and Full Height
Number of drives	1-2
Number of tape cartridges	24
Number of mail slots	1

Physical capacity	Up to 1.6 TB per cartridge compressed; 800 GB native with LTO4 Up to 800 GB per cartridge compressed; 400 GB native with LTO3 Up to 38.4 TB per tape library compressed; 19.6 TB native Up to 19.2 TB per tape library compressed; 9.6 TB native
Data transfer rate	Up to 120 MBps native with LTO Ultrium 4 Up to 80 MBps native with LTO Ultrium 3
Attachment and systems support	The TS3100 Full-Height features LVD SCSI, 4 Gbps fibre- channel and 3 Gbps SAS interfaces. The TS3100 Half-Height features LVD SCSI and 3 Gbps SAS interfaces, both form factors attaching to IBM Power Systems, IBM System x, HP-UX, Sun Solaris, UNIX, Linux and Windows servers and non-IBM servers, workstations and personal computers that support the interface specification

For the most up to date and detailed operating system and attachment requirements, see:

http://www.ibm.com/servers/storage/tape/compatibility/pdf/ts3100_ts3200_interop.pdf

For more information about the IBM System Storage TS3100 Tape Library, see the Redbooks publication, *IBM System Storage Tape Library Guide for Open Systems*, SG24-5946.

For more technical information, see the *IBM System Storage TS3100 Tape Library and TS3200 Tape Library Setup, Operator and Service Guide*, GA32-0545-07, at the website:

<http://www.ibm.com/support/docview.wss?uid=publga32054507>

9.4 IBM System Storage TS3200 Tape Library

Just like the IBM System Storage TS3100 Tape Library, the IBM System Storage TS3200 Tape Library is a dual or single drive tape library, offering high capacity and performance technology for the midrange environments. The TS3200 Tape Library is an external 4U stand-alone or rack-mountable unit that incorporates up to two full height Linear Tape-Open (LTO) IBM TotalStorage Ultrium 3 or Ultrium 4 tape drives or four Ultrium 3 or Ultrium 4 half height tape drives.

The IBM System Storage TS3200 Tape Library Express Model is an excellent tape storage solution for organizations with existing digital linear tape or requiring high-performance automated tape backup. The TS3200 is also designed for organizations that have limited physical space in their IT environments. Operating in a rack environment allows organizations the advantage of placing the TS3200 in a standard 19-inch rack, which provides 76.8 TB of compressed tape storage in just a 4U space. Remote management and a bar code reader are standard in the library, and it can run in sequential or random access mode. Optional features available are rack mount kit, additional power supply, and Path Failover.

The TS3200 has IBM patented Multi-Path Architecture for sharing the library robotics. This allows a library with two drives to be partitioned into two logical libraries, for sharing between servers and/or applications.

Figure 9-3 shows the IBM System Storage TS3200 Tape Library.



Figure 9-3 IBM System Storage TS3200 Tape Library

The IBM System Storage TS3200 Tape Library is available in the following configurations:

- ▶ 3573 4UL: TS3200 Tape Library Model L4U Driveless:
 - Ultrium 4 Full Height: LVD SCSI (95P5002), 3 Gbps SAS (95P5006); Fibre Channel (95P5004)
 - Ultrium 3 Full Height: LVD SCSI (23R7260); 4 Gbps Fibre Channel (23R7261)
 - Ultrium 4 Half-Height: 3 Gbps SAS (45E2243)
 - Ultrium 3 Half-Height: LVD SCSI (95P4998); 3 Gbps SAS (95P5000)
- ▶ Designed to support up to two IBM LTO Ultrium Full Height Tape Drives or up to four IBM LTO Ultrium Half Height Tape Drives, to help increase capacity and performance.

9.4.1 TS3200 features

These are the features for the IBM System Storage TS3200 Tape Library:

- ▶ Designed to support the new IBM Linear Tape-Open (LTO) Ultrium 4 Half-High (HH) Tape Drive and the IBM LTO Ultrium 3 HH Tape Drive, to help increase capacity and performance with Low Voltage Differential (LVD) SCSI and 3 Gb Serial Attached SCSI (SAS) attachments
- ▶ 4U form factor with 48 data-cartridge, four removable magazines, including a three-slot I/O station
- ▶ Sequential or random access mode with a standard bar-code reader
- ▶ Front panel, which integrates the power button, front panel LEDs, control keys, and the operator control panel display, for direct library management
- ▶ Remote library management through a web interface

9.4.2 TS3200 specifications

Table 9-3 shows the specifications of the TS3200.

Table 9-3 IBM System Storage TS3200 Tape Library specifications

Model	Type 3573 Models L4U, L3H, F3H
Drive options	Ultrium 4 Full Height ▶ LVD SCSI (95P5002) ▶ 3 Gbps SAS (95P5006) ▶ Fibre Channel (95P5004) Ultrium 3 Full Height ▶ LVD SCSI (23R7260) ▶ 4 Gbps Fibre Channel (23R7261) Ultrium 4 Half-Height ▶ 3 Gbps SAS (45E2243) Ultrium 3 Half-Height ▶ LVD SCSI (95P4998) ▶ 3 Gbps SAS (95P5000)
Optional feature codes	Path Failover 45E9503 Transparent LTO Encryption 45E3081 Rack mount 23R6998 Right-side magazine set 23R6999 Left-side magazine, upper 45E2225 Left-side magazine, lower 45E2231 Ultrium Tape Cartridges Ultrium 4 media: 95P4278 Ultrium 3 media: 95P2020 Ultrium Cleaning Cartridge 23R7008 Additional Power Supply 23R7263
Tape drive type	IBM LTO Ultrium 3 Full Height and Half-Height IBM LTO Ultrium 4 Full Height and Half-Height
Number of drives	1-4
Number of tape cartridges	48
Number of mail slots	3
Physical capacity	Up to 1.6 TB per cartridge compressed; 800 GB native with LTO Ultrium 4 Up to 800 GB per cartridge compressed; 400 GB native with LTO Ultrium 3 Up to 76.8 TB per tape library compressed; 38.4 TB native with LTO Ultrium 4 Up to 38.4 TB per tape Library compressed; 19.2 TB native with LTO Ultrium 3
Data transfer rate	Up to 120 MBps native with LTO Ultrium 4 Full Height and Half-Height Up to 80 MBps native with LTO Ultrium 3 Full-Height Up to 60 MBps native with LTO Ultrium 3 Half-Height
Attachment and systems support	The TS3200 Full-Height features LVD SCSI, 4 Gbps fibre- channel and 3 Gbps SAS interfaces. The TS3200 Half-Height features LVD SCSI and 3 Gbps SAS interfaces, both form factors attaching to IBM Power Systems, IBM System x, HP-UX, Sun Solaris, UNIX, Linux and Windows servers and non-IBM servers, workstations and personal computers that support the interface specifications.

For the most up to date and detailed operating system and attachment requirements, see:

http://www.ibm.com/servers/storage/tape/compatibility/pdf/ts3100_ts3200_interop.pdf

For more information about the IBM System Storage TS3200 Tape Library, see the Redbooks publication, *IBM System Storage Tape Library Guide for Open Systems*, SG24-5946.

For more technical information, see the IBM System Storage TS3100 Tape Library and TS3200 Tape Library Setup, Operator and Service Guide (GA32-0545-07) at the website:

<http://www.ibm.com/support/docview.wss?uid=publga32054507>

9.5 IBM System Storage TS3310 Tape Library

Designed around a 5U high modular base library unit, the TS3310 is designed to scale vertically with expansion for LTO tape cartridges, drives and redundant power supplies. The base library module, model L5B, is the entry point for the product family. It contains all of the necessary robotics and intelligence to manage the 5U high library system, which houses up to 41 cartridges (35 storage slots and 6 Input/Output slots) and two LTO generation 4 and/or generation 3 tape drives.

Each model E9U contains 92 physical LTO cartridge storage cells and space for up to four LTO 4 and/or LTO 3 tape drives. The TS3310 supports either the standard or WORM LTO data cartridge. Additionally, the E9U has space for up to two (one redundant) power supply modules. (At least one power supply module must be installed if a drive is present in the E9U.)

Figure 9-4 shows the IBM TS3310 base L5B unit with one expansion module.



Figure 9-4 IBM System Storage TS3310 Tape Library Tape Library - Base and Expansion

9.5.1 TS3310 features

These are the features for the IBM System Storage TS3310 Tape Library (Figure 9-5):

- ▶ TS3310 is a modular library, 56 TB up to 641 TB (2:1 compression) (35 to 401 LTO storage slots):
 - From 2 to 18 4 Gbps Fibre Channel and 3 Gbps SAS (LTO4 only), SCSI-U160 (LTO3 only) hot-swappable tape drives
 - Vertical expansion over 400 LTO cartridges
 - Supports Logical Partitioning using Multipath Architecture
 - Optional features; Capacity on demand, Rack Mounting, Path Failover (Data and Control), Redundant power
- ▶ TS3310 Base library module supports, 28 TB native capacity (35 slots), up to two drives, 6 slot cartridge I/O station, 2 logical libraries, robotics and control logic, barcode reader and remote management capability are standard, native SMI-S support, and license key to support initial E9U expansion.

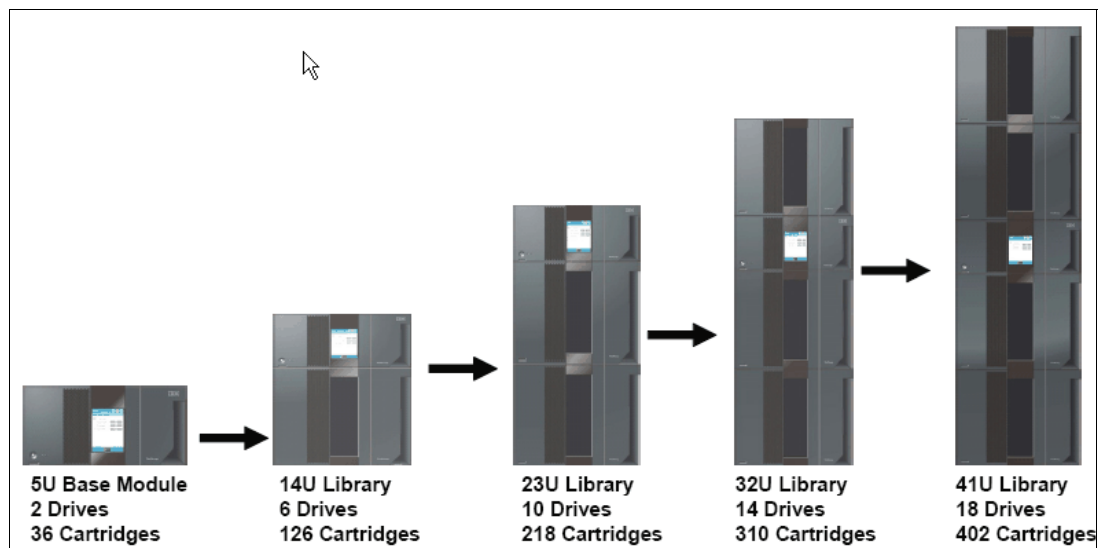


Figure 9-5 TS3310 Upgrade Path

- ▶ One to four expansion modules can be installed, as you can see in Figure 9-5.

The addition of each expansion module provides a total of 92 storage slots, which are enabled through the use of “Capacity on Demand” licenses. Note that the base module ships with licensing for 82 slots. This effectively means that when you add the first expansion module, half the slots are immediately available without the requirement to acquire additional licenses.

You can configure the I/O slots in each module as a group to operate as I/O slots or storage slots (at least one group must operate as I/O slots). This provides a high degree of flexibility. In a library configuration of one L5B base unit and one 9EU expansion unit, for example, there are three possible combinations:

- A total of 6 I/O slots (all in the base module)
- A total of 12 I/O slots (using the slots in the expansion module)
- A total of 18 I/O slots (using the slots in the base module and the expansion module)

For the latest features, functions, and specifications regarding the TS3310, see:

<http://www.ibm.com/systems/storage/tape/ts3310/index.html>

For more technical information, library setup, and configuration guidance, see the website:

http://www.ibm.com/support/docview.wss?rs=1140&context=STCXRH&dc=DA400&q1=ssg1*&uid=ssg1S7001840&loc=en_US&cs=utf-8&lang=en

9.5.2 TS3310 specifications

The specifications for IBM System Storage TS3310 Tape Library are listed in Table 9-4.

Table 9-4 IBM System Storage TS3310 Tape Library Specifications

Type and Model	3576 Model L5B	3576 Model L5B & 4 E9U Models
Configuration	Base library	Base library and 4 expansion units
LTO storage slots (max)	35	401
LTO Input/Output slots (min/max)	6	54
Maximum tape drives	2	18
Total Physical Capacity 2:1 compression	56 TB	641.6 TB
Capacity on Demand increments	n/a	46 cartridges
Maximum logical libraries	2	18
Hot-swap components	Power supplies, tape drives, SMI-S enabled	
Systems management		
Operating systems/servers supported	OS/400, V5R2, V5R3, or later AIX 5L™ V5.1, V5.2, V5.3 or later Sun Solaris 8 and 9 Microsoft Windows 2000 Microsoft Windows 2003 HP-UX 11.0, HP-UX 11.11 and HP-UX 11.231 Linux distributions: Red Hat Enterprise Linux Version 4 SUSE Linux Enterprise Server 9	

9.6 IBM System Storage TS3400 Tape Library

The IBM System Storage TS3400 Tape Library offers the high capacity and performance advantage of the IBM System Storage TS1130 or TS1120 Tape Drives in a smaller automation footprint for IBM Power Systems, System x, System z, and other open systems environments (Figure 9-6).

The TS3400 tape library is an external 5U stand-alone or rack-mountable unit supporting one or two TS1130 tape drives or one to two TS1120 tape drives.

The IBM System Storage TS1130 Tape Drive has a native capacity of 1000 GB (1.0 TB) when using the IBM 3592 Extended Data Cartridge (JB), or 640 GB when using the IBM 3592 Enterprise Cartridge (JA). The IBM TS1130 Tape Drive has a dual-ported switched fabric 4 Gbps Fibre Channel attachment and a data transfer rate of up to 160 MBps per drive.

The IBM System Storage TS1120 Tape Drive has a native capacity of 700 GB when using the IBM Extended Data Cartridge (JB), or 500 GB when using the IBM Data Cartridge (JA). The IBM System Storage TS1120 Tape Drive has a native rate of up to 104 MB/s. The IBM System Storage TS1120 Tape Drive has a dual-ported switched fabric 4 Gbps Fibre Channel attachment. The tape drives must be ordered separately with the final order.

The IBM System Storage TS3400 Tape Library supports the IBM TS1120 and the TS1130 built-in encryption capabilities. The supported encryption methods are Application-Managed Encryption (AME), System-Managed Encryption (SME), and Library-Managed Encryption (LME).

The TS3400 tape library has two removable cartridge magazines, providing 18 data cartridge slots. Up to three slots can be used for I/O slots, and up to two slots can be used as cleaning cartridge slots. The TS3400 tape library provides a media capacity of up to 18 cartridges, allowing for up to 18TB of storage (54TB with 3:1 compression) when using 1 TB extended capacity cartridges.

Remote management by a web browser allows you to communicate directly with the library and perform a wide range of end user, operator and administrator tasks without being at the operator panel. To take further advantage of your investment, you can partition the library into two logical partitions to share the library between various applications. To support continued processing, the TS3400 tape library is equipped with control path and datapath failover.



Figure 9-6 IBM System Storage TS3400 Tape Library

9.6.1 TS3400 features

These are the features for the IBM System Storage TS3400 Tape Library:

- ▶ Supports TS1130 or TS1120 tape drives in a small form factor
- ▶ Supports 1 to 2 TS1130 or TS1120 tape drives
- ▶ Native data rate of up to 160 MB per drive
- ▶ Capacity for growth up to 18 TB in a 5U form factor
- ▶ Media capacity of up to 18 TB (54 TB with 3:1 compression)
- ▶ Control path and data path automatic failover:
 - Enhances the autonomic capabilities of the library
 - Designed to improve the reliability of your backup processes
- ▶ Supports IBM tape drive encryption:
 - Helps secure sensitive at-rest data
 - Built-in encryption is designed to prevent performance drain on host systems

For latest information about the TS3400, see:

<http://www.ibm.com/systems/storage/tape/ts3400/index.html>

You can download the latest *IBM System Storage TS3400 Tape Library Planning and Operator Guide* at the following website:

<http://www.ibm.com/support/docview.wss?uid=ssg1S7001901>

9.6.2 TS3400 specifications

Table 9-5 summarizes the capabilities of the IBM System Storage TS3400 Tape Library.

Table 9-5 IBM System Storage TS3400 Tape Library Specifications

Model	IBM System Storage TS3400 Tape Library
Tape Drive	TS1130 or TS1120 tape drives
Number of Drives	Up to 2, no intermix of TS1130 and TS1120
Number of Tape Cartridges	18
Number of Input/Output slots	3
Number of Logical Libraries	2
Capacity	Up to 18 TB native capacity (54 TB with 3:1 compression)
Data Transfer Rate	Up to 320 MBps
Media Type	IBM 3592 JA/JB/JJ and JW/JR/JX Write Once Read Many (WORM) cartridge
Platform/Operating System Supported	IBM System i i5/OS, IBM i System p AIX and Linux System x See open systems support System z IBM z/OS, IBM VSE and Linux Power Systems AIX, IBM i, Linux

9.7 IBM System Storage TS3500 Tape Library

The IBM System Storage TS3500 Tape Library is a highly scalable, automated tape library for mainframe and open systems backup and archive in midrange to enterprise environments. It is formerly known as IBM TotalStorage Tape Library 3584.

The TS3500 tape library supports System z by the IBM 3953 Tape System (3953 tape system) or the IBM Virtualization Engine TS7740, which enable System z hosts to access the TS3500 tape library cartridge inventory and allow connection to TS1130, TS1120 and IBM TotalStorage 3592 Model J1A tape drives.

The TS3500 tape library can support up to four 3953 tape systems, and up to eight IBM Virtualization Engine TS7740 subsystems per physical library.

The TS3500 tape library's entry base frame provides a more flexible upgrade path for users who want to expand their tape storage at time of need. Capacity on Demand (CoD) configurations for TS3500 tape library L-frame models include the "entry level" configuration, an "intermediate" configuration, and a "full capacity" configuration. S-frame models also allow two CoD configurations.

The TS3500 Model HA1 allows two robotic accessories to operate simultaneously in two to 16 frame configurations. Figure 9-7 shows the minimum and maximum 16 frame TS3500 configuration.



Figure 9-7 IBM System Storage TS3500 Tape Library Minimum and Maximum Configuration

9.7.1 TS3500 features

These are the features for the IBM System Storage TS3500 Tape Library:

- ▶ TS7700 Virtualization Engine:
 - Allows quick access to active data on “virtual volumes” on disk and reduces overall solution cost storing inactive data on “volumes” on tap
 - Optimize data read and write to tape
 - Supports tape automation in a System z environment
- ▶ 3593 tape system:
 - Supports tape automation in a System z environment
 - Also supports the TS7700 Virtualization Engine
- ▶ Dual accessor option:
 - Helps increase the mount performance and overall system availability
 - Two accessors can operate simultaneously in multiple frame configurations
- ▶ Flexible upgrade path:
 - Start with an entry base frame and expand at the time of need
 - Capacity on Demand configurations available
- ▶ Remote, web browser-based management, can access key management functions from any network connected computer
- ▶ Redundant power supplies and AC feeds
- ▶ Control-path and data-path automatic failover
- ▶ Mixed LTO Ultrium and 3592 media
- ▶ Support for:
 - A heterogeneous environment
 - TS1130/TS1120 and TS1040 tape drive encryption
- ▶ Helps you secure sensitive at-rest data

9.7.2 TS3500 specifications

The specifications for IBM System Storage TS3500 Tape Library are listed in Table 9-6.

Table 9-6 IBM System Storage TS3500 Tape Library specifications

Model	IBM System Storage TS3500 Tape Library
Frame definition	L23 – base frame for TS1130, TS1120 or 3592, D23 – drive capable expansion frame for TS1130, TS1120 or 3592, S24 – storage only expansion frame for 3592 cartridges L53 – base frame for LTO, D53 – drive capable expansion frame for LTO, S54 – storage only expansion frame for LTO cartridges HA1 – High Availability service bay frame for use with the dual accessor feature
Tape drive types	TS1130, TS1120 or 3592 tape drives or IBM LTO Ultrium 4, 3, or 2 Tape Drives

Model	IBM System Storage TS3500 Tape Library
Number of frames	One base frame, up to 15 expansion frames The TS3500 Model HA1 installation provides one of the two additional frames required as service bays in a dual accessor library
Number of drives	Up to 12 per frame (up to 192 per library)
Number of tape cartridges	L23 – up to 260; D23 – up to 400; S24 – up to 1,000; Total supported: 15,000 L53 – up to 287; D53 – up to 440; S54 – up to 1,320; Total supported: 20,000
Number of logical libraries	Maximum of 192 (up to number of drives installed)
Number of 3953 Systems	Maximum of four per TS3500 subsystem
Number of TS7700 Virtualization Engine	Maximum of eight per TS3500 subsystem
Capacity	IBM Ultrium 4 Cartridges: up to 16 PB (32 PB with 2:1 compression) 3592 extended capacity cartridges: up to 15 PB (45 PB with 3:1 compression)
Media type	L23/D23/S24: IBM 3592 JA/JJ/JB and JW/JR/JX Write Once Read Many (WORM) cartridges L53/D53/S54: IBM LTO Ultrium 4, 3, 2 Cartridges
Attachment and systems support	TS3500 tape library can attach to IBM Power Systems, System p, System i, and System x servers and non-IBM servers, workstations Attach to System z can be made by the IBM 3953 Tape System (3953 tape system) or the IBM Virtualization Engine TS7740
Operating systems support	Device driver support is available for IBM AIX; IBM OS/400; IBM i; Windows 2000; Windows Server 2003; Linux; Sun Solaris; and HP-UX

To become more familiar with the IBM System Storage TS3500 Tape Library, you can download the *IBM System Storage TS3500 Tape Library Introduction and Planning Guide IBM 3584 Tape Library* at the website:

http://www.ibm.com/support/docview.wss?rs=1159&context=STCMML8&dc=DA400&uid=ssg1S7001944&loc=en_US&cs=utf-8&lang=en

For the latest specifications, features and functions go to:

<http://www.ibm.com/systems/storage/tape/ts3500/index.html>

9.7.3 Advanced Library Management System

The Advanced Library Management System (ALMS) is an available feature that supports dynamic storage management, allowing the user to dynamically create and change logical libraries and configure any drive into any logical library. ALMS is required when attaching to mainframe environments and also when integrating high-density frames into a library configuration. Based on capacity requirements, ALMS can be implemented as an “entry,” “intermediate” or “full” Capacity on Demand level in the TS3500 tape library.

ALMS virtualizes the locations of cartridges in the TS3500. Logical libraries can then consist of unique drives and ranges of volume serial numbers instead of fixed locations.

ALMS is an extension of IBM patented Multi-Path Architecture. With ALMS, the TS3500 is the industry's first standards-based tape library to virtualize the locations of cartridges (called SCSI element addresses) while maintaining native SAN attachment for the tape drives. ALMS enables logical libraries to consist of unique drives and ranges of volume serial (VOLSER) numbers, instead of fixed locations.

When you enable ALMS with its license key, you can assign tape drives to any logical library by using the Tape Library Specialist web interface. Logical libraries can also be added, deleted, or easily changed without disruption. Storage capacity can be changed without impact to host applications.

ALMS offers dynamic management of cartridges, cartridge storage slots, tape drives, and logical libraries. It enables the TS3500 to achieve unprecedented levels of integration for functionality through dynamic partitioning, storage slot pooling, and flexible drive assignment. ALMS eliminates downtime when you add Capacity on Demand (CoD) storage, add or remove logical libraries, or change logical library storage allocation. ALMS also reduces downtime when you add expansion frames, add or remove tape drives, or change logical drive allocation.

9.7.4 Virtual I/O slots

The IBM TS3500 Tape Library has I/O stations and I/O slots that enable you to import and export up to 32 cartridges at any given time. The I/O slots are also known as import/export elements (IEEs). Virtual I/O (VIO) slots increase the quantity of available I/O slots by allowing storage slots to appear to the host as I/O slots. Storage slots that appear to the host as I/O slots are called virtual import/export elements (VIEEs). The goal of virtual I/O slots is to reduce the dependencies between the system administrator and library operator so that each person performs their import and export tasks without needing the other person to perform any actions. With virtual I/O slots, the library automatically moves cartridges from the I/O stations to physical storage slots and from physical storage slots to the I/O stations.

With virtual I/O slots, you can configure up to 255 VIEEs per logical library. Each logical library will have a unique VIEE address space that is not accessible by other logical libraries. New logical libraries will, by default, be assigned the maximum number of virtual I/O slots, while logical libraries that were defined (before ALMS is enabled) initially have the number of physical I/O slots in the library.

For more information covering ALMS and Virtual I/O slots, see IBM System Storage TS3500 Tape Library Introduction and Planning Guide IBM 3584 Tape Library at:

http://www.ibm.com/support/docview.wss?rs=1159&context=STCMML8&dc=DA400&uid=ssg1S7001944&loc=en_US&cs=utf-8&lang=en

9.8 IBM TotalStorage 3953 Tape System for System z

The IBM System Storage TS3500 Tape Library supports the attachment of IBM System z host systems to IBM System Storage TS1130 (IBM 3592-E06/EU6), IBM System Storage TS1120 (IBM 3592-E05), and IBM TotalStorage 3592 Model J1A installed inside the TS3500.

9.8.1 Attachment methods

The System z hosts attach using either Enterprise System Connection (ESCON) or Fiber Connection (FICON) to one of the following products:

- ▶ IBM System Storage TS1120 Model C06 Tape Controller
- ▶ IBM TotalStorage 3592 Tape Controller Model J70
- ▶ IBM System Storage TS7700 Virtualization Engine (FICON only)
- ▶ IBM TotalStorage 3494 Virtual Tape Server (VTS) models B10 and B20

The IBM TS3500 Tape Library is managed on behalf of the System z hosts by an external Library Manager, the IBM 3953 Model L05. This Library Manager, the System z control units, and various switches are housed in a special frame type, the IBM 3953 Model F05 Frame, which is located physically outside the library enclosure. These two components, the Library Manager Model L05 and the 3953-F05 Frame in which it resides, make up the IBM 3953 Tape System. The IBM 3953 Library Manager has common microcode with the Library Manager that is used in the IBM 3494 Tape Library.

Beside the existing TS3500 attachment support for the Open Systems Fibre Channel environments for Linear Tape-Open (LTO) and 3592 drives, the TS3500 together with the IBM 3953 Tape System provides the additional attachment functionality for System z ESCON and FICON environments. Both platforms can be used in an isolated manner, or they can be consolidated into one shared library (Figure 9-8).

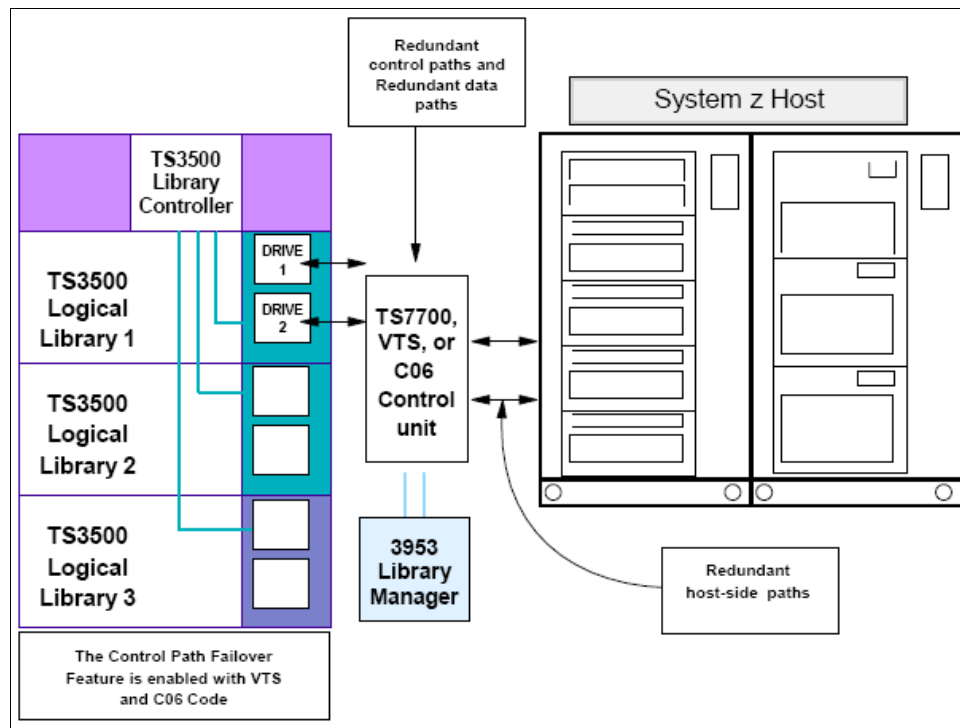


Figure 9-8 TS3500 and 3953 Library Manager Basic Diagram

As shown in Figure 9-8, the System z components are contained in their own logical library within the TS3500, using the multi-path capability of the library. The hosts connect to the TS3500 Tape Library only through the TS1120 Model C06 Controller, the TS7700 Virtualization Engine, or through the IBM VTS Model B10 or B20. The supported connection types are ESCON or FICON except for the TS7700 Virtualization Engine, which supports only FICON connections.

One IBM 3953 Model L05 Library Manager (or a pair of separate Library Managers), which manages the TS3500 logical library on behalf of the System z host, is provided. The IBM 3953-L05 is similar in function to the Library Manager in the IBM 3494 Tape Library.

The controllers, TS7700 Virtualization Engines, and VTSs in the subsystem communicate with the IBM 3953 Model L05 Library Manager or the pair of Library Managers, which control the library on behalf of the System z hosts. The Library Managers send simple SCSI commands to the TS3500 library robotics using the controller, TS7700 Virtualization Engine, or VTS path to the logical unit number 1 (LUN 1) address of the drives.

The communication between the IBM 3953 Library Manager and the tape controllers is through an internal LAN.

For additional information, see the *IBM 3953 Tape System Introduction and Planning Guide* at the website:

<http://www.ibm.com/support/docview.wss?uid=ssg1S7001862>

9.8.2 3953 Tape System features

These are the features for the IBM TotalStorage 3953 Tape System for System z:

- ▶ Builds on recent tape library enhancements:
 - Dual robotics and nondisruptive frame addition
 - Advanced Library Managements System
 - TS1120 and 3592 J1A tape drive support
- ▶ Supports VTS B10 / B20 and TS1120 controller; attached to TS1120 and 3592 J1A tape drives
- ▶ Supports System z operating systems; Linux is supported natively
- ▶ Supports open systems:
 - Coexists with open systems in their own logical library
 - Securely isolated from System z environment
- ▶ 3953 L05 Library Manager:
 - Functionally equivalent to 3494 Library Manager
 - Employs identical microcode and server
- ▶ The 3953 L05 is mounted in the first 3953 F05 frame:
 - Requires dual internal LAN features
 - Requires TotalStorage Master Console (TSMC)
- ▶ Optional 3953 L05 provides redundancy:
 - Mounted in first 3953 F05 frame
 - Facilitates microcode updates
- ▶ Issues TS3500 mount requests:
 - Accepts PLF commands from VTS or TS1120 controller
 - Issues SCSI media manager 'mount request

Figure 9-9 shows a 3593 configuration.

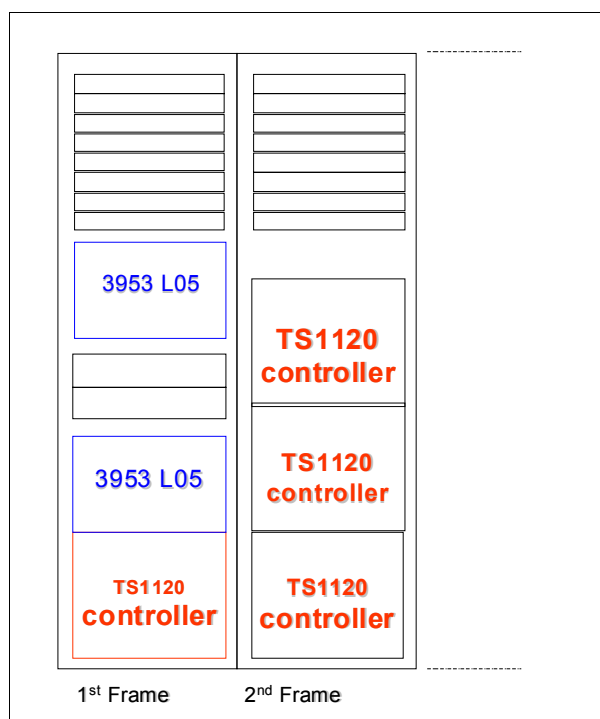


Figure 9-9 IBM TotalStorage 3593 Tape System for System z

9.9 Discontinued products

Here we list the discontinued Tape Libraries and their replacements:

- ▶ IBM TotalStorage 3494 Tape Library: The IBM TotalStorage 3494 Tape Library has been replaced by the IBM Systems Storage TS3500 Tape Library.
- ▶ IBM TotalStorage 3581 Tape Autoloader: The IBM TotalStorage 3581 Tape Autoloader has been replaced by the IBM System Storage TS3100 Tape Library Express Model.
- ▶ IBM TotalStorage 3582 Tape Library: The IBM TotalStorage 3582 Tape Library has been replaced by the IBM System Storage TS3200 Tape Library Express Model.
- ▶ IBM TotalStorage 3583 Tape Library: The IBM TotalStorage 3583 Tape Library has been replaced by the IBM System Storage TS3310 Tape Library.



IBM Tape Virtualization products

In this chapter we discuss tape virtualization and then describe the IBM Tape Virtualization products and solutions:

- ▶ IBM System Storage ProtecTIER products:
 - IBM System Storage TS7650 ProtecTIER Deduplication Appliance
 - IBM System Storage TS7650G ProtecTIER Deduplication Gateway
- ▶ IBM Virtualization Engine TS7700:
 - IBM Virtualization Engine Model TS7740
 - IBM Virtualization Engine Model TS7720
- ▶ IBM TotalStorage 3494 Virtual Tape Server Models B10/B20: Replaced by the IBM Virtualization Engine TS7700
- ▶ DFSMSrmm (formerly known as Enterprise Removable Media Manager (eRMM))

Note: The IBM TotalStorage 3494 Virtual Tape Server has been replaced by the IBM System Storage TS7700 Virtualization product.

10.1 Introduction to Tape Virtualization

In order to maintain continuous business operations, address regulatory requirements and archive business records, you need an infrastructure that enables you to manage your data from online application storage to offline, permanent archive media. Tape is a key part of both the backup and archive life cycle. Tape provides the lowest total cost of ownership alternative for securely storing long-term archives for record keeping and disaster recovery. As data centers and data stores grow, manual tape operations can become more complex. This growth can lead to increased backup and restore times and higher management overhead and costs. A solution to this is virtual tape.

A virtual tape library provides high performance backup and restore by using disk arrays and virtualization software. A virtual tape library is a unique blend of several storage tiers. The life cycle of data from its creation at the server level migrates by backup software to a virtual tape library. There is no better place for the data to be, because the virtual tape library is a combination of high performance SAN-attached disk and high performance servers emulating a tape storage device. At this level, you have many options for the data. For example, the data can remain on the virtual tape library indefinitely, as long as there is enough space, or it can be migrated to tape for off-site storage, archive, or both.

Any business that has a need for high performance backup and restore might need virtualization solutions. Instant access backups and restores are capabilities that only the disk storage tier can provide. All backup software products were designed to back data up to tape drives and libraries, and virtual tape library emulates tape. You can go directly to a virtual tape library and have your backups completed quickly.

Virtual tape libraries fill a void in the backup infrastructure for data that needs to be restored at a certain moment. Many restore requests often happen within six weeks of the data being backed up. Backup software can be configured to back up data to a virtual tape library and then create a virtual tape-to-tape copy for off-site deployment. It is no longer necessary to call the tapes back from an off-site location, unless data is required from years past.

One of the biggest challenges with backup planning today is that the amount of data being backed up is growing, but the time allotted for a backup (the backup window) is shrinking or remaining static. Applications need to be up and operational nearly 24 hours a day. To manage the need for increased data capacity and data protection, customers must find ways to shrink their backup windows and recover as quickly as possible.

In traditional backup environments, as shown in Figure 10-1, the client data is backed up in two ways:

- ▶ LAN clients write their backup data by the backup server to tape
- ▶ LAN-free clients utilize the SAN for direct backup to the tape devices

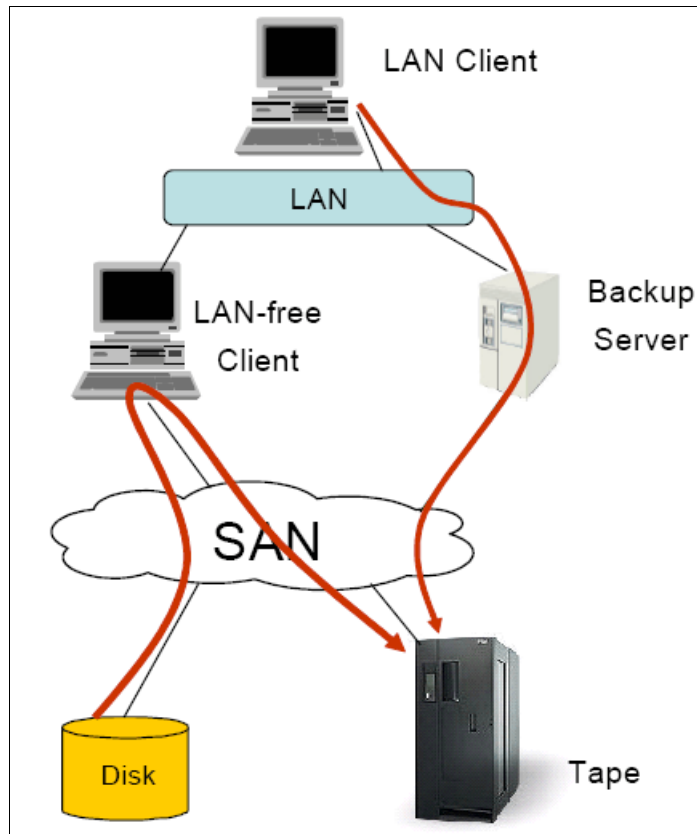


Figure 10-1 Traditional backup

The following problems are encountered with these traditional backups:

- ▶ LAN backups cannot benefit from fast tape drives, because the network transfer rate is the limiting factor; data cannot be delivered fast enough to keep the tape drives streaming.
- ▶ LAN-free clients require dedicated tape drives for every backup session. So, many tape drives are needed during the backup window.

A solution for these problems is to use disk to disk (D2D) or disk to disk to tape (D2D2T) backups.

LAN backups can be done by using a disk buffer, as shown in Figure 10-2. The client writes the backup data by LAN first to a disk buffer on the backup server. From there it can be recalled quite rapidly if needed. After the backup on disk is completed, the backup server can migrate the data from the disk buffer to a physical tape with no impact on the client.

Furthermore, the disk buffer works as a cache. The most recently backed up data resides in the fast disk buffer until its overwritten by newer data.

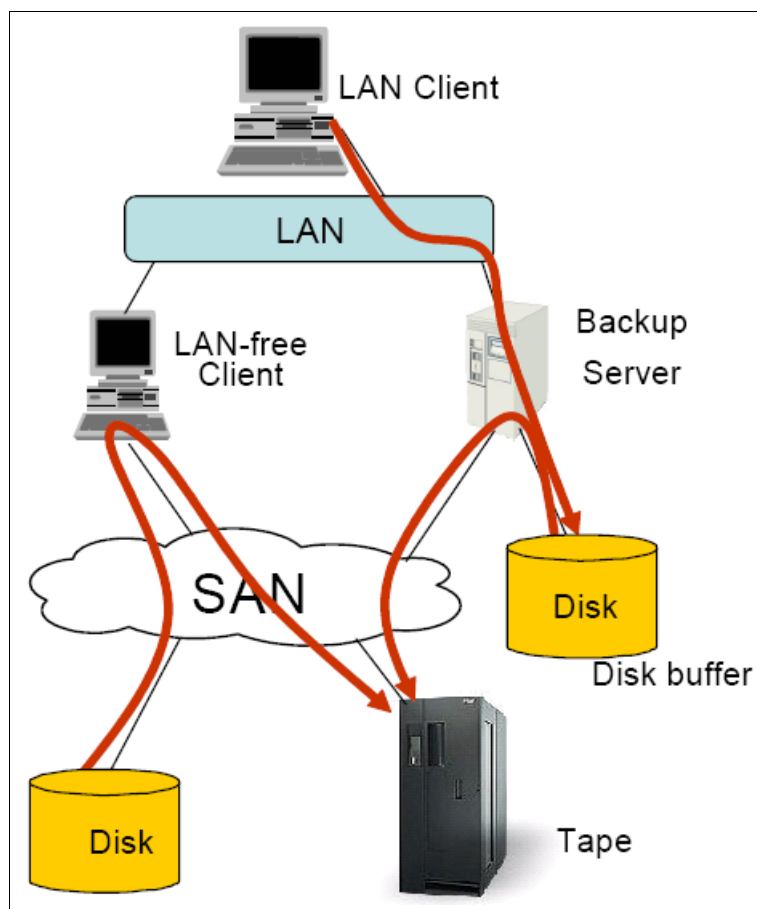


Figure 10-2 Backup to disk buffer

This is a good solution if the backup software supports it. IBM Tivoli Storage Manager supports a hierarchy of storage pools, and can automatically migrate from disk to tape, but other backup applications might not have this capability.

An alternative solution to optimizing backup, is to introduce a virtualization and emulation layer. This layer appears to both the backup server and the LAN-free clients as a tape drive; however internally, the backup data is stored as *logical volumes* on disk space.

The combination of hardware (servers, cache controllers, and cache modules), and software (tape drive and robotic emulation) is called a *Virtual Tape Library* (VTL) as seen in Figure 10-3.

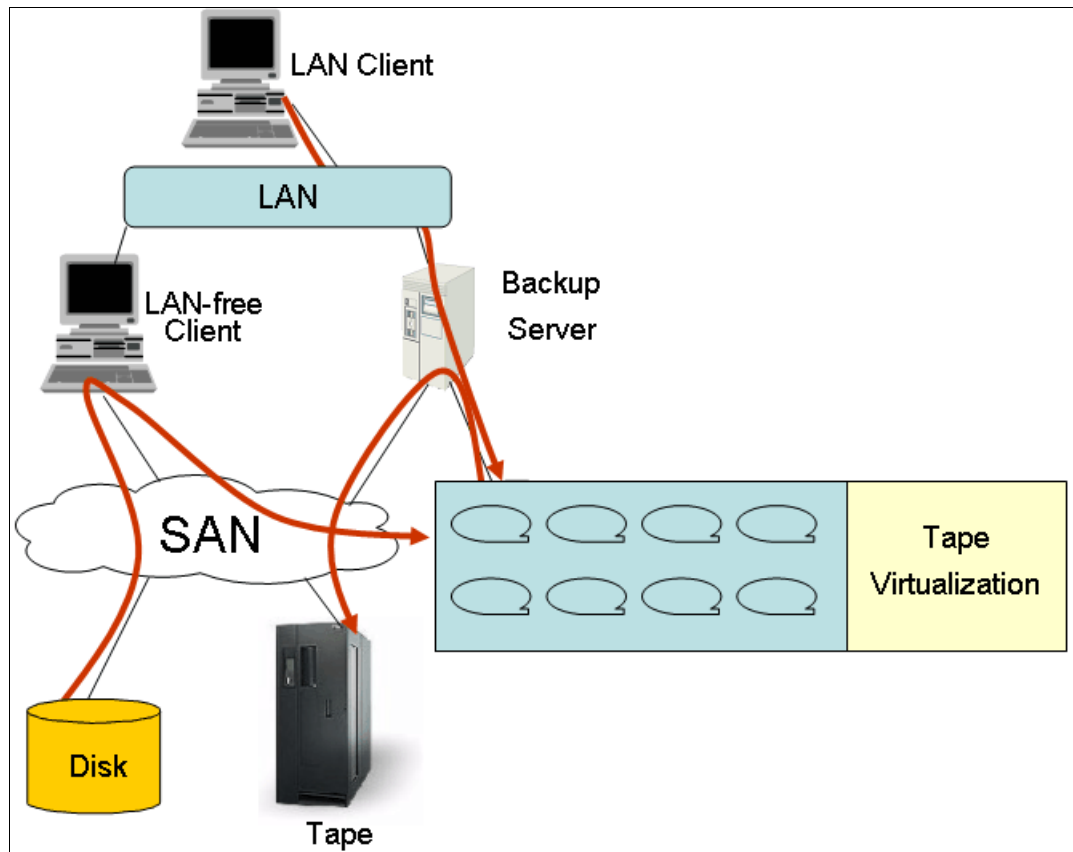


Figure 10-3 Tape virtualization diagram

10.2 TS7650 ProtecTIER Deduplication Appliance

The IBM System Storage TS7650 ProtecTIER Deduplication Appliance (Figure 10-4) is a preconfigured solution of IBM storage, IBM server, and the IBM revolutionary ProtecTIER data deduplication software designed to improve backup and recovery operations. This is not just a bundle of components, but a truly integrated solution that makes it easy to harness the power of deduplication without making radical changes to the existing environment. The solution is available in four configurations designed to meet the disk-based data protection needs of a wide variety of organizations, from mid-sized IT environments to enterprise data centers.

The IBM System Storage TS7650 ProtecTIER Deduplication Appliance provides:

- ▶ Up to 500 MBps or more inline data deduplication performance
- ▶ Up to 25 times or more storage capacity reduction
- ▶ Emulation of up to 12 virtual libraries, 256 virtual drives and 128,000 virtual cartridges
- ▶ Simplified configuration and deployment

ProtecTIER's native replication technology enables virtual tape cartridges to be replicated to a remote location for enhanced disaster recovery and business continuity. By eliminating the need to physically transport actual tape cartridges, systems, storage and data can be recovered quicker and more reliably in the event of a disaster or major system outage. It also lowers the total cost of ownership of backup and recovery by eliminating the costs associated with moving and storing physical tape cartridges.



Figure 10-4 ProtecTIER Deduplication Appliance

10.2.1 TS7650 ProtecTIER Deduplication Appliance features

The IBM System Storage TS7650 ProtecTIER Deduplication Appliance offers many features that can create savings in physical storage, processing and network bandwidth:

- ▶ Emulation of up to 12 virtual libraries, 256 virtual drives and 128,000 virtual cartridges
- ▶ IBM ProtecTIER with patented HyperFactor® data deduplication technology
- ▶ ProtecTIER Native Replication technology
- ▶ IBM System x server for enterprise-level performance
- ▶ IBM Storage Controller with highly reliable Fibre Channel drives
- ▶ Rack, cables, and other components needed to provide a complete solution
- ▶ Up to 500 MBps or more inline data deduplication performance
- ▶ Up to 25 times or more storage capacity reduction
- ▶ Simplified configuration and deployment

10.2.2 TS7650 ProtecTIER Deduplication Appliance configurations

The IBM System Storage TS7650 ProtecTIER Deduplication Appliance is available in the following four configurations:

- ▶ TS7650 ProtecTIER Deduplication Appliance:
 - 7 TB configuration with performance up to 100 MBps or more inline:
 - Inline data deduplication with a physical capacity of 7 TB and usable capacity of 175 TB (based on deduplication ratio of 25:1)
 - 18 TB configuration with performance up to 250 MBps or more inline:
 - Inline data deduplication with a physical capacity of 18 TB and usable capacity of 450 TB (based on deduplication ratio of 25:1)
 - Flexible and scalable growth to 36 TB

- 36 TB Configuration with performance up to 500 MBps or more inline:
 - Inline data deduplication with a physical capacity of 36 TB and usable capacity of 900 TB
 - Up to nine times faster than the competition delivers
- 36 TB Dual-node Cluster Configuration with performance up to 500 MBps or more:
 - Inline data deduplication with a physical capacity of 36 TB and usable capacity of 900 TB
 - High availability

For additional information about IBM System Storage ProtecTIER products, see *The IBM System Storage TS7650G and TS7650 ProtecTIER Servers*, SG24-7652.

For the latest features and support, see the website:

<http://www-03.ibm.com/systems/storage/tape/index.html>

10.3 TS7650G ProtecTIER Deduplication Gateway

The TS7650G ProtecTIER Deduplication Gateway, as seen in Figure 10-5, comprised of the TS7650G Hardware combined with IBM System Storage ProtecTIER Enterprise Edition software, is designed to address the data protection needs of enterprise data centers. The solution offers high performance, high capacity, scalability, and a choice of disk-based targets for backup and archive data.

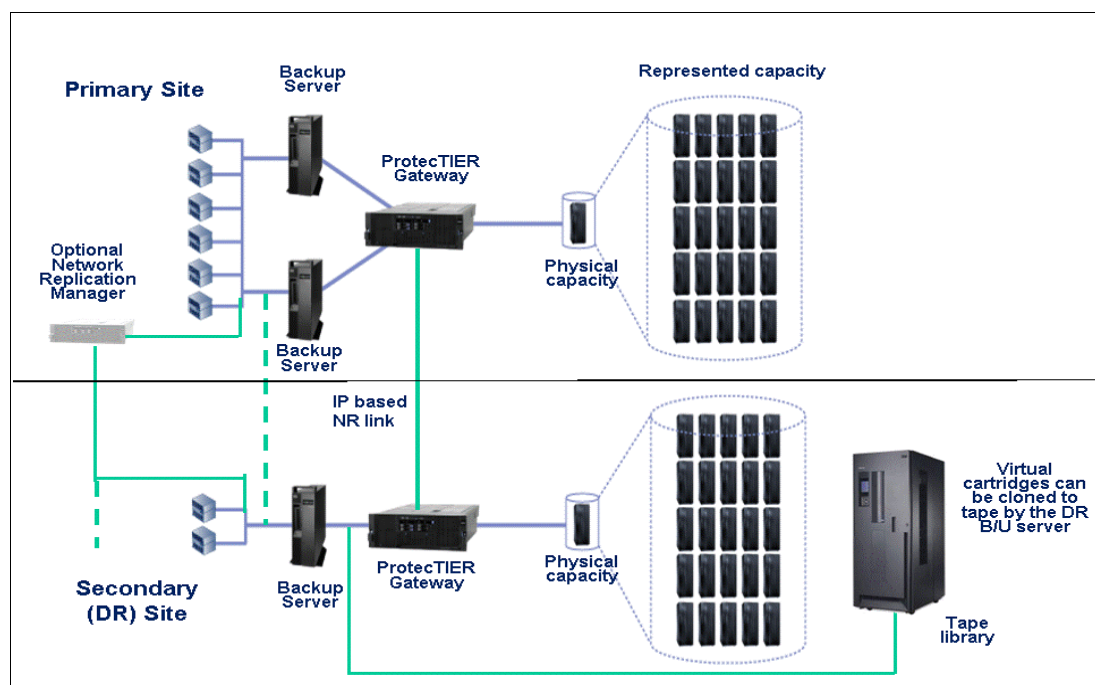


Figure 10-5 ProtecTIER Gateway Scenario with Native Replication

10.3.1 TS7650G ProtecTIER Deduplication Gateway features

The TS7650G ProtecTIER Deduplication Gateway offers:

- ▶ Virtual tape emulation of up to 16 virtual tape libraries per single node or two node cluster configuration and up to 512 per two node cluster or 256 virtual tape drives per TS7650G
- ▶ In-line data deduplication powered by HyperFactor technology
- ▶ Multicore virtualization and deduplication engine
- ▶ Clustering support for higher performance
- ▶ Fibre Channel ports for host and server connectivity
- ▶ Performance: up to 1000 MB/s or more sustained inline deduplication (clustered nodes)
- ▶ Emulation of the IBM TS3500 tape library with IBM Ultrium 2 or Ultrium 3 tape drives
- ▶ Emulation of the Quantum P3000 tape library with DLT tape drives
- ▶ Scales to 1 PB of physical storage over 25 PB of user data

10.3.2 TS7650G ProtecTIER Deduplication Gateway description

The TS7650G ProtecTIER Deduplication Gateway is an enterprise-class data protection platform designed to quickly and safely protect business information while reducing the amount of space required to store it.

Deploying the TS7650G ProtecTIER Deduplication Gateway can help organizations more efficiently protect their corporate data on disk-based storage while helping them manage the exponential growth of new data through reduction and elimination of duplicate data in their backups.

For additional information about IBM System Storage ProtecTIER products, see *The IBM System Storage TS7650G and TS7650 ProtecTIER Servers*, SG24-7652.

10.4 IBM System Storage ProtecTIER Enterprise Edition V2.3

IBM System Storage ProtecTIER Enterprise Edition V2.3 (ProtecTIER) software with HyperFactor offers additional features in addition to providing excellent Virtual Tape Library (VTL) functionality:

- ▶ Data deduplication
- ▶ ProtecTIER Native Replication

10.4.1 Data deduplication

Data deduplication solutions from IBM employ an advanced form of data compression that identifies and eliminates redundant data across the data landscape, making it possible to significantly reduce the amount of data that needs to be protected. This in turn dramatically increases the effective capacity of existing disk storage, so that far less physical disk is required to protect the data. Beyond the direct resource savings associated with needing less disk space, which can be in the hundreds of thousands or millions of dollars. The benefits of data deduplication include:

- ▶ Greater productivity that comes from being able to perform more frequent backups with the same amount of physical disk
- ▶ Increased efficiency because of the greater likelihood that data will be able to be restored from disk rather than a slower medium
- ▶ Reduced energy consumption that results from reducing the amount of disk in operation

Data deduplication uses algorithms to compare data and eliminate duplicated or redundant data. Unlike other methods of compression, however, which work with small amounts of data at the file level, data deduplication can be applied at the subfile level to reduce data size across a much broader landscape.

HyperFactor technology, as seen in Figure 10-6, uses a pattern algorithm that can reduce the amount of space required for storage in the backup environment by up to a factor of 25, based on evidence from existing implementations. The capacity expansion that results from data deduplication is often expressed as a ratio, essentially the ratio of nominal data to the physical storage used.

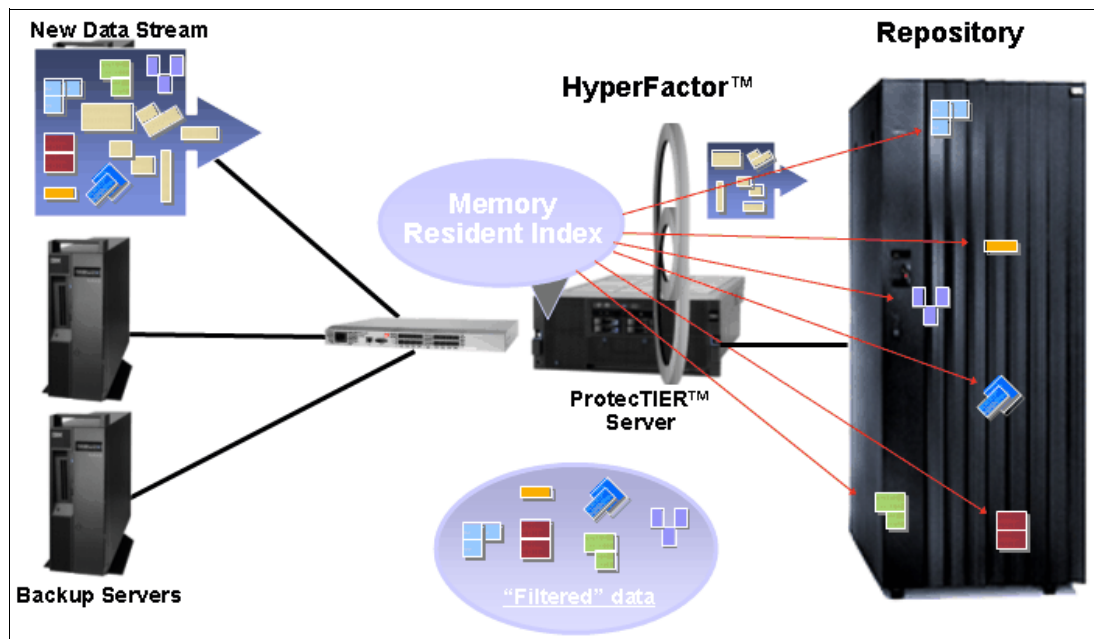


Figure 10-6 HyperFactor Technology

A 10:1 ratio, for example, means that 10 times more nominal data is being managed than the physical space required to store. Capacity savings of 18:1 and greater have been reported from data deduplication, up to 25:1 in the case of IBM solutions as seen in Figure 10-7.

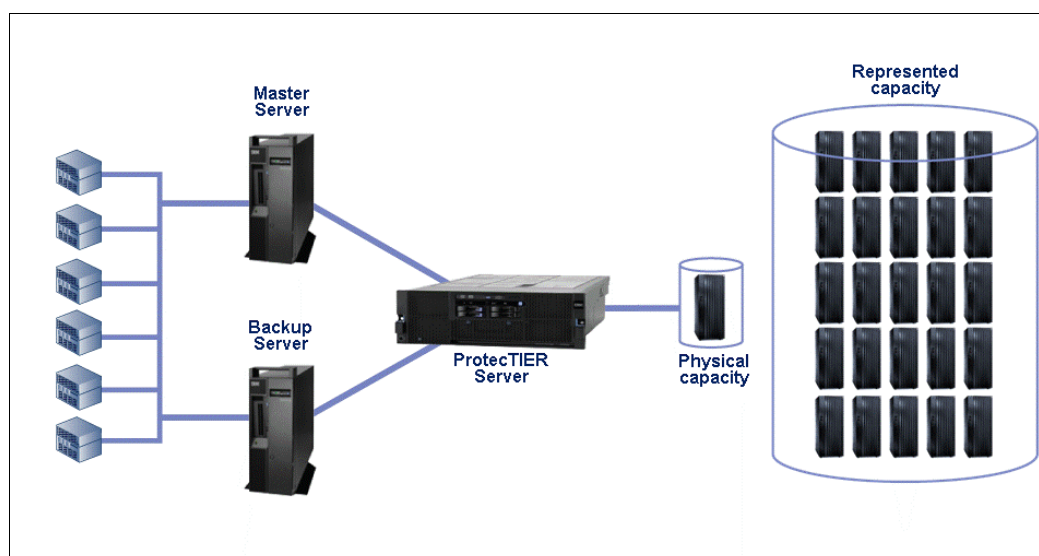


Figure 10-7 Representation Of Capacity Savings using ProtecTIER Technology

10.4.2 ProtecTIER Native Replication

ProtecTIER Native Replication is a logical feature that enables customers to replicate and/or “move” any or all of their virtual tape cartridges from their main site to a remote DR location. The TS7650 appliance and gateway models employ a IP based replication design. Data is replicated between sites from one two node cluster to another. Tape archiving capabilities outside of production environment are also provided with this feature.

The user will be able to create policies at the cartridge level to replicate a single or a range of barcodes in a specified timeframe and with a set priority. After being replicated (or moved) to the DR site, users can choose to clone these cartridges to real physical tape from within their backup application and outside of their main production stream. In case of a “disaster,” the TS7650 or TS7650G located at the DR site can become the production site (failover) until the main site comes back on-line. At that point the user can replicate or move the newly created tapes back to the main production site (failback). ProtecTier and Native Replication enables more frequent testing of Disaster Recovery plans.

Because ProtecTIER deduplicates data before storing it, only the changes of the data are transferred to the DR site over the replication link. This translates into substantial savings in the bandwidth needed for the replication link. Data transfer is started based on several trigger points such as policy based transfer windows, or movement of the virtual tape to a VTL “export slot”. (ProtecTIER’s VTL emulates “import” and “export” slots, as well as opening or closing of the “library” door to “insert” or “eject” a cartridge). Data verification/validation is done at the DR site to ensure integrity of the transferred data prior to make the virtual cartridge/tape available. ProtecTier only replicates a partial set of cartridges. ProtecTIER Native Replication is built on top of the PT two node cluster product and is cluster compliant. Replication can run in parallel to backup and restore activity.

ProtecTIER Native Replication eliminates the risks associated with human intervention and physical tape transportation. Replication occurs automatically (by policy) and simultaneously as data is deduplicated and backed up.

10.5 IBM Virtualization Engine TS7700

Attention: The IBM TotalStorage 3494 Virtual Tape Server has been replaced by the IBM System Storage TS7700 Virtualization Engine.

The IBM Virtualization Engine TS7700 (TS7700 Virtualization Engine) is a family of mainframe virtual-tape solutions that are designed to optimize tape processing. With one solution, the implementation of a fully integrated tiered storage hierarchy of disk and tape utilizes the benefits of both technologies to help enhance performance and provide the capacity needed for today's tape processing requirements. Deploying this innovative subsystem can help reduce batch processing time, total cost of ownership, and management overhead.

10.5.1 TS7700 Virtualization Engine description

The IBM Virtualization Engine TS7700 provides tape virtualization for the IBM System z environment. The Virtualization Engine TS7700, as seen in Figure 10-8, is designed to provide improved performance and capacity to help lower the total cost of ownership for tape processing. It introduces a new modular, scalable, high-performing architecture for mainframe tape virtualization. The TS7700 is available as the TS7740 R1.5, which brings significant enhancements over its previous release and now also comes as a disk-only solution, the TS7720, that does not require physical tape attachment.



Figure 10-8 IBM Virtualization Engine TS7700

The Virtualization Engine TS7700 integrates the advanced performance, capacity, and data integrity design of the 3592 Tape Drives, the IBM industry-leading tape technology, with high-performance disk and a new advanced IBM System p server to form a storage hierarchy managed by robust storage management firmware with extensive self management capability. It includes functions such as advanced policy management to control physical volume pooling, cache management, dual copy, dual copy across a grid network, and copy mode control. The TS7700 offers enhanced statistical reporting and a single, standards-based management interface for TS7700 management.

10.5.2 TS7700 Virtualization Engine highlights

Here we list the TS7700 Virtualization Engine highlights:

- ▶ System z FICON Attachment
- ▶ TS7740 w/ ~14 TB cache
- ▶ TS7720 disk-only w/ ~40 TB or ~70TB cache
- ▶ Support for 2 or 3 cluster Grids
- ▶ Copy Export for stand-alone cluster
- ▶ 3592 drives including native TS1120/TS1130 support
- ▶ 3584 libraries
- ▶ 3494 libraries (w/ 3592) (RPQ required with R1.6 or higher)
- ▶ 256 Virtual Drive addresses
- ▶ 256 Logical Paths per port
- ▶ Out-of-Band Encryption support
- ▶ Broadband Call Home (ECC)

These capabilities are designed to improve performance and capacity to help lower the total cost of ownership for tape processing and avoid the potential impact of human errors. A TS7700 can help improve the efficiency of mainframe tape operations by efficiently using disk storage, tape capacity, and tape speed, and by providing a large number of tape addresses.

The Virtualization Engine TS7700 utilizes outboard policy management functions to manage cache and volume pools, to control selective dual copy, dual copy across a grid network, copy mode control, encryption, and Copy Export. It also includes a new standards-based management interface and enhanced statistical reporting.

IBM Virtualization Engine TS7700 is designed to provide tape virtualization for IBM System z servers, and can help you meet the major needs in your environment, such as reliability and window reduction of the backup/restore process, services downtime caused for physical tape drives and library outages, impacts to the operational staff that manage daily backup/restore processing, and infrastructure simplification through reduction of the number of physical tape libraries, drives, and media.

10.5.3 IBM Virtualization Engine TS7700 Release 1.5 and 1.6

Here we list the support for Release 1.5 and Release 1.6.

Release 1.5

IBM Virtualization Engine TS7700 Release 1.5 includes support for:

- ▶ TS7720 disk-only virtual tape system with up to 70 TB of disk cache
- ▶ Over 13 TB of disk cache capacity for new TS7740 configurations
- ▶ Upgrades to 9 TB total disk cache capacity for existing TS7740 configurations
- ▶ TS1130 Tape Drives as TS7740 back-end drives
- ▶ Integrated Library Manager support
- ▶ Workflow Management controls

Release 1.6

IBM Virtualization Engine TS7700 Release 1.6 has recently been announced and includes support for:

- ▶ 4cluster Grid
- ▶ Grid support with mixed TS7720 and TS7740
- ▶ Cluster families & Cooperative Replication
- ▶ Logical WORM
- ▶ Host override for Cache residency
- ▶ SSPC user authentication and Audit logging
- ▶ 3494 Library support
- ▶ Via RPQ only
- ▶ Code install improvements

10.5.4 Characteristics of components comprising the TS7700

Here we summarize characteristics of these virtualization engines:

- ▶ The Virtualization Engine TS7700 Server provides host connections for up to four FICON channels for attachment to IBM System z servers with the appropriate levels of System z software.
- ▶ In a TS3500 Tape Library, the TS7740 supports from four to 16 IBM 3592 tape drives.
- ▶ In a 3494 Tape Library, the TS7740 supports from four to 12 IBM 3592 tape drives.
- ▶ The TS7720 does not attach to any tape library or drives. A TS7700 with Grid Enablement features can be interconnected with one or two other TS7700s (same model of TS7700) to provide peer-to-peer copy capability between Virtualization Engines for tape using IP network connections.
- ▶ Each TS7700 supports up to a maximum of 256 3490E virtual tape drives.
- ▶ A TS7700 Single or Multi Cluster Grid supports up to 1,000,000 logical volumes, each logical volume having a maximum capacity of 1.2 GB to 12 GB (assuming 3:1 compression and using the 400 to 4000 MB volume sizes).

For additional information about IBM System Storage TS7700 models, see *IBM Virtualization Engine TS7700 with R1.6*, SG24-7712.

10.5.5 IBM Virtualization Engine TS7740

The IBM TS7740 is designed to help customers reduce costs by potentially using fewer tape drives and therefore reducing tape cartridge handling, reducing floor space requirements, and improving tape library slot utilization, by using disk storage and the performance and capacity of TS1130 tape drives. Highlights of the TS7740 virtualization Engine are as follows:

- ▶ Can be attached to supported IBM System z servers
- ▶ Maximum of 768 virtual drives and 1,000,000 virtual volumes
- ▶ Web-based management tools
- ▶ Advanced policy management
- ▶ IBM TS3000 System Console
- ▶ Up to 42 TB native tape volume cache
- ▶ Attach up to 48 TS1100 or 3592 tape drives
- ▶ Up to 12 4 Gbps FICON channel

10.5.6 IBM Virtualization Engine TS7720

IBM TS7720, which addresses the need for a very high volume tape cache capacity for applications that can benefit from having all their data cache resident, uses 1 TB SATA disk drives to provide high capacity cache with rapid recall of data. Highlights of the TS7720 Virtualization Engine are as follows:

- ▶ Can be attached to supported IBM System z servers
- ▶ Maximum of 768 virtual drives and 1,000,000 virtual volumes
- ▶ Web-based management tools
- ▶ Advanced policy management
- ▶ IBM TS3000 System Console
- ▶ Up to 210 TB native tape volume cache using 1 TB SATA drives
- ▶ Up to 12 4 Gbps FICON channels

10.6 IBM Data Facility Storage Management Subsystem

IBM Data Facility Storage Management Subsystem (DFSMS) is a software suite that automatically manages data from creation to expiration. DFSMS provides allocation control for availability and performance, backup/recovery and disaster recovery services, space management, tape management, and reporting and simulation for performance and configuration tuning. It offers the following benefits:

- ▶ Integrated, comprehensive data management solution
- ▶ Delivers managed, predictable service to meet Service Level Agreements for Storage
- ▶ Up to 50% increased effective storage capacity with reduced overall storage cost

10.6.1 The IBM Data Facility Storage Management Subsystem Removable Media Manager

IBM Data Facility Storage Management Subsystem Removable Media Manager (DFSMSRmm) is the functional component of DFSMS for managing all removable media resources including automatic, manual, and virtual tape libraries (such as the IBM TS7700 series), for the z/OS environment.

The following key functions are provided by DFSMSRmm:

- ▶ **Tape Data Set Recording:** DFSMSRmm automatically tracks and records metadata for data written to all types of tape volumes.
- ▶ **Inventory Management:** DFSMSRmm automatically records metadata for tape volumes used on the system, and allows you to define any tape volume, enabling your complete tape inventory be monitored and managed. DFSMSRmm supports tape volumes which are used anywhere in your enterprise, including z/VM, and those managed by the IBM Integrated Removable Media Manager for the enterprise on System z. You can easily partition your tape inventory and tape libraries and use DFSMSRmm to control which system can use the volumes.
- ▶ **Life Cycle Management:** DFSMSRmm provides policies to enable tape data sets and tape volumes to be managed throughout their life. You define your tape data service levels, tape technology retention criteria and DFSMSRmm automatically manages the life cycle.
- ▶ **Scratch Pool Management:** DFSMSRmm supports either a global scratch pool or you can define one or more specific scratch pools for non-system managed tape volumes. DFSMSRmm manages the selection of the pool, and the return of volumes back to the scratch pool after the tape data expires.

10.6.2 DFSMSrmm highlights

Here we list highlights for DFSMSrmm:

- ▶ DFSMSrmm is an integral part of DFSMS and is shipped as part of DFSMS with the z/OS operating system. DFSMSrmm cooperates with and exploits many various system components that are required for the complete security, management, and well being of your tape library.
- ▶ DFSMSrmm volume retention and movement policies, can optionally be specified interactively with ISPF panels. This allows authorized application owners to alter existing values without contacting the tape librarian.
- ▶ DFSMSrmm includes the report generator which is an ISPF based tool for easily building and customizing many various storage management reports. Sample report types and reports enable reporting based on DCOLLECT data, DFSMSshm functional statistics, and DFSMSrmm created data.
- ▶ In addition to the regular release updates, the DFSMSrmm enhancements provide many improvements; which recently include: support for IBM's Integrated Removable Media Manager for the Enterprise (IRMM) on System z, support for D/T3592 Model E06 and larger tape volumes, support for TS7700 Release 1.5 and copy export, and migration health checks for V1R10 and V1R11.
- ▶ DFSMSrmm is a full function removable media management system for your enterprise. DFSMSrmm supports tape volumes which are used anywhere in your enterprise, including z/VM, and those managed by the IBM Integrated Removable Media Manager for the enterprise on System z.
- ▶ DFSMSrmm is functionally compatible with existing tape management systems and runs in parallel during conversion. The easy to use conversion tools and related documentation and Redbooks publications enable the easy migration of your existing tape environment to management by DFSMSrmm. The special expiration date formats of 99nnn and 98nnn are supported by DFSMSrmm.

For DFSMSrmm enhancements, past and present are, see the website:

http://www-01.ibm.com/support/docview.wss?rs=0&q1=T1010391&uid=isg3T1010391&loc=en_US&cs=utf-8&cc=us&lang=en

To learn more about DFSMS, see the *DFSMS V1.10 and EAV Technical Guide*, SG24-7617.

10.7 More information

For more information about the IBM tape virtualization products and service offerings, see:

<http://www-03.ibm.com/systems/storage/tape/index.html>

Also see the following publications:

- ▶ *IBM Virtualization Engine TS7700 Release 1.4a: Tape Virtualization for System z Servers*, SG24-7312
- ▶ *IBM Virtualization Engine TS7700 with R1.6*, SG24-7712
- ▶ *IBM Virtualization Engine TS7510: Tape Virtualization for Open Systems Servers*, SG24-7189
- ▶ *The IBM Virtualization Engine TS7510: Getting Started with i5/OS and Backup Recovery and Media Services*, SG24-7510
- ▶ *IBM Virtualization Engine TS7500: Planning, Implementation, and Usage Guide*, SG24-7520



Complementary storage products

In this chapter we discuss entry-level tape products that are complementary to the LTO series of drives and libraries discussed earlier. The entry-level tape products cover a wide range of formats to suit varied business needs. We also consider drive enclosures for table/shelf and rack mounting.

IBM offers a wide range of external tape drives and external enclosure technology to help support business continuance in the workplace. These models include:

- ▶ IBM 7206 Model 336 External DAT72 (DDS Gen 5) Tape Drive
- ▶ IBM DDS Generation 5 USB Tape Drive
- ▶ IBM DDS Generation 6 USB Tape Drive
- ▶ IBM TotalStorage 7207 External Tape Drive
- ▶ IBM System Storage 7212 Storage Device Enclosure Express Model
- ▶ IBM System Storage 7214 Storage Device Enclosure
- ▶ Half High Tape Drive External Enclosure
- ▶ 1U Tape Drive External Enclosure

11.1 IBM TotalStorage 7206 Model 336 External DDS Gen 5 Tape Drive

The IBM 7206 Model 336 External DAT72 (DDS Gen 5) Tape Drive (as seen in Figure 11-1), is designed to be a cost-effective tape drive featuring the popular DAT72 (DDS) tape technology. It is designed to offer improved data quality and performance and increased capacity compared to the IBM 7206 Model 220. The 7206 Model 336 supports a migration path to greater tape storage capacity at a price point similar to IBM 7206-220 DDS4 tape drives.

The IBM 7206 Model 336 External DDS Gen 5 (DAT72) Tape Drive can be attached to System p servers by an Ultra2, Ultra3, or Ultra320 SCSI LVD interface.

The 7206 Model 336 tape drive can achieve a media capacity up to 72 GB with 2:1 data compression, nearly twice the capacity of the previous IBM 7206-220 DDS4 tape drive. The 7206 Model 336 offers a sustained data transfer rate of up to 6 MB per second (with 2:1 compression).

11.1.1 IBM 7206 Model 336 highlights

Here we list highlights of the IBM TotalStorage 7206 Model 336 External DDS Gen 5 Tape Drive:

- ▶ Designed for improved data rate compared to IBM 7206 Model 220
- ▶ Designed for read and write compatibility with previous generation 4mm tape media
- ▶ Read and write compatible with two previous generations of DDS tape technology
- ▶ Capacity of up to 72 GB per media cartridge
- ▶ Data transfers up to 21 GB per hour
- ▶ Compatible with most IBM System i and IBM System p server models
- ▶ Limited lifetime warranty on IBM DAT72 (DDS Gen 5) media



Figure 11-1 IBM TotalStorage 7206 Model 336 External DDS Gen 5 Tape Drive

11.1.2 More information

For more information, see the website:

<http://www.ibm.com/systems/storage/tape/7206/336/index.html>

11.2 IBM DDS Generation 5 USB Tape Drive

The IBM DDS Generation 5 USB tape drive (as seen in Figure 11-2), is the entry point into the IBM DDS tape drive family and offers dependable capacity and performance for value-conscious small and medium or distributed large enterprise businesses. DDS/DAT technology has the proven reliability and low cost of ownership to effectively meet the demanding backup, archival, and regulatory compliance requirements of your System x environment.



Figure 11-2 IBM DDS Generation 5 USB Tape Drive

IBM DDS Generation 5 technology delivers a capacity of up to 72 GB on a single data cartridge and a transfer rate of up to 22 GB/hr (assuming a 2:1 compression ratio) at an entry price point. Its USB 2.0 interface supports internal plug and play capability with the latest System x systems, as well as external attachment in one of IBM's external tape drive enclosures. DDS Generation 5 technology is also backward read and write compatible with DDS-3 and DDS-4/DAT 40 media to ensure an easy migration or upgrade path for existing DDS/DAT users.

11.2.1 IBM DDS Generation 5 USB Tape Drive highlights

Here we list highlights of the IBM DDS Generation 5 USB Tape Drive:

- ▶ USB interface allowing for direct internal connections with select System x servers
- ▶ Eliminates potential RAID conflicts and extra controllers
- ▶ 5.25-inch half-high form factor
- ▶ Drive can be converted to 3.5 inch form factor either in the factory (CTO) or in the field using the appropriate enablement kit
- ▶ Internal USB interface for simplified plug-and-play installation
- ▶ Optional DDS Tape Enablement kits allow for internal installation in select IBM System x rack-optimized servers
- ▶ Flexible external USB configuration offerings when used with desktop or 1U rack-mount enclosures
- ▶ Support for leading operating systems and popular backup software
- ▶ Up to 36 GB native capacity, and up to 72 GB in compressed mode
- ▶ Up to 3.0 MBps native back-up rate, and up to 6.0 MBps in compressed mode
- ▶ USB 2.0 interface
- ▶ Read/write compatibility with DDS-4, and DDS-3 cartridges
- ▶ Available internal or external USB port

- ▶ Available 5.25-inch half-high bay for internal installations
- ▶ Supported external tape enclosure for external attachment:
 - IBM 1U Rackmount Tape Enclosure
 - IBM Half High Tabletop Tape Enclosure

11.2.2 More information

For more information about the IBM DDS Generation 5 USB Tape Drive, see the website:

<http://www.redbooks.ibm.com/abstracts/tips0755.html?Open>

11.3 IBM DDS Generation 6 USB Tape Drive

The IBM DDS Generation 6 Internal USB Tape Drive provides optimum capacity, performance, and a low total cost of ownership for small and medium business environments or enterprise workgroups that require the highest capacity and performance DDS (DAT) drive available. The drive supports DDS Generation 6 cartridges, which have a native capacity of 80 GB (160 GB compressed) as well as a transfer rate of up to 6.9 MBps native and up to 13.8 MBps compressed as seen in Figure 11-3.



Figure 11-3 IBM DDS Generation 6 USB Tape Drive

The IBM DDS Generation 6 USB Tape Drive maintains the competitive total cost of ownership of Digital Data Storage. Its USB interface supports plug and play capability with the latest System x systems, and can be installed internally in supported servers or supported externally in tape drive enclosures. This option is a higher performance and capacity alternative to the IBM DDS Generation 5 drive.

11.3.1 IBM DDS Generation 6 USB Tape Drive highlights

Here we list highlights of the IBM DDS Generation 6 USB Tape Drive:

- ▶ Highest capacity and performance in the DDS portfolio
- ▶ Up to 80 GB native capacity, and up to 160 GB in compressed mode
- ▶ 6.9 MBps native back-up rate and up to 13.8 MBps in compressed mode
- ▶ USB 2.0 interface
- ▶ Read/write compatibility with DDS Gen-5 and DDS-4 cartridges

- ▶ 5.25 inch half-high form factor; drive can be converted to 3.5 inch form factor either in the factory (CTO) or in the field using the appropriate enablement kit
- ▶ Internal USB interface for simplified plug-and-play installation
- ▶ Optional DDS Tape Enablement Kit allows for internal installation in select IBM System x rack-optimized servers
- ▶ Flexible external USB configuration offerings when used with desktop, 1U rack-mount, or 4U rack-mount enclosures
- ▶ The IBM DDS Generation 6 Media 5 Pack has a limited lifetime warranty

11.3.2 More information

For more information about the IBM DDS Generation 6 USB Tape Drive, see the website:

<http://www.redbooks.ibm.com/abstracts/tips0725.html?Open>

11.4 IBM TotalStorage 7207 External Tape Drive

The IBM TotalStorage 7207 External Tape Drive (as seen in Figure 11-4), offers an affordable back-up, archival storage, and data interchange for your iSeries/AS400 and pSeries/RS6000 systems.

The 7207 Model 330 SLR60 External Tape drive provides up to 37.5 GB of capacity and a data rate of 4 MBps. Media sizes of 30 GB and 37.5 GB are available. Assuming a compression of 2:1, typical of this tape drive, the tape drive reaches capacities of 60 GB and 75 GB respectively and a transfer rate of 8 MBps.



Figure 11-4 IBM TotalStorage 7207 External Tape Drive

11.4.1 IBM TotalStorage 7207 External Tape Drive highlights

Here we list highlights of the IBM TotalStorage 7207 External Tape Drive:

- ▶ SCSI (LVD/SE) attachment to eServer™ iSeries/AS400 and pSeries/RS6000 systems
- ▶ The 7207 model 330 provides Read/Write compatibility with SLR100, MLR3 and MLR1 (QIC) tape formats and read compatible with SLR5 and DC9250 media
- ▶ Sustained Data Rate of 4 MBps (8 MBps with compression) for the 7207 model 330
- ▶ Capacities of up to 37.5 GB (75 GB when compressed) for the model 330, 4 GB (8 GB with compression) for the 122
- ▶ Capacity per cartridge (two options):
 - Up to 30 GB native: Up to 60GB compressed (Part Number: 14P4209)
 - Up to 5 GB native: Up to 10GB compressed (Part Number: 35L0661)

11.4.2 More information

For more information about the IBM System Storage 7207 External Tape Drive, see the website:

<http://www.ibm.com/systems/storage/tape/7212/index.html>

11.5 IBM System Storage 7212 Storage Device Enclosure Express Model

The IBM System Storage 7212 Storage Device Enclosure Express Model (as seen in Figure 11-5 on page 296), allows you to package the latest technology options in tape drives and a DVD-RAM optical drive in a low-profile, modular design for rack-mount or limited-space desktop applications.

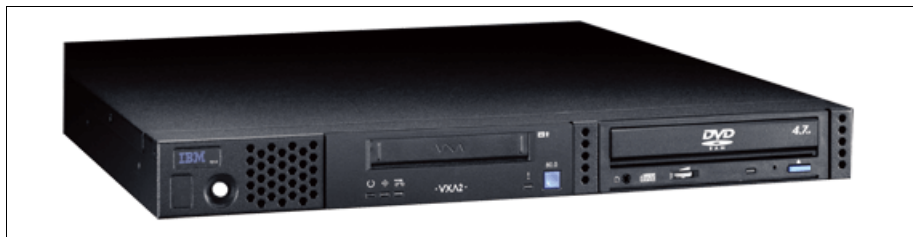


Figure 11-5 IBM System Storage 7212 Storage Device Enclosure Express Model

The IBM System Storage 7212 Storage Device Enclosure Express Model features two popular tape drive technology options. The 7212 Express Model packaging is a low-profile, modular design that is an excellent choice for rack-mount or limited-space desktop applications.

11.5.1 IBM System Storage 7212 Storage Device Enclosure highlights

Here we list highlights of the IBM System Storage 7212 Storage Device Enclosure:

- ▶ Features a compact design that can be configured with up to two storage devices
- ▶ Can be configured for one EIA unit (1U) of a standard 19-inch server rack or as a low-profile desktop solution
- ▶ Offers storage device options for 4mm and 1/4-inch (QIC) tape drive formats
- ▶ Low-profile storage solution for environments in which cabling space and server storage bays are limited
- ▶ Connects to IBM System i and System p workstations and servers
- ▶ Compatible with tape and optical drives
- ▶ Storage options drive types
 - DAT72 Media: 4 mm tape
 - Native physical capacity: 36GB
 - Compatibility: DDS-4, DDS-3
 - SLR60/SLR100 Media: SLR(QIC)
 - Native physical capacity: 37.5/50GB
 - Compatibility: MLR1, MLR3, SLR5, SLR60

11.5.2 More information

For additional information about the IBM System Storage 7212 External Enclosure Device Model, see the website:

<http://www.ibm.com/systems/storage/tape/7212/index.html>

11.6 IBM External Tape Enclosures

IBM External Tape Enclosures provide a space saving and low cost alternative to mounting storage devices internally in your server.

11.6.1 IBM System Storage 7214 Storage Device Enclosure

The IBM System Storage 7214 Storage Device Enclosure (as seen in Figure 11-6), features the latest technology options in tape drives and DVD optical drives. The 7214 Storage Enclosure is a low-profile design that is an excellent choice for mounting in your System p 19" rack.



Figure 11-6 IBM System Storage 7214 Storage Device Enclosure

Despite the decreasing availability of server bays for storage devices, it is still important to consolidate storage devices in a single, convenient location to minimize space and cabling impacts. The IBM System Storage 7214 Tape and DVD Enclosure Express is designed to mount in one EIA unit of a standard IBM Power Systems 19-inch rack, and can be configured with one or two tape or DVD drives. The 7214 Express enclosure attaches to SAS-based models of the IBM Power Systems through external serial attached SAS adapters. The 7214 Express also attaches to certain open system adapters.

The SAS electronic interface on the 7214 Express is designed to provide faster transfer rates, greater convenience and a reduction in space required for system-to-device cabling.

IBM System Storage 7214 Storage Device Enclosure highlights

Here we list highlights of the IBM System Storage 7214 Storage Device Enclosure:

- ▶ 1U slim profile designed for 19-inch rack system environments
- ▶ Features IBM half-high LTO technology designed for reliable performance in small-medium open system environments
- ▶ Supports DAT72, DAT160, DVD-RAM and DVD-ROM drives
- ▶ IBM Half-high LTO 4 supports the encryption of data for increased security with archived data
- ▶ Supports 3 GB/sec IBM SAS interface
- ▶ Dual-drive enclosure, side-by-side
- ▶ Offers choices of HH LTO4 and DAT160 tape and DVD storage technologies

- ▶ Control card sensor designed to track drive function and notify user of maintenance requirements
- ▶ Adheres to LTO specifications
- ▶ HH LTO 4 tape drives can write to LTO 3 media and read LTO 2 media; WORM media also available
- ▶ DAT160 tape drives are read-write compatible with DAT72 and DDS4 media
- ▶ Typical compression
 - 2:1 for tape drives
 - 3:1 for DVD optical drives

Drive options on the 7214 Storage Enclosure

Here we list the drive options on the 7214 Storage Enclosure:

- ▶ DAT72 4 mm tape drive:
Storage capacity of up to 72 GB with a data transfer rate of up to 6 MB per second (assumes 2:1 compression).
- ▶ DAT160 4 mm format tape drive:
Storage capacity of up to 160 GB with a data transfer rate of up to 13.8 MB per second (assumes 2:1 compression).
- ▶ Half-high LTO Ultrium 4 tape drive:
Storage capacity of up to 1.6 TB with a data transfer rate of up to 240 MB per second (assumes 2:1 compression)
- ▶ DVD-RAM drive:
Native capacity of up to 9.4 GB with a data transfer rate of 2.77 MB per second

More information

For additional information about the IBM 7214 Storage Enclosure, see the website:

<http://www.ibm.com/systems/storage/tape/7214/index.html>

11.6.2 Half High Tape Drive External Enclosure

The Half-High SCSI Tape Enclosure (as seen in Figure 11-7), provides a space saving and low cost alternative to mounting a drive internally in your server. In stealth black, this table or shelf unit becomes an ideal solution when you're looking for a home for one of our wide array of Half-Height System x Tape drives.



Figure 11-7 HH Tape Drive External Enclosure

Half High Tape Drive External Enclosure highlights

Here we list the drive options on the 7214 Storage Enclosure:

- ▶ Single Half-Height Bays
- ▶ Either Table Top or Shelf Form Factor
- ▶ Drive dependant Interface of either Single SCSI Enclosure Adaptor Kit (42C3910) or SAS Enclosure Adaptor Kit (40K2599) required
- ▶ SAS Enclosure Adaptor Kit (40K2599) supports both SAS and SATA Based Drive

More information

For more information about the HH Tape Drive External Enclosure, see the website:

<http://www.ibm.com/systems/storage/tape/halfhightapedrive/index.html>

Note: The IBM Full High Tape Drive External Enclosure has been withdrawn from the market as of January 2, 2009.

11.6.3 1U Tape Drive External Enclosure

The IBM 1U rack-mount external tape enclosure (as seen in Figure 11-8 on page 299), provides a cost-effective solution for adding up to two half-height tape drives externally to your IBM rack-optimized system x server environments.



Figure 11-8 1U Tape Drive External Enclosure

1U Tape Drive External Enclosure highlights

Here we list highlights of the IBM 1U Tape Drive External Enclosure:

- ▶ Two Half-Height Bays
- ▶ 1U Rack-Mount Form Factor
- ▶ Interface of either Single SCSI Enclosure Adaptor Kit (42C3910) or SAS Enclosure Adaptor Kit (40K2599) required (Up to 2 Max)
- ▶ SAS Enclosure Adaptor Kit (40K2599) supports both SAS & SATA Based Drives

More information

For additional information about the 1U Tape Drive External Enclosure, see the website:

<http://www.ibm.com/systems/storage/tape/1utapedrive/index.html>

Note: The IBM 4U Tape Drive External Enclosure has been withdrawn from the market as of July 3, 2008.



Part 3

Storage networking

In Part 3 we introduce important technologies and protocols that can be used to interconnect storage devices. These include SAN, NAS, and iSCSI. Then we provide information about the recently introduced IBM Ethernet switches and routers.



Introduction to storage networking

Variations for storage networking seem to be materializing faster than they can be implemented. Storage networking offers significant capabilities and flexibility for accessing stored data that was not previously available. Understanding the technology basics is essential to making the best choices.

In this chapter we introduce various storage networking options that you can choose to build the infrastructure for accessing your stored data. We mainly cover the fundamentals of SAN and various SAN expansion technologies.

12.1 Overview

Why are there so many forms of storage networking? One reason is that new technologies emerge and evolve but do not replace the investment in previous technologies overnight. And no single storage networking approach solves all problems, or optimizes all variables. There are trade-offs in cost, ease-of-management, performance, distance, and maturity, to name but a few. For the foreseeable future, multiple storage networking alternatives will coexist, often within the same organization.

Three popular models for attaching storage for the network are Storage Area Network (SAN), Network Attached Storage (NAS), and iSCSI. These methods help to remove direct attachments between storage and server, giving more flexibility in storage access.

The SAN can be viewed as an extension to the storage bus concept that enables storage devices and servers to be interconnected using similar elements as in LAN: routers, hubs, switches, directors and gateways. Storage resides on this dedicated network, providing an any-to-any connection for processors and storage on that network. The data traffic is based on a block level I/O (requests access devices directly). The most common media is Fibre Channel. Fibre Channel Protocol (FCP) is the I/O protocol for open systems, and Fibre Connectivity (FICON) replaces it for mainframe environments. Ethernet-based SANs are also emerging. Today, there are several popular protocols used to build or extend a SAN besides FCP like Internet SCSI (iSCSI), Fibre Channel over IP (FCIP) and Internet Fibre Channel Protocol (iFCP).

The NAS device is attached to a TCP/IP-based network (LAN or WAN), and accessed using CIFS, NFS or specialized I/O protocols for file access and file sharing. The data traffic is based on a file level I/O. It receives an NFS or CIFS request over a network and has an internal processor which translates that request to the block-I/O commands to access the appropriate device only visible to the NAS product itself.

12.2 Storage Area Network

A Storage Area Network (SAN) is a dedicated network for storage devices and the processors that access those devices. Figure 12-1 shows a picture of a SAN. SANs today are usually built using Fibre Channel technology, but the concept of a SAN is independent of the underlying type of network. I/O requests to disk storage on a SAN are called “block I/Os” because, just as for direct-attached disk, the read and write I/O commands identify a specific device (disk drive or tape drive) and, in the case of disks, specific block (sector) locations on the disk.

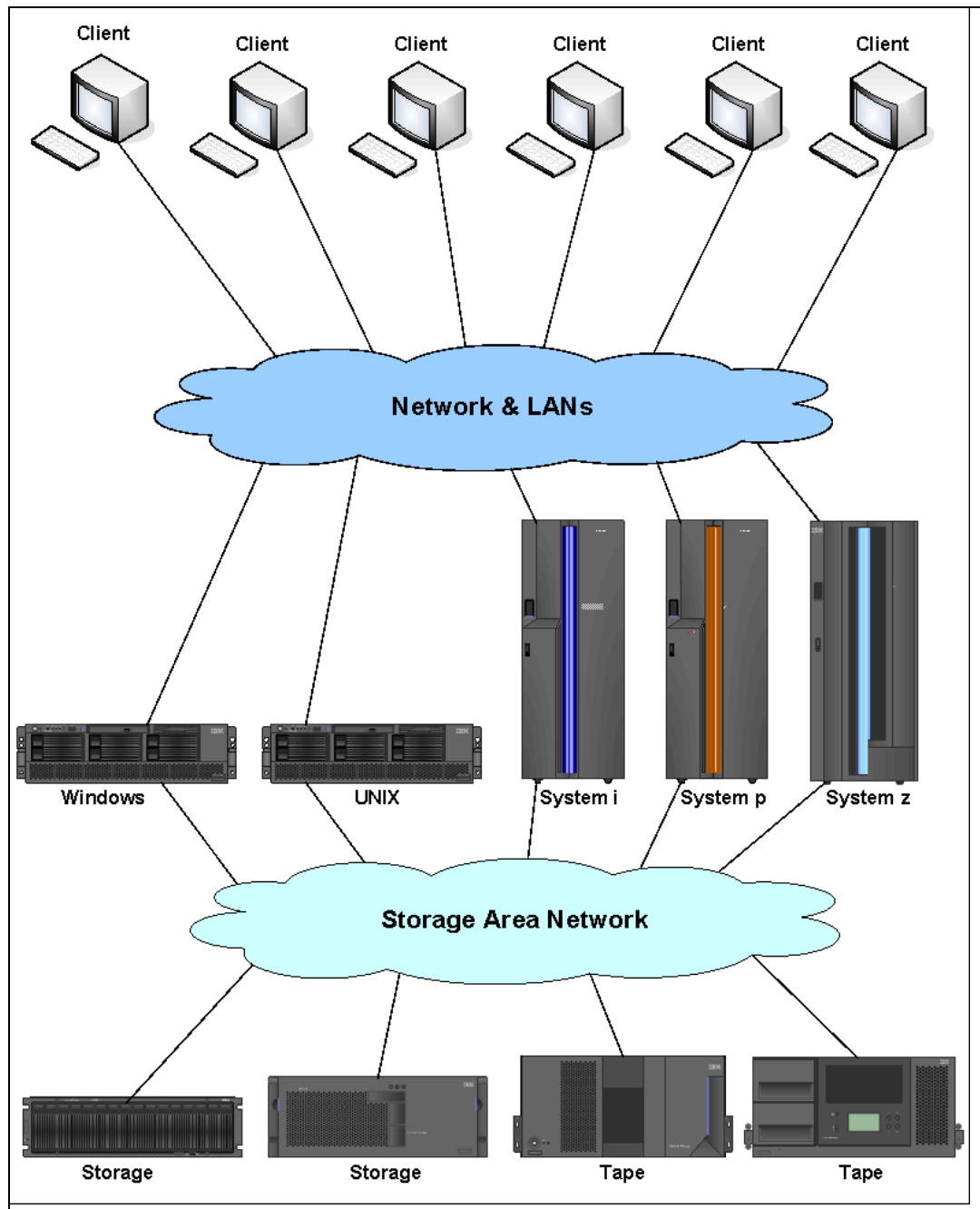


Figure 12-1 A picture of SAN

The major potential benefits of a SAN can be categorized as follows:

- ▶ Improvement to application availability:
Storage is independent of applications and accessible through multiple data paths for better reliability, availability and serviceability.
- ▶ Higher application performance:
Storage processing is off-loaded from servers and moved onto a separate network, for example, LAN-free and server-free backups.
- ▶ Centralized and consolidated storage:
Simpler management, scalability, flexibility, and availability are provided.
- ▶ Data transfer and vaulting to remote sites:
Remote copy of data is enabled for disaster protection and against malicious attacks.
- ▶ Simplified centralized management:
Single image of storage media simplifies management.

In this section, we explain important concepts involved in SAN.

12.2.1 Fibre Channel

Fibre Channel is well established in the open systems environment as the underlying architecture of the SAN. Fibre Channel is a technology standard that allows data to be transferred from one network node to another at very high speed. The interconnections between nodes are not necessarily based on *fiber* optics, but can also be based on copper cables. In current implementations, Fibre Channel standard speed is generally available from 2 Gbps to 10 Gbps, however most older 2 Gbps equipment is being replaced by faster kit.

This standard is backed by a consortium of leading vendors and has been accredited by the American National Standards Institute (ANSI).

Fibre Channel is structured with independent layers, as are other networking protocols. There are five layers, where 0 is the lowest layer. The physical layers are 0 to 2. These layers carry the physical attributes of the network and transport the data created by the higher level protocols, such as SCSI, TCP/IP, or FICON.

As shown in Figure 12-2, the top two layers (the session and transport layers) can be used by these protocols to move data segments. These segments are then rolled into a packet, which in turn are rolled into a frame. The originator creates the frame and sends it to the destination, which unravels the frame back to a segment.

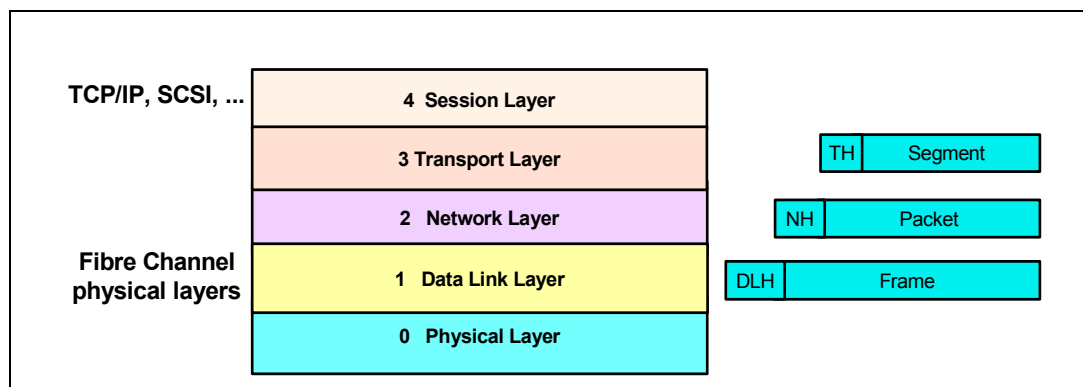


Figure 12-2 Fibre Channel protocol layers

12.2.2 SAN topologies

Fibre Channel based networks support three types of topologies:

- ▶ Point-to-point
- ▶ Loop (arbitrated)
- ▶ Switched

These can be implemented separately or interconnected to form a *fabric*. The fabric can also be extended to cover even greater distances. Currently, the distance limitation for synchronous data transfers is about 100 km and virtually unlimited for asynchronous data transfers (see 12.3, “IP storage networking technologies” on page 315).

Synchronous data transfers require an acknowledgement that the data arrived at the receiving end before the data is discarded at the transmitting end, whereas asynchronous data transfers do not require acknowledgement before the data is discarded.

Point-to-point topology

The point-to-point topology is the easiest Fibre Channel configuration to implement, and it is also the easiest to administer. This simple link can be used to provide a high-speed interconnection between two nodes as shown in Figure 12-3. A node is any device with one or more Fibre Channel ports.

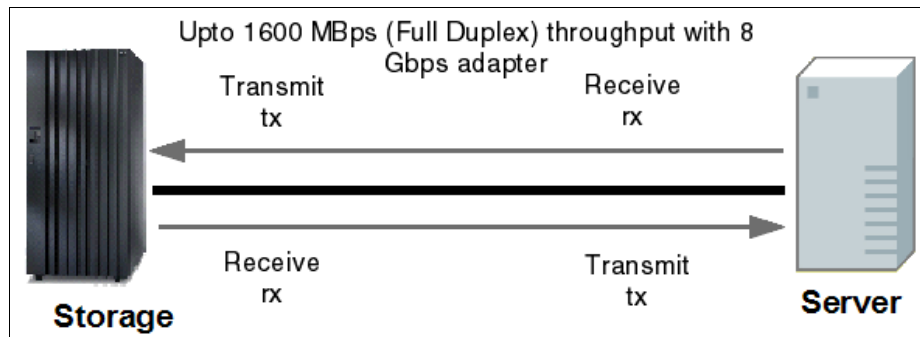


Figure 12-3 Fibre Channel point-to-point topology

Because connectivity is limited to two nodes, the exploitation of point-to-point in tape environments is limited; however the distance between nodes can be up to 10 km, which enables a tape library to be located at another site.

When greater connectivity and performance are required, each device can be connected to a fabric without incurring any additional expense beyond the cost of the fabric itself.

Loop (arbitrated) topology

The Fibre Channel arbitrated loop offers relatively high bandwidth and connectivity at a low cost. For a node to transfer data, it must first arbitrate to win control of the loop. When the node has control, it is now free to establish a virtual point-to-point connection with another node on the loop. After this point-to-point (virtual) connection is established, the two nodes consume all of the loop's bandwidth until the data transfer operation is complete. After the transfer is complete, any node on the loop can now arbitrate to win control of the loop.

A Fibre Channel arbitrated loop includes the following characteristics:

- ▶ Support of up to 126 devices is possible on a single loop.
- ▶ Devices can be hot-swapped with the implementation of hubs and bypass ports.
- ▶ A loop is self-discovering; it finds out who is on the loop and tells everyone else.
- ▶ Logic in the port allows a failed node to be isolated from the loop without interfering with other data transfers.
- ▶ Virtual point-to-point communications are possible.
- ▶ A loop can be interconnected to other loops, essentially forming its own fabric.
- ▶ A loop can be connected to a suitable Fibre Channel switch to create fan-out, or the ability to increase the size of the fabric even more. Note that not all switches support direct attachment of loops.

FC hub devices support FC loop connections while offering several of the benefits of switches. Figure 12-4 shows an FC loop using a hub. Be aware that FC hub technology is no longer marketed and is now rare in production environments.

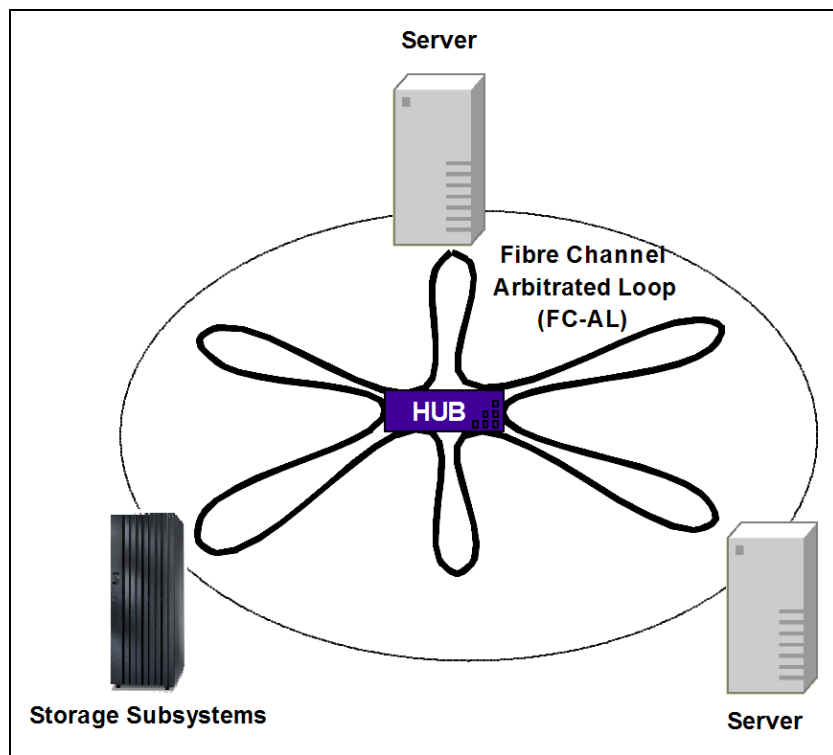


Figure 12-4 Fibre Channel loop topology

Switched topology

Fibre Channel switches function in a manner similar to traditional network switches to provide increased bandwidth, scalable performance, an increased number of devices, and, in certain cases, increased redundancy. Fibre Channel switches vary in the number of ports and media types they support.

Note: This topology is the most widely used in current SAN implementations.

Multiple switches can be connected to form a switch fabric capable of supporting a large number of host servers and storage subsystems, as shown in Figure 12-5. When switches are connected, each switch's configuration information has to be copied (*cascaded*) into all the other participating switches.

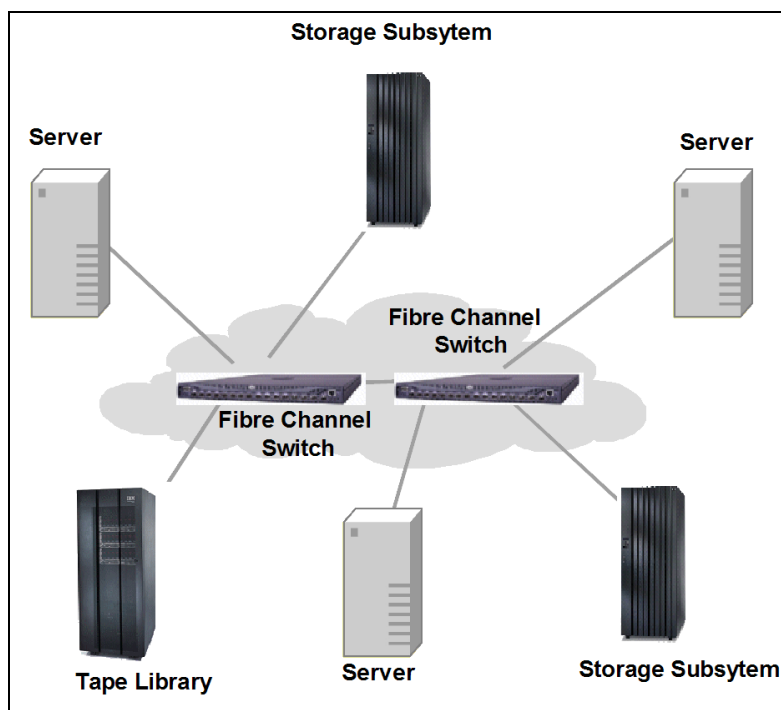


Figure 12-5 Fibre Channel switched topology

Switched fabrics provide the richest function for the tape environment, providing distance for remote tape vaulting and disaster tolerance, multi-server connectivity for tape library sharing, and multiple paths for redundancy. Switches can be connected together (Figure 12-5) to provide extra paths and larger networks.

Single switched fabric

A switched point-to-point fabric without redundancy can support a large number of servers. A configuration can be set up to allow every server to have access to every switch, and every controller to be connected to at least two switches. This allows any server to get to any controller or device, and it allows for continuous operation (although with degraded performance) in the event that a switch fails.

An example of a non-redundant fabric is shown in Example 12-6.

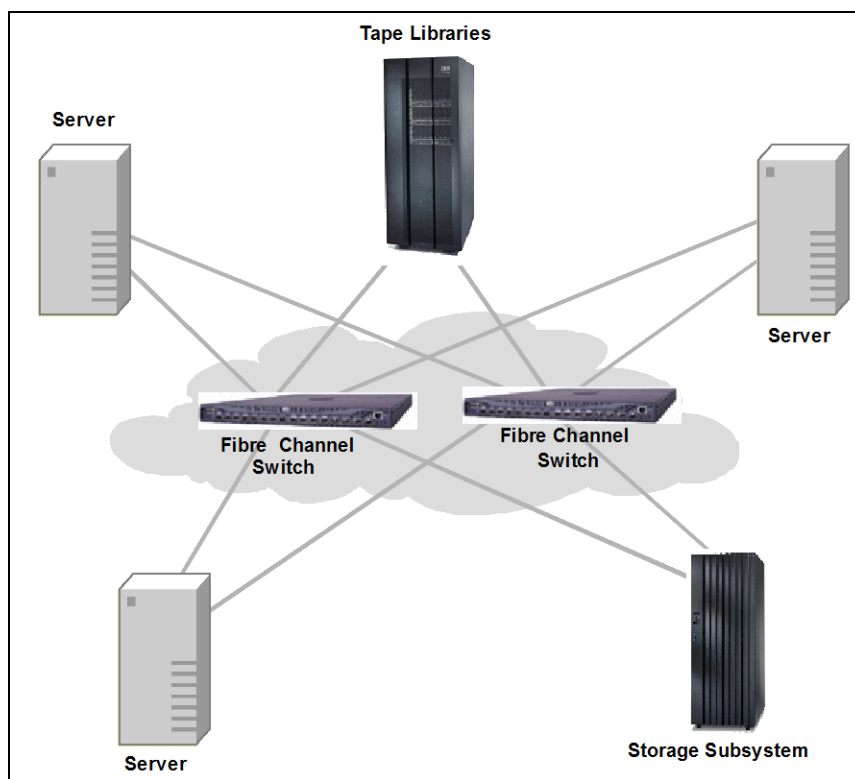


Figure 12-6 Fibre Channel single switched topology

Switched fabric with redundancy

A switched fabric with redundancy provides interconnections between switches, so that the collection of switches looks like one large, any-to-any switch. Fabric management becomes more complex than with basic switched point-to-point configurations and there are often limits imposed by product vendors on the number of switch-to-switch hops permitted when negotiating the fabric. Inter-switch links can fail and must be identified (many switch vendors do not yet support any reporting on inter-switch links).

Traffic can be routed in many ways. For technical, security, or other reasons, various levels of zoning or other mechanisms can be used to restrict the any-to-any access. Performance monitoring and configuration changes or upgrades needed to keep the network performing adequately are more complex.

Tip: Zoning splits the SAN into subnetworks. The servers within a zone have any-to-any connectivity, but anything outside the zone is not visible to them. Virtual SANs (VSANs) are used by Cisco, where the SAN is partitioned at the hardware level, into logical sections, for the purpose of isolating traffic, security, and error containment.

The primary advantage of a switched redundant fabric is that it looks like a very large logical switch, where a single connection provides access to any other port on the total set of switches, as shown in Figure 12-7.

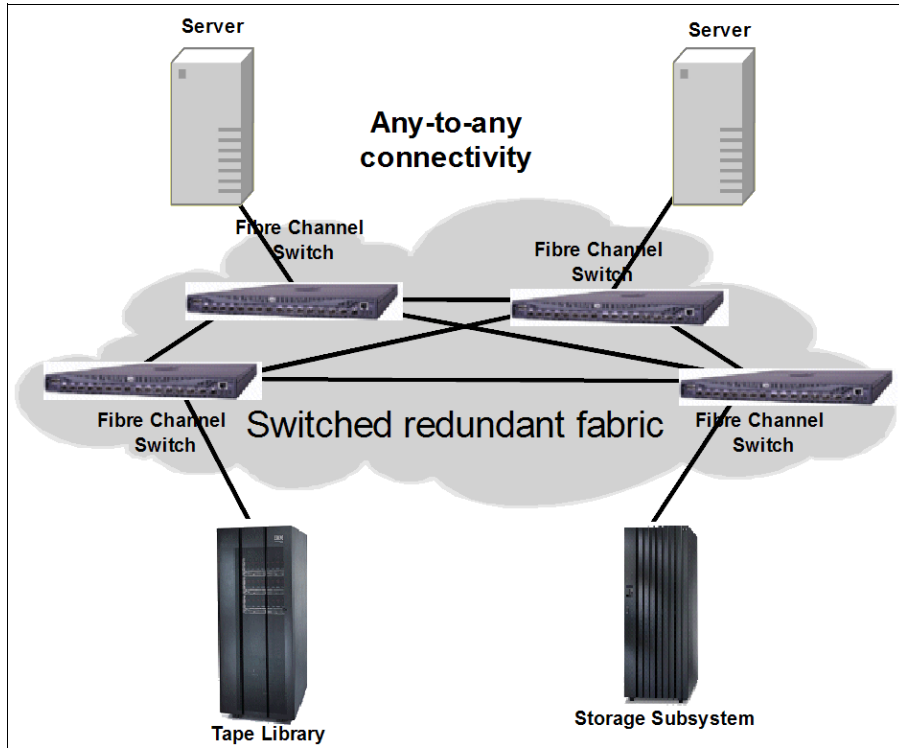


Figure 12-7 Fibre Channel switched topology (redundant switches)

12.2.3 Physical components of the SAN infrastructure

Here we take a brief look at components that are commonly encountered in the FC SAN implementations.

Cables and connectors

FC connectors come mainly in two types as LC and SC connectors. Figure 12-8 shows FC connectors and cables. LC connectors are now the standard.

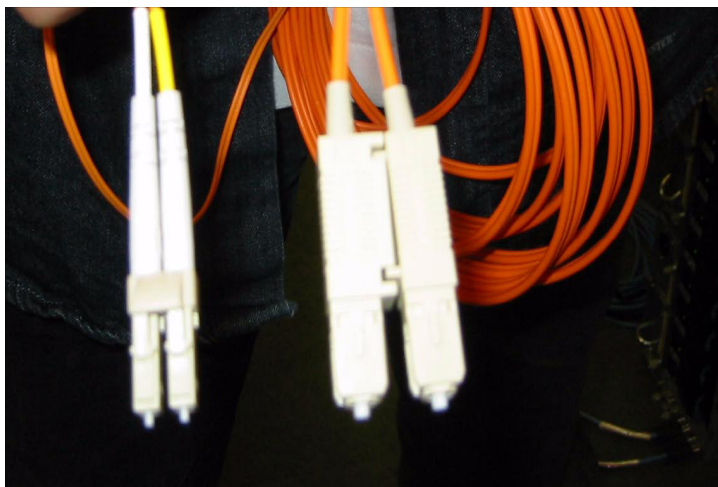


Figure 12-8 LC and SC connectors

Transceivers

Transceivers are plugged on switches or directors on each port and used to convert the internal communication transport to gigabit transport. For 1 Gbps, the transceivers are called Gigabit Interface Converters (GBICs) and for 2 Gbps, 4 Gbps, 8 Gbps and 10 Gbps, the transceivers are called SFPs (Small Form Factor Pluggable Media (SFPs)). Figure 12-9 shows SFPs and a GBIC.

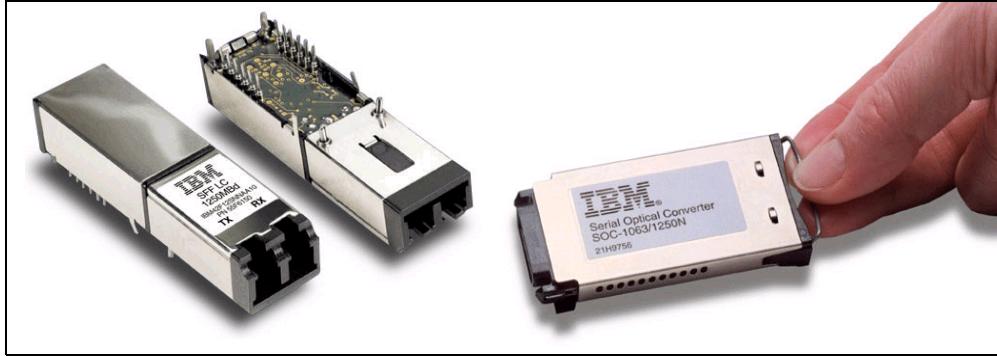


Figure 12-9 SFP (Small Form Factor Pluggable Media) and GBIC (Gigabit Interface Converter)

Fibre Channel Host Bus Adapters

The Host Bus Adapter (HBA) connects to the bus of the host or storage system. It can connect to the cable leading to the SAN. The function of the HBA is to convert the parallel electrical signals from the bus into a serial signal to pass to the SAN. Figure 12-10 shows an HBA.

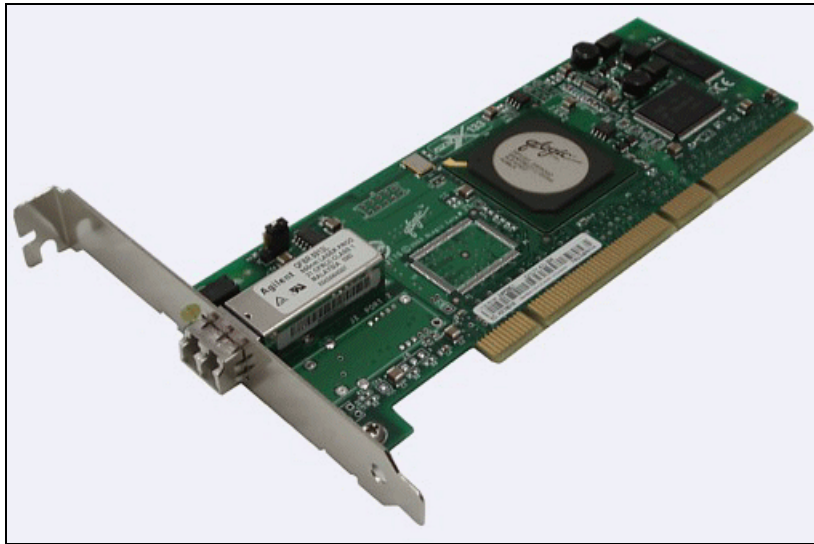


Figure 12-10 Host Bus Adapter with single Fibre port

HBAs with dual Fibre Channel ports are also available. Figure 12-11 shows a dual port HBA.

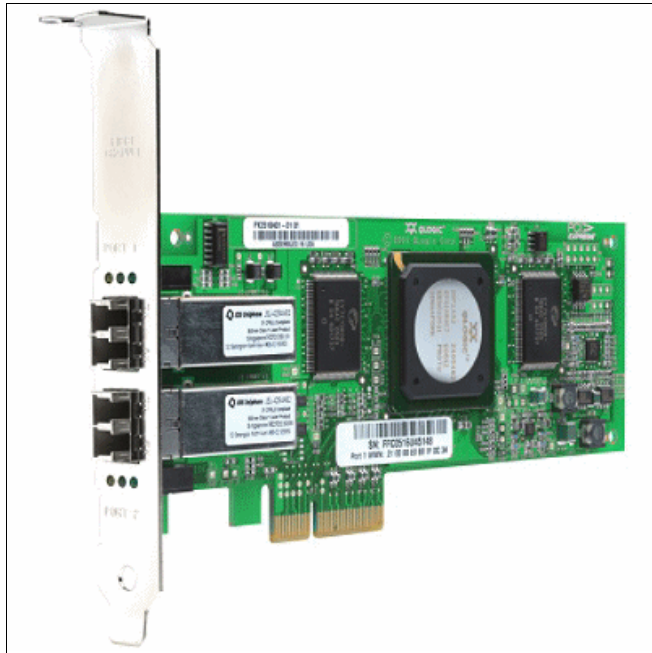


Figure 12-11 Dual port HBA

iSCSI HBAs

For iSCSI based SAN solutions, iSCSI HBAs are available. iSCSI uses TCP/IP and allows two hosts to exchange iSCSI commands using an IP network. Single and dual port iSCSI HBAs are available in the market. Figure 12-12 shows an iSCSI adapter with dual ports.

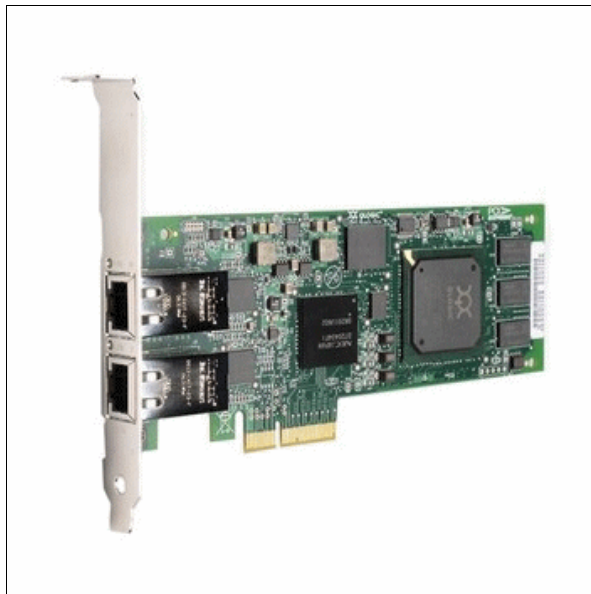


Figure 12-12 Dual port iSCSI HBA

FCoE adapters

Fibre Channel over Ethernet is one of the new technologies in the market that you can use to make the SAN infrastructure available on Ethernet while retaining your existing Fibre Channel SAN. Figure 12-13 shows an FCoE adapter.

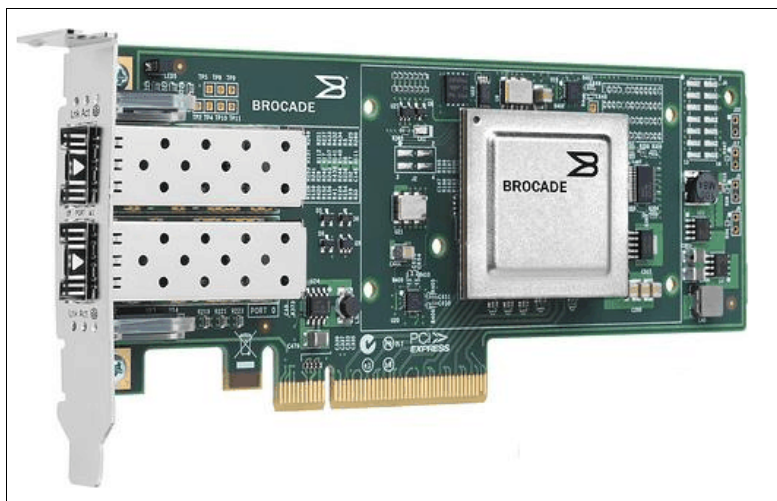


Figure 12-13 FCoE adapter with dual port

SAN switches

SAN switches help your servers, tape drives, libraries, and your SAN storage box to connect in switched fabric. This provides connectivity and easy access to the data stored on your SAN storage devices such as DS5000. SAN switches provide faster speeds up to 10 Gbps with new products in market. Figure 12-14 shows a SAN40B-4 model switch. For more details about the current SAN switch offering by IBM, see Chapter 13, “Storage Area Network products” on page 325.



Figure 12-14 IBM System Storage SAN40B-4

12.2.4 Naming and addressing

Because there are various kinds of Fibre Channel ports and addresses, it is important to understand a few important concepts involved in naming and addressing:

- ▶ A node is a communicating device.
- ▶ A node can contain one or more ports called N_Ports. Each N_port has a unique 8-byte port name or World-Wide-Name.
- ▶ During communication, each N_Port is assigned a 24-bit port address called an N_Port ID that is used for frame routing.
- ▶ A Fibre Channel fabric switch port is called an F_Port. It is for attachment to N_Ports.
- ▶ A standard Fibre Channel mechanism enables switches to network with each other.

12.2.5 Types of ports

Here we describe various types of ports:

- ▶ E_Port: A port used to connect a Multiprotocol Router to an edge fabric. An EX_Port follows standard E_Port protocols and supports FC-NAT but does not allow fabric merging across EX_Ports.
- ▶ FL_Port: A fabric loop port to which a loop attaches; it is the access to the fabric for NL_Ports on a loop.
- ▶ FX_Port: A fabric port that can operate as either an F_Port or an FL_Port.
- ▶ G_Port: A generic port that supports either E_Port or F_Port functionality.
- ▶ L_Ports: A node port (NL_Port) or fabric port (FL_Port) that has arbitrated loop capabilities
- ▶ L_Port: A loop port supporting the Arbitrated Loop protocol. It appears as the part of the output of a switchShow command.
- ▶ NL_Port: A node loop port, supporting the Arbitrated Loop protocol.
- ▶ NPIV: Refers to N_Port ID Virtualization. It allows a single FC_Port to appear as multiple, distinct ports providing separate port identification and security zoning with the fabric for each operating system image as if each one had its own physical port.
- ▶ Nx_Port: A node port that can operate as either an N_Port or NL_Port.

Note: For more information about naming and addressing, see the Redbooks publication, *Introduction to Storage Area Networks*, SG24-5470.

12.2.6 FICON

FICON (Fibre Connectivity) is a high speed input/output interface for mainframe computer connections to storage devices. A FICON channel is a high bandwidth connection between processor and storage device within a relatively close proximity.

FICON channels increase I/O capacity through the combination of a new architecture and faster physical link rates to make them more efficient than ESCON (Enterprise System Connection), IBM previous fiber optic channel standard.

FICON is based on FCP, and runs on an FCP infrastructure - including all the cabling, switched, and directors.

Note: For more information about FICON, see the Redbooks publication, *FICON Native Implementation and Reference Guide*, SG24-6266.

12.3 IP storage networking technologies

iSCSI, FCIP, and iFCP are SAN extension technologies ideal for connecting smaller departmental and less I/O intensive servers into a SAN.

SAN deployment and its resulting benefits have primarily been focused on mission-critical islands of application servers within individual data centers. The difficulty and cost associated with migrating the large number of data center midrange servers to Fibre Channel have made it impractical for IT managers to extend the benefits of SAN to midrange applications. A basic diagram of these technologies is shown in Figure 12-15.

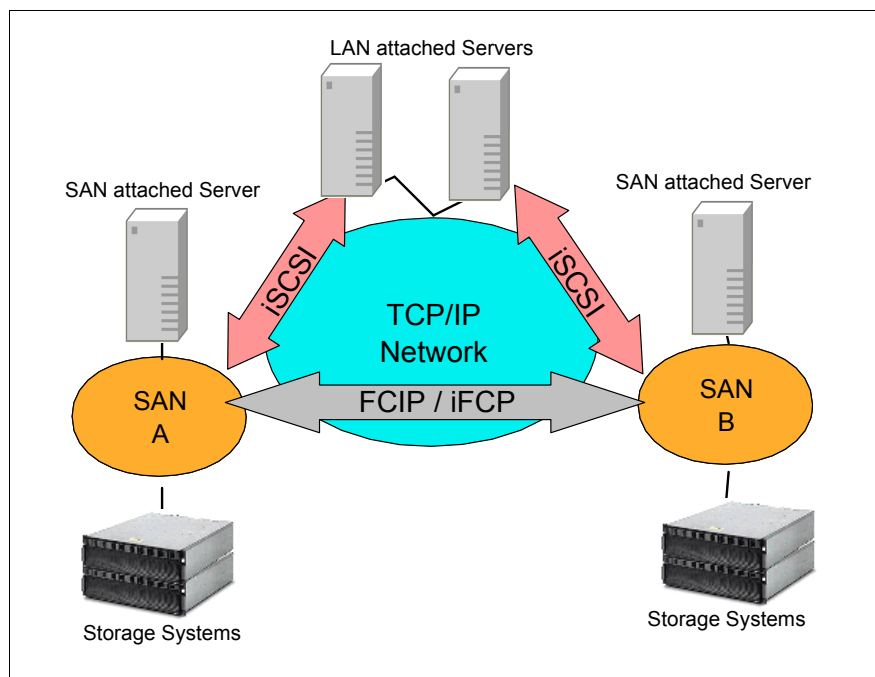


Figure 12-15 IP storage networking technology

12.3.1 iSCSI

iSCSI allows storage to be accessed over a TCP/IP network as though it was locally attached. The server pushes SCSI commands out through the Ethernet NIC. As the SCSI commands exit out through the server, they are encapsulated within IP packet and are forwarded across the LAN to a LAN/SAN gateway interface. These iSCSI packets are then translated onto the Fibre Channel SAN through the TCP/IP transport and conversion protocol.

iSCSI is an ideal point to multipoint solution for connecting dispersed SAN islands. Using iSCSI, midrange servers can gain access to consolidated storage while retaining their existing IP infrastructure, allowing a cost-effective extension of SAN benefits to midrange applications residing on servers within the data center and on departmental servers located throughout the enterprise.

SCSI uses TCP/IP for reliable data transmission over potentially unreliable networks. The iSCSI layer interfaces to the operating system's standard SCSI set. The iSCSI layer includes encapsulated SCSI commands, data and status reporting capability. When, for example, the operating system or application requires a data write operation, the SCSI CDB (Command Descriptor Block) must be encapsulated for transport over a serial gigabit link and delivered to the target.

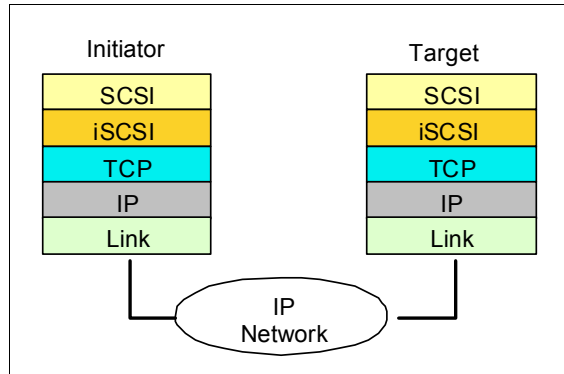


Figure 12-16 iSCSI protocol layering model

The iSCSI protocol monitors the block data transfer and validates completion of the I/O operation. This occurs over one or more TCP connections between initiator and target. In practical applications, an initiator can have multiple target resources over an IP network, and consequently multiple concurrent TCP connections active.

12.3.2 FCIP

Fibre Channel over IP (FCIP) uses a tunneling protocol to transport Fibre Channel frames over an existing IP infrastructure, and therefore is better suited for point-to-point solutions (Figure 12-17). Using FCIP, customers can utilize their current wide-area networking infrastructure for connecting remote SAN islands over long distances.

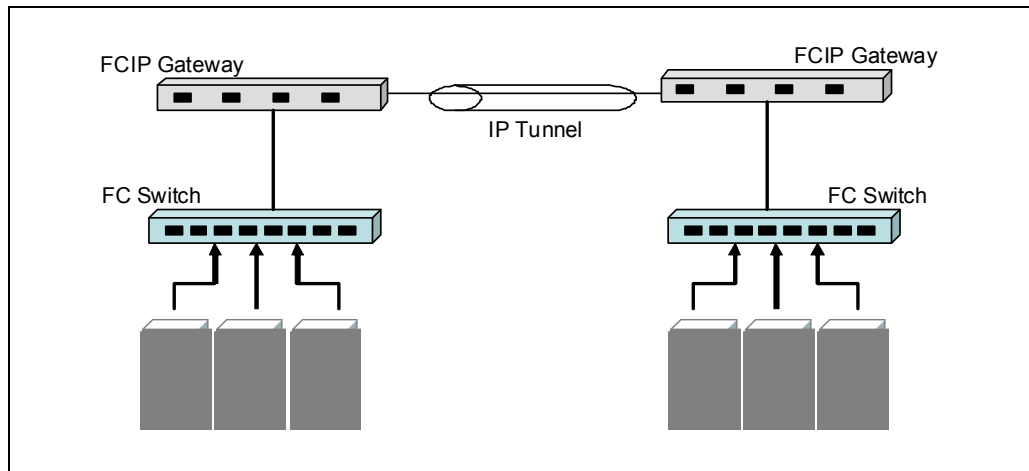


Figure 12-17 FCIP with IP tunneling

Fibre Channel over IP is an ideal combination of technologies to address the dual requirements of storage networking and networking over distance. Fibre Channel is a mature technology optimized for storage data movement within the campus and data center. It represents a major investment in software compatibility, interoperability, and proven applications for campus-based storage networking. Likewise, IP is a mature technology optimized for data movement across WAN distances. It represents a major investment in software compatibility, equipment interoperability, and proven applications for WAN-based data networking.

Today's Fibre Channel-over-IP solutions encapsulate Fibre Channel and transport it over a TCP socket (Figure 12-18). As in all IP networks, performance can vary based on the types of switches and routers, the number of hops the packets must traverse, and the level of congestion in the network. Today, storage transport performance over IP networks-especially over public networks- is limited due to the variable latency of service provider networks.

As IP and Ethernet equipment continues to evolve, higher levels of Quality of Service (QoS), Cost of Service (CoS), provisioning, and circuit emulation must provide the latency guarantees required by synchronous storage applications. In controlled environments, these technologies might even improve the performance of IP networks. Regardless, Fibre Channel over IP is currently a very cost-effective technology for asynchronous applications such as remote data backup.

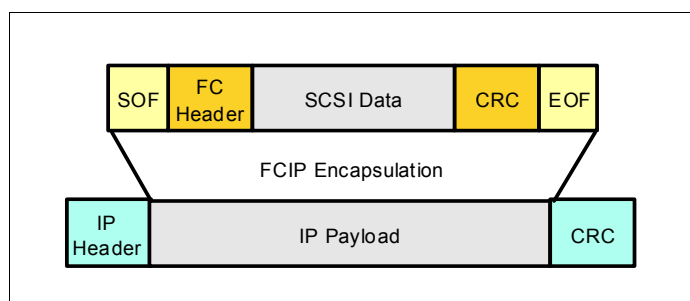


Figure 12-18 FCIP Encapsulation model

12.3.3 iFCP

Internet Fibre Channel Protocol (iFCP) is an emerging standard and TCP/IP-based method for interconnecting Fibre Channel SANs and SAN devices and IP networks. iFCP technology provides multipoint access to Fibre Channel devices. iFCP capitalizes on Internet protocol network services while utilizing the performance and interoperability capabilities of the Fibre Channel network. With iFCP, existing SCSI and Fibre Channel networks can be interconnected into the existing Internet Protocol environment. iFCP can be used in conjunction with Fibre Channel switching and routing protocols, or it can completely replace them.

iFCP is designed for customers who might have a wide range of Fibre Channel devices (that is, Host Bus Adapters, Subsystems, Hubs, Switches, and so on) and want the flexibility to interconnect these devices with IP network. iFCP can interconnect Fibre Channel SANs with IP, as well as allow customers the freedom to use TCP/IP networks in place of Fibre Channel networks for the SAN itself. Through the implementation of iFCP as a gateway-to-gateway protocol, these customers can maintain the benefit of their Fibre Channel devices while utilizing a highly scalable, manageable, and flexible enterprise IP network as the transport medium of choice.

iFCP enables Fibre Channel device-to-device communication over an IP network, providing more flexibility compared to only enabling SAN-to-SAN communication. For example, iFCP has a TCP connection per N_Port to N_Port couple, and such a connection can be set to have its own Quality of Service (QoS) identity (Figure 12-19).

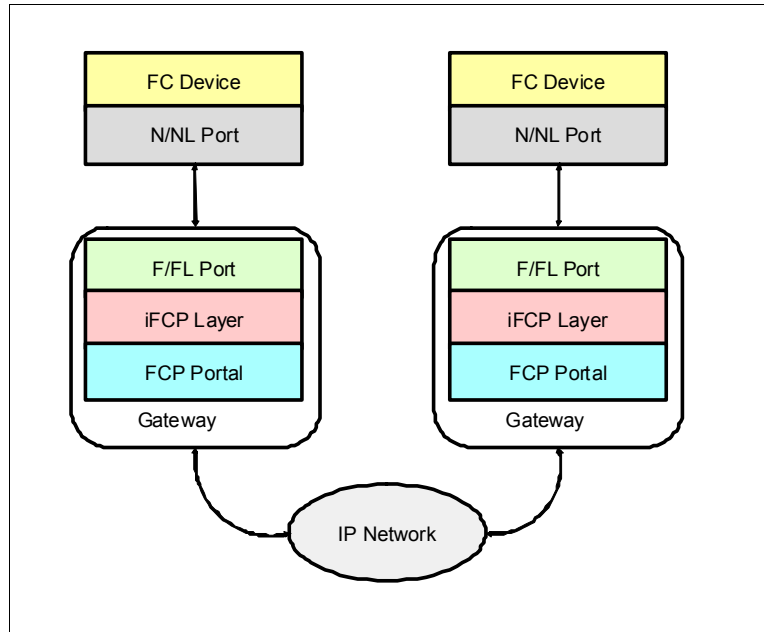


Figure 12-19 iFCP Gateway architecture

A drawback of the FCIP protocol which makes iFCP more attractive is that FCIP is a protocol that uses tunneling to encapsulate Fibre Channel data packets for forwarding over the TCP/IP network. This means that FCIP only works within a Fibre Channel infrastructure, whereas, iFCP can handle both iSCSI and FCIP traffic. Applications developed for Fibre Channel SAN environments are supported over iFCP.

Benefits of Internet Fibre Channel Protocol (iFCP) for storage networks are that scalability, distance, and connectivity issues are virtually eliminated. Existing Ethernet protocol structure allows for rapid deployment of applications and solutions that already make use of the TCP/IP protocol layers. With one less routing protocol to support, network complexity and management is also potentially reduced. More importantly, the lower cost of Gigabit Ethernet switches as opposed to Fibre Channel switches enables for a lower total cost of ownership (TCO) of the enterprise SAN.

12.3.4 Comparison of iSCSI, FCIP, iFCP

Here we provide a short comparison to show the differences between iSCSI, FCIP, and iFCP. In Figure 12-20, you can see the unique structure of the protocol stacks as well as the various key attributes listed in Table 12-1.

iSCSI:

- ▶ A transport protocol for SCSI that operates on top of TCP
- ▶ A new mechanism for encapsulating SCSI commands on an IP network
- ▶ A protocol for a new generation of storage end-nodes that natively use TCP/IP

FCIP:

- ▶ FCIP is a tunneling protocol for connecting geographically distributed Fibre Channel SANs transparently over LANs, MANs, or WANs
- ▶ Relies upon TCP for congestion control and management and upon both TCP and FC for data error and data loss recovery
- ▶ Uses TCP/IP as the transport while retaining Fibre Channel services intact

iFCP:

- ▶ A gateway-to-gateway protocol for the implementation of a Fibre Channel fabric in which TCP/IP switching and routing elements supplement or replace Fibre Channel fabric components
- ▶ The protocol enables the attachment of existing Fibre Channel storage devices or Fibre Channel SANs to an IP network

The following benefits are common to all IP storage protocols:

- ▶ Built on SCSI and Ethernet technologies
- ▶ Provides more affordable SAN infrastructure
- ▶ Increases operating distance
- ▶ Improves availability of storage systems

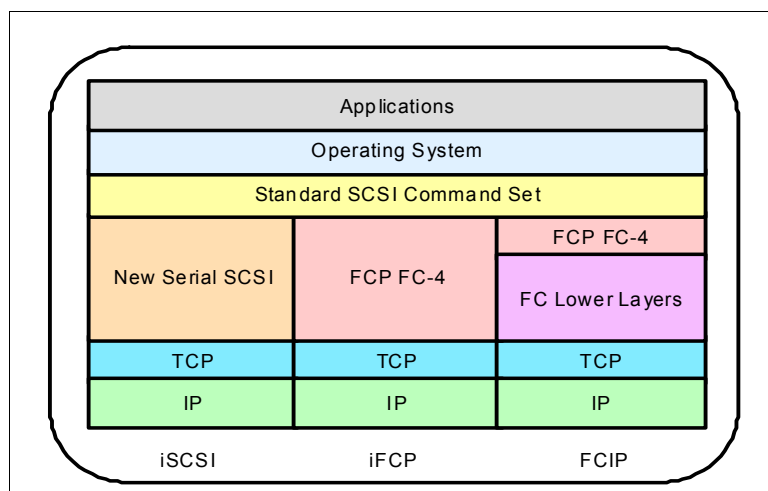


Figure 12-20 iSCSI, FCIP, and iFCP protocol stacks

Table 12-1 summarizes the key attributes of iSCSI, FCIP, and iFCP protocols.

Table 12-1 Key attributes iSCSI, FCIP, iFCP

Protocol Attributes	iFCP	iSCSI	FCIP
Implementation	native IP transport	native IP transport	encapsulation, tunneling
SCSI encapsulation	FCP	new iSCSI layer	FCP
Prioritization based on port identification	yes	yes	no
End device interface	FC/FCP	IP/iSCSI	FC/FCP
End device routing	RIP, OSPF, BGP	RIP, OSPF, BGP	FSPF
Fibre Channel device support	yes	no	yes

12.3.5 FCoE

FCoE is an enhancement that expands FC into the Ethernet by combining two leading-edge technologies (FC and the Ethernet).

FCoE is the standard that is driving convergence and the emergence of FC over the Ethernet. However, in and of itself (FC over Ethernet), FCoE will not be enough to allow for fabric convergence. Without question, it is a move in the correct direction but without any enhancements. Ethernet itself does not meet the requirements for data center convergence. FCoCEE (pronounced eff-see-oh-see) is Fibre Channel over Convergence Enhanced Ethernet, which is enabled by FCoE.

It is the transport, or mapping, of encapsulated FC frames over the Ethernet. Very simply, the Ethernet provides the physical interface, and FC provides the transport protocol, giving us an FC frame delivered in an Ethernet frame.

Mixing FCoE and existing FC and Ethernet networks is easy and expected; the most probable implementation is “from the edge” ... adding FCoE with new equipment while keeping existing Ethernet and FC hardware/cabling in place until it make sense to replace with FCoE.

Here we list highlights of FCoE:

- ▶ FCoE uses Converged Network Adapters (CNA)
- ▶ CNA run either Ethernet NIC traffic AND/OR Fibre Channel traffic
- ▶ Enhanced Ethernet protocol supports FC traffic
- ▶ Enhancements adds loss-less data transmission and additional management functions
- ▶ FCoE also called FCoCEE – Fibre Channel over Converged Enhanced Ethernet
- ▶ Physically CNA use 10Gb Ethernet ports
- ▶ Each port can run all NIC, all FC, or mixed NIC/FC traffic

Note: For more information about FCoE, see the Redpaper publication, *An Introduction to Fibre Channel over Ethernet, and Fibre Channel over Convergence Enhanced Ethernet*, REDP-4493.

12.4 Network Attached Storage

Network Attached Storage (NAS) is a device that resides on a network that can be shared with non-storage traffic. Today, the network is usually an Ethernet LAN, but can be any network that supports the IP-based protocols that NAS uses.

Figure 12-21 shows a diagram of an NAS appliance. In contrast to “block I/O” used by SANs, NAS I/O requests are called “file I/Os”. File I/O is a higher-level type of request that, in essence, specifies the file to be accessed, an offset into the file (as though the file was a set of contiguous bytes), and a number of bytes to read or write beginning at that offset. File I/O requests are mainly of the type CIFS, NFS or specialized I/O protocols for file access and file sharing. Unlike block I/O, there is no awareness of a disk volume or disk sectors in a file I/O request. Inside the NAS product, an operating system tracks where files are located on disk, and issues a block I/O request to the disks to fulfill the file I/O read and write requests it receives.

In contrast to SAN devices that can usually also be direct-attached (for example, by point-to-point Fibre Channel) as well as network-attached by SAN hubs and switches, an NAS device is generally only an NAS device and attaches only to processors over a LAN or WAN.

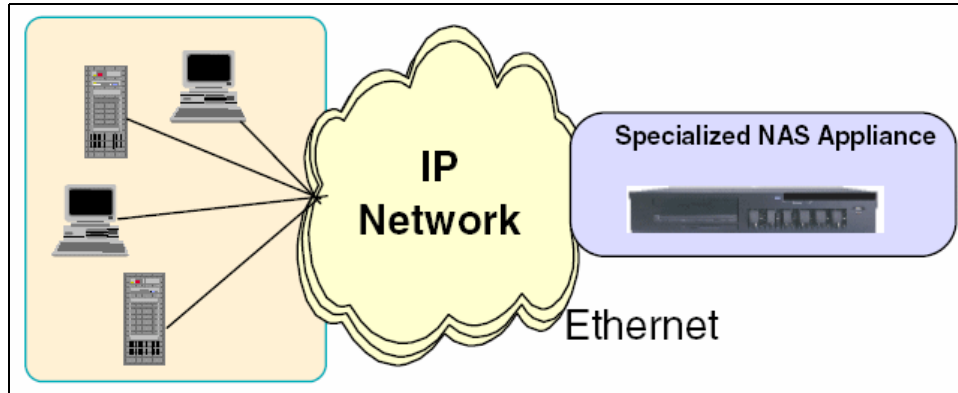


Figure 12-21 A picture of NAS appliance

12.4.1 NAS benefits

The major potential benefits of an NAS device can be categorized as follows:

- ▶ Ease-of-installation:

NAS is generally easier to install and manage than a SAN. An NAS appliance can usually be installed on an existing LAN/WAN network. NAS manufacturers often cite “up and running” times of 30 minutes or less. Hosts can potentially start to access NAS storage quickly, without needing disk volume definitions or special device drivers. In contrast, SANs take more planning, including design of a Fibre Channel network and selection/installation of SAN management software.

- ▶ Resource pooling:

NAS allows capacity within the appliance to be pooled. That is, the NAS device is configured as one or more file systems, each residing on a specified set of disk volumes. All users accessing the same file system are assigned space within it on demand. That is certainly more efficient than buying each user their own disk volumes (DAS), which often results in certain users having too much capacity and others too little. So NAS pooling can minimize the need to manually reassign capacity among users. However, NAS pooling resides within an NAS appliance, and there is little if any sharing of resources across multiple appliances. This raises costs and management complexity as the number of NAS nodes increases. In contrast, an advantage of a SAN is that all devices on a SAN can be pooled, multiple disk and tape systems. So, eventually as total capacity grows, a SAN might be easier to manage and more cost-effective.

- ▶ File sharing:

NAS provides file sharing using the NFS and CIFS protocol.

12.4.2 NAS gateways

An NAS gateway provides the function of a conventional NAS appliance but without integrated disk storage. The disk storage is attached externally to the gateway, possibly sold separately, and can also be a stand-alone offering for direct or SAN attachment. The gateway accepts a file I/O request (for example, using the NFS or CIFS protocols) and translates that to a SCSI block-I/O request to access the external attached disk storage.

The gateway approach to file sharing offers the benefits of a conventional NAS appliance, with additional potential advantages:

- ▶ Increased disk capacity scalability (compared to the capacity limits of an integrated NAS appliance).
- ▶ Ability to offer file sharing and block-I/O on the same disk system. Disk capacity in the SAN can be shared (reassigned) among gateway and non-gateway use. So a gateway can be viewed as an NAS/SAN hybrid, increasing flexibility and potentially lowering costs.

For more information about NAS, see Chapter 5, “IBM System Storage N series” on page 173.

12.5 Wavelength Division Multiplexing

Wavelength Division Multiplexing (WDM) or Multiplexing by itself is not a storage networking protocol. Because it is used also to create storage networking infrastructure, we briefly explain the technology here.

Wavelength Division Multiplexing is deployed as part of the physical layer. It is therefore independent of protocol, simply passing signal information in the format it is received. WDM allows the simultaneous transmission of a number of data streams over the same physical fiber, each using a unique optical wavelength.

WDM receives incoming optical signals from many sources (Fibre Channel, IP, ESCON, FICON) which it converts to electrical signals, it then assigns them a specific wavelength (or lambdas) of light and retransmits them on that wavelength. This method relies on the large number of wavelengths available within the light spectrum. You can think about WDM as though each channel is a particular color of light; several channels then make up a rainbow (Figure 12-22). In summary WDM enables many signals to be concentrated into a single fibre, all being sent at various wavelengths.

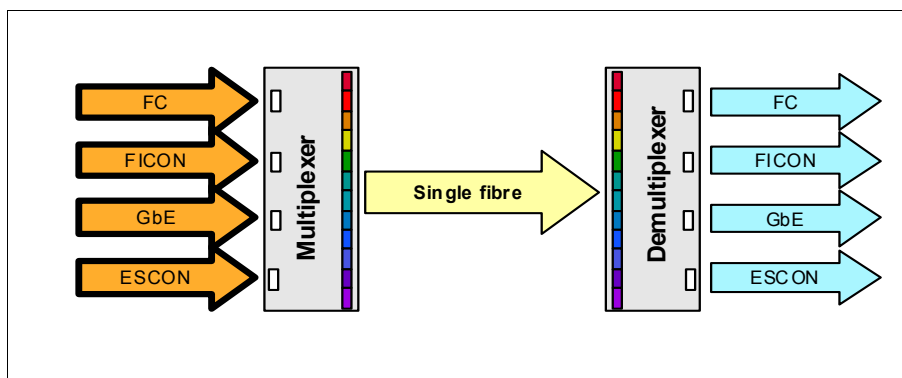


Figure 12-22 Wave Division Multiplexer concept

Each wavelength can carry a signal at any bit rate less than an upper limit defined by the electronics, typically up to several gigabits per second. Due to the nature of these boxes, they are often considered transparent to protocol and bit rate.

Today, WDM of itself, is rarely seen. Coarse Wave Division Multiplexing (CWDM), which combines up to 16 wavelengths onto a single fiber, and Dense Wave Division Multiplexing (DWDM) which combines up to 64 wavelengths onto a single fiber are the standard.

12.6 Selecting the best alternative

Which storage networking alternative is best for a given organization might be obvious based on organizational objectives, current storage infrastructure, and what the alternatives provide. Or, it might be a totally open question. Storage technology has clearly become more varied and sophisticated, and accordingly decisions have become more complex than ever. Choice means flexibility, and that is good, but which choice to make is not always clear.

- ▶ If a group of individual users with PCs needs to share disk storage capacity and perhaps also share files in that storage, then NAS might be easiest to install and manage.
- ▶ If application servers need to share disk storage, and are each accessing independent (block I/O) databases, an FC based SAN might be appropriate.
- ▶ For a small number of servers where no SAN exists, iSCSI might be less expensive and less complex.
- ▶ If required, FCIP or iFCP provide a cost-effective way to achieve business protection, enabling such solutions as remote tape archiving.

12.7 More information

For a better understanding of Fibre Channel, SAN technology, FICON and NAS technology see these Redbooks publications:

- ▶ *Introduction to Storage Area Networks*, SG24-5470
- ▶ *Designing an IBM Storage Area Network*, SG24-5758
- ▶ *IBM SAN Survival Guide*, SG24-6143
- ▶ *FICON Implementation Guide*, SG24-6497
- ▶ *FICON (FCV Mode) Planning Guide*, SG24-5445
- ▶ *Introduction to IBM S/390 FICON*, SG24-5176
- ▶ *IBM S/390 FICON Implementation Guide*, SG24-5169
- ▶ *The IBM TotalStorage NAS Integration Guide*, SG24-6505



Storage Area Network products

Companies are searching for more efficient ways to manage ever-expanding volumes of data and to make that data accessible throughout the enterprise - this is propelling the move of storage into the network. The Storage Area Network (SAN) infrastructure offers simplified storage management, scalability, flexibility, availability, and improved data access, movement, and backup.

SAN switches interconnect multiple host servers with storage servers and devices to create a SAN. The SAN switches can be used either as stand-alone devices to build a simple SAN fabric, or they can be interconnected with other switches to build a larger SAN fabric.

In this chapter we present the IBM System Storage SAN product portfolio, which includes Fibre Channel switches and directors with Fibre Channel and FICON ports.

The following entry, midrange, and enterprise level switch products are presented:

- ▶ IBM System Storage SAN b-type switches and director
- ▶ Cisco switches and directors

Note: These products can be used with disk storage systems, which are explained in chapters for Disk Systems. Additionally, these products can be used with tape drives and libraries, as discussed in Part 2, "Tape systems" on page 225.

For the latest SAN products and information, see the website:

<http://www.ibm.com/servers/storage/san/>

13.1 SAN switches and connectivity products

IBM System Storage SAN switch offerings provide integrated small and medium business (SMB) and enterprise solutions with multiprotocol local, campus, metropolitan, and global storage networking. IBM provides the choice of Brocade (b-type), and Cisco switches and directors.

IBM System Storage SAN switch products provide a broad range of storage networking options designed with a common architecture and integrated enterprise SAN management capabilities, and supported by the broadest range of IBM open server and storage devices. The interconnection of IBM and compatible switches can support the creation of scalable, dual redundant core-to-edge SAN fabrics that can support high performance, scalability, and fault tolerance required by e-business applications and enterprise storage management applications.

IBM System Storage SAN switch offerings can be positioned into four broad product groups: entry, midrange, enterprise, and multiprotocol router. Suitable products for every SAN project can be found among these.

13.1.1 Common characteristics

With all IBM System Storage SAN products, you get the following capabilities:

- ▶ Industry standard performance with 1, 2, 4, 8, 10 Gigabit per second (Gbps) throughput
- ▶ Broadest range of IBM open server and storage support including fabric, loop, and private loop attachments
- ▶ Intelligent Fabric Services Architecture providing switch interoperability
- ▶ Enterprise-level scalability with fault-tolerant core-to-edge SAN fabrics containing thousands of devices
- ▶ Open fabric management allowing support for the widest range of solutions, from very small workgroup SANs up to very large enterprise SAN fabrics with thousands of devices
- ▶ Flexible management options including Tivoli Ready certification for centralized management of very large enterprise SAN fabrics
- ▶ Manageable by IBM Tivoli Productivity Center for Fabric
- ▶ Common enterprise SAN fabric which simplifies deployment, management and network growth
- ▶ Common firmware which enables introduction of new switch technologies while protecting prior switch investments
- ▶ Flexible Fibre Channel connectivity provides connectivity to a host of IBM and non-IBM servers and storage products
- ▶ Pay-as-you-grow scalability provides scalable network growth in a modular, cost-effective, and nondisruptive manner

13.1.2 Other switch features

Here we cover other features available in selected IBM System Storage SAN switch products.

Advanced Web Tools

Advanced Web Tools is an intuitive graphical user interface (GUI) that allows network managers to monitor and manage SAN fabrics consisting of switches using a Java-capable web browser from standard desktop workstations.

Advanced Performance Monitoring

Advanced Performance Monitoring enables SAN administrators to monitor transmit and receive traffic from the source device all the way to the destination device. This end-to-end visibility into the fabric enables SAN administrators to identify bottlenecks and optimize fabric configuration resources.

Advanced Security

Advanced Security (AS) significantly reduces the security holes left by traditional SAN implementations and greatly improves the ability to minimize SAN-specific vulnerabilities by providing a comprehensive, policy based security system for IBM SAN Switch fabrics.

Advanced Zoning

Advanced Zoning segments a fabric into virtual private SANs. It provides data exchange between devices in the same zone and prohibits exchange to any device not in the same zone. Advanced zoning of 1, 2, and 4 Gbps switches separately enlarges the range of hardware enforcement and so provides the switch with more security access control functions as before, preventing unauthorized devices from accessing the fabric.

Common IBM SAN Switch capabilities

IBM System Storage SAN Switches include universal ports that can determine the port type when connected to a fabric port (F_port), fabric loop port (FL_port) or expansion port (E_port).

Enterprise Fabric Connectivity Manager

Enterprise Fabric Connectivity Manager (EFCM) software is designed to support interconnection of multiple IBM System Storage SAN m-type switches and directors for the creation of enterprise-to-edge SANs.

Extended Fabric

The Extended Fabric feature provides extensions within the internal switch buffers. This maintains performance with distances greater than 10 km, and up to 120 km, by maximizing buffering between the selected switch interconnect links. With the Extended Fabric feature, the ISLs are configured with up to 60 buffer credits and optimize buffers for up to 120 km on 1 Gbps fiber optic link, and up to 60 km on 2 Gbps fiber optic link.

Fabric Manager

Fabric Manager provides a Java-based application that can simplify management of a multiple switch fabric. It administers, configures, and maintains fabric switches and SANs with host-based software.

Fabric Watch

Fabric Watch enables switches to continuously monitor the health of the fabrics, watching for potential faults based on defined thresholds for fabric elements and events, so making it easy to quickly identify and escalate potential problems.

FlexPort scalability

FlexPort technology is designed to support scalable switch upgrades in four-port increments without fabric disruption.

ISL Trunking

The ISL Trunking feature allows up to four ISLs between the same pair of switches to be grouped and to act as a single, high speed “pipe” or trunk with a capacity of up to 8 Gbps.

Remote Switch

The Remote Switch feature enables two fabric switches to be connected over an asynchronous transfer mode (ATM) connection. This requires a compatible Fibre Channel to ATM gateway, and can have a distance of up to 10 km between each switch and the respective ATM gateway.

More information

- For the interoperability matrix, see the website:

<http://www.ibm.com/servers/storage/support/san/index.html>

Select the product from list, then click *Plan and Upgrade* under the desired switch product.

- For technical support information, see the website:

<http://www.ibm.com/servers/storage/support/san/index.html>

- SAN-related Redbooks publications include:

- *Introduction to Storage Area Networks*, SG24-5470
- *IBM/Cisco Multiprotocol Routing: An Introduction and Implementation*, SG24-7543
- *Implementing an IBM/Cisco SAN*, SG24-7545
- *IBM SAN Survival Guide*, SG24-6143
- *SAN Storage Performance Management Using TotalStorage Productivity Center*, SG24-7364
- *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116

13.2 Entry SAN switches

The IBM System Storage SAN switch entry-level products are designed specifically to address the needs of small to medium-size SAN environments. They can be used to create a wide range of high-performance SAN solutions, from simple single-switch configurations to larger multi-switch configurations which support fabric connectivity and advanced business continuity capabilities.

Infrastructure simplification solutions for IBM System x, IBM System i, and IBM System p include storage consolidation and high-availability server clustering with IBM System Storage disk storage arrays. Business continuity solutions include data protection with IBM System Storage tape libraries and devices and IBM Tivoli Storage Manager data protection software. IBM entry fabric switches provide up to 8 Gbps, 24 ports for fully non-blocking performance, and advanced intelligence features.

Entry-level products include:

- IBM System Storage SAN24B-4 Express
- Cisco MDS 9124 Express for System Storage

In this section, we introduce IBM System Storage Switch solutions ideal for entry-level applications.

13.2.1 IBM System Storage SAN24B-4 switch

The SAN24B-4 is a high performance scalable switch that provides 24 fabric with a “ports on demand” feature, so you can scale your network as you grow. You can get the switch with 8 port licence and then can upgrade it conveniently up to 24 by just enabling the licence. With auto-sensing link speeds at 1, 2, 4, and 8 Gbps and a flexible design to configure this switch as a fabric switch or an Access Gateway, it is suitable for small to mid-sized businesses.

A single SAN24B-4 Express switch can serve as the cornerstone of a Storage Area Network for those who want to obtain the benefits of storage consolidation and are just beginning to implement Fibre Channel storage systems. Such an entry-level configuration can consist of one or two Fibre Channel links to a disk storage array or to an LTO tape drive. An entry-level eight-port storage consolidation solution can support up to seven servers with a single path to either disk or tape. The “Ports on Demand” feature is designed to enable a base switch to grow to sixteen or twenty-four ports to support more servers and more storage devices without taking the switch offline. Figure 13-1 shows the SAN24B-4 fabric switch.



Figure 13-1 IBM System Storage SAN24B-4 Express

The SAN24B-4 fabric switch requires Fabric OS v6.1.0 or later. The switch offers easy to use Web Tools, 8 Gb FC, Long Distance support, Advanced Zoning, Full-Fabric support, Fiber Watch, Advanced Performance Monitoring, Enhanced Group Management, and ISL Trunking. The base switch also offers eight default ports and Ports on Demand (POD) licenses are available in eight-port increments. With flexible architecture based on GoldenEye2 ASIC, the switch supports F, FL,E, and M Ports at 8 Gbps. The switch also has USB port support for firmware download, configuration upload and download, and supportsave.The SAN24B-4 has a 1U form factor and is a single FRU with no field replaceable parts. The switch has one power supply and three integrated fans.

Note: Access Gateway mode is supported only in 24-port configurations, and *only* 2 GB Brocade branded USB drives are supported on the USB port. The 4 and 8 Gbps link speeds are supported only with Brocade branded SFPs

Highlights

Here we list highlights of the IBM System Storage SAN24B-4 Express (Figure 13-1):

- Foundation for new infrastructure simplification and business continuity solutions for servers running Microsoft Windows, Unix, Linux and IBM AIX and OS/400 operating systems
- “Ports on Demand” scalability from 8 to 16 to 24 ports
- Provides new levels of performance with 8 Gbps Fibre Channel (FC) technology

More information

- For additional product details, see the website:

<http://www.ibm.com/systems/storage/san/b-type/san24b-4/express>

13.2.2 Cisco MDS 9124 Express for IBM System Storage

Cisco MDS 9124 SAN switch provides 24 auto-sensing Fibre Channel ports capable of speeds of 4, 2, and 1 Gbps in a compact 1RU form-factor chassis and is designed to meet the performance and scalability requirements of the most demanding environments.

The base MDS 9124 comes with eight ports activated, two redundant hot swappable power supplies, and eight shortwave SFPs. Enhanced flexibility of MDS 9124 is provided by the 9124 8-port activation license. Using this functionality, customers can start with a base configuration of 8 ports and upgrade to 16 and 24 ports. Figure 13-2 shows the MDS 9124 Multilayer Fabric Switch.

Higher availability solutions can be created using multiple Cisco MDS 9124 switches. Such implementations can be well-suited to server clustering environments. Such a configuration can support from six to 22 servers, each with dual Fibre Channel adapters cross-connected to redundant 9124 switches, which are cross-connected to a dual-controller storage system.

While the focus of the Cisco MDS 9124 is as the foundation of medium-sized SMB SANs, it can also be configured to participate as a component of a tiered enterprise SAN with other members of the Cisco MDS 9000 for IBM System Storage family. Cisco MDS 9000 SAN-OS firmware provides enterprise-class capabilities such as virtual SANs (VSANs), Port Channels, quality of service (QoS) and security for deployment in core-edge enterprise SAN solutions. These capabilities help provide investment protection as SAN requirements evolve and grow over time (Figure 13-2).



Figure 13-2 Cisco MDS 9124 Express

Highlights

Here we list highlights of the Cisco MDS 9124 Express (Figure 13-2):

- ▶ Foundation for new infrastructure simplification and business continuity solutions for servers running Microsoft Windows, UNIX, Linux, NetWare and IBM OS/400 operating systems
- ▶ Designed for high availability with hot-swappable, dual power supplies and non-disruptive firmware upgrades
- ▶ Cisco MDS 9000 family compatibility supports scalability and consistent service as the SAN grows
- ▶ Enterprise Package and Fabric Manager Server Package provide added intelligence and value

Note: The MDS 9124 Multilayer switch supports N-Port identifier virtualization (NPV) to reduce the number of Fibre Channel domain IDs in SANs.

More information

For additional product details, see the website:

<http://www.ibm.com/systems/storage/san/ctype/9124>

13.3 Midrange SAN switches

IBM System Storage midrange SAN solutions provide additional capability, features, and benefits beyond the simple entry solutions. The IBM System Storage midrange SAN switches provide 1, 2, 4 and 8 Gbps port-to-port non-blocking throughput with auto-sensing capability for connecting to older 1 Gbps host servers, storage, and switches. They come in 40 to 80 ports with “ports on demand” scalability models.

IBM System Storage SAN switches can be used to create dedicated, reliable, and high performance networks for storage products like disk subsystems, tape drives, and tape drive libraries. In addition, all of these models are fully interoperable with the previous IBM System Storage SAN Switches, and can be added to existing fabrics, enabling transition from existing Fibre Channel storage networks to the faster technology.

Midrange SAN switches include:

- ▶ IBM System Storage SAN80B-4
- ▶ IBM System Storage SAN40B-4
- ▶ Cisco MDS 9134 for IBM System Storage

In this section, we introduce IBM System Storage Switch solutions ideal for midrange applications.

13.3.1 IBM System Storage SAN40B-4

The IBM System Storage SAN40B-4 SAN fabric switch provides 24, 32, or 40 active ports and is designed for high performance with 8 Gbps link speeds and backward compatibility to support links running at 4, 2, and 1 Gbps link speeds. High availability features make it suitable for use as a core switch in midrange environments or as an edge-switch in enterprise environments where a wide range of SAN infrastructure simplification and business continuity configurations are possible. IBM POWER Systems, System x, System z and many non-IBM disk and tape devices are supported in many common operating system environments. Optional features provide specialized distance extension, dynamic routing between separate or heterogeneous fabrics, link trunking, FICON, performance monitoring and advanced security capabilities.

The SAN40B-4 fabric switch requires Fabric OS v6.1.0 or later. The base model of the switch has 24 ports enabled, and the POD licenses are available in 8 port increments. Integrated Routing is a licensed feature which is supported on every port of the switch and requires a POD license for all 40 ports. The ports on the switch are grouped in 8-port groups matching the trunk group, and with ISL Trunking speeds of up to 64 Gbps can be achieved per trunk.

Dynamic Path selection can be used for optimizing the performance and load balancing, and the switch can be managed using Web Tools. The built-in USB port can be used for firmware download, configuration upload and download, and supports save, and the switch supports non-disruptive firmware downloads. New features in Fabric OS v6.2.0 make the switch Virtual Fabric capable. A single physical chassis can be subdivided into two or more logical switches creating a logical fabric with other switches.

Integrated Routing is a licensed feature that is supported on every port of the switch and requires the POD license for all 40 Ports. Two hot-swappable, redundant 125W power supply

and fan assemblies are included with the switch and these are field replaceable units (FRU). Each FRU has an ON/OFF switch AC plug and a power supply/fan status LED, and the switch has a 1U form factor.

Highlights

Here we list highlights of the IBM System Storage SAN40B-4 (Figure 13-3):

- ▶ High port density design with up to 40 ports in an efficient, space saving 1U height
- ▶ Simple-to-use midrange and enterprise SAN fabric switch for IBM POWER Systems, System x, System z and other server environments
- ▶ New levels of performance with 8 Gbps Fibre Channel (FC) technology
- ▶ Non-disruptive capacity activation from 24 to 32 to 40 ports with Ports on Demand scalability
- ▶ High availability with redundant, hot-swappable fans and power supplies and non-disruptive software upgrades



Figure 13-3 IBM System Storage SAN40B-4

More information

For additional product details, see the website:

<http://www.ibm.com/systems/storage/san/b-type/san40b-4>

13.3.2 IBM System Storage SAN80B-4

The IBM System Storage SAN80B-4 SAN fabric switch provides 48, 64, or 80 active ports and is designed for high performance with 8 Gbps link speeds and backward compatibility to support links running at 4, 2, and 1 Gbps link speeds. High availability features make it suitable for use as a core switch in midrange environments or as an edge-switch in enterprise environments where a wide range of SAN infrastructure simplification and business continuity configurations are possible. IBM Power Systems, System x, System z and many non-IBM disk and tape devices are supported in many common operating system (OS) environments. Optional features provide specialised distance extension, dynamic routing between separate or heterogeneous fabrics, link trunking, Fibre Connection (FICON), performance monitoring and advanced security capabilities.

The SAN80B-4 fabric switch requires Fabric OS v6.1.0 or later. Port hardware is based on the GoldenEye2 ASIC. Each ASIC can support 32 ports at 1, 2, 4, and 8 Gbps link speeds. The base model of the switch comes with 48 ports enabled, and the POD licenses are available in 16-port increments. New features in Fabric OS v6.2.0 make the switch Virtual Fabric capable. A single physical chassis can be subdivided into two or more logical switches creating a logical fabric with other switches. Integrated Routing is a licensed feature that is supported on every port of the switch and requires the POD license for all 80 Ports. The ports on the switch are grouped in 8-port groups matching the trunk group, and with ISL Trunking speeds of up to 64 Gbps can be achieved per trunk. Dynamic Path selection can be used for optimizing the performance and load balancing, and the switch can be managed using Web Tools.

Important: The USB port supports only a 2 GB Brocade branded USB drive. The 4 and 8 Gbps link speeds are supported only with Brocade branded SFPs.

IBM/Brocade 8 Gbps Directors and Switches The built-in USB port can be used for firmware download, configuration upload and download, and support save, and the switch supports non-disruptive firmware downloads. The switch has two hot-swappable, redundant 300W power supplies and three hot-swappable fan assemblies. Both the power supplies and the fan assemblies are field replaceable units, and they have a status LED on them.

Highlights

Here we list highlights of the IBM System Storage SAN80B-4 (Figure 13-4):

- ▶ High port density design with up to 80 ports in an efficient, compact 2U height helps save rack space
- ▶ Robust midrange and enterprise SAN fabric switch for IBM POWER Systems, System x, System z and other server environments
- ▶ Provides new levels of performance with 8 Gbps Fibre Channel (FC) technology
- ▶ Ports-on-Demand scalability supports non-disruptive capacity activation from 48 to 64 to 80 ports
- ▶ Designed to support high availability with redundant, hot-swappable fans and power supplies and non-disruptive software upgrades



Figure 13-4 IBM System Storage SAN80B-4

More information

For additional product details, see the website:

<http://www.ibm.com/systems/storage/san/b-type/san80b-4>

13.3.3 Cisco MDS 9134 for IBM System Storage

A wide range of IBM System Storage medium-size and enterprise storage area network (SAN) IT simplification and business continuity solutions can be created with the Cisco MDS 9134 for IBM System Storage stackable fabric switch. Infrastructure simplification solutions for the IBM System i, System p, System x and System z families of servers include storage consolidation and high-availability server clustering with IBM System Storage disk storage arrays. Business continuity solutions include data protection with IBM System Storage tape libraries and devices and IBM Tivoli Storage Manager data protection software; and disaster protection with IBM System Storage disk metro and global mirroring disaster recovery solutions.

The Cisco MDS 9134 for IBM System Storage is designed to address the needs of medium-sized businesses and large enterprises with a wide range of Storage Area Network (SAN) capabilities. It can be used as part of a high performance simple SAN with

single-switch or stacked switch configurations for business-class customers in support of IT simplification and business continuity solutions. It can also be used as an edge switch for device aggregation with 10 Gbps core director configurations for large enterprise customers.

Fabric connectivity capabilities can be the basis for IT simplification solutions for IBM System i, System p, System x and System z servers and storage consolidation and high-availability server clustering with IBM System Storage disk storage arrays. Business continuity capabilities can help businesses protect valuable data with IBM System Storage tape libraries and IBM Tivoli Storage Manager data protection software. Advanced connectivity capabilities can help businesses protect against major disasters with IBM System Storage disk metro and global mirroring disaster recovery solutions.

Highlights

Here we list highlights of the Cisco MDS 9134 for IBM System Storage (Figure 13-5):

- ▶ Simple-to-use 4 and 10 Gigabit per second performance for simplification and business continuity solutions with Windows, UNIX, Linux, NetWare, IBM OS/400 and IBM z/OS servers
- ▶ Cost-effective “green” switch requires up to 50 percent less power per port
- ▶ Stackable design provides pay-as-you-grow scalability and flexibility with On-Demand port activation features
- ▶ Designed to support high availability with redundant, hot swappable power supplies and fans and non-disruptive firmware upgrades
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Enterprise, Mainframe, and Fabric Manager Server Packages provide added intelligence and value

Note: The MDS 9134 Multilayer Switch supports N-Port identifier virtualization (NPV) to reduce the number of Fibre Channel domain IDs in SANs.



Figure 13-5 Cisco MDS 9134 for IBM System Storage

More information

For additional product details, see the website:

<http://www.ibm.com/systems/storage/san/ctype/9134>

13.4 Enterprise SAN switches

The IBM System Storage SAN switches and directors are well suited to address enterprise SAN customer requirements for infrastructure simplification and improved Business Continuity.

The SAN Director is designed to be interoperable with other members of the IBM System Storage switch family. You can configure a wide range of highly scalable solutions that address demands for integrated, heterogeneous mainframe and open server enterprise SANs.

Enterprise products include:

- ▶ IBM System Storage SAN768B
- ▶ IBM TotalStorage SAN384B
- ▶ IBM TotalStorage SAN256B
- ▶ Cisco MDS 9513 for IBM System Storage
- ▶ Cisco MDS 9506 for IBM System Storage
- ▶ Cisco MDS 9509 for IBM System Storage

In this section, we introduce IBM System Storage Switch solutions ideal for enterprise applications. Advanced security with comprehensive, policy-based security capabilities can improve availability and simplify operation.

13.4.1 IBM TotalStorage SAN256B

The IBM TotalStorage SAN256B is designed to provide outstanding performance, enhanced scalability and a design ready for high-performance 4 Gbps, 8 Gbps, and 10 Gbps capable hardware and expanded capability features. The SAN256B is well suited to address enterprise SAN customer requirements for infrastructure simplification and improved business continuity.

The SAN256B director interoperates with other members of the IBM TotalStorage SAN b-type and m-type families. It can be configured with a wide range of highly scalable solutions that address demands for integrated IBM System z and open system server enterprise SANs.

Below is a summary of the hardware features of the IBM TotalStorage SAN256B:

- ▶ 8, 4, 2, and 1 Gbps server and device interfaces with auto-sensing capability
- ▶ Brocade Fabric OS is common across all members of the IBM TotalStorage SAN b-type family
- ▶ Supports advanced Inter-Switch Link (ISL) Trunking, Load Balancing and Advanced Zoning
- ▶ Web Tools, Fabric Watch, Hot Code Activation and Performance Monitor
- ▶ Optional Adaptive Networking Services, Extended Fabric Activation, Remote Switch Activation, FICON Accelerator, FICON with CUP Activation, Advanced Security Activation, FCIP Activation
- ▶ Available blades include:
 - 8 or 4 Gbps 16-port, 32-port and 48-port switch blades
 - 16-port FC/2-port IP routing blades
 - 8-port FC plus 8-port iSCSI blade
 - 6-port 10 Gbps FC ISL blade
- ▶ 8 Gbps and 4 Gbps shortwave and longwave Small Form-factor Pluggable (SFP) optical transceivers

- ▶ Available 10 Gbps XFPs support high-performance ISLs and distance extension requirements
- ▶ Copper GbE SFPs are available for connection to Ethernet fabrics
- ▶ Full Fabric Operation and Universal Port (E, F and FL port) operation
- ▶ Many non-disruptive software upgrades and hot-swappable switch blades, power supplies and fans
- ▶ Fabric Shortest Path First (FSPF) designed to reroute around failed links
- ▶ Chassis includes two control processor blades plus space for one to eight port blades, dual power supplies and fans in a 14U rack height and can be installed in IBM TotalStorage SAN Cabinet C36.

Highlights

Here we list highlights of the IBM TotalStorage SAN256B (Figure 13-6):

- ▶ High availability with built-in redundancy designed to avoid single points of failure
- ▶ Highly scalable director with 16, 32 or 48 Fibre Channel (FC) ports per port switch blade and from 16 to 384 ports in a single domain
- ▶ Multiprotocol router blade with sixteen FC ports and two Internet Protocol (IP) ports for SAN routing and distance extension over IP
- ▶ iSCSI blade enables servers to access storage over IP (Ethernet)
- ▶ 10 Gbps Fibre Channel blade provides extended distance ISL connectivity between directors over dark fibre or DWDM
- ▶ Fibre Channel switch blades support either 4, 2, and 1 Gbps link speeds or 8, 4 and 2 Gbps link speeds
- ▶ Sixteen and thirty-two port switch blades support IBM FICON
- ▶ Interoperable with other IBM TotalStorage SAN b-type and m-type switches and directors
- ▶ Offers advanced security with comprehensive policy-based security capabilities
- ▶ Offers advanced fabric services such as end-to-end performance monitoring and fabric-wide health monitoring



Figure 13-6 IBM TotalStorage SAN256B

More information

For additional product details, see the website:

<http://www.ibm.com/systems/storage/san/b-type/san256b>

13.4.2 IBM System Storage SAN384B

The IBM System Storage SAN384B fabric backbone is designed to be the premier platform for consolidation of your data center connectivity, providing high-performance and highly available data networking. Providing new levels of performance with industry-leading 8 Gbps Fibre Channel (FC) technology, it is also one of the first members of the IBM System Storage b-type SAN family designed to exploit Brocade's new Data Center Fabric architecture.

The SAN384B interoperates with other members of the IBM System Storage b-type SAN family as well as other fabrics:

- ▶ It can be configured with a wide range of connectivity options, including 10, 8, 4, 2, and 1 gigabit per second (Gbps) Fibre Channel, up to 4 Gbps Fibre Connectivity (FICON), and Fibre Channel over Internet Protocol (FCIP) over 1 gigabit per second Ethernet (GbE).
- ▶ It is also designed to enable support for emerging high-performance and high-function network protocols, including Fibre Channel over Converged Enhanced Ethernet (FCoCEE).
- ▶ The SAN384B is designed to serve as the basis for transforming existing networks into a unified, high-performance data center fabric, connecting applications with their data and virtual servers with virtual storage.

As a member of the IBM System Storage family of b-type SAN products, the SAN384B is designed to participate in fabrics containing other b-type SAN devices manufactured by Brocade. This versatile hardware can serve as a new top tier (or backbone) in a complex fabric and provide connections to other b-type SAN directors, switches and routers.

Here is a summary of the hardware features of the IBM System Storage SAN384B:

- ▶ Chassis includes two control processor blades and two core module blades plus space for up to four port blades, dual power supplies, and two fan modules in an 8U rack height (plus 1U exhaust shelf).
- ▶ Available blades include
 - 8 Gbps 16-port, 32-port and 48-port switch blades
 - 24-port FCoCEE blade
 - 6-port 10 Gbps Inter Switch Link (ISL) blades
 - FC routing blade with sixteen 4 Gbps FC ports and two Gigabit Ethernet (GbE) ports
 - 8 Gbps Extension blade containing twelve 8, 4, 2, and 1 Gbps FC ports, ten Ethernet ports supporting 1 GbE and up to two optional 10 GbE ports
- ▶ Available optical transceivers include:
 - 10 Gbps, 8 Gbps and 4 Gbps shortwave small-form-factor pluggable (SFPs)
 - 10 Gbps and 8 Gbps 10 km and 4 Gbps 4 km and 10 km longwave SFPs
 - 8 Gbps 25 km and 4 Gbps 30 km extended distance longwave SFPs
- ▶ Full Fabric and Universal Port operation (E_Port, F_Port and FL_Port)
- ▶ Can be installed in IBM TotalStorage SAN Cabinet (2109-C36)

Highlights

Here we list highlights of the IBM System Storage SAN384B (Figure 13-7):

- ▶ Drive new levels of performance with 8 Gbps Fibre Channel (FC) technology in a compact design
- ▶ Reduce total cost of ownership (TCO) through consolidation of network resources
- ▶ Protect existing infrastructure investment while positioning for future technologies Manage your infrastructure with greater flexibility and scalability
- ▶ Unify management framework for consolidated and virtualized resources
- ▶ Improve energy efficiency by combining higher band-width with reduced power consumption



Figure 13-7 IBM System Storage SAN384B

More information

For additional product details, go to:

<http://www.ibm.com/systems/storage/san/b-type/san384b>

13.4.3 IBM System Storage SAN768B

The IBM System Storage SAN768B fabric backbone is designed to be the premier platform for consolidation of your data center connectivity, providing high-performance and highly available data networking. Providing new levels of performance with industry-leading 8 Gbps Fibre Channel (FC) technology, it is also the first member of the IBM System Storage b-type family designed to exploit Brocade's new Data Center Fabric architecture.

The SAN768B interoperates with other members of the IBM System Storage b-type and m-type families as well as other fabrics.

- ▶ It can be configured with a wide range of connectivity options, including 10, 8, 4, 2, and 1 gigabits per second (Gbps) Fibre Channel, up to 4 Gbps Fibre Connections (FICON), and Fibre Channel over Internet Protocol (FCIP) over 1 gigabit per second Ethernet (GbE).
- ▶ It is also designed to enable support for emerging high-performance and high-function network protocols, including Fibre Channel over Converged Enhanced Ethernet (FCoCEE).
- ▶ The SAN768B is designed to serve as the basis for transforming existing networks into a unified, high-performance data center fabric, connecting applications with their data and virtual servers with virtual storage.

As a member of the IBM System Storage family of b-type products, the SAN768B is designed to participate in fabrics containing other b-type and m-type devices manufactured by Brocade. This versatile hardware can serve as a new top tier (or backbone) in a complex fabric and provide connections to other b-type and m-type directors, switches and routers

Here is a summary of the hardware features of the IBM System Storage SAN768B:

- ▶ Chassis includes two control processor blades and two core module blades plus space for up to eight port blades, dual power supplies with an option for two additional power supplies, and three fan modules in a 14U rack height
- ▶ Available Blades include:
 - 8 Gbps 16-port, 32-port, and 48-port switch blades
 - 24-port FCoCEE blade
 - 10 Gbps Inter Switch Link (ISL) blades
 - FC routing blade with sixteen 4 Gbps FC ports and two Gigabit Ethernet (GbE) IP ports
 - 8 Gbps Extension blade containing twelve 8, 4, 2, and 1 Gbps FC ports, ten Ethernet ports supporting 1 GbE and up to two optional 10 GbE ports
- ▶ Available optical transceivers include:
 - 10 Gbps, 8 Gbps, and 4 Gbps shortwave small-form-factor pluggable (SFPs)
 - 10 Gbps and 8 Gbps 10 Km and 4 Gbps 4 Km and 10 Km longwave SFPs
 - 10 Gbps XFPs, and GbE optical SFPs for IP connectivity
- ▶ Full Fabric and Universal Port operation (E_Port, F_Port and FL_Port)
- ▶ Can be installed in IBM TotalStorage SAN Cabinet (2109-C36)

Highlights

Here we list highlights of the IBM System Storage SAN768B (Figure 13-8):

- ▶ Drive new levels of performance with 8 Gbps Fibre Channel (FC) technology
- ▶ Reduce total cost of ownership (TCO) through consolidation of network resources
- ▶ Protect existing infrastructure investment while positioning for future technologies
- ▶ Manage your infrastructure with greater flexibility and scalability
- ▶ Unify management framework for consolidated and virtualized resources
- ▶ Improve energy efficiency by combining higher band-width with reduced power consumption



Figure 13-8 IBM System Storage SAN768B

More information

For additional product details, go to:

<http://www.ibm.com/systems/storage/san/b-type/san768b>

13.4.4 Cisco MDS 9506 for IBM System Storage

The Cisco MDS 9506 for IBM System Storage supports 1, 2, 4, 8, and 10 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance and manageability required to consolidate geographically dispersed storage devices into a large enterprise SAN.

The Cisco MDS 9506 for IBM System Storage utilizes two Supervisor-2 Modules designed for high availability and performance. The Supervisor-2 Module combines an intelligent control module and a high-performance crossbar switch fabric in a single unit. It uses Fabric Shortest Path First (FSPF) multipath routing, which provides intelligence to load balance across a maximum of 16 equal-cost paths and to dynamically reroute traffic if a switch fails.

Each Supervisor-2 Module provides the necessary crossbar bandwidth to deliver full system performance in the Cisco MDS 9506 director with up to four Fibre Channel switching modules. It is designed to eliminate the impact on system performance of the loss or removal of a single crossbar module.

Highlights

Here we list highlights of the Cisco MDS 9506 for IBM System Storage (Figure 13-9):

- ▶ Provides Fibre Channel throughput of up to 4 Gbps per port and up to 64 Gbps with each PortChannel Inter-Switch Link connection
- ▶ Offers scalability from 12 to 192 Fibre Channel ports
- ▶ Offers 10 Gbps ISL ports for inter-Data Center links over metro optical networks
- ▶ Offers Gigabit Ethernet IP, GbE ports for iSCSI or FCIP connectivity over global networks
- ▶ High-availability design with support for non-disruptive firmware upgrades
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Enterprise, SAN Extension over IP, Mainframe and Storage Services Enabler and Fabric Manager Server Packages provide added intelligence and value



Figure 13-9 Cisco MDS 9506 for IBM System Storage

More information

For additional product details, go to:

<http://www.ibm.com/systems/storage/san/ctype/9506>

13.4.5 Cisco MDS 9509 for IBM System Storage

The Cisco MDS 9509 for IBM System Storage provides 1, 2, 4, 8, and 10 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance and manageability required to consolidate geographically dispersed storage devices into a large enterprise SAN.

The Cisco MDS 9509 for IBM System Storage utilizes two Supervisor-2 Modules to support high availability and performance. The Supervisor-2 Module combines an intelligent control module and a high-performance crossbar switch fabric in a single unit. It uses Fabric Shortest Path First (FSPF) multipath routing, which supports load balancing across a maximum of 16 equal-cost paths that dynamically reroute traffic if a switch fails

Highlights

Here we list highlights of the Cisco MDS 9509 for IBM System Storage (Figure 13-10):

- ▶ Provides Fibre Channel throughput of up to 8 gigabits per second (Gbps) per port and up to 64 Gbps with each PortChannel Inter-Switch Link connection
- ▶ Offers scalability from 12 to 336 1, 2, 4 and 8 Gbps Fibre Channel ports
- ▶ Offers 10-Gbps ISL ports for inter-data center links over metro optical networks
- ▶ Offers Gigabit Ethernet IP (GbE) ports for iSCSI or FCIP connectivity over global networks
- ▶ High-availability design with support for non-disruptive firmware upgrades
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Enterprise, SAN Extension over IP, Mainframe and Storage Services Enabler and Fabric Manager Server packages provide added function and value



Figure 13-10 Cisco MDS 9509 for IBM System Storage

More information

For additional product details, go to:

<http://www.ibm.com/servers/storage/san/ctype/9509>

13.4.6 Cisco MDS 9513 for IBM System Storage

The Cisco MDS 9513 for IBM System Storage supports 1, 2, 4, 8, and 10 Gbps Fibre Channel switch connectivity and intelligent network services to help improve the security, performance and manageability required to consolidate dispersed SAN islands into a large-enterprise SAN.

The Cisco MDS 9513 for IBM System Storage utilizes two Supervisor-2 modules to support high availability. The Supervisor-2 Module provides industry-leading scalability, intelligent SAN services, non-disruptive software upgrades, stateful process restart, and failover and redundant operation. Dual crossbar switching fabric modules provide a total internal switching bandwidth of 2.4 terabytes per second (Tbps) for interconnection of up to 11 Fibre Channel switching modules.

Fibre Channel switching modules improve performance, flexibility and density. The Cisco MDS 9513 for IBM System Storage requires a minimum of one Fibre Channel switching module and allows a maximum of 11. These modules are available in 12-, 24- or 48-port 4 and 8 Gbps configurations, enabling the Cisco MDS 9513 to support 12 to 528 Fibre Channel ports per chassis. Optionally, a 4-port 10-Gbps Fibre Channel module is available for high-performance Inter-Switch Link (ISL) connections over metro optical networks.

Highlights

Here we list highlights of the Cisco MDS 9513 for IBM System Storage (Figure 13-11):

- ▶ Provides Fibre Channel throughput of up to 8 gigabits per second (Gbps) per port and up to 64 Gbps with each PortChannel Inter-Switch Link connection
- ▶ Offers scalability from 12 to 528 1, 2, 4 and 8 Gbps Fibre Channel ports
- ▶ Offers 10-Gbps ISL ports for inter-data center links over metro optical networks
- ▶ Offers Gigabit Ethernet (GbE) IP ports for iSCSI or FCIP connectivity over global networks
- ▶ Features high-availability design with support for non-disruptive firmware upgrades
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN “islands” on a single physical fabric
- ▶ Provides added function and value through Enterprise, SAN Extension over IP, Mainframe, Storage Services Enabler and Fabric Manager Server packages



Figure 13-11 Cisco MDS 9513 for IBM System Storage

More information

For additional product details, go to:

<http://www.ibm.com/systems/storage/san/ctype/9513>

13.5 Multiprotocol routers

Since the introduction of Storage Area Networks, customers have built multiple SAN networks (or islands) for various applications, often with fabric switch components from various manufacturers. Certain islands were built by various departments within a company, while other islands resulted from mergers, acquisitions, or reorganizations. Dissimilar SAN equipment with various capabilities or a desire to isolate important applications has constrained opportunities for enhanced infrastructure simplification and vital business continuity solutions.

The IBM System Storage multiprotocol router provides Fibre Channel FC-FC Routing Service, which allows the interconnection of multiple SAN islands without requiring that the separate fabrics be merged into a single large SAN. This capability can help create a tiered or extended enterprise SAN infrastructure without having to redesign or reconfigure the entire environment.

Figure 13-12 shows a conceptual view of SAN multiprotocol routing, where SANs from various vendors can be connected.

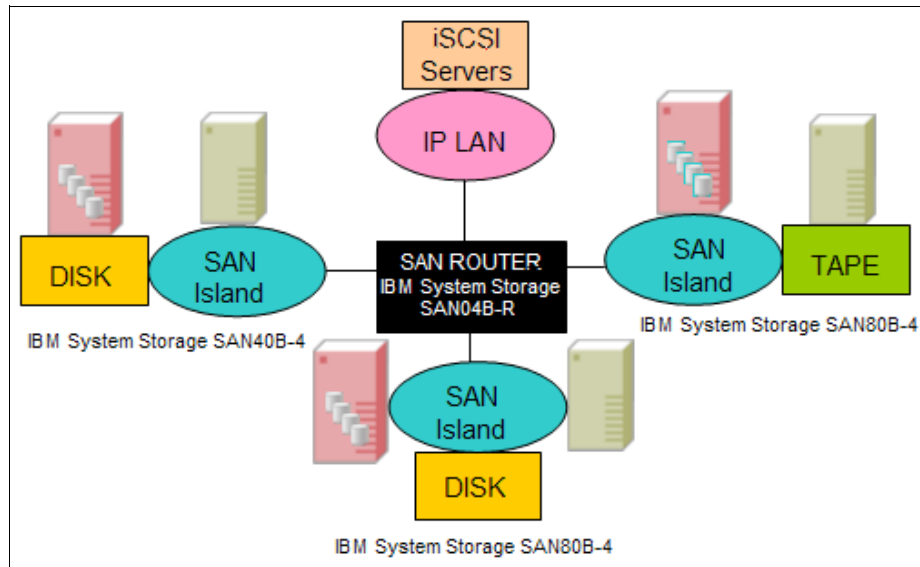


Figure 13-12 SAN multiprotocol routing

Enterprise products include:

- ▶ IBM System Storage SAN04B-R multiprotocol router
- ▶ IBM System Storage SAN06B-R multiprotocol router
- ▶ Cisco MDS 9222i for IBM System Storage

In this section, we introduce IBM System Storage SAN Multiprotocol Router solutions ideal for enterprises with a business need to route SAN traffic between various SAN islands.

13.5.1 IBM System Storage SAN04B-R multiprotocol router

A wide range of IBM System Storage midrange and enterprise Storage Area Network (SAN) infrastructure simplification and business continuity solutions can be created with the IBM System Storage SAN04B-R multiprotocol router. Infrastructure simplification solutions for the IBM Power Systems and System x families include disaster tolerance over metropolitan and global IP networks with IBM System Storage disk arrays, tape libraries, and IBM Tivoli Storage Manager data protection software. Separate SAN islands can also be consolidated using Fibre Channel routing. Support for System z servers is provided by the optional High-Performance Extension and FICON CUP Activation features.

Local site infrastructure simplification solutions can be extended to one or more remote sites for enhanced data protection and disaster tolerance. The IBM System Storage SAN04B-R multiprotocol router provides Fibre Channel over IP and FCIP Tunneling Service for distance extension, which can enable cost-effective and manageable metro and global business continuity solutions. This extended distance connectivity can help create consolidated remote tape vaulting data protection plus metro mirror and global mirror disk-based, disaster-tolerant solutions.

Since the introduction of Storage Area Networks, customers have built multiple SAN networks (or islands) for various applications, often with fabric switch components from various manufacturers. Certain islands were built by various departments within a company, while other islands resulted from mergers, acquisitions or reorganizations. Dissimilar SAN equipment with various capabilities or a desire to isolate important applications has constrained opportunities for enhanced infrastructure simplification and vital business continuity solutions.

The IBM System Storage SAN04B-R multiprotocol router can provide Fibre Channel FC-FC Routing Service, which allows the interconnection of multiple SAN islands without requiring that the separate fabrics be merged into a single large SAN. This capability can help create a tiered or extended enterprise SAN infrastructure without having to redesign or reconfigure the entire environment.

Highlights

Here we list highlights of the IBM System Storage SAN04B-R multiprotocol router (Figure 13-13):

- ▶ Designed for high performance with 4 Gigabit per second (Gbps) Fibre Channel (FC) ports and hardware-assisted traffic processing for line-rate performance across Ethernet Internet Protocol (IP) ports
- ▶ Utilizes existing Internet, IP-based Metropolitan Area Network (MAN) or Wide Area Network (WAN) infrastructures for metro and global SAN extension for business continuity solutions
- ▶ Enables consolidation of Storage Area Network (SAN) islands for infrastructure simplification without compromising security
- ▶ Hardware-based compression, large window sizes and selective acknowledgement of IP packets designed to optimize performance of SAN extension over IP networks
- ▶ Eight virtual FCIP tunnels per IP port are enabled by the High-Performance Extension feature to help maximize scalability and utilization of MAN/WAN resources
- ▶ Integrated IBM System Storage SAN b-type (Brocade) switch management helps simplify installation and administration and helps provide fabric investment protection
- ▶ FICON Accelerator uses special emulation techniques to reduce or eliminate degradation for selected applications such as IBM z/OS Global Mirror (XRC) and tape pipelining

Figure 13-13 shows IBM System Storage SAN04B-R multiprotocol router.



Figure 13-13 IBM System Storage SAN04B-R multiprotocol router

IBM System Storage SAN04B-R provides Fibre Channel FC-FC Routing Service, which allows the interconnection of multiple SAN islands without requiring that the separate fabrics be merged into a single large SAN. As shown in Figure 13-14, this capability can help create a tiered or extended enterprise SAN infrastructure without having to redesign or reconfigure the entire environment.

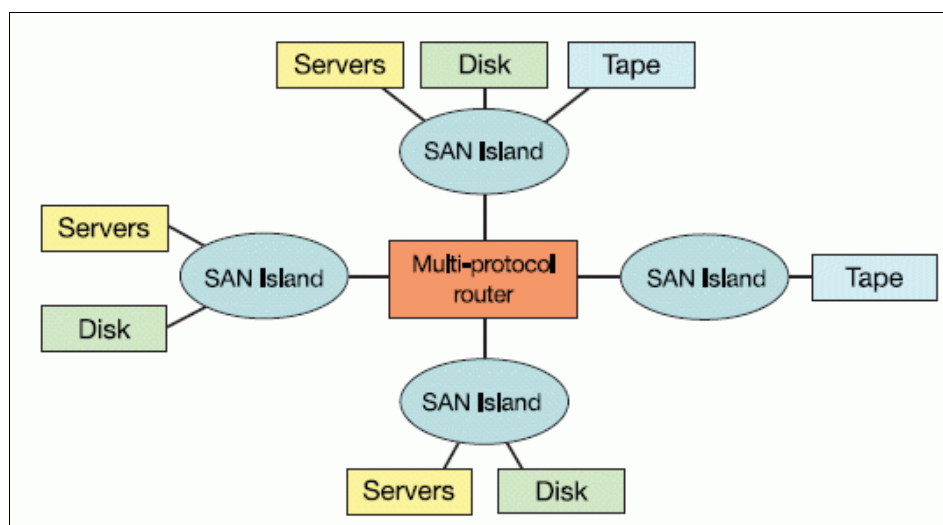


Figure 13-14 Inter-SAN routing

FC-FC Routing Service

FC-FC Routing Service feature is designed to allow devices located on separate SAN fabrics to communicate without merging the fabrics into a single large SAN environment. This routed network consists of multiple individual SAN fabrics that form one storage network connectivity model, known as a MetaSAN.

FCIP Tunneling Service

FCIP Tunneling Service feature is designed to allow organizations to extend Fibre Channel SANs over longer distances, using IP-based MAN/WAN infrastructure. This service can be integrated with the FC-FC Routing Service.

iSCSI Gateway Service

iSCSI Gateway Service is designed to allow the integration of low-cost Ethernet-connected servers into a Fibre Channel SAN by using iSCSI protocol, enabling IP-based servers to utilize Fibre Channel SAN resources. It enhances asset utilization and can help enable new applications such as centralized backup. This level of integration reduces the expense of connecting low-cost servers directly into centrally managed SAN storage.

More information

For additional product details, see the website:

<http://www.ibm.com/systems/storage/san/b-type/san04b-r>

13.5.2 IBM System Storage SAN06B-R multiprotocol router

In addition to the features of the SAN04B-R router, SAN06B-R supports 8, 4, 2 or 4, 2, and 1 Gbps FC link speeds.

Highlights

Here we list highlights of the IBM System Storage SAN06B-R multiprotocol router (Figure 13-15):

- ▶ Designed for high performance with 8 Gigabit per second (Gbps) Fibre Channel (FC) ports and hardware-assisted traffic processing for line-rate performance across 1 GbE Ethernet Internet Protocol (IP) ports
- ▶ Utilizes existing Internet, IP-based Metropolitan Area Network (MAN) or Wide Area Network (WAN) infrastructures for metro and global SAN extension for business continuity solutions
- ▶ Hardware-based compression, large window sizes and selective acknowledgement of IP packets designed to optimize performance of SAN extension over IP networks
- ▶ Up to eight virtual FCIP tunnels are available to help maximize scalability and utilization of MAN/WAN resources
- ▶ Enables consolidation of Storage Area Network (SAN) islands for infrastructure simplification without compromising security
- ▶ Integrated IBM System Storage SAN b-type (Brocade) switch management helps simplify installation and administration and helps provide fabric investment protection

Figure 13-15 shows IBM System Storage SAN06B-R multiprotocol router.



Figure 13-15 IBM System Storage SAN06B-R multiprotocol router

More information

For additional product details, go to:

<http://www.ibm.com/systems/storage/san/b-type/san06b-r>

13.5.3 Cisco MDS 9222i for IBM System Storage

The Cisco MDS 9222i for IBM System Storage is designed to address the needs of medium-sized businesses and large enterprises with a wide range of Storage Area Network (SAN) capabilities. It can be used as a cost-effective high performance SAN extension over IP router switch for midrange SMB customers in support of IT simplification and business continuity solutions. It can also provide remote site device aggregation and SAN extension connectivity to large customer data center directors.

A wide range of IBM System Storage midrange and enterprise Storage Area Network (SAN) IT simplification and business continuity solutions can be created with the Cisco MDS 9222i for IBM System Storage multiservice modular switch. IT simplification solutions for IBM Power Systems, System i, System p, System x and System z servers include storage consolidation and high-availability server clustering with IBM System Storage disk storage arrays. Business continuity capabilities can help businesses protect valuable data with IBM System Storage tape libraries and IBM Tivoli Storage Manager data protection software. Advanced connectivity capabilities can help business protect against major disasters with IBM System Storage disk metro and global mirroring disaster recovery solutions.

Highlights

Here we list highlights of the Cisco MDS 9222i for IBM System Storage (Figure 13-16):

- ▶ Multiservice design for high performance business continuity solutions with Windows, UNIX, Linux, NetWare, IBM OS/400 and IBM z/OS servers
- ▶ Cost-effective “green” switch design requires up to 27 percent less power per port
- ▶ Modular design provides scalability and configuration flexibility
- ▶ Excellent availability with redundant, hot swappable components and non-disruptive firmware upgrades
- ▶ Includes Virtual SAN (VSAN) capability for SAN consolidation into virtual SAN islands on a single physical fabric
- ▶ Enterprise, Mainframe and Fabric Manager Server Packages provide added intelligence and value



Figure 13-16 Cisco MDS 9222i for IBM System Storage

More information

For additional product details, go to:

<http://www.ibm.com/systems/storage/san/ctype/9222>

13.6 SAN solutions

IBM System Storage SAN solutions integrate IBM and other vendor server, storage, SAN switches and software components into solution templates which have been extensively tested to provide high availability, scalability, security, and simplicity of management. IBM SAN solutions are offered with worldwide IBM service and end-to-end solution support. These solutions templates might be customized by IBM Business Partners or IBM Global Services to address individual customer requirements.

IBM System Storage Infrastructure Simplification solutions and Business Continuity solutions are offered for Small and Medium Business (SMB) customers who require simplicity and affordability and for large enterprise customers who demand highest availability, scalability, security and management intelligence.

- ▶ Infrastructure simplification solutions consist of storage consolidation with shared disk storage and storage pooling with IBM SAN Volume Controller software.
- ▶ Business Continuity solutions consist of data protection with shared local or remote tape library technology and disk storage with point-in-time copy technology and automated backup and recovery software; and disaster tolerance with disk storage mirroring technologies and SAN fabric extension technologies which can provide metro and global distance solutions.

The following examples show solution possibilities using IBM System Storage SAN switches.

For additional IBM SAN solution information, go to:

<http://www.ibm.com/servers/storage/solutions/san/>

13.6.1 Infrastructure simplification solutions

IBM offers solutions to simplify IT infrastructures, a key first step in freeing the resources you need to begin developing an on demand operational model.

Complexity can prevent even the best organization from acting nimbly to meet ever-changing market and client demands. Complexity inherent in typical IT infrastructures can stall your business goals by making execution on new tasks too difficult or time consuming.

IBM solutions for infrastructure simplification are designed to help you improve efficiency, lower total cost of ownership, and reduce time-consuming and costly errors. These solutions can help you lower the cost of storing and managing data. IBM System Storage solutions for infrastructure simplification are designed to improve the effectiveness of your storage infrastructure and reduce long-term costs, while reducing the time and effort spent on managing data.

An *entry-level Infrastructure simplification solution* consists of up to twenty two servers attached to one 24-port IBM System Storage SAN24B-4 Express switch with an IBM System Storage disk array.

An example *Infrastructure simplification with high-availability server solution* is shown on Figure 13-17. The topology is created with redundant switches, and each server, with dual Fibre Channel adapters is cross-connected by the separate switches to the IBM System Storage DS5000 disk array.

An easy way to migrate your SAN from *entry-level* to *midrange* products is to shut down one 24-port SAN24B-4 fabric switch and replace it with one 40 port SAN40B-4 fabric switch.

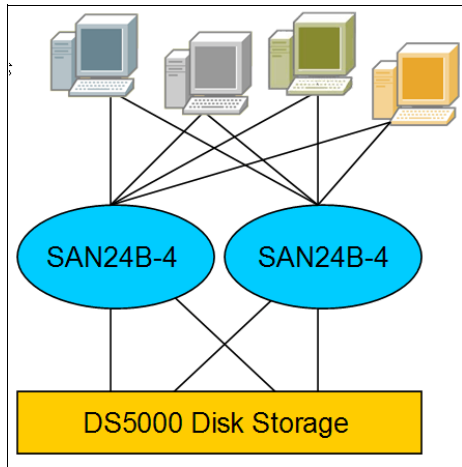


Figure 13-17 Infrastructure simplification with high-availability server solution

An expanded solution provides a higher data transfer rate by using 8 Gbps throughout - servers with 8 Gbps HBA card, for example: IBM BladeCenter, the IBM System Storage SAN fabric switches with 1, 2, 4, and 8 Gbps auto-sensing capabilities, and the IBM System Storage DS5000 Express with 8 Gbps host connection function, as shown in Figure 13-18.

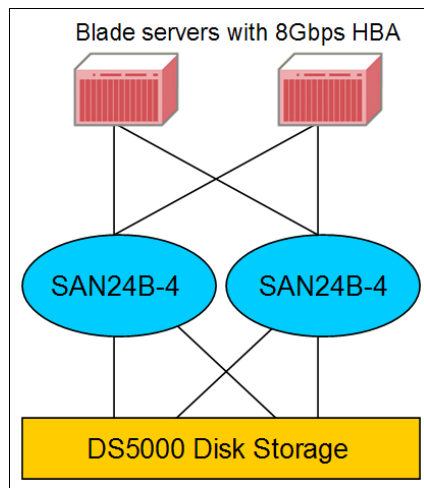


Figure 13-18 Infrastructure simplification with 8Gbps solution

Local site infrastructure simplification solutions can be extended to one or many remote sites for enhanced data protection and disaster tolerance. The IBM System Storage multiprotocol routers provide Fibre Channel over IP and FCIP Tunneling Service for distance extension that can enable cost-effective and manageable metro and global business continuity solutions. This extended distance connectivity can help create consolidated remote tape vaulting data protection plus metro mirror and global mirror disk-based disaster tolerant solutions.

13.6.2 Business continuity solutions

Today's customers, employees, suppliers and business partners expect to be able to tap into your information any time of day, from any location. At the same time, your business must be increasingly sensitive to issues of customer privacy, data security and regulatory requirements.

To keep your operations running, and your business competitive, you need a comprehensive strategy that addresses three primary aspects of business continuity: high availability, disaster recovery, and continuous operations.

IBM System Storage SAN Family solutions can help you protect critical business assets, and align recovery costs based on patented business impact and information value. A typical IBM System Storage SAN switch *Business Continuity solutions* example of backup site solution is shown in Figure 13-19.

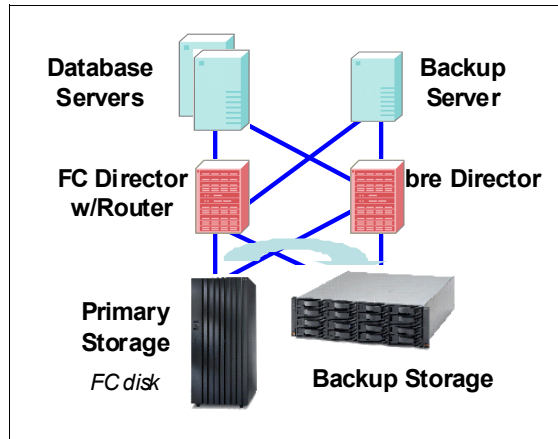


Figure 13-19 Typical backup site solution

When considering total cost of ownership, there are two choices: FCIP or FC direct route by DWDM or CWDM. See Figure 13-20 for FC over IP and Figure 13-21 for FC route through DWDM or CWDM.

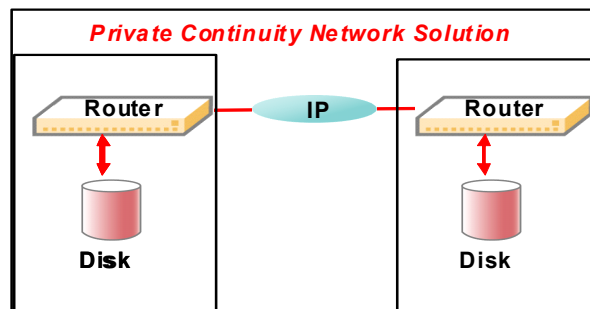


Figure 13-20 FC over IP

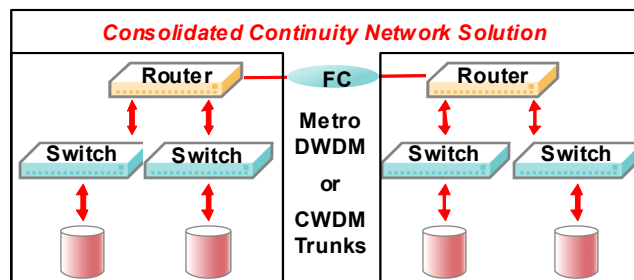


Figure 13-21 FC over DWDM or CWDM

Another IBM System Storage SAN switches *Business Continuity solutions* example is remote tape backup with the IBM SAN Multiprotocol router: As Figure 13-22 shows, an enterprise with a remote backup site can build a tape backup solution over distance with an IBM Multiprotocol router. An offsite Tivoli Storage Manager, by SAN routing, can pull data from a production disk storage array then copy it to the tape drives. This allows a minimal number of tape drives to be used and also utilizes electronic production of offsite media.

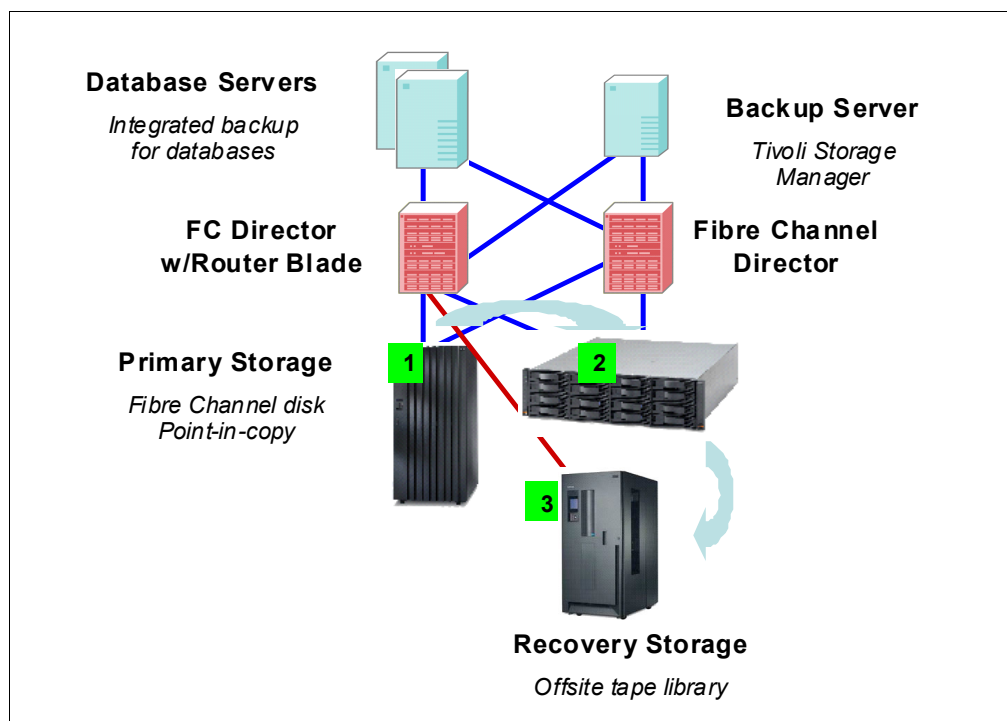


Figure 13-22 Remote tape backup solution

13.7 More information

For more information about all the SAN related topics and the products discussed in this chapter, see the following Redbooks publications:

- ▶ *Introduction to Storage Area Networks*, SG24-5470
- ▶ *IBM/Cisco Multiprotocol Routing: An Introduction and Implementation*, SG24-7543
- ▶ *Implementing an IBM/Cisco SAN*, SG24-7545
- ▶ *IBM SAN Survival Guide*, SG24-6143
- ▶ *SAN Storage Performance Management Using TotalStorage Productivity Center*, SG24-7364
- ▶ *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116



Data Center Networking

As you consolidate and virtualize your IT environment, capitalizing on the latest networking technologies and approaches can help you save money, improve service, and reduce the risk of security breaches. IBM Networking Strategy and Optimization Services, a network infrastructure optimization for consolidation and virtualization, provides services to plan and justify a networking infrastructure that supports and fully contributes to a dynamic, flexible IT environment.

Expanded network technology choices with the new IBM Ethernet switches and routers provide an integrated end-to-end resiliency and security framework. The big product range from IBM makes data center networking a cost efficient solution. The product range includes:

- ▶ IBM b-type Ethernet switches and routers
- ▶ IBM j-type Ethernet switches and routers
- ▶ Cisco Nexus

For more information about the Network Technologies, Network Management, and Networking Services, see the website:

<http://www.ibm.com/systems/networking>

14.1 IBM b-type Ethernet switches and routers

IBM b series switches encompass a wide range of variety, including Ethernet switches, converged switch, routers, and so on. Here are the model types available in the b-type family:

- ▶ IBM Converged Switch B32
- ▶ IBM Ethernet Switch B24X
- ▶ IBM r-series Ethernet Switches
- ▶ IBM c-series Ethernet Switches
- ▶ IBM g-series Ethernet Switches
- ▶ IBM m-series Ethernet/IP Routers
- ▶ IBM Ethernet Switch B08S and IBM Ethernet Switch B16S

14.1.1 IBM Converged Switch B32

As organizations evolve towards a dynamic infrastructure, they need new ways to reduce the complexity of their environments. To address this challenge, IBM offers a versatile switch that supports both Fibre Channel (FC) and Ethernet LAN data in one fabric using FCoE to help data centers simplify their growing infrastructures. The IBM Converged Switch B32 (Figure 14-1) is designed to provide outstanding performance with a reliable platform that helps reduce cabling complexity, equipment acquisition costs, and operational costs associated with space, power consumption and cooling. This multiprotocol top-of-rack switch features excellent space efficiency and low power consumption, leading the way toward a “greener” data center.



Figure 14-1 IBM Converged Switch B32

Key features: Here are the key features of the IBM Converged Switch B32:

- ▶ Designed for outstanding performance with eight Fibre Channel (FC) ports concurrently active at 8 Gbps and twenty-four CEE ports concurrently active at 10 Gbps link speeds
- ▶ High density design with 32 ports in a 1 U enclosure
- ▶ Exceptional “green” energy efficiency significantly reduces power consumption while generating less heat
- ▶ Enterprise-class availability features such as hot-swappable, redundant, and integrated fan and power supply assemblies
- ▶ Streamlines management by utilizing IBM System Storage Data Center Fabric Manager (DCFM) and extensions for FCoE and CCE
- ▶ CEE ports are capable of transporting both FC storage and Ethernet LAN traffic, eliminating the need for separate server SAN and LAN adapters and cables
- ▶ Advanced ASIC technology provides FC port trunking and Ethernet link aggregation
- ▶ Common Brocade Fabric OS supports the entire IBM System Storage SAN b-type family helps protect SAN infrastructure investments with backward compatibility

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/ethernet/b-type/b32>

14.1.2 IBM Ethernet Switch B24X

The IBM Ethernet Switch B24X (Figure 14-2) is a compact, high-performance, highly-available, and high-density 10 Gigabit Ethernet (GbE) solution designed for mission critical data centers, large enterprises, and High-Performance Computer (HPC) requirements. With an ultra-low latency, cut-through, non-blocking architecture, full wire-speed throughput, this switch provides a cost-effective solution for server or compute-node connectivity.



Figure 14-2 IBM Ethernet Switch B24X

Key features: Here are the key features of the IBM Ethernet Switch B24X:

- ▶ 24x ports of dual-speed 10/1 GbE SFP+ ports plus 4x 10/100/1000 MbE RJ45 ports
- ▶ Capable of line speed, full-duplex throughput on all ports: 488 Gbps total switch capacity and ultra-low latency (1.5 micro-seconds)
- ▶ 1 + 1 redundant, hot-swappable and load-sharing 300 W AC Power Supplies for high availability
- ▶ Hot-swappable, resilient triple-fan assembly with automatic speed control for efficient front-to-back cooling: 2+1 redundant fans in assembly
- ▶ Variety of SFP+ transceivers available including 1 GbE and 10 GbE Short Reach and Long Reach and available low-power consuming (1.0 watts) SFP+ optics up to 300 meters
- ▶ Supports low-cost, low-latency (0.25 micro-seconds), low-energy-consuming (0.1) direct attached SFP+ copper (Twinax) cabling up to 5 meters
- ▶ A 1U, high-density top-of-rack data center switch for 10 GbE server access and aggregation with 24x 10 GbE/1 GbE dual-speed (SFP+) ports plus 4x 10/100/1000 MbE (RJ45) ports
- ▶ Flexibility to mix 10 GbE and 1 GbE servers, protecting investments and streamlining migration to 10 GbE-capable server farms

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/ethernet/b-type/x-series>

14.1.3 IBM r-series Ethernet Switches

IBM r-series Ethernet Switches (Figure 14-3) are available in four chassis models and allow network designers to standardize on a single product family for end-of-row, aggregation, and backbone switching. In addition to their data center and enterprise role, the switches, with their high-density and compact design, are an ideal solution for High-Performance Computing (HPC) environments and Internet Exchanges and Internet Service Providers (IXPs and ISPs) where non-blocking, high-density Ethernet switches are needed.

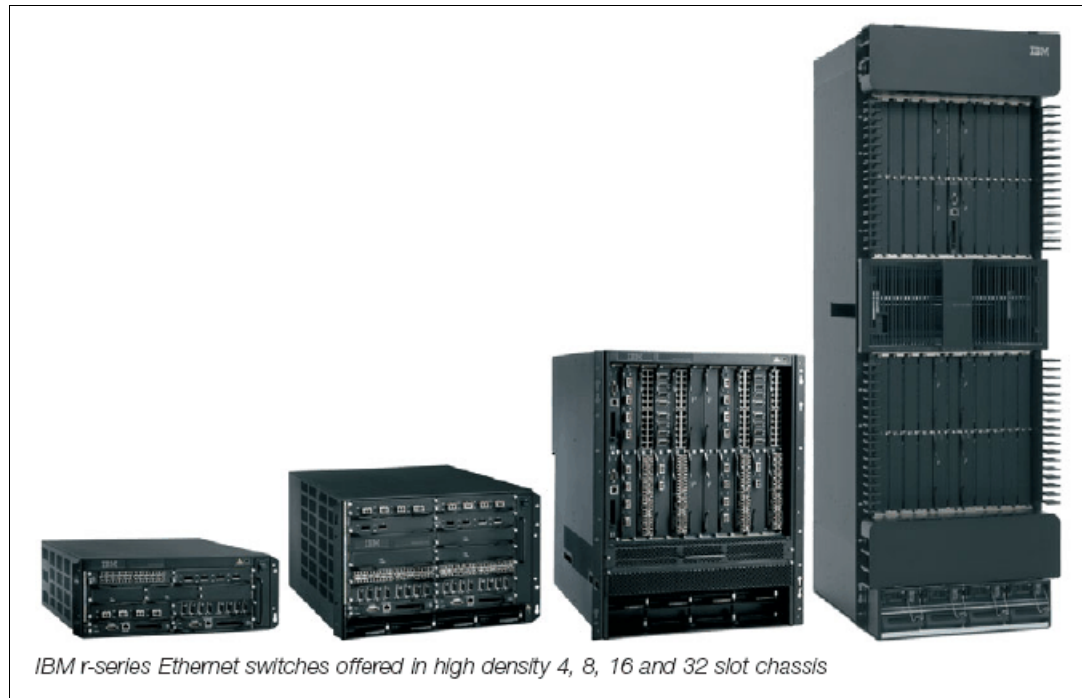


Figure 14-3 IBM r Series Ethernet Switches

Key features: Here are the key features of the IBM r-series Ethernet Switches:

- ▶ 4, 8, 16 and 32-slot high-capacity modular switches for end-of-row, aggregation, and core switching in data centers, large enterprises, HPC, IXP, and ISP networks
- ▶ Powerful suite of unicast and multicast IPv4 and IPv6 protocol support
- ▶ Interchangeable half-height line modules reduce sparing costs, TCO, and provide cost-effective modular growth
- ▶ Highly-dense chassis design supports up to 512 10 GbE or 1,536 wire-speed 1 GbE ports in a single 32-slot chassis
- ▶ Highly-available design features redundant and hot-pluggable hardware, hitless Layer 2 software upgrades, and graceful BGP and OSPF restart
- ▶ Advanced non-blocking Clos fabric features adaptive self-routing with graceful system degradation in the event of two or more module failures
- ▶ End to End QoS supported with hardware based honoring and marking and congestion management
- ▶ Scalable hardware-based IP routing to 512,000 IPv4 routes per line module.
- ▶ Highly available chassis design supports hot-swappable 1:1 management module redundancy, N+1 switch fabric module redundancy, M+N power module redundancy, and N+1 fan redundancy

- ▶ Interface modules include 16-port 10 GbE (SFP+), 2-port and 4-port 10 GbE (XFP), 24-port 1 GbE (RJ45 and SFP), and 48-port 1 GbE (MRJ21) modules

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/ethernet/b-type/r-series>

14.1.4 IBM c-series Ethernet Switches

IBM c-series Ethernet Switches (Figure 14-4) are powerful enablers of advanced converged enterprise backbones. Featuring state-of-the-art Quality of Service and wire-speed unicast/multicast routing, they enable the efficient rollout of converged backbones, providing reliable transport of Voice over IP (VoIP), video services, and mission critical data. The platforms low latency, high-touch processing and deep buffering makes it an ideal fit for Data Center top-of-the-rack server access switch.

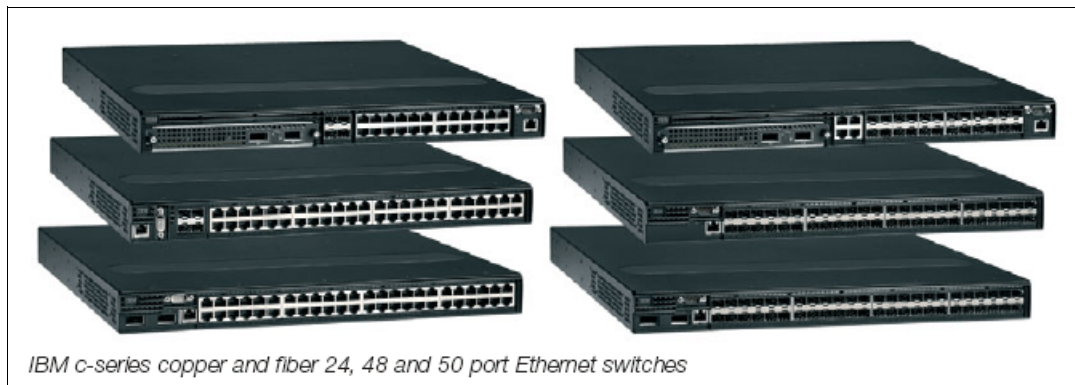


Figure 14-4 IBM C Series Ethernet Switches

Key features: Here are the key features of the IBM c-series Ethernet Switches:

- ▶ Flexible model configurations include 24x ports of 1 GbE or 48x ports of 1 GbE with SFP and RJ-45 options and 2x 10 GbE XFP uplinks available
- ▶ The systems offers wire-speed, non-blocking performance in all configurations
- ▶ Hot-swappable, resilient, six-fan, fan tray with sensors to automatically regulate speed
- ▶ Comprehensive hardware-based security and policies including hardware-based Layer 3 and Layer 2 ACLs (both inbound and outbound) with logging
- ▶ 64 MB of deep buffering for each 24-port 1 GbE group or 2-port 10 GbE group (64 MB to 192 MB total) excels at handling transient bursts in traffic.
- ▶ Advanced Carrier-Grade Ethernet services include platform capable of MPLS (VPLS, VLL), Multi-VRF, MPLS VPNs, along with Provider Bridges, Provider Backbone Bridges
- ▶ Comprehensive IPv4 unicast and multicast routing support based on the rich feature set of Brocade Multi-Service IronWare
- ▶ Full Layer 2 switching capabilities facilitate network resiliency include STP/RSTP/MSTP, PVST/PVST+ compatibility, Metro Ring Protocol (MRP), IEEE 802.3ad Link Aggregation (LACP), Virtual Switch Redundancy Protocol (VSRP), and jumbo frames.

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/ethernet/b-type/c-series>

14.1.5 IBM g-series Ethernet Switches

IBM g-series Ethernet Switches (Figure 14-5) provide enterprise organizations with a flexible and feature-rich solution for building a secure and converged network edge. Upgradeable with 10-Gigabit Ethernet, PoE, and IronStack stacking technology, the switches provide enterprises with the cost and operational benefits of a “pay-as-you-grow” architecture.



Figure 14-5 IBM g series Ethernet Switch

Key features: Here are the key features of the IBM g-series Ethernet Switches:

- ▶ High density, 48x 1 GbE Class 3 (15.4 watts) Power over Ethernet (PoE) capable ports with redundant removable load-sharing power supplies
- ▶ 4x 100/1000 Mbps Ethernet combination SFP ports available for connectivity using SFP transceivers for extended distances (total 48x 1 GbE ports active on a system concurrently)
- ▶ Flexible, pay-as-you-grow architecture with modular 2-port 10 GbE option (B48G)
- ▶ IBM Ethernet Switch B48G: 48x 10/100/1000 Mbps Ethernet RJ45 ports including 4x 100/1000 Mbps Ethernet SFP combination ports and optional 2-port 10 GbE module
- ▶ IBM Ethernet Switch B50G: 48x 10/100/1000 Mbps Ethernet RJ45 ports including 4x 100/1000 Mbps Ethernet SFP combination ports and 2x 10 GbE CX4 stacking ports for connectivity to another B50G
- ▶ Flexible and resilient design offering redundant, removable, load-sharing power supplies, field-upgradeable 10 GbE module (B48G) and IronStack stacking capabilities (B50G)
- ▶ Advanced Layer 2 unicast and multicast support including IGMP and PIM snooping to improve bandwidth utilization
- ▶ Flexible options to upgrade the software to Edge Layer 3, adding support for IP routing protocols such as RIPv1/v2 and OSPFv2
- ▶ Highly available system design featuring N+1 redundant, hot-swappable power supplies

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/ethernet/b-type/g-series>

14.1.6 IBM m-series Ethernet/IP routers

High performance multi-service IP/MPLS routers (Figure 14-6) help virtualize data center core routing and are available in four high density chassis. Featuring state-of-the-art QoS and wire-speed unicast/multicast routing for IPv4, IPv6, and MPLS services, they enable the efficient rollout of converged backbones, providing reliable transport of Voice over IP (VoIP), video services, and mission-critical data.

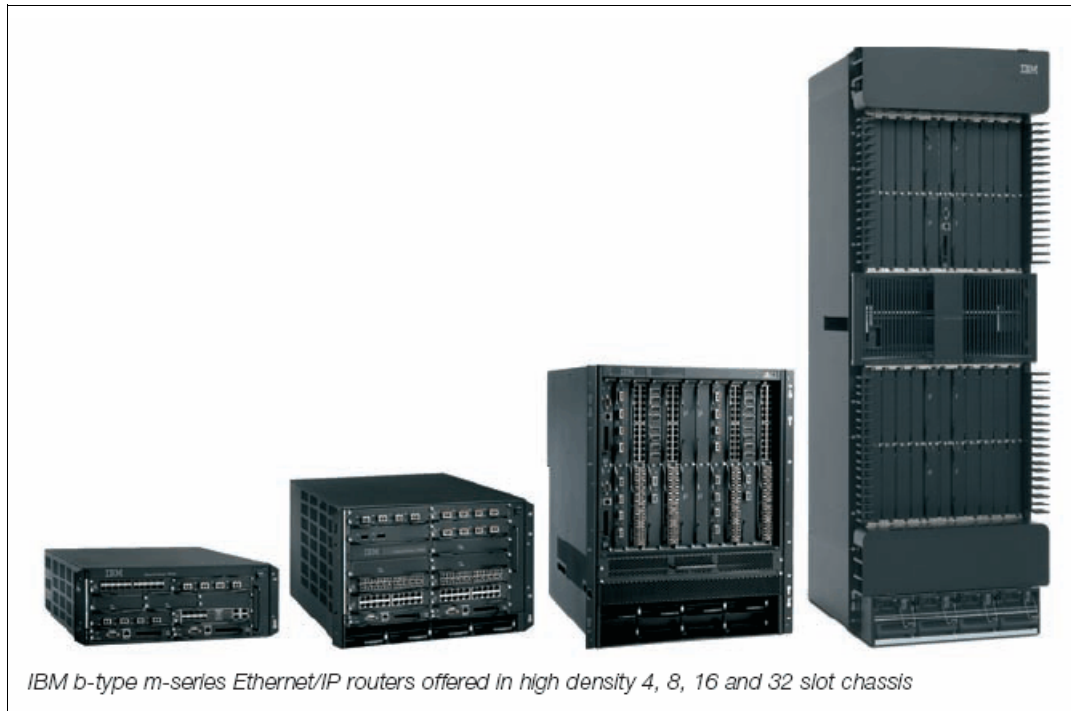


Figure 14-6 IBM m Series Ethernet/IP Routers

Key features: Here are the key features of the IBM m-series Ethernet/IP routers:

- ▶ 4, 8, 16, and 32-slot high-capacity modular routers ideal for a wide range of advanced applications in the high-security data center core, Internet edge/aggregation routing, large-enterprise backbone, high-performance computing (HPC), and Metropolitan Area Networks (MAN)
- ▶ Wire-speed IPv4, IPv6, and MPLS routing featuring full Forwarding Information Base (FIB) programming in hardware
- ▶ High-performance, wire-speed, non-oversubscribed port density featuring up to 128 10 GbE, 1,536 1 GbE, 256 OC-12/48, or 64 OC-192 ports in a single 32-slot chassis
- ▶ Fully distributed, non-blocking architecture with up to 3.2 Tbps data capacity (~2 billion pps) per system
- ▶ Exceptional 320 Gbps cross-module link aggregation supports up to 32 10 GbE or OC-192 links enabling high-bandwidth inter-switch trunking in the backbone.
- ▶ Scalable hardware-based IP routing to 512,000 IPv4 and 112,000 IPv6 routes in hardware, 2 million IPv4 BGP routes and up to 256 BGP peers H
- ▶ Highly available chassis design supports hot-swappable 1:1 management module redundancy, N+1 switch fabric module redundancy, M+N power module redundancy, and N+1 fan redundancy.
- ▶ Interface modules include 4-, 2-port 10 GbE (XFP), 48-port 1 GbE (MRJ21), 20-port 1 GbE (RJ-45 and SFP), 2-port OC-192 PoS/SDH (XFP), and 8-, 4-, 2-port OC-12/48 PoS/SDH (SFP) modules.

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/ethernet/b-type/m-series>

14.1.7 IBM s Series Ethernet Switch

IBM s-series Ethernet Switch modular PoE capable switches (Figure 14-7) provide an industry-leading scalable, low-latency, fault-tolerant infrastructure for cost-effective deployment of Voice over IP (VoIP), wireless, and high-capacity data services throughout the enterprise. Designed to extend control from the network edge to the backbone, the switches provide intelligent network services, including superior quality of service (QoS), predictable performance, advanced security, comprehensive management and integrated resiliency. A common operating system and shared interface and power supply modules between the Ethernet Switch B08S and B16S help reduce the cost of ownership by minimizing operational expenses and improving return on investment (ROI).



Figure 14-7 IBM s Series Ethernet Switches

Key features: Here are the key features of the IBM s Series Ethernet Switch:

- ▶ Advanced chassis-based convergence solution provides a scalable, secure, low-latency and fault-tolerant infrastructure for cost-effective deployment of Voice over IP (VoIP), wireless, and high-capacity data services throughout the enterprise
- ▶ 8- and 16-slot systems for maximum deployment versatility. Highly available 1+1 management, fabric, and N+1 power redundant architecture enables resilient and fault-tolerant network infrastructures
- ▶ The B16S can scale to 384 IEEE 802.3af Class 3 PoE 10/100/1000 Mbps ports, each capable of delivering 15.4 watts to provide customers with a convergence-ready infrastructure that will scale to support future growth
- ▶ Robust PoE auto-detection with IEEE 802.1AB LLDP and LLDP-MED enables support for PoE and non-PoE devices along with auto-configuration of VoIP endpoints, simplifying device deployment
- ▶ Resilient Advanced Layer 2 protocol support including Protected Link Groups, IEEE 802.3ad Link Aggregation (LACP), UDLD, Virtual Switch Redundancy Protocol (VSRP), Metro Ring Protocol (MRP), and STP/RSTP/MSTP to build a highly tolerant network infrastructure
- ▶ Advanced Layer 2 unicast and multicast support including IGMP and PIM snooping to improve bandwidth utilization, with flexibility to upgrade to Full Layer 3 and IGMPv1/v2/v3, IGMP Proxy, and PIM-SM/-SSM/-DM multicast routing optimizes network traffic

- ▶ Flexible options to upgrade the software to Full IPv4 Layer 3, including support for IP routing protocols such as RIPv1/v2, OSPFv2, BGP-4, and support for multicast routing, and Full IPv4+IPv6 Layer 3, adding support for RIPvng, OSPFv3, 6-to-4 static tunnels, and IPv6 ACLs
- ▶ Future-proof with available IPv6 management and interface modules that can integrate today with existing IPv4 switches within and across the network.

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/ethernet/b-type/s-series>

14.2 IBM j-type Ethernet switches and routers

IBM j-type Ethernet switches and routers are designed to deliver the performance, scalability and high availability required for today's high-density data center and cloud computing environments

The IBM j-type family includes following switches and router:

- ▶ IBM Ethernet Switch J08E/J16E
- ▶ IBM Ethernet Switch J48E
- ▶ IBM Ethernet Router J02M/J06M/J011M

14.2.1 IBM Ethernet Switch J08E and IBM Ethernet Switch J16E

The IBM Ethernet Switch J08E (Figure 14-8) and IBM Ethernet Switch J16E are designed to deliver the performance, scalability and high availability required for today's high-density data center and cloud computing environments. Supporting a variety of GbE and 10 GbE line cards and featuring a built-in migration path to 100 GbE deployments, the J08E and J16E provide the highly scalable solution required by today's high-performance data centers and are positioned to support even more demanding network environments in the future.

The high-density, high-performance J08E and J16E are also used for aggregating access switches deployed in data center top-of-rack or end-of-row applications, as well as for supporting Gigabit Ethernet server access in data center end-of-row deployments. The J08E delivers up to 960 million packets per second (Mpps) of high-density, wire-speed 10 GbE performance, while the J16E delivers approximately 1.9 billion packets per second (Bpps) of 10 GbE performance. Both systems are designed to provide sufficient capacity to support the most demanding data center networks.



Figure 14-8 IBM j-type Ethernet Switch

Key features: Here are the key features of the IBM Ethernet Switch J08E and IBM Ethernet Switch J16E:

- ▶ High-performance 8-slot (J08E) and 16-slot (J16E) switches support data center as well as campus LAN core and aggregation layer deployments
- ▶ Scalable switch fabric delivers up to 320 Gbps per slot
- ▶ 48-port 10/100/1000BASE-T and 100BASE-FX/1000BASE-X line cards support up to 384 (J08E) or 768 (J16E) GbE ports per chassis
- ▶ Eight-port 10GBASE-X line cards with SFP+ interfaces deliver up to 64 (J08E) or 128 (J16E) 10 GbE ports per chassis
- ▶ Carrier-class architecture includes redundant internal Routing Engines, switch fabrics, power and cooling, ensuring uninterrupted forwarding and maximum availability
- ▶ All IBM j-type switches and routers run the modular, fault-tolerant Juniper Networks JUNOS Software operating system.
- ▶ Fourteen (J08E) rack-unit (RU) and 21 (J16E) RU chassis Eight (J08E) and sixteen (J16E) dedicated I/O slots
- ▶ 6.2 (J08E) and 12.4 (J16E) Tbps backplane capacity
- ▶ 320 Gbps (full duplex) per slot fabric capacity
- ▶ Full 10 GbE line-rate forwarding (even under failure conditions)
- ▶ Dedicated data, control and management planes
- ▶ LCD panel for system monitoring Energy efficient: 200,000 packets per second per watt

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/j-type/j08e>

14.2.2 IBM Ethernet Switch J48E

Running Juniper Networks' JUNOS Software operating system, the IBM Ethernet Switch J48E (Figure 14-9) was designed for high-performance server access deployments. A single switch can be deployed initially; as requirements grow, Virtual Chassis technology allows up to nine additional switches to be interconnected over a 128 gigabit-per-second (Gbps) backplane and managed as a single device, with a single configuration file and OS image. Modular Gigabit Ethernet (GbE) and 10-Gigabit Ethernet (10 GbE) uplink module options enable Virtual Chassis technology to be extended to switches in separate racks or even in separate data centers.



Figure 14-9 IBM Ethernet Switch J48E

Key features: Here are the key features of the IBM Ethernet Switch J48E:

- ▶ Full Layer 2 and Layer 3 Ethernet switching Virtual Chassis technology enables up to 10 switches to be interconnected as a single logical device supporting up to 480 ports in data center top-of-rack or end-of-row applications
- ▶ Interconnected switches in a Virtual Chassis configuration share a single control plane and operating system, and automatically assign master (active) and backup (hot-standby) Routing Engines
- ▶ Integrated application-specific integrated circuit (ASIC)-based Packet Forwarding Engine and integrated Routing Engine (RE) deliver switching and routing functionality at wire rates
- ▶ Front-panel LCD display offers flexible interface for performing device bring-up and configuration rollbacks, reporting switch alarm and LED status, or restoring the switch to its default settings
- ▶ Eight ports of IEEE 802.3af-compatible Class 3 Power over Ethernet (PoE)
- ▶ Single rack unit-height device (1RU) 48 10/100/1000BASE-T front-panel ports
- ▶ Optional GbE or 10 GbE front-panel uplink modules for connecting to other switches or upstream devices
- ▶ Scalable to 480 server access ports with up to 20 10 GbE uplinks to the core
- ▶ Back-panel Virtual Chassis ports interconnect up to 10 switches over 128 Gbps backplane
- ▶ Redundant, internal hot-swappable power supplies
- ▶ Hot-swappable fan tray with redundant blowers
- ▶ Dual Routing Engines with graceful Routing Engine switchover (GRES)
- ▶ Single management interface.

Note: For more information about specification, model number, product number, and type, see: <http://www.ibm.com/systems/networking/hardware/j-type/j48e>

14.2.3 IBM Ethernet Router J02M/J06M/J11M

IBM Ethernet Routers J02M, J06M and J11M (Figure 14-10) deliver high-port density as well as performance of up to 960 Gbps throughput, scalability, and reliability in a space-efficient package. The routers offer fully redundant hardware that includes a redundant Switch Control Board (SCB) and Routing Engines (REs) plus fan trays and power supplies designed to increase system availability.



Figure 14-10 IBM Ethernet Router J02M/J06M/J11M

Key features: Here are the key features of the IBM Ethernet Router J02M/J06M/J11M:

- ▶ High performance Ethernet L2/L2VPN/L3/L3VPN router, designed to deliver advanced routing features, including network virtualization with MPLS, low-latency Multicast, and Quality of Service, without compromising performance
- ▶ Redundant hardware and high availability features designed to ensure that the network is always up and running with:
 - Redundant hardware: cooling, power supplies, Routing Engines, and Switch Control Boards
 - Modular operating system
 - Separate data and control planes
 - Graceful restart
 - Non-stop routing
 - MPLS FRR
 - VPLS multi-homing
- ▶ Excellent performance to ensure that services and customers stay connected:
 - Powered by custom-designed application-specific integrated circuits (ASICs)
 - Enhanced QoS capabilities
 - Additional packet processing flexibility
 - Scaling enhancements including route lookup, next hop, IFL scaling, and interface accounting
 - Additional control storage capacity to provide headroom for future JUNOS Software operating system features
 - Enhanced multicast

- ▶ Operational efficiencies and simplicity delivered through JUNOS Software operating system
- ▶ Each chassis scales in size with choices of 2, 6 or 11 slots that can be populated with line cards for access or network interfaces
- ▶ Flexible port densities and WAN interfaces (Dense Port-Count Line Cards)
- ▶ Simultaneous support for Layer 2 and Layer 3 Ethernet services
- ▶ Can be deployed in two-tier collapsed architectures, which can yield operational savings through reduced power, cooling and space requirements.
- ▶ 2, 6 and 11-slot systems with full duplex throughput per slot provide versatile port densities and WAN interfaces
- ▶ Each chassis scales in size and can be populated with line cards for access or network interfaces
- ▶ Ideal fit for high-density deployments including the data center, campus and WAN gateways
- ▶ Flexible Architecture with IPv6, IPv4, industry leading IP/MPLS and VPLS
- ▶ Powered by the I-chip ASIC, the Ethernet Router family includes features such as enhanced QoS capabilities and scaling enhancements and advanced packet processing performance. Each slot provides line-rate 40 1-Gbps packet forwarding.

Note: For more information about specification, model number, product number, and type, see the website:

<http://www.ibm.com/systems/networking/hardware/j-type/m-series>

14.3 Cisco Nexus 5000

In this section we discuss the Cisco Nexus 5000 systems in regard to reducing the complexity of an organization's environment.

14.3.1 Cisco Nexus 5000 overview

As organizations evolve towards a dynamic infrastructure, they need new ways to reduce the complexity of their environments. To address this challenge, IBM offers versatile Cisco switches that support Fibre Channel, Converged Enhanced Ethernet and Fibre Channel over Ethernet (FCoE) to help data centers simplify their growing infrastructures.

14.3.2 Cisco Nexus 5010/5020

The Cisco Nexus 5010 for IBM System Storage and Cisco Nexus 5020 for IBM System Storage FCoE switches help reduce costs through data center infrastructure simplification. A unified fabric over 10 Gigabit Ethernet (GbE) for server LAN and SAN traffic enables consolidation of server adapters, cables and top-of-rack (TOR) switches by up to 50 percent.

Key features: Here are the key features of the Cisco Nexus 5010/5020:

- ▶ Ports and power connections are at the rear, closer to server ports, helping keep cable lengths as short and efficient as possible
- ▶ IEEE Data Center Bridging features for lossless transmission, priority flow control and enhanced transmission selection
- ▶ Standards-based FCoE protocol preserves existing Fibre Channel network management models and tools, helping protect investments in software and staff training
- ▶ Unified Fabric designed to consolidate all data center I/O onto Layer 2 Ethernet and reduce capital and operating costs by reducing the number of server adapters, cables, and upstream switches needed
- ▶ Server I/O consolidation reduces energy consumption by eliminating the need for separate Fibre Channel adapters, cables, and switches
- ▶ Consistent management is provided through consistency of both Cisco NX-OS Software and Cisco MDS 9000 SAN-OS Software management models and tools
- ▶ Multiprotocol switches with 10 Gbps Fibre Channel over Ethernet (FCoE), 10 Gbps Converged Enhanced Ethernet, traditional 1, 10 Gbps Ethernet ports and optional 1, 2 and 4 Gbps Fibre Channel (FC) ports
- ▶ Cisco Nexus 5010 for IBM System Storage provides eight 1, 10 GbE and twelve 10 GbE fixed ports with one expansion module slot
- ▶ Cisco Nexus 5020 for IBM System Storage provides sixteen 1, 10 GbE and twenty-four 10 GbE fixed ports with two expansion module slots
- ▶ Expansion modules include eight 1, 2, 4 Gbps FC ports; four 10 GbE and four 1, 2, 4 Gbps FC ports; six 10 GbE ports
- ▶ 10 GbE ports are capable of transporting both storage and LAN traffic, eliminating the need for separate server SAN and LAN adapters and cables
- ▶ Dual speed 1, 10 GbE ports help provide a smooth transition to 10 Gigabit Ethernet

Note: For more information about specification, model number, product number, and type, see: <http://www.ibm.com/systems/networking/hardware/ethernet/c-type/nexus>



Part 4

IBM System Storage software

In Part 4, we discuss the IBM System Storage software products, which are designed with the goal of helping customers drive down the cost and complexity of storage management while providing greater flexibility to address rapidly changing storage needs.

The IBM Storage Software portfolio consists of:

- ▶ IBM System Storage Virtualization:
 - IBM System Storage SAN Volume Controller Entry Edition
 - IBM System Storage SAN Volume Controller
- ▶ IBM Tivoli Storage Manager:
 - IBM Tivoli Storage Manager
 - IBM Tivoli Storage Manager Extended Edition
 - IBM Tivoli Storage Manager for Mail
 - IBM Tivoli Storage Manager for Databases
 - IBM Tivoli Storage Manager for Microsoft SharePoint
 - IBM Tivoli Storage Manager HSM for Windows
 - IBM Tivoli Storage Manager for Advanced Copy Services
 - IBM Tivoli Storage Manager for Copy Services
 - IBM Tivoli Storage Manager for SAP V6.1
 - IBM Tivoli Storage Manager for Space Management
 - IBM Tivoli Storage Manager for Storage Area Networks
 - IBM Tivoli Storage Manager for System Backup and Recovery
 - IBM Tivoli Storage Manager FastBack
 - IBM Tivoli Storage Manager FastBack for Microsoft Exchange
 - IBM Tivoli Storage Manager FastBack for Bare Machine Recovery
 - IBM Tivoli Continuous Data Protection for Files

- ▶ IBM Tivoli Storage Productivity Center (TPC) Suite:
 - IBM Tivoli Storage Productivity Center Basic Edition
 - IBM Tivoli Storage Productivity Center for Data
 - IBM Tivoli Storage Productivity Center for Disk
 - IBM Tivoli Storage Productivity Center for Replication
 - IBM Tivoli Storage Productivity Center Standard Edition
 - IBM Tivoli Storage Productivity Center with Advanced Provisioning
 - IBM System Storage Productivity Center

Attention: With Tivoli Storage Productivity Center V4.1, Tivoli Storage Productivity Center and Tivoli Storage Productivity Center for Replication, which were previously separate products, are now integrated and you can start the IBM Tivoli Storage Productivity Center for Replication user interface from within the Tivoli Storage Productivity Center user interface.

IBM Tivoli Storage Productivity Center for Replication is still a separate, stand-alone product on z/OS.

Also with Tivoli Storage Productivity Center V4.1, there is no longer a separate Tivoli Storage Productivity Center for Fabric - only Disk and Data.

- ▶ IBM System Storage Multilevel Grid Access Manager Software (GAM)
- ▶ IBM General Parallel File System (GPFS)
- ▶ The IBM scale out file services (SOFS)



IBM System Storage Virtualization

IBM has more than thirty years experience in virtualization technologies, with thousands of virtualization systems installed worldwide.

In this chapter, we describe what storage virtualization is and the need for it in today's environment. We look at the IBM approach to both in-band and out-of-band storage virtualization. The fundamental differences between the two architectures are described and examples are given of IBM products in these categories. We then explain virtualization in relation to storage networking and discuss benefits of implementing an IBM System Storage virtualization solution.

15.1 What is storage virtualization?

Storage virtualization products have become widespread in the IT industry today. They can exist in one or more layers in a storage network and can be described as *the abstraction of the physical location of the data from the logical representation of it*.

Storage virtualization is a general approach to decoupling logical resources from their physical elements, so that those resources can be allocated faster, more cost-effectively, and more dynamically. It is also the pooling of physical storage resources, from one or more storage devices, which are then presented and managed through an abstraction layer in order to present a unified view of a “virtual” storage resource or pool on what appears to be a single device to users and attached hosts (Figure 15-1).

The abstraction can be made on several levels of the components of storage networks and is not limited to the disk subsystem. The virtualization layer provides the same kind of services to the layer above (as the hidden layer below provides). Storage virtualization software separates the representation of storage to the operating system (and its users) from the actual physical components.

Note: Storage virtualization existed, and was taken for granted in the mainframe computing environment for many years.

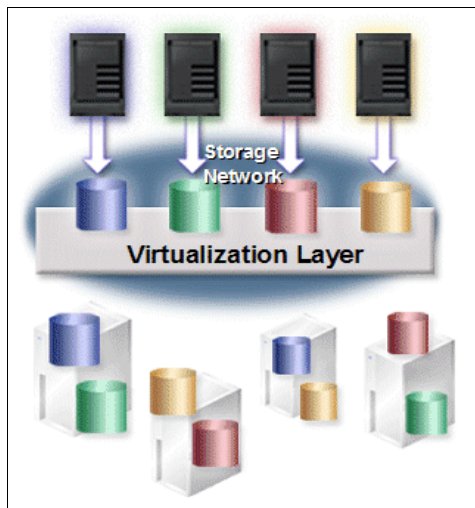


Figure 15-1 SAN with virtualization

15.2 The need for storage virtualization

Today, information has become the lifeline for business sustainability and competitive advantage. It shapes every idea and every decision. But as the world becomes more instrumented, interconnected, and intelligent, the ability to manage and deliver trusted information as a strategic asset becomes a critical enabler to making organizations smarter and increasingly more competitive.

15.2.1 Inefficiencies in sharing storage

The following problems arise with the increasing complexity of storage environments:

- ▶ Too many various storage platforms and server operating systems.
- ▶ Multiple management consoles
- ▶ IT staff skill levels
- ▶ Availability requirements
- ▶ Poor utilization levels in traditional environments
 - Typically 30%-50% for disk in a non-virtualized environment¹
 - Typically 20%-40% for tape in a non-virtualized environment¹

SAN is making it easier for customers to spread their IT systems out geographically, but even in networks, various types of servers that use various operating systems do not automatically get the full benefit of sharing storage. Instead, the storage is partitioned to each particular type of server, this creates complex management and inefficient use of storage. When storage is added, upgraded, moved or, removed applications can often be disrupted. At the same time, the reduced cost of storage and the technology of storage networks, with ever faster data transfer rates, have enabled customers to use increasingly sophisticated applications. This has caused even greater complexity and difficulty of management as the amount of storage required grows at unprecedented rates.

15.2.2 Benefits of implementing storage virtualization in your SAN

The case for using storage virtualization to reduce or even eliminate the problems listed previously becomes stronger as the problems themselves increase:

- ▶ Reduced storage management complexity
- ▶ Significantly reduced downtime, both planned and unplanned.
- ▶ Enhanced productivity through improved data availability
- ▶ Consolidation to a single pool of storage
- ▶ Heterogeneous storage is presented as homogenous storage
- ▶ Better choice of storage, switch, and SAN technologies
- ▶ Removal of "Tie-in" to a specific vendor or platform
- ▶ Improved storage resource utilization
- ▶ A single, cost-effective set of advanced copy services
- ▶ Simplification and standardization of the IT infrastructure
- ▶ Ease transition to an on demand IT infrastructure

These benefits lead to lower TCO and better flexibility to react to business changes.

¹ taken from an SNIA virtualization education document written in 2008.

15.3 Levels of storage virtualization

Virtualization can be achieved at various levels in a storage network (Figure 15-2).

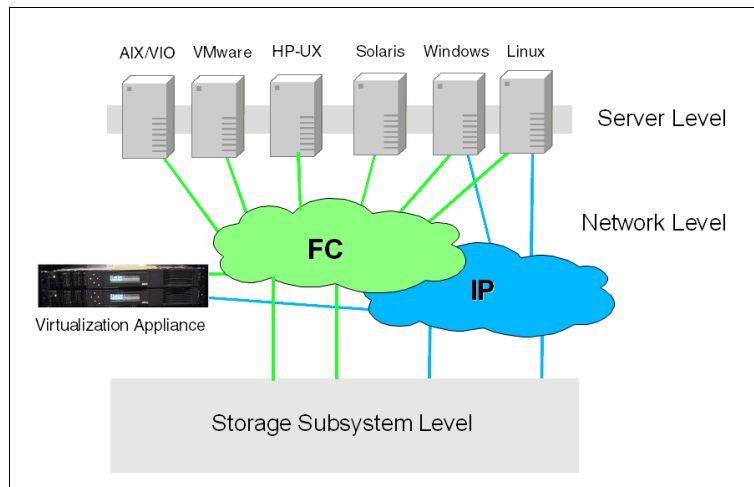


Figure 15-2 Levels of application of storage virtualization

15.3.1 Server level

Abstraction at the server level is achieved with logical volume management on the server operating systems. At first sight, increasing the level of abstraction on the server seems well suited for environments without storage networks, but this can be important in storage networks too.

15.3.2 Network level

At the network level, virtualization can enable the independence of storage pools from heterogeneous servers. The network (either FC SAN fabric or IP) is configured to allow the virtualization appliances to see the storage subsystems, and for the servers to see the virtualization appliances. Servers are not able to directly see, or operate on, the storage subsystems.

15.3.3 Storage subsystem level

Intelligent disk storage systems provide virtualization already by consolidating multiple physical storage devices into larger logical drive groups. They then add disk failure protection and subdivide these large, protected, groups into smaller virtual drives. These RAID storage controllers are commonplace and available for almost all requirements. Storage virtualization takes this to the next level by enabling the presentation, and management, of disparate storage systems by one virtualization appliance.

15.4 Server, network, or storage subsystem virtualization

As we stated, virtualization can be implemented at any of these levels. The IBM strategy is to move the storage device management intelligence out of the server, reducing the dependency of having to implement specialized software, like Logical Volume Managers (LVM), at the server level.

15.4.1 Implementing virtualization at a fabric or network level

By implementing virtualization at a fabric or network level, storage control is moved into the network layer, which gives the virtualization opportunity to all, and at the same time reduces complexity by providing a single view of storage. The storage network can be used to utilize all kinds of services across multiple storage devices, including virtualization.

15.4.2 Fabric level virtualization models

In-band and out-of-band are two implementations of virtualization at the fabric level. Models can be drawn for these methods of storage virtualization, as illustrated in Figure 15-3. These models are not mutually exclusive. In many environments a combination of both might be desired.

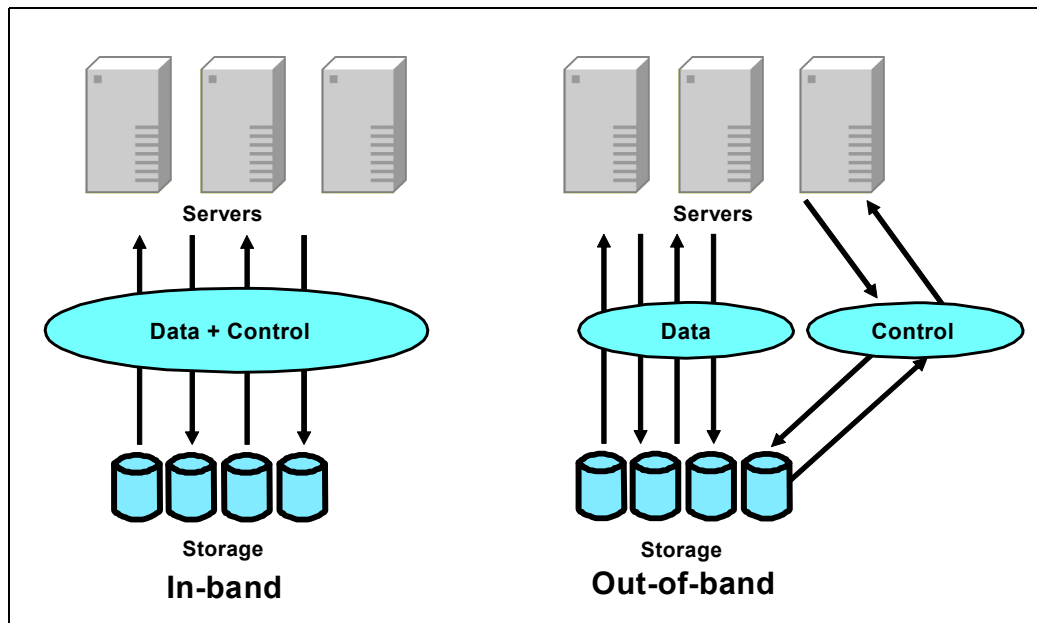


Figure 15-3 In-band and out-of-band virtualization models

In-band or block aggregation

When we implement an in-band virtual storage network, both data and control flow over the same path. Levels of abstraction exist in the data path, and storage can be pooled under the control of a domain manager. In general, in-band solutions are perceived to be simpler to implement, especially because they do not require special software to be installed in servers (other than conventional multipathing software). In-band solutions can also provide caching and advanced functions within the storage network. This can help to improve the performance of existing disk systems and can extend their useful life, and reduce the cost of new storage capacity by enabling the use of lower function, lower cost disk systems, without the loss of performance.

Other advantages include:

- ▶ Ability to off load function from the host
- ▶ Providing storage management for the SAN
- ▶ Performing performance optimizations in the data path
- ▶ Supporting host systems not in a cluster
- ▶ Supporting multiple heterogeneous hosts
- ▶ Releasing the customer from a particular vendor's storage

- ▶ Integrating with storage to create a better management picture
- ▶ Offering excellent scalability

The IBM System Storage SAN Volume Controller is an in-band virtualization appliance solution that maps virtualized volumes visible to hosts and applications to physical volumes on storage devices.

Out-of-band or file aggregation

In an out-of-band implementation, the data flow is separated from the control flow. This is achieved by separating the data and metadata (data about the data) into separate places. Out-of-band virtualization involves moving all mapping and locking tables to a separate server (the metadata controller) that contains the metadata of the files.

In an out-of-band solution the clients request authorization to data from the metadata controller, which grants it, handles locking, and so on. After they are authorized, clients access the data directly without any metadata controller intervention. After a client has obtained access to a file, all I/O will go directly over the SAN to the storage devices. For many operations, the metadata controller does not even intervene.

Separating the flow of control and data in this manner allows the I/O to use the full bandwidth that a SAN provides, while control can go over a separate network or routes in the SAN that are isolated for this purpose.

Other advantages include:

- ▶ Releasing the customer from a particular vendor's storage
- ▶ Providing storage management for the SAN
- ▶ Offering excellent scalability
- ▶ Off loading host processing
- ▶ Supporting storage management from multiple vendors
- ▶ Integrating well with storage management software
- ▶ Supporting multiple heterogeneous hosts
- ▶ Relatively low overhead in the data path

IBM GPFS is an example of an out of band virtualization system.

15.5 IBM System Storage approach to virtualization

The IBM approach to helping you with your storage needs is to address the entire problem. From a TCO perspective, the initial purchase price is becoming an increasingly small part of the equation, in the order of 10%. As the cost per megabyte of disk drives continues to decrease, the focus is shifting away from hardware toward software value add functions, storage management software, and services. This is not to downplay the importance of a highly reliable, high performance hardware solution such as the IBM System Storage DS5000 or DS8000. But software is emerging as the differentiating factor and offers a highly competitive advantage. The advantage is due to the fact that advanced functionality provided by software and storage management software play a vital role in administering distributed IT assets, maintaining high availability, and minimizing downtime.

IBM has identified three major product areas where significant improvements can be achieved. The SNIA Storage Model describes these areas as *Block Aggregation*, *File Aggregation*, and *Management and Productivity*, as depicted in Figure 15-4.

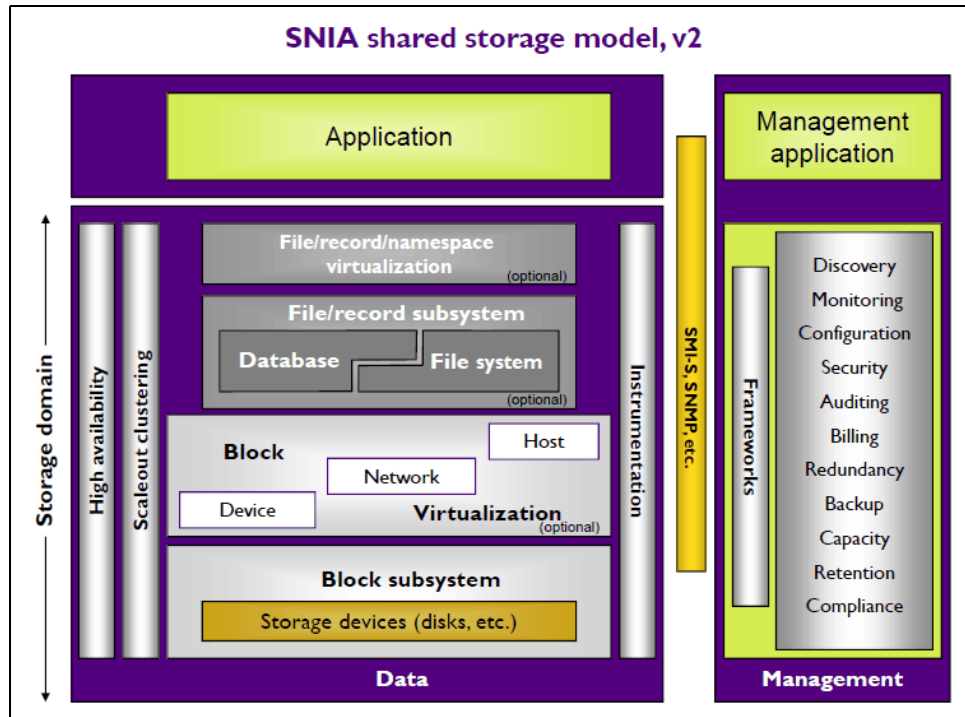


Figure 15-4 SNIA storage model v2

For more information about the SNIA Shared Storage Model, see the website:

http://www.snia.org/education/storage_networking_primer/shared_storage_model/

15.5.1 IBM block aggregation characteristics

Block aggregation in the SNIA model provides what is also referred to in the industry as block virtualization. Block level virtualization provides servers with a logical view of physical storage. The SNIA Block Aggregation Model, as seen in Figure 15-4, specifies that block aggregation can be performed within hosts (servers), in the storage network (storage routers, storage controllers) or in storage devices (intelligent disk arrays). The IBM System Storage SAN Volume Controller is a block aggregation product within the storage network.

15.5.2 IBM file aggregation characteristics

For file/record aggregation in the SNIA model, hosts get file metadata from file system or Network Attached Storage (NAS) controllers, then access the data directly. File aggregation can be used in conjunction with or independent from block aggregation.

15.6 IBM System Storage Virtualization family

IBM Storage virtualization products exist in all areas of the storage product portfolio. They are designed to help simplify storage infrastructure, optimize storage utilization, and enable business to adapt quickly and dynamically to variable environments.

IBM System Storage Virtualization family of products can help you with the following tasks:

- ▶ Simplify deployment and administration while increasing utilization of physical storage asset:
- ▶ Improve business application availability.
- ▶ Manage all of your storage assets with a single, comprehensive, management tool set.

15.6.1 IBM SAN Volume Controller

The IBM SAN Volume Controller (SVC) is the world's leading storage virtualization system with more than 4500 clusters and 15000 storage engines (nodes) shipped to date. It is designed to combine storage capacity from multiple disk systems into a reservoir of capacity that can be better managed as a business resource and not as separate storage islands. The abstraction of the physical location of the data from the logical representation of it that the application on a server sees allows better utilization, flexibility, efficiency, and performance.

For more information, see Chapter 16, "IBM System Storage SAN Volume Controller" on page 383 for a brief overview of what SVC is and what business benefits it can deliver.

15.6.2 IBM XIV Storage System

The IBM XIV Storage System is a next-generation high-end open disk storage system. An integral part of the IBM broad spectrum of system storage and SAN offerings, the XIV system is based on a grid of standard, off-the-shelf hardware components connected in any-to-any topology by means of massively paralleled, non-blocking Gigabit Ethernet. Its innovative architecture is designed to deliver the highest levels of reliability, performance, scalability and functionality at low overall cost, while providing unprecedented ease of management. Software virtualization features of the XIV system include snapshotting, thin provisioning, space reclamation, and remote mirroring.

For more information about the XIV, see 4.8, "IBM XIV Storage System" on page 137.

15.6.3 IBM System Storage DS8000 Turbo

The IBM System Storage DS8000 Turbo is a high-performance disk storage subsystem for business-critical enterprise workloads. Virtualization features offer support for advanced IBM deduplication technology, flexible system logical partitions (LPARs) and Metro/Global Mirror Incremental Resync.

For more information about the DS8000, see 4.4, "The DS8000 series" on page 101

15.6.4 The IBM Virtualization Engine

The IBM Virtualization Engine (TS7530 TS7740 and TS7720 Virtualization Engines) are a family of mainframe virtual-tape solutions that are designed to optimize tape processing. With one solution, the implementation of a fully integrated tiered storage hierarchy of disk and tape utilizes the benefits of both technologies to help enhance performance and provide the capacity needed for today's tape processing requirements. Deploying this innovative subsystem can help reduce batch processing time, total cost of ownership and management overhead.

For more information about the IBM Virtualization Engine, see 10.5, "IBM Virtualization Engine TS7700" on page 285.

15.6.5 IBM System Storage TS7650 ProtecTIER Appliance

The IBM System Storage TS7650 ProtecTIER Deduplication Appliance is a preconfigured solution of IBM storage, IBM server, and the IBM revolutionary ProtecTIER data deduplication software designed to improve backup and recovery operations. This is not just a bundle of components, but a truly integrated solution that makes it easy to harness the power of deduplication without making radical changes to the existing environment.

For more information about the TS7650, see 10.3, “TS7650G ProtecTIER Deduplication Gateway” on page 281.

15.6.6 IBM System Storage TS7650G ProtecTIER Deduplication Gateway

The IBM System Storage TS7650G ProtecTIER Deduplication Gateway is designed to meet the disk-based data protection needs of the enterprise data center while enabling significant infrastructure cost reductions. The solution offers industry leading inline deduplication performance and scalability up to 1 petabyte (PB) of physical storage capacity per system that can provide up to 25 PB of storage capacity*. Combined with IBM storage, the ProtecTIER Gateway solution provides a powerful disk-based repository to improve the retention and availability of backup and archive data.

For more information about the TS7650, see 10.3, “TS7650G ProtecTIER Deduplication Gateway” on page 281.

15.6.7 IBM Tivoli Storage Productivity Center

The IBM Tivoli Storage Productivity Center suite of storage infrastructure management tools can help customers reduce the complexity of managing their storage environments by centralizing, simplifying and automating storage tasks associated with storage systems, storage networks, replication services and capacity management. A brief summary of these products is in Table 15-1.

Table 15-1 Tivoli Storage Productivity Center product features.

Component	Use
IBM Tivoli Storage Productivity Center for Data	Manage the capacity utilization of storage systems, file systems and databases and to automate file-system capacity provisioning
IBM Tivoli Storage Productivity Center for Disk	Perform device configuration and management of multiple devices from a single user interface, tune and proactively manage the performance of storage devices on the Storage Area Network (SAN)
IBM Tivoli Storage Productivity Center for Replication	Control and monitor copy services operations (like FlashCopy, Metro Mirror and Global Mirror capabilities) and use data replication for data protection and disaster recovery

For more information about the IBM Tivoli Storage Productivity Center products, see Chapter 15, “IBM System Storage Virtualization” on page 371.

15.6.8 IBM System Storage N series

Wide ranging suites of virtualization functions and options are available within the N series product line. For more information, see Chapter 5, “IBM System Storage N series” on page 173.

Virtualization features and functions in the N series product line are listed in Table 15-2.

Table 15-2 N series virtualization features

Feature	Use
Snapshot	Almost instantaneous point in time, read only copies of data. Almost zero overhead during creation and space efficient copy using little extra space in most cases.
Virtual File Manager (VFM)	Provide out of band data management functionality for server and storage consolidation, migration, remote office data management and disaster recovery features; functionality accomplished through automated policy-based data management utilizing a global namespace
MultiStore®	Quickly and easily create separate, private logical partitions in filer network and storage resources.
SnapMirror	Mirror data to one or more network filers. Constant updates to the mirrored data keep it current and available for disaster recovery
SnapMover	A fast and simple solution for migrating data among IBM System Storage N series systems while avoiding impact on data availability and disruption to users.
FlexClone	Ability to generate nearly instantaneous real time replicas of data sets and storage volumes.

15.6.9 IBM General Parallel File System

The IBM General Parallel File System (GPFS) is more than just a file system; it is a file management infrastructure. It provides unmatched performance and reliability with scalable access to critical file data. GPFS distinguishes itself from other cluster file systems by providing concurrent high-speed file access to applications executing on multiple nodes of an AIX cluster, a Linux cluster, or a heterogeneous cluster of AIX and Linux nodes.

New with GPFS 3.2.1.5, you can add Windows client nodes to an AIX, Linux, or mixed cluster. In addition to providing file storage capabilities, GPFS provides storage management, information life cycle management tools, and centralized administration, and allows for shared access to file systems from remote GPFS clusters. Highlights of the comprehensive features and benefits list of the GPFS file system are listed in Table 15-3.

Table 15-3 GPFS features and benefits.

Feature	Benefit
Highly Scalable File System: Supports thousands of nodes and thousands of disks	GPFS enables the scale out expansion of the file system using a building block approach with commodity components as service demands and data volumes increase.
Data Protection: Both user data and metadata can be replicated (mirrored) in the GPFS installation.	Improves availability by ensuring reliable access to data and metadata.
Parallel File System: Divides individual files into multiple blocks and stores these blocks across multiple disks or disk arrays in parallel.	Higher performance and reliability by eliminating the bottlenecks that typically arise when an entire file resides on a single node.
Storage Pools: User-defined grouping of storage resources.	Storage can be grouped or tiered based on factors such as location, performance and reliability.
External Storage Pools: Policy driven interactions with external applications for near-line data storage.	External pools utilize the high performance policy engine to expand your data storage options. For example, an external pool data can be stored on a tape library.
Policy-Driven Automation	Simplifies data management and allows matching the cost of storage to the value/importance of the data. For example, critical data can be stored on the fastest storage.
Concurrent File Access and Block-Level Locking	Enables multiple users, programs or nodes to access a single file simultaneously – accelerates processing for parallel applications that share data and eliminates the need for multiple copies of data, reducing data storage and management costs.
High Availability	Automatically recovers from events that can normally interrupt data availability. Applications and users can continue without interruption in the event of a node, path or storage system failure.
Clustered NFS: Tools to deploy and manage a clustered NFS file serving solution.	Provides an integrated solution enabling scalable NFS file serving from a cluster of Linux servers.
Large Blocksize Options	Data blocks of 16K, to 4MB can be used to optimize throughput for various applications. Large blocks allow more data to be written or read in a single operation. The wide range of block sizes enables more efficient use of disk bandwidth and available storage space.

For more information about GPFS, see these links.

- ▶ GPFS product home page:
<http://www.ibm.com/systems/clusters/software/gpfs>
- ▶ GPFS 3.2 information flyer:
<ftp://ftp.software.ibm.com/common/ssi/pm/fy/n/clf03001usen/CLF03001USEN.PDF>
- ▶ Introduction to GPFS version 3.2.1 (current version at the time of writing):
<ftp://ftp.software.ibm.com/common/ssi/sa/wh/n/clw03005usen/CLW03005USEN.PDF>

These virtualization products are part of an IBM commitment to the open standards adopted by the Storage Networking Industry Association (SNIA). They implement standard common information model (CIM) based APIs to allow management applications from IBM, and other vendors to administer and monitor their activities.

Storage software solutions are a key element in the IBM overall storage strategy. This strategy addresses several of the most pressing needs currently facing Chief Information Officers (CIO) and IT managers. As part of its strategy, IBM intends to deliver industry-leading technologies that will help dramatically reduce total cost of ownership (TCO) for storage.

15.7 More information

For more information about the IBM System Storage Virtualization family, see the web page:

<http://www.ibm.com/systems/storage/virtualization>

For more information about SAN virtualization, see the Redpaper publication, *Virtualization in a SAN*, REDP-3633, which can be found at:

<http://www.redbooks.ibm.com/redpapers/pdfs/redp3633.pdf>

For more information about the SNIA Shared Storage Model, see the web page:

http://www.snia.org/education/storage_networking_primer/shared_storage_model/

For an Introduction to Virtualization tutorial from SNIA, see the web page:

http://www.snia.org/images/tutorial_docs/VirtualStnds/Peglar_Virtualization_I.pdf



IBM System Storage SAN Volume Controller

The IBM System Storage SAN Volume Controller (SAN Volume Controller) is an IBM storage virtualization solution that is designed to reduce the complexity and costs of managing storage networks. It provides storage virtualization and helps increase the utilization of existing capacity and centralize the management of multiple controllers in an open-system SAN environment.

The San Volume Controller supports attachment to IBM and non-IBM storage systems. It enables storage administrators to reallocate and scale storage capacity and make changes to underlying storage systems without disrupting applications.

In this chapter we provide an overview of the San Volume Controller (SVC) and discuss a few of its features and capabilities.

16.1 The need for storage virtualization

At the business level, clients are faced with three major storage challenges:

- **Managing storage growth:** Storage needs continue to grow at a rate that is normally higher than what had been planned for each year. As an example, storage subsystems might be purchased to last for 3 to 5 years; however, organizations are finding that they are filling to capacity much earlier than that.

In order to accommodate this growth, administrators are either extending their current storage subsystems in chunks, or buying various types of storage subsystems to match their storage needs and budget.

- **Increasing complexity:** As storage needs grow, this need might be filled by multiple disk subsystems, which might not even be from the same vendor.

Together with the variety of server platforms and operating systems in a customer's environment, customers might have SANs with multiple and diverse storage subsystems and host platforms, as shown in Figure 16-1.

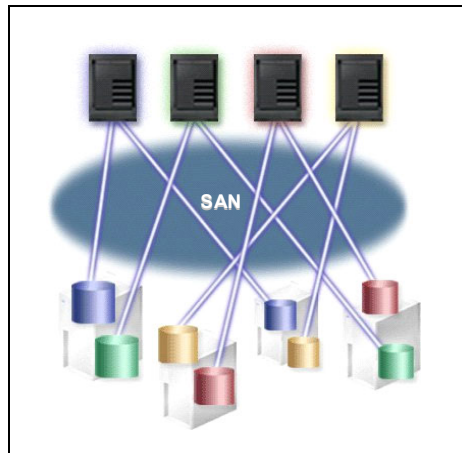


Figure 16-1 Today's environment

Combining this situation with the shortage of skilled storage administrators, the cost and risk of storage increases as the environment becomes more complex.

- **Maintaining availability:** With the increased range of storage options available, the storage growth rate, and no similar increase in storage budget, businesses have to manage more storage with no staff or minimal additional staff.

With the complexity highlighted before, and with requirements for higher business system availability, the potential for errors increases as each new storage subsystem is added to the infrastructure. Making changes to the storage infrastructure to accommodate storage growth runs the risk of outages that might not be acceptable to the enterprise.

Storage needs are increasing at an ever faster rate, and the challenge of managing disparate storage systems is growing. The IBM TotalStorage SAN Volume Controller brings storage devices together in a *virtual pool* to make all storage appear as follows:

- One “logical” device to centrally manage and to allocate capacity as needed
- One solution to help achieve the most effective use of key storage resources on demand

Virtualization solutions can be implemented in the storage network, in the server, or in the storage device itself. The IBM storage virtualization solution is SAN-based, or in-band, which helps allow for a more open virtualization implementation. Locating virtualization in the SAN, and therefore in the path of input/output (I/O) activity, helps to provide a solid basis for policy-based management. The focus of IBM on open standards means that its virtualization solution supports freedom of choice in storage-device vendor selection.

The IBM System Storage SAN Volume Controller solution helps administrators:

- ▶ Simplify storage management
- ▶ Reduce IT data storage complexity and costs while enhancing scalability
- ▶ Extend on demand flexibility and resiliency to the IT infrastructure
- ▶ Increase application availability by making changes in the infrastructure without having to shutdown hosts

16.2 IBM System Storage SAN Volume Controller

The IBM System Storage SAN Volume Controller (SVC), machine type 2145, provides the ability to simplify the storage infrastructure, utilize storage resources more efficiently, improve personnel productivity and increase application availability. The SAN Volume Controller pools storage volumes from IBM and non-IBM disk arrays into a single reservoir of capacity which can be managed from a central point. The SAN Volume Controller allows data to be migrated between disk arrays without disrupting the applications and moves copy services into the network where the same suite of copy services tools can be applied across the entire storage configuration irrespective of storage vendor restrictions that normally apply for the individual disk controllers in use.

16.2.1 Overview

Storage area networks (SAN) enable companies to share homogeneous storage resources across the enterprise. But for many companies, information resources are spread over a variety of locations and storage environments, often with products from various vendors. To achieve higher utilization of resources, companies need to be able to share their storage resources from all of their environments, regardless of vendor. While storage needs rise rapidly, and companies operate on lean budgets and staffing, the best solution is one that utilizes the investments already made and that provides growth when needed. The IBM System Storage SAN Volume Controller offers a solution that can help strengthen existing SANs by increasing storage capacity, efficiency, uptime, administrator productivity and functionality.

16.2.2 Design intent

The IBM System Storage SAN Volume Controller is designed to handle the following tasks:

- ▶ Combine storage capacity from multiple vendors into a single repository of capacity with a central management point
- ▶ Help increase storage utilization by providing host applications with more flexible access to capacity
- ▶ Help improve productivity of storage administrators by enabling management of combined storage volumes from a single interface
- ▶ Support improved application availability by insulating host applications from changes to the physical storage infrastructure

- Enable a tiered storage environment, in which the cost of storage can be better matched to the value of data
- Support advanced copy services, from higher- to lower-cost devices and across subsystems from multiple vendors

The SAN Volume Controller combines hardware and software into a comprehensive, modular appliance. Using IBM eServer System x server technology in highly reliable clustered pairs, the SAN Volume Controller is designed to avoid single points of failure. SAN Volume Controller software is a highly available cluster optimized for performance and ease of use.

Storage utilization

The SAN Volume Controller is designed to help increase the amount of storage capacity that is available to host applications. By pooling the capacity from multiple disk arrays within the SAN, it helps enable host applications to access capacity beyond their island of SAN storage. SNIA estimates are that open systems disk utilization in a non virtualized environment is only between 30%-50% currently.

Scalability

A base configuration SAN Volume Controller consists of a single I/O group. An I/O group is a pair of high performance, redundant Intel processor-based servers (referred to as *nodes* or *storage engines*). Highly available I/O groups are the basic configuration of a cluster. Adding another I/O group can help increase cluster performance and bandwidth.

The SAN Volume Controller can scale *out* to support up to four I/O groups, and it can scale *up* to support 1024 host servers. For every cluster, the SVC supports up to 4096 *virtual disks*. For more information, see 16.3.1, “SVC 5.1: New SVC engine hardware” on page 388.

Personnel productivity

The SAN Volume Controller is managed at the cluster level, providing a single point of control over all the managed storage. The SAN Volume Controller provides a comprehensive, easy-to-use graphical interface for central management. This simple interface incorporates the Storage Management Initiative Specification (SMI-S) API, and further demonstrates the IBM commitment to open standards. With this single interface, administrators can perform configuration, management and service tasks over storage volumes from all the attached storage controllers. The GUI allows administrators to map disk storage volumes to virtual pooled volumes to help better use existing storage.

Link infrastructure performance to business goals

By pooling storage into a single reservoir, the SAN Volume Controller helps insulate host applications from physical changes to the storage pool, in order to minimize disruption. The SVC simplifies the storage infrastructure by including a dynamic data-migration function, which migrates storage from one device to another, without taking it offline. Using this function, administrators can reallocate and scale storage capacity without disrupting applications, increasing application availability.

With the SVC, your business can build an infrastructure from existing assets that is simpler to manage, easier to provision, and can be changed without impacting application availability. Businesses can use their assets more efficiently and actually measure the improvements. They can allocate and provision storage to applications from a single view and know the effect on their overall capacity situation instantaneously. They can also quantify improvements in their application availability to enable better quality of service goals. These benefits help businesses manage their costs and capabilities more closely, linking the performance of their infrastructure to their individual business goals.

Tiered storage

In most IT environments, inactive data makes up the bulk of stored data. The SAN Volume Controller helps administrators control storage growth more effectively by moving low-activity or inactive data into a hierarchy of lower-cost storage. Administrators can free disk space on higher-value storage for more-important, active data. This is achieved by creating various groups of storage, corresponding to underlying storage with various characteristics, for example, speed and reliability. With the SAN Volume Controller, you can better match the cost of the storage used to the value of data placed on it.

Copy services

With many conventional SAN disk arrays, copy services can be performed within the array, or between identical arrays. The SAN Volume Controller enables administrators to apply a single set of advanced copy services, such as the IBM FlashCopy and Global Mirror service across multiple storage subsystems from various vendors.

With the previous software release 4.3 and now additional features in the new 5.1 release copy services functionality is significantly improved. SVC now supports remote mirroring of data volumes at both “metropolitan” and “global” distances (IBM Metro Mirror and Global Mirror) with cluster relationships possible in a variety of combinations between 4 separate clusters. This provides the ability to have a consolidated DR site replicating data for up to 3 production sites. This capability provides organizations more flexibility and choice to create solutions that meet their unique requirements for disaster recovery capabilities.

New reverse FlashCopy functionality enables FlashCopy targets to become restore points for the source without breaking the FlashCopy relationship and without having to wait for the original copy operation to complete. SVC supports multiple targets and thus multiple rollback points.

For more information about these advanced copy services facilities, see the sections beginning with 16.6.1, “Metro Mirror and Global Mirror” on page 395.

Technology for an on demand environment

Businesses are facing growing, critical application data supported by complex heterogeneous storage environments, while their staffs are overburdened. The SAN Volume Controller is one of many offerings in our System Storage portfolio that are essential for an on demand storage environment. These offerings can help you simplify your IT infrastructure, manage information throughout its lifecycle and maintain business continuity.

16.3 New features in recent SAN Volume Controller versions

SVC 5.1 is a major release featuring new node (engine) hardware models and software features. This release also incorporates functionality improvements and increased maximum configuration limits. Prior to SVC 5.1, release 4.3 also included significant changes that were not in the previous version of this Redbooks publication and so are included here for completeness.

16.3.1 SVC 5.1: New SVC engine hardware

In this section we describe features of the new SAN Volume Controller engine hardware:

CF8 premium node: This node has the following features:

- ▶ Double the throughput (IOPs and GB/s) capability of the existing (already industry leading) 8G4 nodes
- ▶ Upgrade option to add SAS SSD drives to each node by an IBM developed very high throughput SAS HBA
- ▶ Triple the cache capacity per node - 24GB - 192GB in an 8 node cluster.
- ▶ 4x 8Gbit Fabric ports.
- ▶ Dual-redundant internal power supplies
- ▶ Native iSCSI host attachment. (also available for existing hardware platforms.)

8A4 EE node: This node has the following features:

- ▶ Available with SVC entry level edition software (EE) 8A4 is the only node hardware currently available for EE.
- ▶ Maximum 250 disks configurable in 5.1 (up from 64 in 4.3)
- ▶ Single I/O group

16.3.2 SVC 5.1: New SVC software features

In this section we introduce significant new features of the SVC 5.1 release. More detail on these features can be found in this Redbooks publication and in *Implementing the IBM System Storage SAN Volume Controller V5.1*, SG24-6423:

- ▶ Multiple Cluster Mirroring (support for up to four clusters in various disaster recovery configurations)
- ▶ Native iSCSI host attachment for all node models except 4F2
- ▶ Split-cluster GA Support (ability to split the nodes in the cluster across two sites as an HA solution)
- ▶ Improvements in I/O handling for VDisk mirroring, reducing failover times
- ▶ Multi-target Reverse FlashCopy (fast recovery from backups, before they have completed and without disrupting backup cascades)
- ▶ Space-efficient Enhancements
- ▶ VDisk Mirroring to migrate a “thick” to a “thin” virtual disk
- ▶ Zero Impact Zero Detection on incoming writes; do not allocate any space when an incoming write contains all zeros
- ▶ Cache Algorithm Optimizations: Enhanced cache management algorithms provide response time reductions and increased throughput, even on existing node hardware

16.3.3 What was new with SVC 4.3.1

Here we present a recap of the features that were added with release 4.3.1:

- ▶ New node type: 2145-8A4:

The Entry Edition hardware comes with identical functionality as the 2145-8G4 nodes, 8 GB memory and four 4 Gbps FC interfaces. They provides approximately 60% of the performance of the actual 2145-8G4 nodes. This an ideal choice for entry level solutions with reduced performance requirements, but without any functional restrictions. It comes with a physical disk based licensing.

- Embedded CIMOM:

The CIMOM and associated SVC CIM Agent is the software component that provides the industry standard CIM protocol as a management interface to the SAN Volume Controller. Up to SVC 4.3.0, the CIMOM ran on what was initially known as the SVC Master Console, but was replaced in SVC 4.2.0 by the SSPC based management console. The SSPC is an integrated package of hardware and software that provides all the management software (SVC CIMOM, SVC GUI) required to manage the SAN Volume Controller as well as components for managing other storage systems.

Customers might continue to use either the Master Console or SSPC to manage SVC 4.3.1. In addition, the software components required to manage the SAN Volume Controller (SVC CIMOM and SVC GUI) are provided by IBM in software form allowing customers who have a suitable hardware platform to build their own Master Console.

Note: With SVC 5.1, the usage of the embedded CIMOM is mandatory. Therefore, in an upgrade scenario, it is best to first switch the existing configurations from the Master Console/SSPC based CIMOM to the embedded CIMOM (remember to update the TPC configuration if it is in use). In a second step, upgrade the Master Console/SSPC, and finally do the upgrade of the SAN Volume Controller cluster.

- Windows 2008 support for the SAN Volume Controller GUI and Master Console
- SSPC 1.3 support
- NTP synchronization: The SVC cluster time will operating in one of two exclusive modes:
 - Default mode: In this mode, the cluster uses the configuration node's system clock.
 - NTP mode; the cluster uses an NTP time server as its time source, and adjusts the config node's system clock according to time values obtained from the NTP server. When operating in NTP mode, the SVC cluster will log an error if an NTP server is unavailable.
- Performance enhancement for overlapped Global Mirror writes

16.4 SAN Volume Controller architecture

The IBM System Storage SAN Volume Controller is based on the COMmodity Parts Storage System (*Compass*) architecture developed at the IBM Almaden Research Center. The overall goal of the Compass architecture is to create storage subsystem software applications which minimize the changes required when new storage hardware is deployed. The approach is to minimize the dependency on unique hardware, and to allow exploitation of, or migration to, new SAN interfaces simply by plugging in new commodity adapters. Compass solutions can scale and expand over time, because they can be ported to just about any platform and remain current with the latest processor and chipset technologies on each. The SAN Volume Controller compass architecture implementation uses Linux as a convenient development platform to deploy this function. This has, and will continue to, enhance the ability of IBM to deploy robust function in a timely way.

The SAN Volume Controller is a modular solution that consists of a management console (SSPC) for management, up to eight cluster nodes (added in pairs), and external UPS, one for each node, for write cache data protection. The nodes are the hardware elements of the SAN Volume Controller. SVC combines these nodes (engines) to create a high availability cluster. Each of the engines in the cluster is populated with up to 24GB of high-speed memory, which serves as the cluster cache. A management card installed in each server monitors various parameters used by the cluster to determine the optimum and continuous data path.

The storage engines (or storage nodes) are always installed in pairs and combined into a high availability cluster. In simple terms, a cluster is a collection of servers that, together, provide a set of resources to a client. The key point is that the attached hosts (clients) using the SAN Volume Controller to access storage have no knowledge of the underlying physical hardware of the cluster. This means that the host client is isolated and thus protected from changes to the physical storage hardware. This brings a number of benefits. Perhaps the most important of these benefits is high availability. Resources on clustered servers act as highly available versions of unclustered resources. If a node (an individual computer) in the cluster is unavailable, or too busy to respond to a request for a resource, the request is transparently passed to another node capable of processing it. Clients are, therefore, unaware of the exact locations of the resources they are using, as can be seen in the diagram in Figure 16-2.

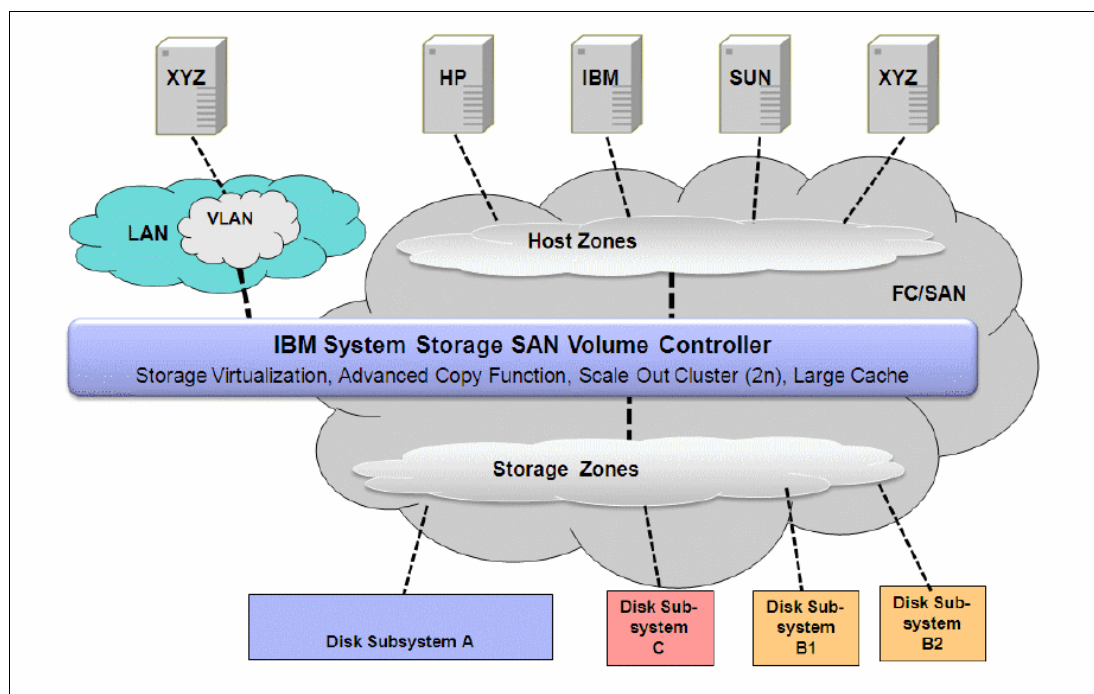


Figure 16-2 SAN Volume Controller Architecture overview

During normal cluster operation, the nodes keep in touch with each other. If a node is idle for a while (a few seconds) then a heartbeat signal is sent to assure connectivity with the cluster. If a node fails for any reason, the workload intended for it will be taken over by another node until the failed node has been restarted and re-admitted to the cluster (which happens automatically). If the microcode on a node becomes corrupted, resulting in a failure, the workload is transferred to another node. The code on the failed node is repaired, and the node is re-admitted to the cluster (again, all automatically).

Note: Although the SAN Volume Controller code is based on a Linux kernel, the clustering feature is not based on Linux clustering code. The clustering failover and failback feature is part of the SAN Volume Controller application software itself.

SAN Volume Controller nodes within the cluster are grouped into pairs (called I/O groups), with a single pair being responsible for serving I/O on a given VDisk (a VDisk is the SAN Volume Controller device that appears to a host system as a SAN attached disk). One node within the I/O Group represents the preferred path for I/O to a given VDisk; the other node represents the non-preferred path. This preference will alternate between nodes as each VDisk is created within an I/O Group to balance the workload evenly between the two nodes.

Each I/O group contains up to 48 GB of mirrored cache (on SAN Volume Controller model CF8 nodes). Each node also includes a Host Bus Adapter (HBA) providing 4 FC SAN connections, and 2 1Gb Ethernet connections for iSCSI and management interface connectivity.

The HBA allows the SAN Volume Controller to connect and operate at up to 8 Gbps SAN speed.

- ▶ 8Gbps for CF8 and 8A4 nodes.
- ▶ 4 Gbps (Standard for 8G4 and 8F4 and option for 8F2)
- ▶ 2 Gbps (standard for model 4F2 and 8F2)

Note: The 2145-8G4 SAN Volume Controller was withdrawn from marketing in December 2009.

Beyond automatic configuration and cluster administration, data from attached storage clients is also transmitted reliably. When data is written by the host, the preferred node within the I/O Group stores a write in both its own write cache and the write cache of its partner (non-preferred) node before sending an “I/O complete” status back to the host application. To ensure that data is written if a node fails, the surviving node empties its write cache and proceeds in write-through mode until the cluster is returned to a fully operational state.

As yet another data protection feature, the SAN Volume Controller is supplied with UPS. In addition to voltage regulation to protect valuable electronic components within the SAN Volume Controller configuration, in the event of a main power outage, the UPS provides enough power to destage data to the SAN Volume Controller internal disk and shut down the nodes within the SAN Volume Controller cluster gracefully. This is a feature found in most high-end disk subsystems.

Important: The SAN Volume Controller UPS is essentially an NVram destage backup battery. This UPS does not provide “normal” UPS facilities to allow the SAN Volume Controller to continue operation in the event of external power failure. Ensure that an external UPS is provided if this functionality is required.

16.5 SSD support in the SAN Volume Controller

The optional solid state drives (SSDs) available in the new 2145-CF8 nodes provide a new ultra-high-performance storage option. They are available in the 2145-CF8 nodes only. SSDs can be pre-installed in the new nodes or installed on a per disk basis at a later point in time without interrupting service, as a field hardware upgrade.

The performance improvements made in CPU cache/memory access times have been far greater than the physical disk performance improvements made during the same period of time. The disks are limited by the rotational speed and mechanical positioning technology used. As a result there is now a huge difference between time to access data resident in cache or memory (typically in the region of 10 - 50 nanoseconds for L2 cache and DRAM) and data on even the fastest disks (5 -10 ms).

SSD technology or Storage Class Memory (SCM) can help in overcoming the problem that has grown over the last few decades, and is known as the *Memory/Storage Bottleneck*. The diagram in Figure 16-3 shows a graphical representation of the scale of this problem and how SCM can help. A more comprehensive explanation of this technology and its application in the SAN Volume Controller can be found in the *Implementing the IBM System Storage SAN Volume Controller V5.1*, SG24-6423.

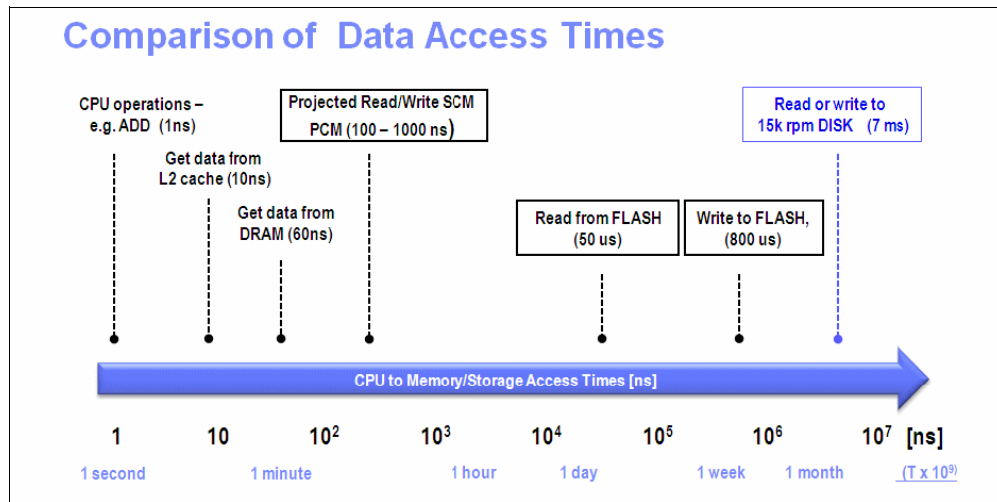


Figure 16-3 The Memory/Storage bottleneck

SVC SSD performance

The scale out performance provided by the unique SVC SSD implementation will deliver ultra-high throughput and fast response time for hard-to-tune workloads, up to 800,000 read IOPS. Also, SSD is fully integrated within the SAN Volume Controller architecture. Replication, data movement and management are the same for the SSD backed MDisk/VDisk storage as for any other managed storage in the SAN Volume Controller.

SVC SSD features

Here we highlight a few of the SSD features available:

- ▶ Up to four SSDs can be installed on each SAN Volume Controller 2145-CF8 node
- ▶ Each SSD provides up to 140 GB of capacity
- ▶ SSDs are hot-pluggable and hot-swappable
- ▶ Up to 560 GB usable SSD capacity per node, 2.4TB (mirrored) per cluster

16.6 SAN Volume Controller virtualization

The SAN Volume Controller provides block aggregation and volume management for disk storage within the SAN. Essentially this means that the SVC manages a number of back-end storage controllers and maps the physical storage within those controllers to logical disk images that can be seen by application servers and workstations in the SAN. The SAN is zoned in such a way that the application servers cannot see the back-end storage, preventing any possible conflicts caused by the SVC and the application servers both trying to manage the same back-end storage devices or LUNs.

The SVC I/O Groups are connected to the SAN so that all back-end storage and all application servers are visible to all of the I/O Groups. The SVC I/O Groups see the storage presented to the SAN by the back-end controllers as a number of disks, known as Managed Disks or MDisks. Because the SVC does not attempt to provide recovery from physical disk failures within the back-end controllers MDisks are usually, but not necessarily, part of a RAID array. The application servers on the other hand do not see the MDisks at all.

The zoning layout is shown in Figure 16-4. Host zones connect the hosts to the I/O Groups, and the disk zone connects the I/O Groups to the storage controllers. Typically there are several host zones, because hosts can be grouped together in “host groups” to better manage access.

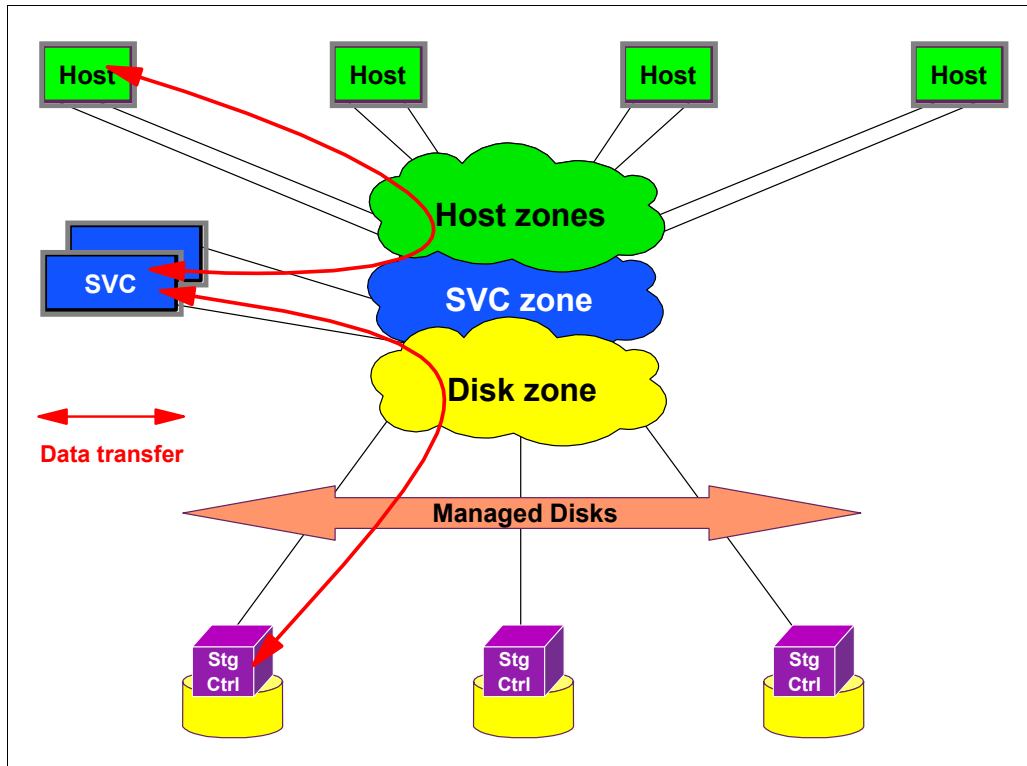


Figure 16-4 How to zone for correct SAN Volume Controller disk access

The MDisks are collected into groups, known as managed disk groups (MDGs). The MDisks that are used in the creation of a particular VDisk must all come from the same MDG. Each MDisk is divided up into a number of extents (default minimum size 16 MB, maximum size of 2GB), which are numbered sequentially from the start to the end of each MDisk. Conceptually, this might be represented as shown in Figure 16-5.

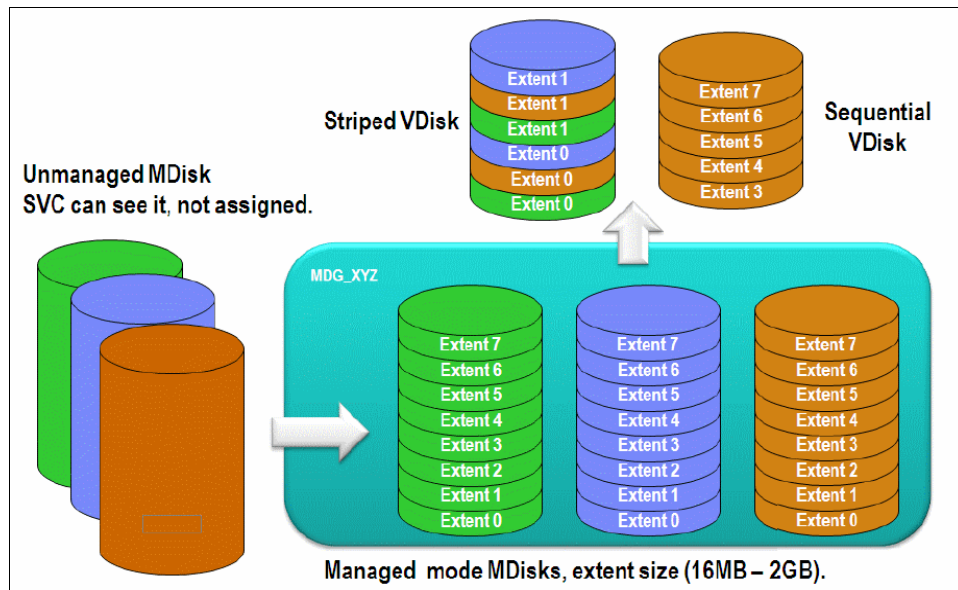


Figure 16-5 Extents being used to create a VDisk

The virtualization function in the SAN Volume Controller maps the VDisks seen by the application servers on to the MDisks provided by the back-end controllers. I/O traffic for a particular VDisk is, at any one time, handled exclusively by the nodes in a single I/O Group. Therefore, although a cluster can have many nodes within it, the nodes handle I/O in independent pairs. This means that the I/O capability of the SVC scales well (almost linearly), because additional throughput can be obtained by simply adding additional I/O Groups.

Each SVC node presents a VDisk to the SAN through multiple paths, usually four. In normal operation, two nodes are used to provide redundant paths to the same storage, meaning that depending on zoning, a single host HBA will see up to eight paths to each LUN presented by the SVC. Because most operating systems are not able to resolve multiple paths back to a single physical device, IBM provides a multipathing device driver.

The multipathing driver supported by the SVC is IBM Subsystem Device Driver (SDD). Note that the SDD code supports both the SVC and the disk systems such as the ESS, DS6000, and DS8000. Provided the latest version is used, a host can connect concurrently to both an SVC and a “native” DS6000 or Ds8000 environment. See Figure 16-6.

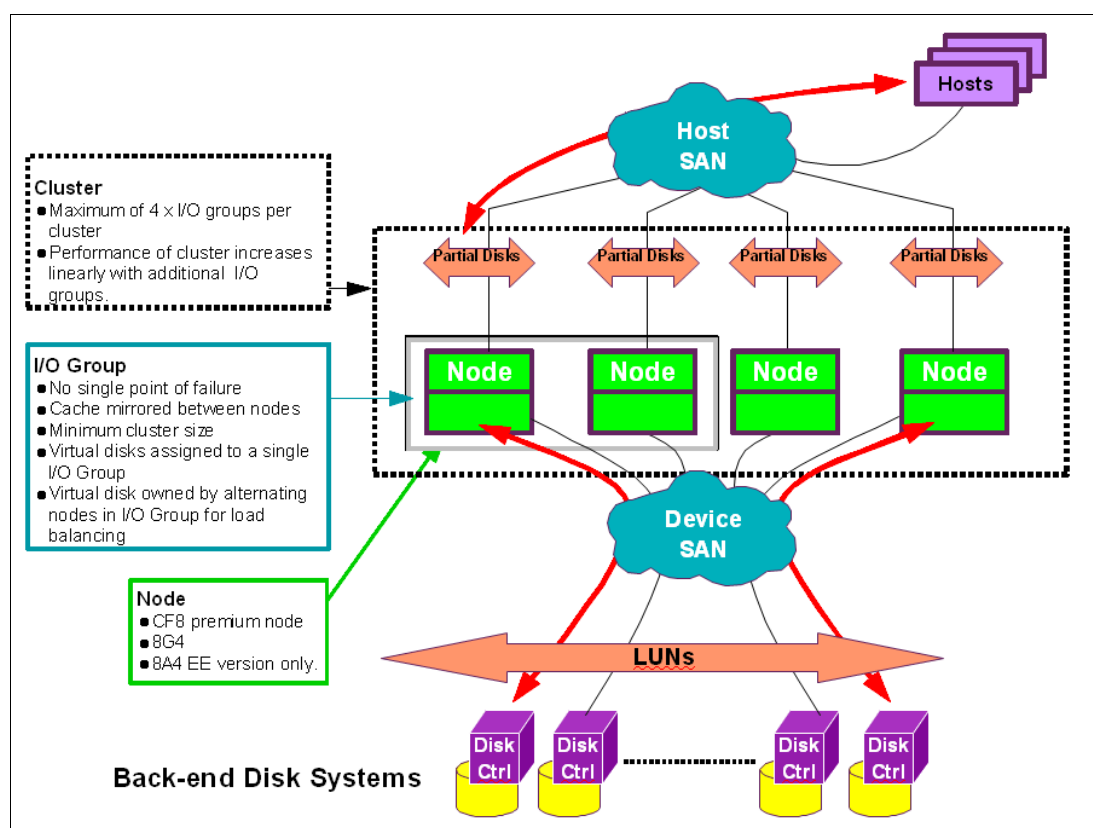


Figure 16-6 SAN Volume Controller logical view

Notes:

- ▶ Nodes auto-negotiate the fabric speed on a per-port basis and are not restricted to run at the same speed as other node pairs in the cluster.
- ▶ To provide high availability, the SAN Volume Controller nodes must be configured in redundant SAN fabrics.
- ▶ The Fibre Channel switches need to be zoned to permit the hosts to see the SVC nodes and the SVC nodes to see the storage Controllers. The SVC nodes within a cluster must be able to see each other and the master console. In addition, if there are two SVC clusters with Metro Mirror (formerly referred to as PPRC) services between them, zoning must be set so that all the nodes in both clusters see all the other nodes in both clusters.
- ▶ As well as a Fibre Channel connection or connections, each device has an Ethernet connection for configuration and error reporting, though only one of the nodes, the configuration node, binds an IP address to its Ethernet connection.

16.6.1 Metro Mirror and Global Mirror

Metro Mirror (MM) for SAN Volume Controller is, like its equivalents on DS5000, DS6000, and DS8000, a synchronous remote replication function that operates over “metropolitan” distances. Global Mirror (GM) is an asynchronous replication over much longer distances than those supported by MM.

The SVC Metro Mirror function supports replication at distances up to 300 km, although as with any synchronous remote replication function, the performance impact on host applications is related to the distance between source and target. Accordingly, application performance requirements might limit the usable distance between sites.

Like other synchronous remote replication functions, the SVC Metro Mirror function indicates that a host server I/O is complete only after data has been stored (usually in cache) at both locations.

After the initial copy operation from the source location to the target location is complete, the SVC Metro Mirror function is designed to maintain a fully synchronized copy of the source data at the target site at all times.

We are delivering MM and GM as a single feature, which gives great flexibility to customers to use one or both functions as they see fit. Existing MM customers will be able to use GM at no additional license charge. The functions are licensed by capacity of source and target volumes. The capacity used for MM and GM is added together to determine the correct charging tier. A single vDisk can use MM or GM but not both.

The SVC Metro Mirror function operates between SVC clusters at each site. Unlike conventional Metro Mirror functions, with SVC, the source and target volumes can be in the same or separate disk systems, which can come from any supported SVC storage.

Multiple Cluster Mirroring enables Metro Mirror and Global Mirror relationships to exist between a maximum of four SVC clusters.

Figure 16-7 shows an example of a configuration with a consolidated DR site (A) providing DR resources for SVC clusters on production sites B, C, and D (B to A, C to A, D to A), but other configurations (such as triangle-shaped or square-shaped) are also supported.

Both Metro Mirror and Global Mirror are supported within the same overall configuration (for example, B to A as Metro Mirror, and D to A as Global Mirror).

The SVC clusters can take advantage of the maximum number of remote mirror relationships, which has been increased to 8192 in SVC 5.1.

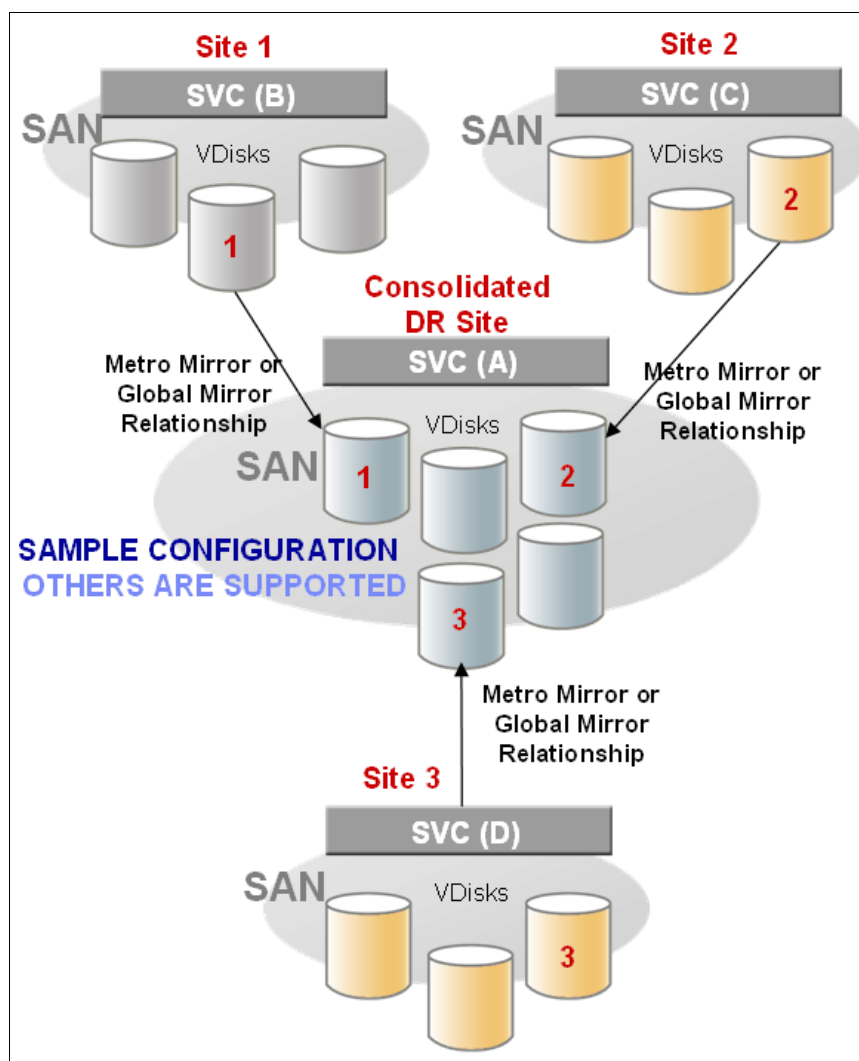


Figure 16-7 Multiple Cluster Mirroring configuration example

16.6.2 FlashCopy

The FlashCopy (FC) function of the IBM System Storage SAN Volume Controller (SVC) provides the capability to perform a point-in-time (PiT) copy of one or more VDisks or groups of VDisks (consistency groups). This technique is used to help solve the problem where it is difficult to make a consistent copy of a data set that is being constantly updated. The FlashCopy source is frozen for several seconds during the PiT copy process. It will be able to accept I/O when the PiT copy bitmap is set up and the FlashCopy function is ready to intercept read/write requests in the IO path.

Although the “real” background copy operation takes time, the resulting data at the target appears as though the copy were made instantaneously. These FlashCopy targets can be used for application data backup, thus reducing backup windows to near zero time from an application perspective. The SVC FlashCopy service also provides the facility to make multiple copies of the same source volume or make copies at various times while preserving the earlier copies (restore points). See Figure 16-8.

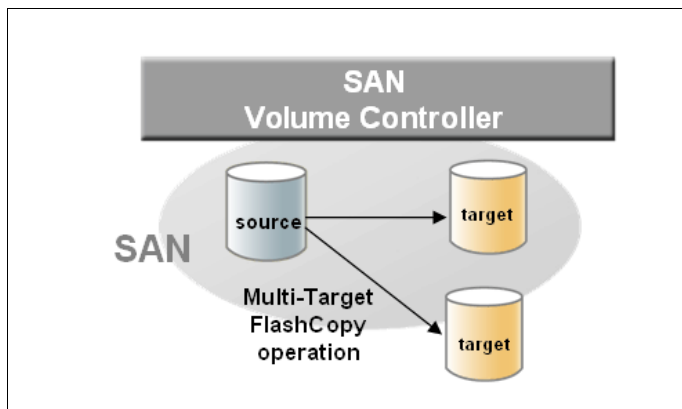


Figure 16-8 Multi target FlashCopy

With the Space Efficient (SE) VDisk and VDisk mirroring functions introduced recently in the SAN Volume Controller, it becomes possible to use FlashCopy and SE VDisks to perform space reclamation within the managed disk space of the SVC. See 16.6.6, “Space reclamation”.

16.6.3 Reverse FlashCopy

With SVC version 5.1.x, Reverse FlashCopy support is available (Figure 16-9). Reverse FlashCopy enables FlashCopy targets to become restore points for the source without breaking the FlashCopy relationship and without having to wait for the original copy operation to complete. It supports multiple targets (up to 256 per source volume) and thus multiple rollback points.

A key advantage of SVC Multi-Target Reverse FlashCopy function is that the reverse does not destroy the original target; thus anything using the target, such as a tape backup process, will not be disrupted; multiple separate recovery points can be tested.

The SAN Volume Controller is also unique in that an optional copy of the source VDisk can be made before starting the reverse copy operation in order to diagnose problems.

For more details, see 16.6.4, “Mirrored VDisks” on page 398.

Figure 16-9 provides an illustration of Reverse FlashCopy

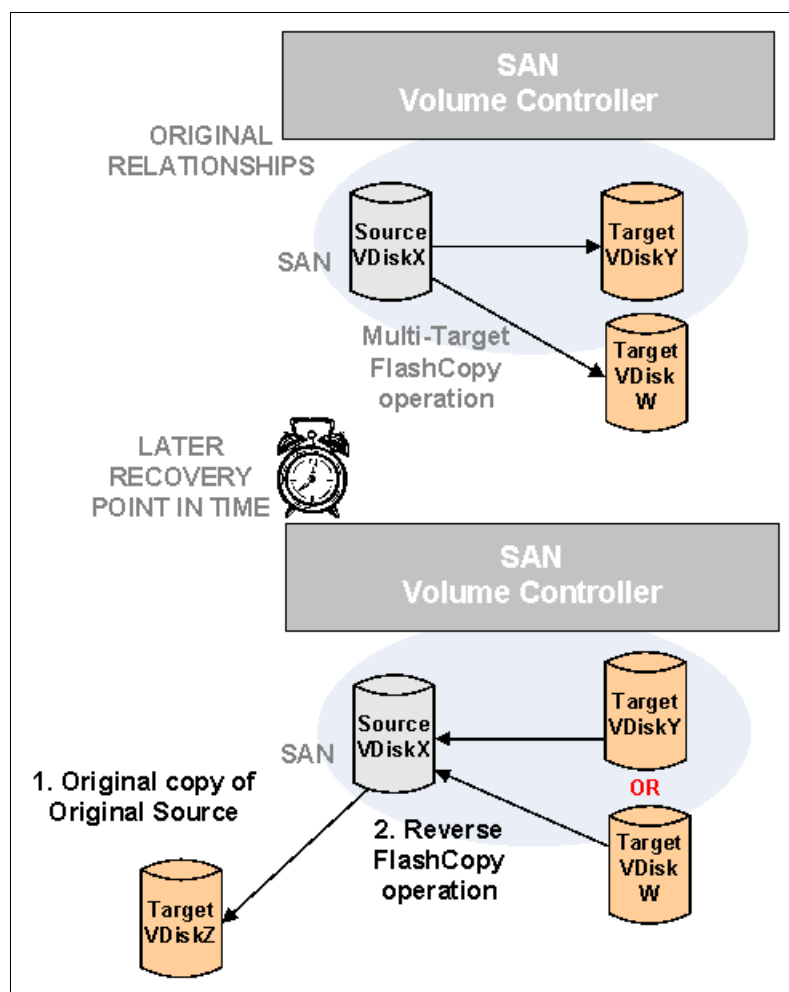


Figure 16-9 Reverse FlashCopy

When a user suffers a disaster and needs to restore from an on-disk backup, the procedure is as follows:

1. (Optionally) Create a new target VDisk (VDisk Z) and FlashCopy the production VDisk (VDisk X) onto the new target for later problem analysis.
2. Create a new FlashCopy map with the backup to be restored (VDisk Y) or (VDisk W) as the source VDisk and VDisk X as the target VDisk if this does not already exist.
3. Start the FlashCopy map (VDisk Y -> VDisk X) with the new "-restore" option to copy the backup data onto the production disk.

The production disk will be instantly available with the backup data.

16.6.4 Mirrored VDisks

Starting with SVC 4.3, the Mirrored VDisk feature provides a simple RAID1 function which allows a VDisk to remain accessible even when an MDisk on which it depends has become inaccessible. This is achieved using two copies of the VDisk typically allocated from various MDisk Groups or using image-mode copies.

A copy is not a separate object and cannot be created or manipulated except in the context of the VDisk. Copies are identified by the configuration interface with a copy ID of their parent VDisk.

The feature does provide a “point in time” copy functionality achieved by “splitting” a copy from the VDisk. The feature does not address other forms of mirroring based on Remote Copy (sometimes called “Hyperswap”), which mirrors VDisks across I/O Groups or clusters, nor is it intended to manage mirroring or remote copy functions in back-end controllers.

Figure 16-10 gives an overview of how VDisk mirroring works.

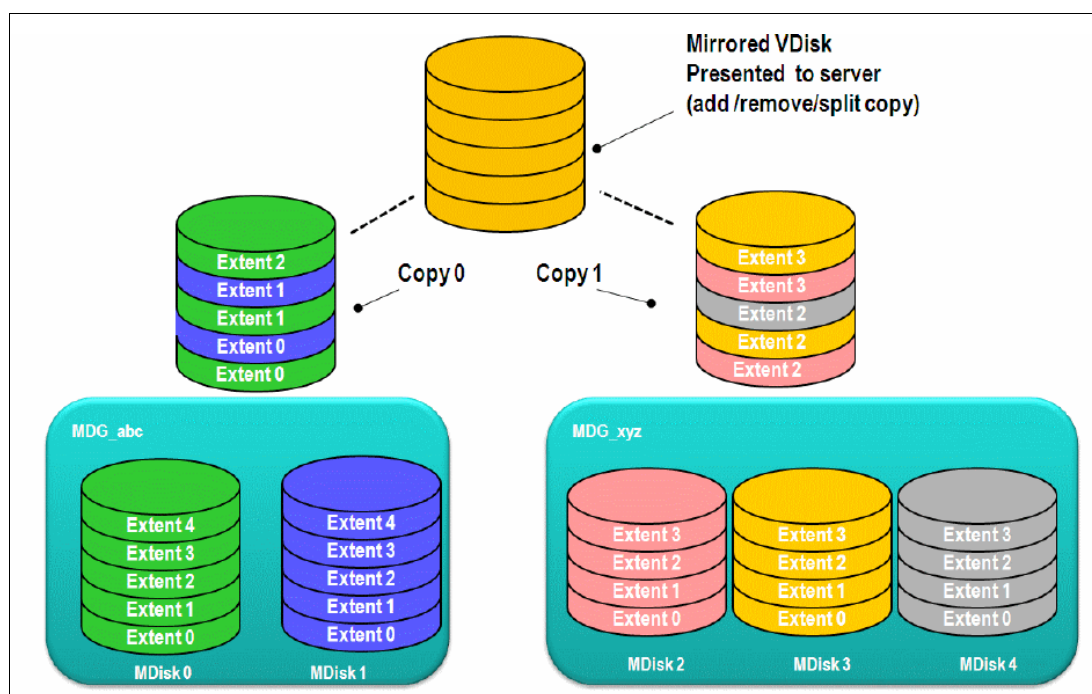


Figure 16-10 VDisk mirroring

16.6.5 Space-Efficient VDisks

Starting with SVC 4.3, VDisks can be configured to either be *Space-Efficient* or *Fully Allocated*. A Space-Efficient VDisk (SE VDisk) will behave with respect to application reads and writes, as if it were fully allocated. When an SE VDisk is created, the user specifies two capacities, the real capacity of the VDisk and its virtual capacity. This practice is also referred to as *Thin Provisioning*.

The *real capacity* determines the quantity of MDisk Extents that will be allocated for the VDisk. The *virtual capacity* is the capacity of the VDisk reported to other SAN Volume Controller components (such as FlashCopy, Cache, Remote Copy) and to the host servers.

The real capacity is used to store both the user data and the metadata for the SE VDisk. The real capacity can be specified as an absolute value or a percentage of the virtual capacity. The Space-Efficient VDisk feature can be used on its own to create over-allocated or late-allocation VDisks, or can be used in conjunction with FlashCopy to implement Space-Efficient FlashCopy. SE VDisks can also be used in conjunction with the Mirrored VDisks feature.

16.6.6 Space reclamation

SE VDisk and VDisk mirroring enable customers to reclaim unused allocated disk space (zeros) when converting a fully allocated VDisk to a Space-Efficient Virtual Disk (SEV) using VDisk Mirroring. To migrate from a fully-allocated to Space-Efficient VDisk, you simply need to add the target space-efficient copy, wait for synchronization to complete, then remove the source fully-allocated copy. This procedure for migrating from fully allocated “thick” VDisks to space efficient “thin” VDisks can be seen in Figure 16-11.

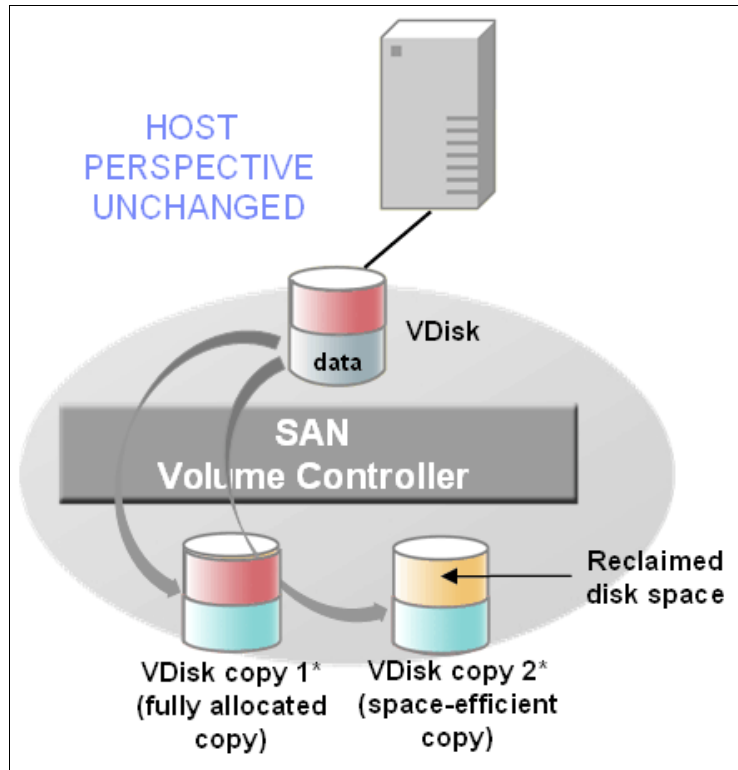


Figure 16-11 Thick to thin VDisk migration.

16.7 SAN Volume Controller iSCSI host attachment

SVC 5.1.0 now supports host access by iSCSI (as an alternative to Fibre Channel). The iSCSI attachment feature reduces the requirement for expensive host HBAs and FC ports on the SAN and so provides a lower cost attachment alternative. All SVC functionality available to FC attached hosts is also available to any iSCSI attached hosts with the exception of host boot support in the initial release. This new functionality is available at no extra cost to all SVC customers, the iSCSI attachment feature is part of the SVC 5.1 base software.

SVC uses the node built-in Ethernet ports for iSCSI traffic. All node types which can run SVC v5.1.0 can use the iSCSI feature (that is all nodes apart from model 4F2).

SVC supports the CHAP authentication methods for iSCSI, IP addresses can fail over to the partner node in the IO group if a node fails. This reduces the need for multipathing support in the iSCSI host.

Support for iSCSI introduces one additional ipv4 and one additional ipv6 address for each Ethernet port on every node. These IP addresses are independent of the cluster configuration IP addresses.

The following diagrams indicate a sample of the available configurations for iSCSI and management interface connections. We only show IPv4 addresses for clarity, but IPv6 addresses can also be assigned in addition to, or instead of, IPv4 addresses.

Figure 16-12 indicates the simplest configuration with all interfaces on the same subnet. Figure 16-13 shows a three subnet configuration with redundant iSCSI connections and a single management subnet. Figure 16-14 shows a fully redundant configuration.

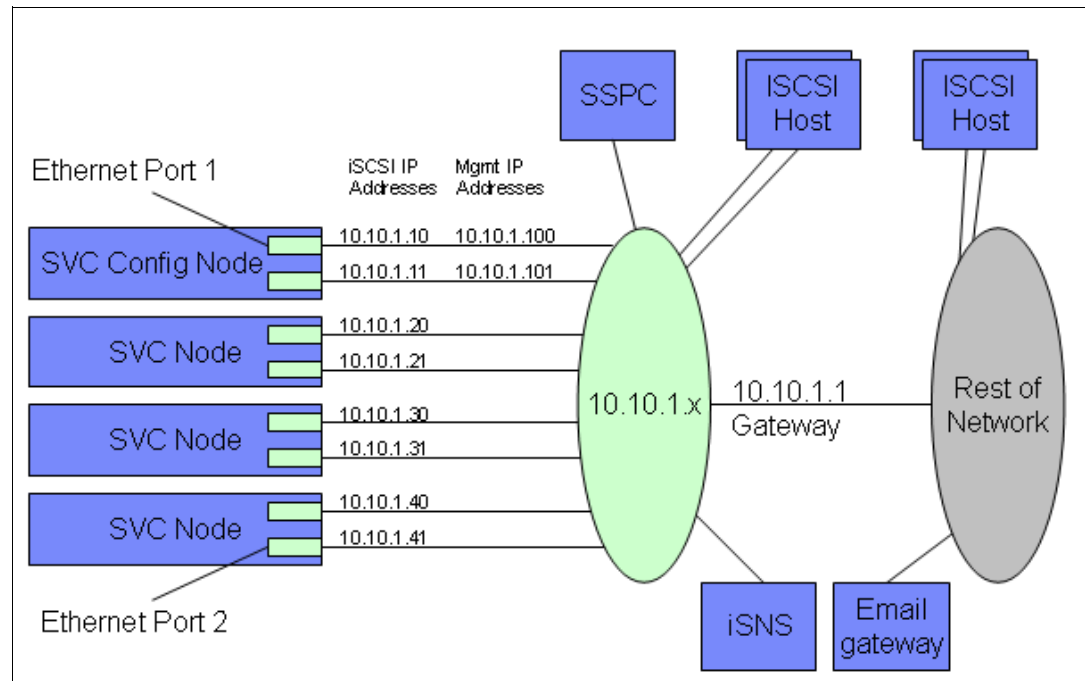


Figure 16-12 Simple single subnet config

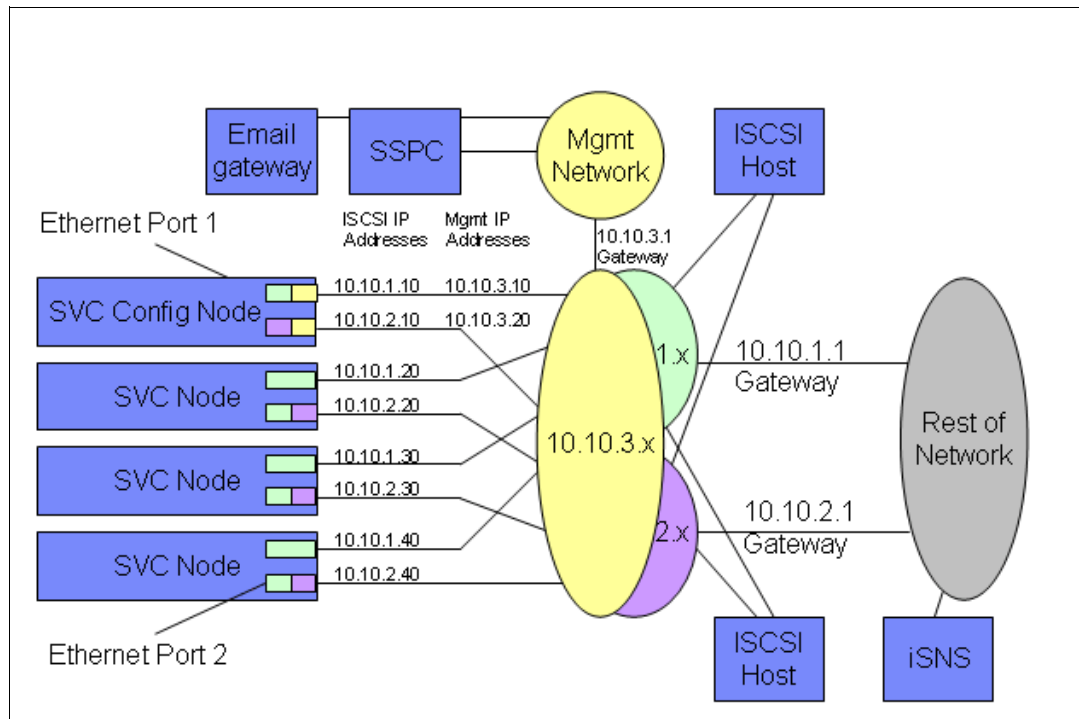


Figure 16-13 Redundant iSCSI subnets

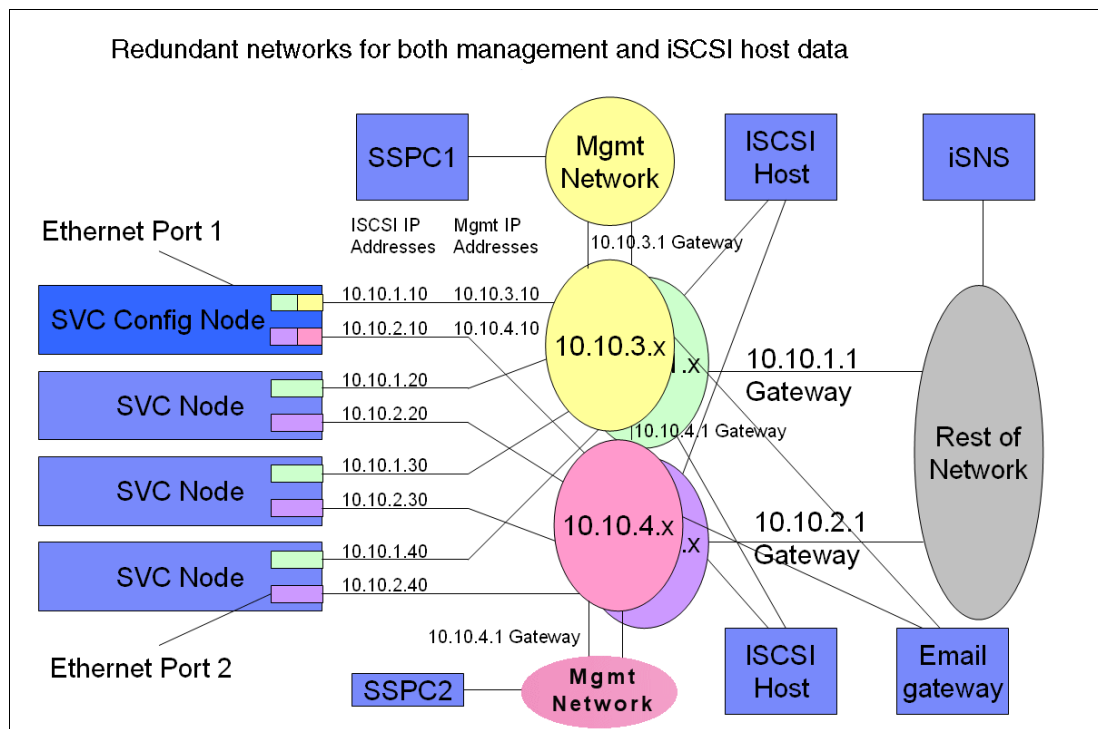


Figure 16-14 Fully redundant iSCSI and management subnets

16.7.1 Supported iSCSI host attachments

The following list details the available iSCSI host attach options with the initial release:

- ▶ RHEL SP 5.3, RHEL 4 update 6 (32 and 64-bit)
- ▶ SLES10 SP2 (32 and 64-bit)
- ▶ Windows 2003 SP1, SP2
- ▶ Windows 2008 SP1, SP2
- ▶ AIX 5.3, 6.1
- ▶ Sun Solaris 10
- ▶ HP-UX 11i V3

16.7.2 Additional Ethernet changes

The cluster configuration node might now be accessed on both Ethernet ports. This means the cluster might have two ipv4 and two ipv6 addresses that are used for configuration purposes. The cluster might therefore be managed by SSPCs on separate networks – this provides redundancy in the event of a failure of one of these networks, however there is no automatic failover capability, so if one network is down, manually switch to the other IP address.

16.8 SAN Volume Controller software licensing

There are three parameters to consider when licensing the SAN Volume Controller software:

- ▶ The amount of storage managed by the SAN Volume Controller cluster:

The total amount of storage that is managed by the SAN Volume Controller cluster might be greater than the amount of storage that is virtualized by the cluster. Usage is charged based on the amount of Managed disk and not the amount of storage virtualized by a cluster, as shown in Figure 16-15.

- Usage is charged based on the amount of Managed disk and not the amount storage virtualized by a cluster

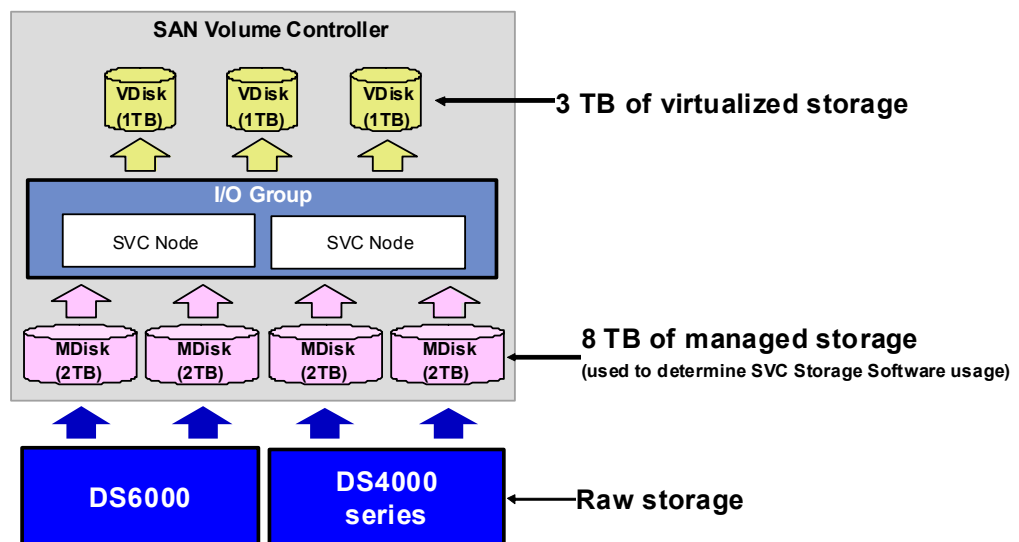


Figure 16-15 SAN Volume Controller managed storage-based licensing

- The amount of storage participating in FlashCopy:

The amount of virtualized storage that you want to participate in simultaneous active FlashCopy relationships is priced based on the amount of source storage using FlashCopy in a cluster. You might want to manage 8 TB of storage, but FlashCopy only 1 TB of storage. In this case, you need a 1 TB FlashCopy license even though there is 1 TB of source and 1 TB of target FlashCopy volumes. See Figure 16-16.

A single mapping from a source Virtual Disk of 1 TB to a target of 1 TB consumes 1 TB of license quota, further multiple target mappings from the same source consume no extra license quota. A mapping from a 1 TB target of an existing mapping to another 1 TB target (a cascaded mapping) consumes a further 1 TB of license quota because the target data from the first mapping becomes source data for the cascaded mapping.

- Priced based on the amount of storage using FlashCopy in a cluster
- Storage using FlashCopy is derived from the total size of all source/target virtual disks in active FlashCopy relationships

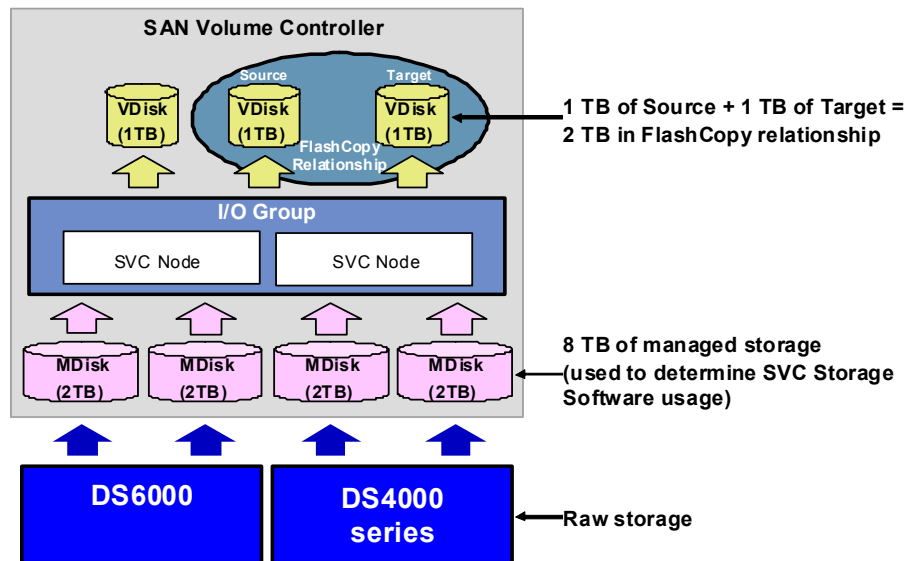


Figure 16-16 FlashCopy license illustration

- The amount of virtualized storage in Metro or Global Mirror relationships:

You might be managing 8 TB of storage and cascading FlashCopies of 1 TB of data (2 TB FlashCopy license required) and Metro Mirror copying 1 TB. In this case, both clusters that are taking part in the Metro Mirror relationship require a 1 TB Metro Mirror license. In the case of intracluster Metro Mirror, where both primary and secondary volumes are in the same cluster, then the license must be large enough to cover both, as shown in Figure 16-17.

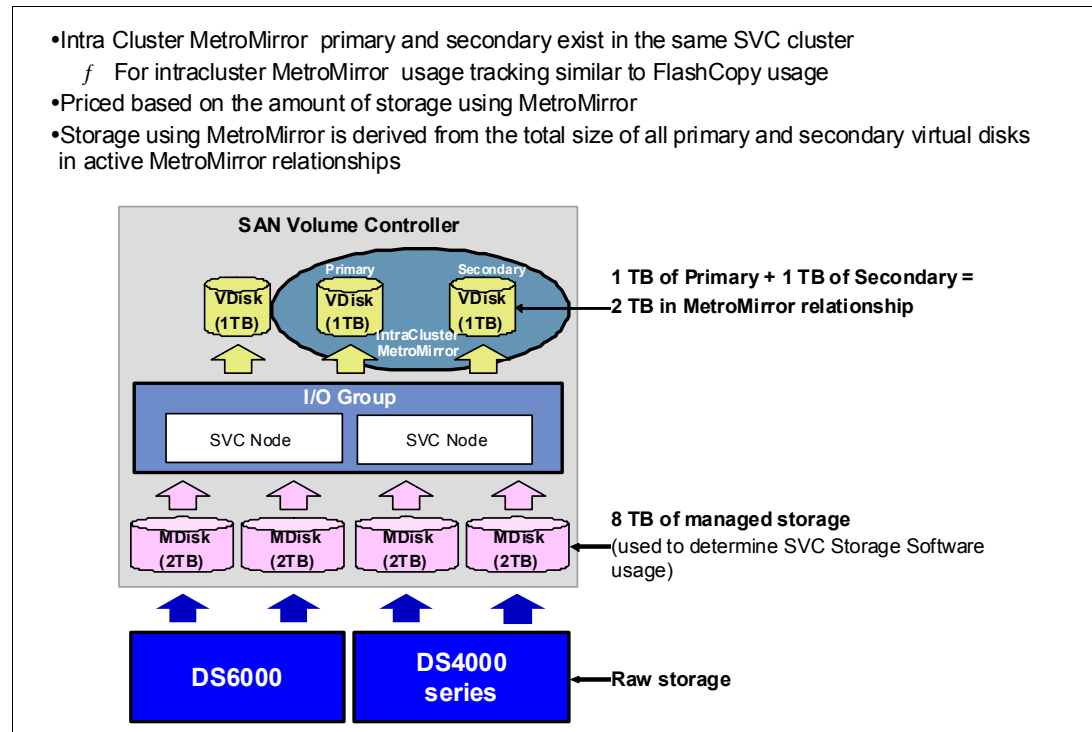


Figure 16-17 Metro Mirror and Global Mirror license illustration

As another example including Global Mirror, consider at site A: 10 TB of virtualization, 2 TB source for Metro Mirror, 2 TB target for Metro Mirror, and 2 TB source for Global Mirror; at Site B: Base virtualization of 4 TB, and 2 TB of target for Global Mirror (source is at Site A). This requires a Site A order for 10 TB base virtualization with 6 TB Metro/Global Mirror and a Site B order for 4 TB base virtualization and 2 TB of Metro/Global Mirror.

It is possible to increase any one of these three licenses independently of the other. That is, you can increase the total amount of managed storage without increasing the other licenses if the amounts of storage being copied remains unchanged. Similarly it is possible to change the copy licenses independently of each other.

16.9 SAN Volume Controller solutions

The SAN Volume Controller's unique features give it real advantages when used for storage tasks such as replication, data migration and storage capacity allocation.

16.9.1 Business continuity

Replicating data in a heterogeneous SAN environment from a high performance storage type at a central site to a lower cost set of storage at a remote site is always a performance concern. There are limits to the array based replication ability to have consistency groups across unlike devices. Also host level replication consumes precious server resources. SAN Volume Controller with its Metro and Global Mirror functions eliminates this problem making it significantly simpler to deploy for disaster recovery.

The latest SVC V5.1 software provides Global Mirror (consistent asynchronous secondary copy of data) at nearly unlimited distances with minimal performance impact on production servers. Figure 16-18 provides an overview of how this process works.

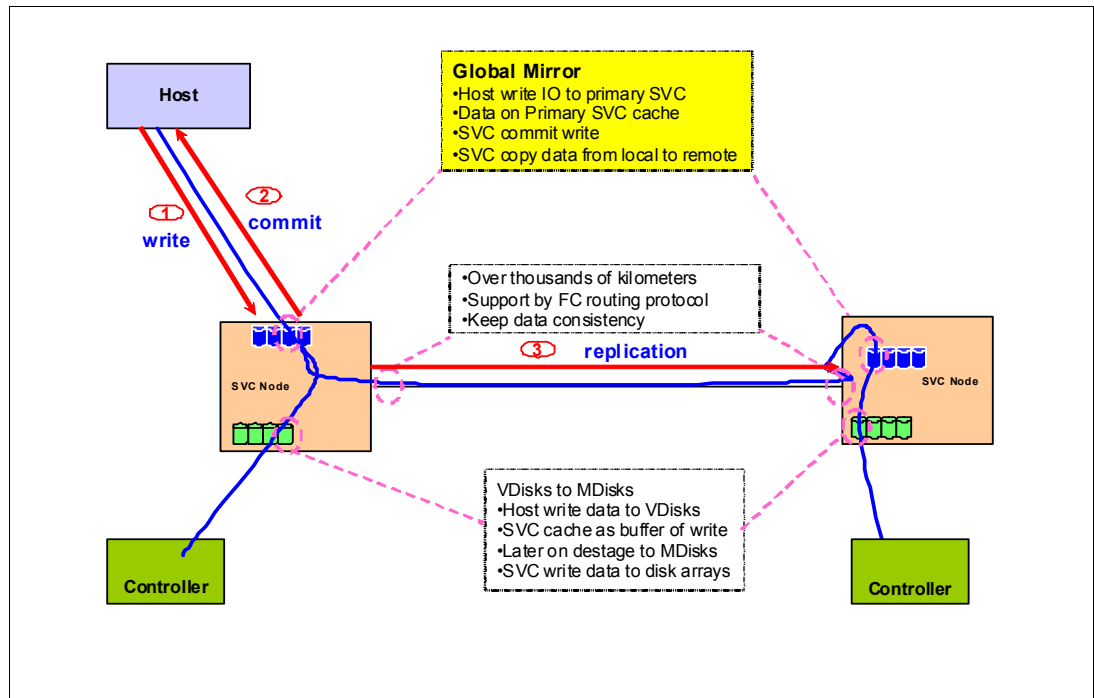


Figure 16-18 Business continuity with SAN Volume Controller Global Mirror

Metro Mirror provides synchronous data copy over shorter distances.

Both Metro Mirror and Global Mirror on SVC are supported by IBM TotalStorage Productivity Center for Replication.

16.9.2 Infrastructure simplification

The SAN Volume Controller helps simplify the storage infrastructure by data migration and storage capacity utilization features.

Data migration

Data migration is the process of moving data from one storage volume to another. This might be required because you want to move less valuable data to cheaper storage, or perhaps a new faster storage system has been purchased, and you want to move critical data onto it. In non-virtualized environments, this is a disruptive, time-consuming operation, requiring application downtime. Certain storage systems do not support a direct remote copy function - which forces host-based volume drain, back up to tapes or restore from tapes during the migration. In addition data is unavailable during the migration process.

Figure 16-19 shows a typical migration process without the SAN Volume Controller:

- ▶ Step 1: Take the storage offline.
- ▶ Step 2: Migrate data from old storage array to new storage array.
- ▶ Step 3: After migration, re-establish connections to host and bring the new storage system online.
- ▶ Step 4: Return or dispose of old array.

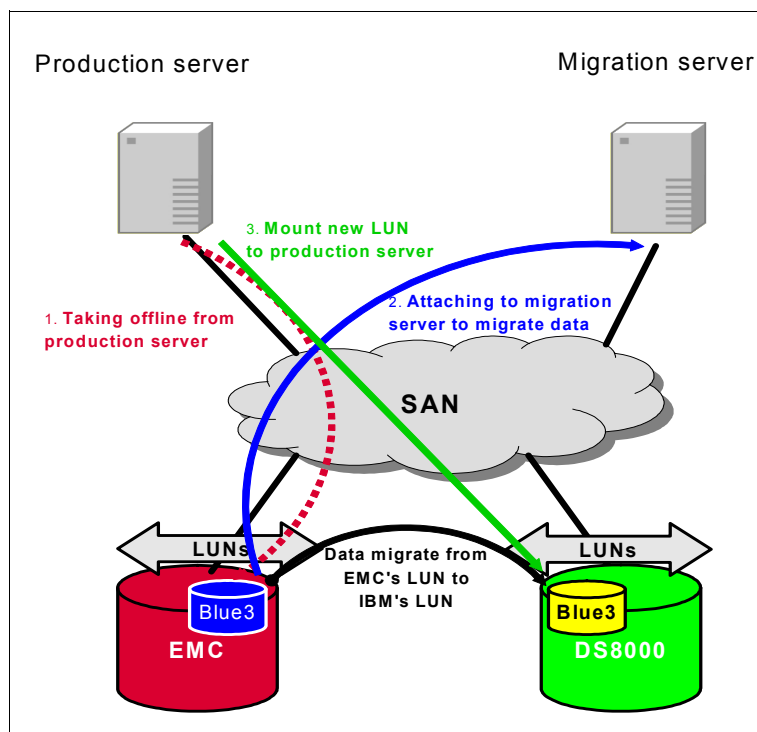


Figure 16-19 Data migration using SAN

Figure 16-20 shows how you can migrate from a non-virtualized environment, to a SAN Volume Controller virtual environment:

1. The existing storage, applications and hosts are identified.
2. Next the SVC is installed, and the original volume to be migrated is mapped as a “hypothetical” or unmanaged SAN volume on the SVC. The application server now sees the virtual volume with all the existing data mapped to it. Operations can proceed as before.
3. Finally, if for example, more space is required for that data, you can nondisruptively move the data from the original volume to a new virtual volume. This will allow dynamic expansion of the volume and frees up the space used by the original application disk.

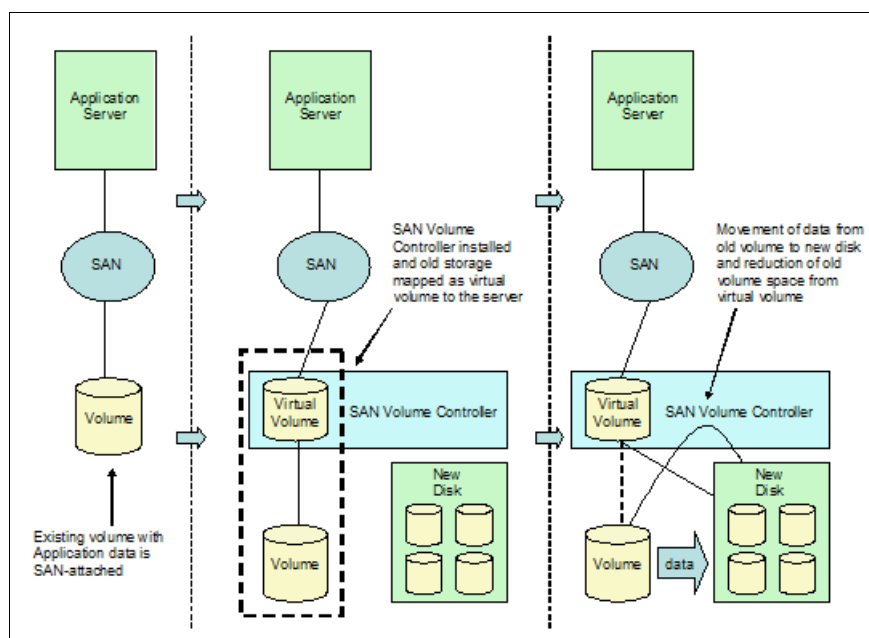


Figure 16-20 Data migration using SAN Volume Controller

To migrate existing SAN Volume Controller data to other storage, you simply tell SVC to remap the data on the virtual disk from one MDisk (managed disk) to another, online, without disrupting the application. The SVC manages all the physical infrastructure, and the attached hosts have no awareness of this, because the virtual disk (VDisk) mapping and structure remains unchanged.

Storage capacity utilization

Although traditional SAN heterogeneous environments share the physical network, but it offers very limited sharing capacity resulting in poor utilization of storage assets. It has procurement limitations due to the various purchasing requirements for each storage array island. Configuration changes are very complex as well as consolidating storage across data centers. It is hard to identify where storage pools capacity imbalances to move blocks around to improve and balance utilization.

The SAN Volume Controller enables hosts to use virtual disks which consist of managed disks from any device within the storage pools. Unused storage capacity can be easily reallocated to help improve capacity utilization. The procurement of extra storage (in an environment using the SAN Volume Controller) can be deferred until the total SAN capacity approaches utilization trigger points. Storage capacity balancing done at the virtualization layer transparently to servers and applications.

16.10 SAN Volume Controller compatibility

The SAN Volume Controller supports a wide range of host operating systems, server platforms, FC SAN switches and storage controllers. A list of the more recent additions to this host platform interoperability matrix can be seen in Table 16-1.

Table 16-1 SAN Volume Controller interop recent operating system additions.

Host platform	Added in:
SUSE Linux enterprise server V11	New in SVC 5.1
VMware Vsphere 4	New in SVC 5.1
MAC OS 10.5	New in SVC 4.3
Citrix Xen server	New in SVC 4.3
Windows 2008 server Hyper-V	New in SVC 4.3

Supported disk subsystems include: IBM System Storage DS4000, DS5000, DS6000, DS8000 series storage systems, and N series. It also supports several models of Hitachi Thunder, Lightning, and TagmaStore, EMC CLARiiON and Symmetrix, Network Appliances models, and Hewlett-Packard StorageWorks Modular Arrays and Enterprise Virtual Array. See Table 16-2 for a list of the recent additions to the supported disk controller subsystems.

Table 16-2 Recently added storage subsystems.

Storage subsystem	Added in:
IBM System Storage DS5020 DS3950	New in SVC 5.1
IBM System Storage DCS9550 DCS9900	New in SVC 5.1
Hitachi Adaptable Modular Storage (AMS) 2100, 2300, 2500	New in SVC 5.1
HP Storageworks Enterprise Virtual Array (EVA) 8400, 6400	New in SVC 5.1
HP Storageworks Modular Smart Array 2000 (MSA2000)	New in SVC 5.1
EMC Clariion CX4-960	New in SVC 5.1
Pillar Axiom 600	New in SVC 5.1
Nexsan SATABeast	New in SVC 5.1
Texas Memory Systems RamSan-500	New in SVC 5.1
Xiotech Emprise 5000	New in SVC 5.1
Fujitsu ETERNUS8000 models 2000 & 1200	New in SVC 5.1
Fujitsu ETERNUS4000 models 600 & 400	New in SVC 5.1

For complete up-to-date SAN Volume Controller supported operating systems, hosts, HBAs, SAN switches and storage subsystems, see the links for *Supported Hardware and Software Levels* for the appropriate SVC code level under the Migration and upgrade session at the following website:

<http://www.ibm.com/servers/storage/support/software/sanvc/planning.html>

16.11 SAN Volume Controller commonly used terms

Here is a short glossary of the most commonly used SAN Volume Controller terms (in alphabetic order).

Boss node	A single node acts as the boss node for overall management of the cluster. If the boss node fails, another node in the cluster will take over the responsibilities.
Configuration node	At any one time, a single node in the cluster is used to manage configuration activity. This configuration node manages a cache of the configuration information that describes the cluster configuration and provides a focal point for configuration commands. Similarly, at any one time, a single node acts as the boss node for overall management of the cluster.
Extent	An extent is a fixed size unit of data that is used to manage the mapping of data between MDisks and VDisks.
Front-end and back-end	SVC takes managed disks and represents these to application servers (hosts). The managed disks are looked after by the “back-end” application of the SAN Volume Controller. The virtual disks presented to hosts are looked after by the “front-end” application in the SVC.
Grain	A grain is the unit of data represented by a single bit in a FlashCopy bitmap, 256 K in SVC.
I/O group	An input/output (I/O) group contains two SVC nodes defined by the configuration process. Each SVC node is associated with exactly one I/O group. The nodes in the I/O group provide access to the VDisks in the I/O group.
LU and LUN	Strictly speaking, there is a difference between a logical unit (LU) and a logical unit number (LUN). A LUN is a unique identifier used on a SCSI bus that enables it to differentiate between up to eight separate devices (each of which is a logical unit). In practice, the two terms are used interchangeably. In this book, when we talk about a LUN, we mean the unit of storage that is defined in a storage subsystem such as an IBM System Storage DS4000, DS6000 and DS8000 Storage Server or storage servers from other vendors.
Managed disk	Managed disk (MDisk) is a SCSI disk presented by a RAID controller and managed by the SAN Volume Controller. The MDisk must not be configured to be visible to host systems on the SAN.
Managed disk group	The managed disk group (MDG) is a collection of MDisks that jointly contain all the data for a specified set of VDisks.
Master console	The master console is the platform which runs the software used to manage the SVC.
Node	Node is the name given to the individual servers in an SVC cluster that run the SVC software.
SAN Volume Controller	The SAN Volume Controller (SVC) is a SAN appliance designed for attachment to a variety of host computer systems, which carries out block level virtualization of disk storage.
Virtual disk	Virtual disk (VDisk) is an SVC device that appears to host systems attached to the SAN as a SCSI disk. Each VDisk is associated with exactly one I/O group.

16.12 More information

For more information, see the following Redbooks publication:

- *Implementing the IBM System Storage SAN Volume Controller V5.1*, SG24-6423
<http://www.redbooks.ibm.com/redbooks/pdfs/sg246423.pdf>

Also see the following websites:

- For the latest SAN Volume Controller product information:
<http://www.ibm.com/servers/storage/software/virtualization/svc/>
- For the full SAN Volume Controller compatibility matrix:
http://www.ibm.com/support/docview.wss?rs=591&context=HW29F&uid=ssg1S1003554&loc=en_US&cs=utf-8&lang=en
- For support for the SAN Volume Controller:
<http://www.ibm.com/servers/storage/support/software/sanvc/>



IBM Tivoli Storage Manager

IBM Tivoli Storage Manager is the number one product of choice for an efficient and effective enterprise wide storage solution. It provides a data protection solution for backup, archiving, disaster recovery planning, space management, database and application protection, bare machine recovery, and record retention. More than 44 operating platforms are supported, using a consistent graphical user interface.

Tivoli Storage Manager provides:

- ▶ Centralized administration for data and storage management
- ▶ Fully automated data protection
- ▶ Efficient management of information growth
- ▶ High-speed automated server recovery
- ▶ Full compatibility with hundreds of storage devices, and local area network (LAN), wide area network (WAN), and storage area network (SAN) infrastructures
- ▶ Optional specifically designed backup solutions for major groupware, enterprise resource planning (ERP) applications, and database products.

Tivoli Storage Manager is the premier storage management solution for mixed platform environments.

17.1 New for IBM Tivoli Storage Manager V6.1.2

New features and enhancements are available in Tivoli Storage Manager Version 6.1.2 and subsequent fix packs, for the server, backup-archive clients, and related products. For information about how to use the features, see the documentation listed in 17.10, “More information” on page 471.

These features and enhancements are summarized in short descriptions, so that you can start thinking about the potential benefits to your storage-management operations:

17.1.1 Server updates

Many features in the Tivoli Storage Manager Version 6.1 server are new for previous Tivoli Storage Manager users. The most significant change is the new database for the server:

- ▶ Server database:

Tivoli Storage Manager version 6.1 provides a new DB2 server database. Advantages include automatic statistics collection and database reorganization, full-function SQL queries, and elimination of the need for offline audits of the database.

Upgrading to V6.1 requires that data in a current Tivoli Storage Manager server database be extracted and then inserted into the new database structure. Tivoli Storage Manager provides utilities to perform the process.

- ▶ Changes to the Version 6.1 Administration Center:

Many features in the Tivoli Storage Manager Administration Center Version 6.1 are new for previous users:

- Updated Integrated Solutions Console:

In V6.1, the Administration Center is hosted by the IBM Integrated Solutions Console (ISC) Advanced Edition Version 7.1. After installation of the Integrated Solutions Console installation completes, open a web browser and enter the following URL, which will display the logon screen for the Integrated Solutions Console: https://local_host:9043/ibm/console. This screen indicates a successful installation of the Integrated Solutions Console.

- WebSphere® Windows service:

In V6.1, the WebSphere Windows service is named TSM Administration Center - TsmAC.

- Identify managing servers:

The table of servers that is the hub of the enterprise-management work page has a column that identifies the managing server, if one exists, for each listed server. By sorting or filtering on the column, you can display the set of servers that are managed by a given server.

- Hover help for table links:

The Administration Center typically displays Tivoli Storage Manager objects in a table. In V6.1, when the cursor hovers over an object image, hover-help text is displayed. The hover help identifies the default action that results when you click the link that is associated with the object.

- Links to information about server messages and Administration Center messages:

When a problem or issue occurs with the server or Administration Center, you are immediately notified and provided with a brief message about the problem or issue. The message number is also provided. In V6.1, you can obtain detailed information

about a message by clicking the link that is associated with the message number. The information is displayed in a new browser window.

- Maintenance script enhancements:

Tivoli Storage Manager utilizes a maintenance script to perform scheduled maintenance tasks. In V6.1, you can generate a maintenance script in one of two styles: *predefined* and *custom*.

A *predefined* maintenance script is one that is generated through a wizard. This script contains standard commands that cannot be altered. A predefined script can only be modified in the wizard.

A *custom* maintenance script is created using the Administration Center maintenance script editor. To have more control of your maintenance tasks, you can modify the commands that you specify. You can also use the editor to update your custom maintenance script.

- Client nodes and backup sets enhancements:

The redesigned Administration Center displays information about backup sets, client nodes, and client-node groups in one portlet. The design includes search functions that you can use to find and display information more quickly. When you select a client node, a summary panel is displayed with the current operation status, server actions, and client-node actions.

- Session and process information available in the health monitor:

The Administration Center health monitor now includes information about server processes and sessions. The information is also available in the properties notebooks for servers.

- Centralized server-connection management:

In V6.1, server-connection tasks, such as adding a server connection, changing a password, and creating a server instance, are consolidated in a single location: the *Manage Servers* work item, located in the ISC navigation tree.

- Changes to management-class activation:

In the V6.1, Tivoli Storage Manager no longer activates changes to existing management classes automatically. You must activate the changes manually. Before the changes take effect, they are validated. Results of the validation are displayed. You or another administrator can review them, and then either confirm or cancel the activation.

- Data deduplication:

Data deduplication is a method of eliminating redundant data in sequential-access disk (FILE) primary, copy, and active-data storage pools. One unique instance of the data is retained on storage media, and redundant data is replaced with a pointer to the unique data copy. The goal of deduplication is to reduce the overall amount of time that is required to retrieve data by letting you store more data on disk, rather than on tape.

Data deduplication in Tivoli Storage Manager is a two-phase process. In the first phase, duplicate data is identified. During the second phase, duplicate data is removed by certain server processes, such as reclamation processing of storage-pool volumes. By default, a duplicate-identification process begins automatically after you define a storage pool for deduplication. (If you specify a duplicate-identification process when you update a storage pool, it also starts automatically.) Because duplication identification requires extra disk I/O and CPU resources, Tivoli Storage Manager lets you control when identification begins as well as the number and duration of processes.

You can deduplicate any type of data except encrypted data. You can deduplicate client backup and archive data, Tivoli Data Protection data, and so on. Tivoli Storage Manager can deduplicate whole files as well as files that are members of an aggregate. You can deduplicate data that has already been stored. No additional backup, archive, or migration is required.

For optimal efficiency when deduplicating, upgrade to the version 6.1 backup-archive client.

Important: You can use the data-deduplication feature with Tivoli Storage Manager Extended Edition only.

Figure 17-1 shows an overview of the deduplication process.

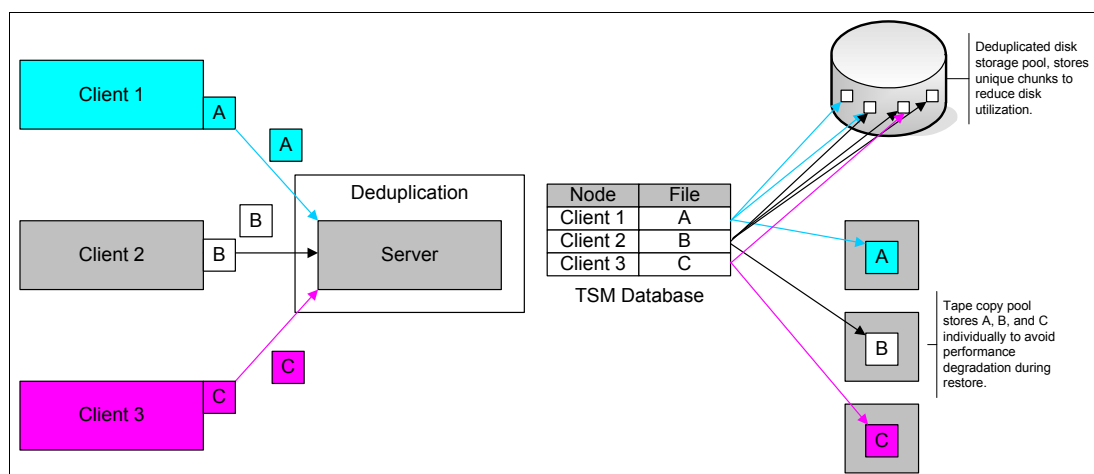


Figure 17-1 Deduplication overview

► Storage devices:

New device support and other changes to storage devices are available in Tivoli Storage Manager Version 6.1:

- ACSLS functionality for Windows systems
- Passthru device driver for HP-UX 11i v2 and v3 on the IA64 architecture
- Support for HP and Quantum DAT160 drives and media
- Support for Sun StorageTek T10000 drives, T10000B drives, and T10000 media.

► Disaster recovery manager support for active-data pools:

To restore your client systems more quickly and efficiently, you can now use active-data pools in your recovery plans and procedures.

► EXPIRE INVENTORY command enhancements:

The EXPIRE INVENTORY command is now enhanced with new functionality.

The additional parameters that you can now use are NODE, DOMAIN, TYPE, DURATION, AND RESOURCE. You can use these parameters to target specific client nodes and domains, and also to determine the type of data to be processed. You can use the RESOURCE parameter to specify the number of parallel processes that you want to run within the single EXPIRE INVENTORY process. You can run up to ten threads at one time, but if you are processing one node, only one thread is utilized.

- ▶ No-query restore changes:
The no-query restore (NQR) function and the internal algorithms responsible for NQR were changed to take advantage of DB2 capabilities and to improve performance.
- ▶ Support for NetApp® SnapMirror to Tape feature:
With Tivoli Storage Manager you can create SnapMirror to Tape images of file systems on NetApp file servers.
- ▶ ODBC driver support:
Tivoli Storage Manager Version 6.1 uses the DB2 open database connectivity (ODBC) driver to query the database and display the results.
- ▶ Backup sets and client node enhancements:
The Administration Center now displays the backup sets, client nodes, and client node groups from one portlet.
- ▶ Enabled functions:
Functions that were disabled in Tivoli Storage Manager V6.1.0 and V6.1.1 are now enabled in Version 6.1.2.

Until Tivoli Storage Manager V6.1.2, a database that contained backup sets or tables of contents (TOCs) cannot be upgraded to V6. These restrictions no longer exist.

In addition, the following commands have been enabled in Version 6.1.2:
 - BACKUP NAS client command if the TOC parameter specifies PREFERRED or YES
 - BACKUP NODE if the TOC parameter specifies PREFERRED or YES
 - DEFINE BACKUPSET
 - GENERATE BACKUPSET
 - GENERATE BACKUPSETTOC.
- ▶ Licensing changes:
Following the release of Tivoli Storage Manager Version 6.1.2, Tivoli Storage Manager Version 6.1.0 will no longer be available for download or purchase. Due to this unique circumstance, certain 6.1.2 packages will be available with a license module. See the following information for details on how this situation affects your environment:
 - Existing Version 6.1.0 and 6.1.1 users:
If you have installed version 6.1.0 and are using a version 6.1.0 license, you can download the 6.1.2 package from the Service FTP site:
<ftp://ftp.software.ibm.com/storage>
 - Version 5 users:
If you have not yet installed a version of the V6.1 server, when you upgrade, you must upgrade directly to version 6.1.2. Version 6.1.2 is available with a license module from Passport Advantage® or from your Tivoli Storage Manager sales representative.
 - New users:
Version 6.1.2 is available from Passport Advantage or from your Tivoli Storage Manager sales representative.
- ▶ TSMDLST utility for AIX:
The tsmdlst utility is now available for AIX beginning in Tivoli Storage Manager fix pack level 6.1.2.
- ▶ Device support for AIX:
The number of devices per driver is extended for AIX. The Tivoli Storage Manager device driver can configure 1024 devices for each driver.

- ▶ **PREVIEW parameter for DSMSEV INSERTDB:**
A PREVIEW parameter is available for the DSMSEV INSERTDB utility in Tivoli Storage Manager fix pack level 6.1.2. The DSMSEV INSERTDB utility is used only as part of the process for upgrading a V5 Tivoli Storage Manager server to V6.1.
- ▶ **SAN discovery support for Linux on zSeries®:**
Support for SAN discovery functions is now available for the Linux on zSeries operating system beginning in Tivoli Storage Manager fix pack level 6.1.2.
- ▶ **New server option SANDISCOVERYTIMEOUT:**
A new server option is available for SAN discovery functions beginning in Tivoli Storage Manager fix pack level 6.1.2.

17.1.2 Reporting and monitoring feature

The reporting and monitoring feature uses a combination of the Tivoli Common Reporting tool, IBM Tivoli Monitoring, and the IBM Tivoli Data Warehouse to offer you reports and real time monitoring information about Tivoli Storage Manager servers and client activity.

17.1.3 Backup-archive client updates

The following features are new for IBM Tivoli Storage Manager in Version 6.1:

- ▶ **Mac OS X-specific information is now included in the UNIX and Linux publication**
The Tivoli Storage Manager for UNIX and Linux Backup-Archive Clients Installation and User's Guide includes the Mac OS X-specific information.
- ▶ **Mac OS X NLS support:**
The Tivoli Storage Manager for Mac OS X Backup-Archive Client now supports the NLS languages supported by the Tivoli Storage Manager.
- ▶ **Mac OS X API support:**
The Tivoli Storage Manager API is now supported on Mac OS X.
- ▶ **Improved memory usage for backup of Hierarchical Storage Management (HSM) managed file systems.**
The Tivoli Storage Manager Backup-Archive client can do a full incremental backup of very large HSM managed file systems containing 100,000,000 files or more.
- ▶ **UTF-8 encoding support is added for Tivoli Storage Manager UNIX and Linux clients:**
UTF-8 is added for the language locales that are already supported by the Tivoli Storage Manager Backup-Archive Client.

The Tivoli Storage Manager Version 6.1 client message catalogs and help files are encoded in UTF-8. If you are installing Tivoli Storage Manager client message catalogs for languages other than English, you must also have the appropriate iconv UTF-8 converters installed on your system. If the appropriate iconv UTF-8 converters are not installed, all Tivoli Storage Manager client messages will be displayed in English.
- ▶ **Availability of 64-bit binaries:**
The client packages for Linux on POWER, Linux zSeries, and one of the AIX clients contain 64-bit binaries.

- ▶ Snapdiff option for NFS data stored on NetApp filers:
When used with the incremental command, the snapdiff option streamlines the incremental process by performing an incremental backup of the files reported as changed by NetApp, instead of scanning the volume looking for files that have changed.
- ▶ Snapdiff option for CIFS data stored on NetApp filers:
When used with the incremental command, the snapdiff option streamlines the incremental process by performing an incremental backup of the files reported as changed by NetApp, instead of scanning the volume looking for files that have changed.
- ▶ Full VM backup and restore support:
The Tivoli Storage Manager **backup vm** command has been enhanced to provide full VM backup capabilities in addition to the file-level backup capabilities previously provided. You can copy full VM virtual disk exports to the backup proxy from snapshots. The entire virtual machine is backed up as a single item, similar to a Tivoli Storage Manager image backup. The exports are broken into 2 GB file chunks that the Tivoli Storage Manager backup-archive client backs up at a file level.
- ▶ Support for restoring Active Directory individual objects:
You can use Active Directory individual object recovery and item-level restore during normal day-to-day operations to recover from accidental corruption or deletion of Active Directory objects by restoring one or more individual Active Directory objects. This feature does not require you to shut down or restart the Active Directory server.
- ▶ The Windows Native GUI has been replaced with the Java GUI:
The Windows GUI is a Java application. The non-Java Windows Native GUI is installed as dsmmfc.exe in the installation directory. However, it has not been updated with the new Tivoli Storage Manager Version 6.1 features.
- ▶ Support for **querysummary** option:
You can use the **querysummary** processing option to extend the **query archive** and **query backup** commands. This support provides a restore preview so that you can determine whether to use the classic or no-query restore method.
- ▶ Support for **srvoptsetencryptiondisabled** option:
You can use the **srvoptsetencryptiondisabled** processing option to ignore encryption options in a client options set from a Tivoli Storage Manager server.
- ▶ Enhanced help facilities:
The command-line client help command is enhanced so that you can specify the command, option, or message on which you want help information.

17.1.4 Application program interface updates

Several features in IBM Tivoli Storage Manager Version 6.1 are new for previous Tivoli Storage Manager users.

The following features are new for Tivoli Storage Manager in Version 6.1:

- ▶ The **dsmUpdateObjEx** function call updates the meta information that is associated with backup or archive objects on the server. This provides the ability to select from several archive objects with the same name
- ▶ Mac OS X support was added in Release 5.5.1
- ▶ Novell Netware is not supported in this release.

17.1.5 Space Management for UNIX and Linux client updates

The following features are new for IBM Tivoli Storage Manager in Version 6.1:

- ▶ GPFS 3.2 storage pool support:
 - Tivoli Storage Manager 6.1 allows multiple General Parallel File System (GPFS) storage pools in one file system. Monitoring a file system includes monitoring each storage pool in the file system.
 - The automigration command, **dmsautomig**, permits automatic migration of storage pools and file systems.
- ▶ Hierarchical Storage Management (HSM) for AIX and Linux for GPFS does not require RSCT fileset for cluster support:

A new responsiveness service function is being added that provides:

- Node response monitoring
- Node failure detection and initiate failover actions
- Event notification processing

This function replaces the requirement of installing the RSCT Group Services. Do not use RSCT fileset for cluster support.

- ▶ Partial file recall enhancements for optimal tape access:

The **dsmrecall** command will recall partial files with the **-OFFset** and **-SIZE** options. With this command, you can specify the portion of a file that is to be recalled.

17.1.6 HSM for Windows updates

Several features in the Hierarchical Storage Management (HSM) for Windows client Version 6.1 are new for previous HSM for Windows client users.

The following features are new for IBM Tivoli Storage Manager HSM for Windows client in Version 6.1:

- ▶ Threshold migration:

Threshold migration provides the capability to automatically maintain a certain amount of free space on protected file systems. No further interaction and manual inspection or monitoring is needed to prevent out-of-space conditions.
- ▶ Support for IPv6:

You can configure a connection with the Tivoli Storage Manager server that uses only IPv4, or a connection that uses either IPv4 or IPv6. In the latter case, the version depends on the network configuration of the file server where HSM for Windows client is installed, and the capabilities of the Tivoli Storage Manager server. The default setting is **commethod v6tcpip**.
- ▶ Support for Microsoft Windows Server 2008:

With the HSM for Windows client, you can use 64-bit versions of Microsoft Windows Server 2008, Standard Edition, and Enterprise Edition.

17.1.7 IBM Tivoli Storage FlashCopy Manager V2.1

IBM Tivoli Storage FlashCopy Manager V2.1 is designed to deliver high levels of data protection for business-critical IBM DB2, SAP, Oracle, Microsoft Exchange, and Microsoft SQL applications through snapshot backup and restore capabilities that are integrated with the applications.

Tivoli Storage FlashCopy Manager uses advanced IBM storage hardware snapshot technology to help create a high performance, low impact application data protection solution. It integrates with IBM System Storage DS8000, IBM SAN Volume Controller, and IBM XIV Storage System products. For Microsoft Windows environments, Tivoli Storage FlashCopy Manager also supports other hardware that is capable of Microsoft Volume Shadow Copy Services (VSS) functions, such as IBM System Storage DS3000, DS4000, and DS5000.

Tivoli Storage FlashCopy Manager can replace the following products:

- ▶ IBM Tivoli Storage Manager for Copy Services
- ▶ IBM Tivoli Storage Manager for Advanced Copy Services.

17.1.8 Data Protection for Snapshot Devices updates

These components are offered as part of Tivoli Storage Manager for Advanced Copy Services, V6.1, Program Number 5608-E10, which is available as a licensed program product, and to all subsequent releases and modifications until otherwise indicated in new editions. They are successors to the following components, respectively:

- ▶ Data Protection for Snapshot Devices for DB2 Advanced Copy Services, V5.5
- ▶ Data Protection for Snapshot Devices for Oracle, V5.4
- ▶ Data Protection for Disk Storage and SAN Volume Controller for Oracle, V5.5.1.

17.1.9 Data Protection for Mail updates

Data Protection for Exchange 6.1 provides the new mailbox restore feature.

With the Data Protection for Exchange 6.1 mailbox restore feature, you can perform individual mailbox recovery and item-level recovery operations in Microsoft Exchange Server 2003 or Microsoft Exchange Server 2007 environments using Data Protection for Exchange backups.

Tip: Mailbox restore tracks and stores mailbox location history, which is used to automate mailbox restore operations. This causes a slight delay before each backup.

Mailbox restore applies to backups that are taken with Data Protection for Exchange:

- ▶ For Exchange Server 2003 environments, mailbox restore applies to Data Protection for Exchange Legacy backups only. For Exchange Server 2003, mailbox restore operations cannot be performed using VSS Backups.
- ▶ For Exchange Server 2007 environments, mailbox restore applies to any Data Protection for Exchange Legacy backups or VSS Backups.
- ▶ Data Protection for Exchange 6.1 (and later) maintains mailbox location history. No mailbox location history is available for backups taken with prior versions. When restoring from these prior version backups, if the mailbox to be restored from has been moved or deleted since the time of the backup, the */mailboxoriglocation* parameter is necessary.

17.1.10 Data Protection for SAP updates

The following new functionality has been added to Version 6.1 of Data Protection for SAP Oracle or DB2:

- ▶ Executable files on Windows platforms (except Java applets) now bear a digital signature.
- ▶ Install Anywhere has replaced Install Shield as the installation vehicle.

- ▶ As of version 7.1, the SAP BR*Tools components have a facility for invoking snapshot (in SAP terminology, volume) backups and restores. Such requests received by TSM for ERP are redirected to the Tivoli Storage Manager for Advanced Copy Services (TSM for ACS) product (if it is installed). To facilitate the interaction of TSM for ACS with TSM for ERP when the user wants to perform a Tivoli Storage Manager backup of the snapshots produced, certain parameters have been added to the TSM for ERP profile for use by TSM for ACS. For more information, see the TSM for ACS documentation.
- ▶ AIX 6.1 is now supported.

Support for 32-bit platforms has been discontinued.

17.1.11 Storage agent updates

With Tivoli Storage Manager for Storage Area Networks Version 6.1, you can use the file-device-sharing software IBM General Parallel File System. You can also continue to use either Tivoli SANergy® or IBM TotalStorage SAN File System. IBM General Parallel File System is the preferred option for the operating systems on which it is supported.

17.1.12 Products and components not updated in V6.1

Products and components that are not updated in V6.1 might be available in earlier versions.

The following products and components are not updated in V6.1:

- ▶ IBM Tivoli Storage Manager server for z/OS operating systems
- ▶ IBM Tivoli Storage Manager for Data Retention for z/OS
- ▶ IBM Tivoli Storage Manager backup-archive client for Novell NetWare
- ▶ IBM Tivoli Storage Manager for Copy Services
- ▶ IBM Tivoli Storage Manager for Mail: Data Protection for Lotus® Domino® for UNIX, Linux, and i5/OS
- ▶ IBM Tivoli Storage Manager for Databases: Data Protection for Oracle
- ▶ IBM Tivoli Storage Manager for Databases: Data Protection for Microsoft SQL Server
- ▶ IBM Tivoli Storage Manager for System Backup and Recovery
- ▶ IBM Tivoli Storage Manager FastBack
- ▶ IBM Tivoli Storage Manager FastBack for Bare Machine Recovery
- ▶ IBM Tivoli Storage Manager FastBack for Microsoft Exchange.

For information about these products, see the website:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v1r1>

17.1.13 Information center updates

The information center is refreshed as required to update technical content and improve the information.

17.2 IBM Tivoli Storage Manager overview

Businesses face a tidal wave of information and data that seems to increase daily. The ability to successfully manage information and data has become imperative. The IBM Tivoli Storage Manager 6.1 family of products help businesses successfully gain better control and efficiently manage the information tidal wave with significant enhancements in multiple facets of data protection.

From its inception, Tivoli Storage Manager has been a highly scalable and available data protection solution. Tivoli Storage Manager 6.1 takes data protection scalability to the next level with a new relational database, based on IBM DB2 technology, designed to store many more objects and manage more data. Greater availability is delivered through enhancements such as online, automated database reorganization. In addition, the increased scalability and the ability to utilize the latest in server technology helps deliver increased performance of backup and recovery processes.

Tivoli Storage Manager 6.1 delivers new near real-time monitoring and operational reporting capabilities.

Tivoli Storage Manager 6.1 helps decrease the rate of storage capacity required to contain data growth with a new built-in data deduplication feature that helps eliminate redundant data. This can enable significantly more backup data to be stored on disk.

For enterprise IT departments, the ability to work around the clock anywhere in the world, can translate to increasingly stringent recovery point objectives and recovery time objectives, and drives the need for more effective business continuity, planning, and execution. Tivoli Storage Manager 6.1 helps address these challenges with several new features and enhancements, including:

- ▶ Individual mailbox and item level recovery for Microsoft Exchange users
- ▶ Up to two times throughput improvement for a single Tivoli Storage Manager server for operations, such as backing up small files from multiple clients
- ▶ Enhancements designed to improve performance for IBM System Storage N series and NetApp Network Attached Storage (NAS)
- ▶ Disaster recovery preparedness enhanced for VMware, IBM System Storage N series, NetApp NAS environments, and Microsoft Windows Active Directory environments.

17.3 IBM Tivoli Storage Manager architecture

This topic provides general information about IBM Tivoli Storage Manager architecture. For more details, see the *Tivoli Storage Manager - Administrator's Guide*.

IBM Tivoli Storage Manager clients are the workstations, file servers, mobile computers, and other machines that must have their data protected. IBM Tivoli Storage Manager client software is installed on these systems.

IBM Tivoli Storage Manager architecture is shown in Figure 17-2.

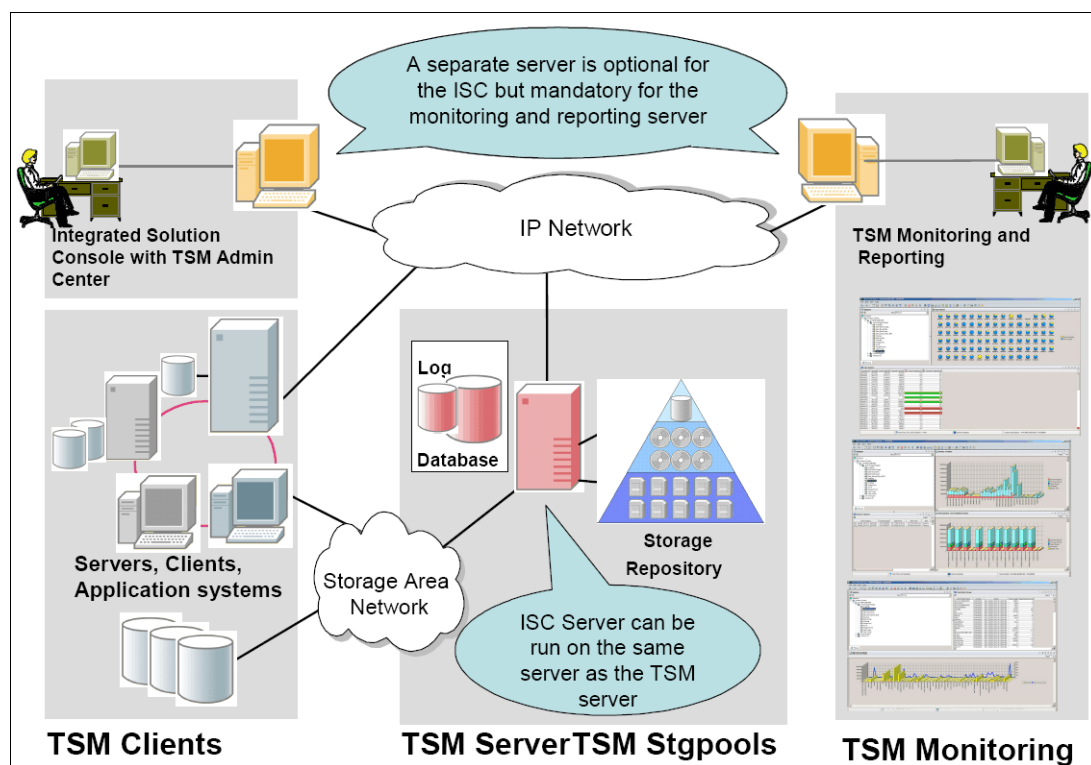


Figure 17-2 IBM Tivoli Storage Manager architecture

Tivoli Storage Manager is based on a DB2 database. The database tracks what is backed up, where it is stored, how long it will be stored for and how many versions of the data are kept and the number and length of time a copy of the file is kept after it is deleted from the original file system.

The relational database allows Tivoli Storage Manager to perform tasks that are not possible when you use a flat file master catalog to track metadata and enables a two-phase commit, which protects the integrity of the database and allows interrupted backups and restores to be restarted. For example, the relational database can:

- ▶ Move data from one type of storage pool to another
- ▶ Retroactively update backed-up data when a policy changes
- ▶ Track individual files
- ▶ Schedule any type of client or administrative process
- ▶ Reclaim expired dead space on tapes.

The Tivoli Storage Manager client sends its data to the Tivoli Storage Manager server either by the LAN or by the SAN. Most backups occur through schedules, but clients can perform on demand backups whenever they want. Clients can also perform their own restores. Tivoli Storage Manager has an improved Administration Center that was introduced in version 5.3 and can now be installed either on the same machine as the Tivoli Storage Manager Server or on a separate machine.

The actual data that the client sends is stored in storage pools. Tivoli Storage Manager is unique in the fact that the storage pools can form a storage hierarchy made up of more than 500 supported devices. This allows for flexibility, longevity, and, most important, fast backups and fast restores.

Most businesses back up their data initially to disk storage. This allows for hundreds of clients to back up at the same time. Then, based on policies, the data migrates in a fashion that expedites restores to tape or optical disk. When the data migrates, all data belonging to one client is moved together to the next pool. By keeping all of that data together, restores are faster because not as much tape positioning is required. This migration process can also accommodate data movement to collocated tapes, which further expedites restores by having data for just one user on them. Collocation can be used to group important functions of your business to ensure rapid recovery in the event of disaster.

The environment can be firewall protected and Tivoli Storage Manager allows individual configuration of nearly every TCP port that it uses for communication:

- ▶ **TCP/IP port:**

To enable the Backup/Archive client, command-line Administrative Client, and the scheduler to run outside a firewall, the port specified by the `tcpport` server option must be opened by the firewall administrator. This port is set on the client and the server using the `tcpport` option. The setting must be the same on the client and server. The default TCP/IP port is 1500.

- ▶ **TCP/IP ports for the remote workstation:**

The two TCP/IP ports for the remote workstation client must be opened. Use the `WEBPORTS` option in the remote workstations option file to specify these ports. If you do not specify the values for the `WEBPORTS` option, the default zero (0) causes TCP/IP to randomly assign two free port numbers.

17.3.1 IBM Tivoli Storage Manager architectural components

The IBM Tivoli Storage Manager consists of several key components. This topic identifies these components and shows how they are related.

For more information, see the documentation listed in 17.10, “More information” on page 471.

IBM Tivoli Storage Manager Server

The principal architectural component of the IBM Tivoli Storage Manager server is its DB2 database. The IBM Tivoli Storage Manager database was especially designed for the task of managing data and implements zero-touch administration.

All policy information, logging, authentication and security, media management, and object inventory are managed through this database.

Most of the fields are externalized through IBM Tivoli Storage Manager high-level administration commands, SQL `SELECT` statements, or, for reporting purposes, by using an ODBC driver. Obviously, this database is fully protected with software mirroring, roll-forward capability, and its own management, online backup and restore functions.

For storing the managed data, the IBM Tivoli Storage Manager server manages a storage repository. The storage repository can be implemented in a hierarchy using any combination of supported media or magnetic or optical disk, tape, and robotic storage devices, which are locally connected to the server system or are accessible through a SAN. To take advantage of SAN technology, the IBM Tivoli Storage Manager server has features implemented that dynamically share SAN-connected, automated tape library systems among multiple IBM Tivoli Storage Manager servers, as well as provide (as an option) local area network, or LAN-free and server-free backup.

Figure 17-3 shows a Tivoli Storage Manager server storage example.

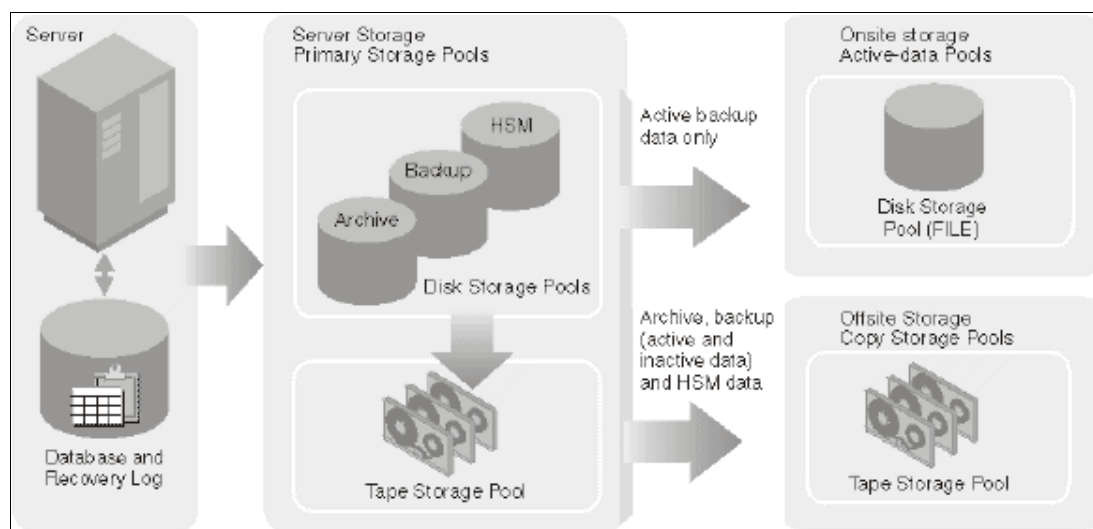


Figure 17-3 Example of server storage

IBM Tivoli Storage Manager Backup/Archive client

Data is sent to the IBM Tivoli Storage Manager server using the IBM Tivoli Storage Manager Backup/Archive client and complementary products. These products work together with the IBM Tivoli Storage Manager Server base product to ensure that any data stored is managed as defined.

The IBM Tivoli Storage Manager Backup/Archive client, included with the server, provides the operational backup and archive functions. The client implements the patented progressive backup methodology, adaptive sub-file backup technology, and unique record retention methods for backup and archive functions.

The Backup/Archive clients are implemented as multi-session clients, which means that they are able to take advantage of the multi-threading capabilities of modern operating systems.

IBM Tivoli Storage Manager for Storage Area Networks

IBM Tivoli Storage Manager for Storage Area Networks allows client systems to write data directly to, or read data directly from, storage devices attached to a SAN. This is called *LAN-free data movement*.

LAN-free data movement makes LAN bandwidth available for other uses and decreases the load on the Tivoli Storage Manager server, allowing it to support a greater number of concurrent client connections.

The key component of Tivoli Storage Manager for Storage Area Networks is the storage agent. You install the storage agent on a client system that shares storage resources with the Tivoli Storage Manager server, as shown in Figure 17-4.

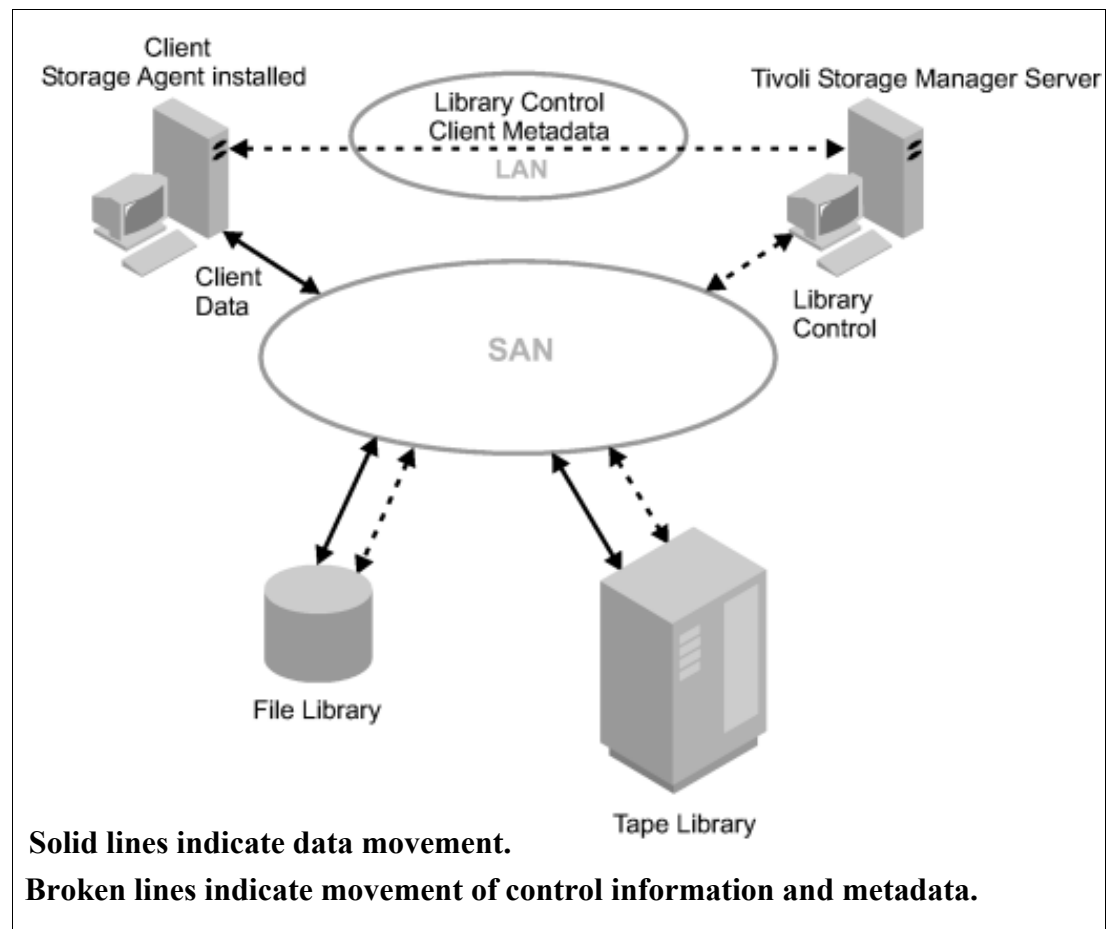


Figure 17-4 SAN data movement

The storage agent can support several clients while installed on only one of the clients. You can also install the storage agent on a client system that does not share storage resources with the Tivoli Storage Manager server, but that is connected to a client system that does share storage resources. The LANFREEMETHOD option allows a client system that shares storage resources to communicate with the storage agent. The LANFREEMETHOD option also allows the storage agent to support several clients while installed on only one of the clients.

Figure 17-5 shows SAN data movement with the LANFREECOMMMETHOD option.

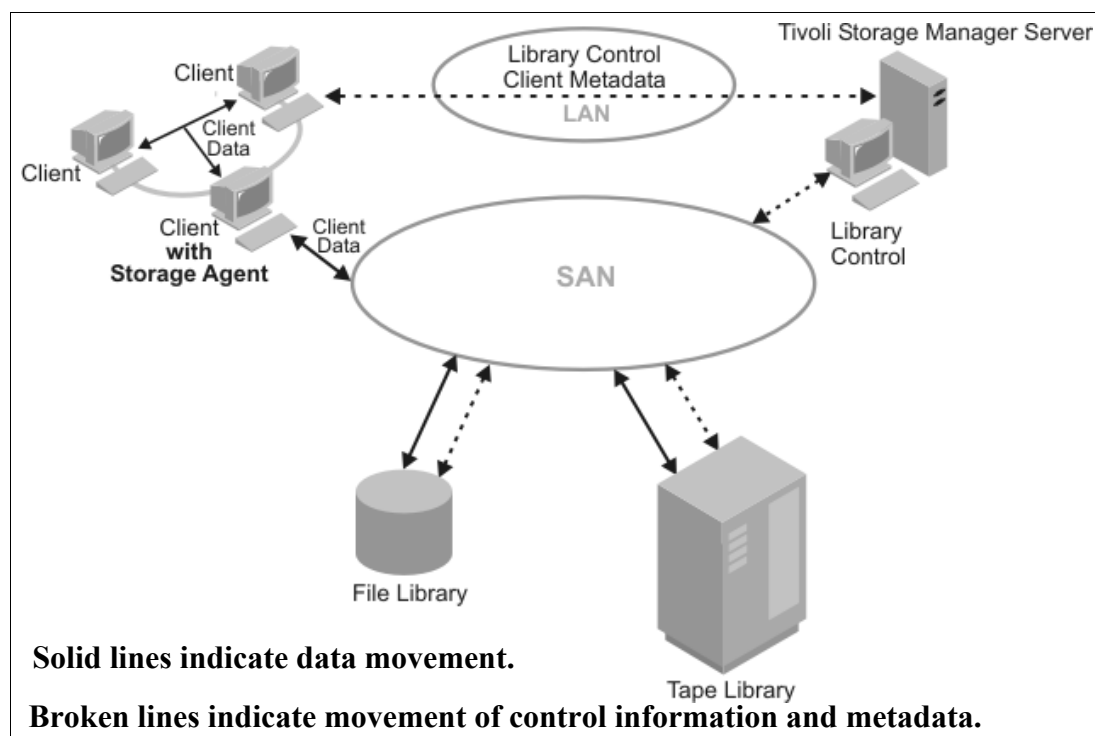


Figure 17-5 SAN data movement with the LANFREECOMMMETHOD option

A Tivoli Storage Manager server, acting as a library manager, controls the storage devices. This server can be the server working in conjunction with the storage agent or another Tivoli Storage Manager server in the enterprise. The Tivoli Storage Manager server keeps track of the metadata that the client has stored. The metadata, such as policy information and file name and size, is passed over the LAN connection between the storage agent and server.

The storage agent communicates with the server to obtain and store database information and to coordinate device and volume access. The server and client coordinate and transfer data access through the SAN. The client uses the storage agent for operations where appropriate. For example, if a SAN path is defined, the client (by means of the storage agent) transfers data on that path. If a failure occurs on the SAN path, failover occurs and the client uses its LAN connection to the Tivoli Storage Manager server and moves the client data over the LAN.

The storage agent can send the data directly to the server using the LAN control paths between the storage agent and the server. An example is a LAN-free storage pool that is updated to read-only after the client connects to the server and obtains its initial policy information. The storage agent, instead of failing the operation, sends the data to the server. If the storage hierarchy is configured so that the Next storage pool destination is available, the server performs the operation.

Tivoli Storage Manager supports SAN-attached device sharing in the following environments:

- ▶ Tivoli Storage Manager native library management support consisting of an ACSLS, SCSI, or 349X library manager and library clients or just a library manager
- ▶ Shared disk storage using a FILE library and the integration of IBM General Parallel File System, Tivoli SANergy, or IBM TotalStorage SAN File System. IBM General Parallel File System is the preferred option for operating systems on which it is supported
- ▶ External libraries.

17.3.2 How client data is stored

Tivoli Storage Manager policies are rules that determine how the client data is stored and managed. The rules include where the data is initially stored, how many backup versions are kept, how long archive copies are kept, and so on.

You can have multiple policies and assign the various policies as needed to specific clients, or even to specific files. Policy assigns a location in server storage where data is initially stored. Server storage is divided into storage pools that are groups of storage volumes.

Server storage can include hard disk, optical, and tape volumes.

Clients use Tivoli Storage Manager to store data for any of the following purposes:

- Backup and restore:

The backup process copies data from client workstations to server storage to ensure against loss of data that is regularly changed. The server retains versions of a file according to policy, and replaces older versions of the file with newer versions. Policy includes the number of versions and the retention time for versions.

A client can restore the most recent version of a file, or can restore earlier versions.

- Archive and retrieve:

The archive process copies data from client workstations to server storage for long-term storage. The process can optionally delete the archived files from the client workstations. The server retains archive copies according to the policy for archive retention time. A client can retrieve an archived copy of a file.

- Instant archive and rapid recovery:

Instant archive is the creation of a complete set of backed-up files for a client. The set of files is called a backup set. A backup set is created on the server from the most recently backed-up files that are already stored in server storage for the client. Policy for the backup set consists of the retention time that you choose when you create the backup set.

You can copy a backup set onto compatible portable media, which can then be taken directly to the client for rapid recovery without the use of a network and without having to communicate with the Tivoli Storage Manager server.

- Migration and recall:

Migration, a function of the Tivoli Storage Manager for Space Management program, frees up client storage space by copying files from workstations to server storage. On the client, the Tivoli Storage Manager for Space Management program replaces the original file with a stub file that points to the original in server storage. Files are recalled to the workstations when needed.

This process is also called Hierarchical Storage Management (HSM). After being configured, the process is transparent to the users. Files are migrated and recalled automatically.

Policy determines when files are considered for automatic migration. On the UNIX or Linux systems that support the Tivoli Storage Manager for Space Management program, policies determine whether files must be backed up to the server before being migrated. Space management is also integrated with backup. If the file to be backed up is already migrated to server storage, the file is backed up from there.

Figure 17-6 shows policy as part of the Tivoli Storage Manager process for storing client data.

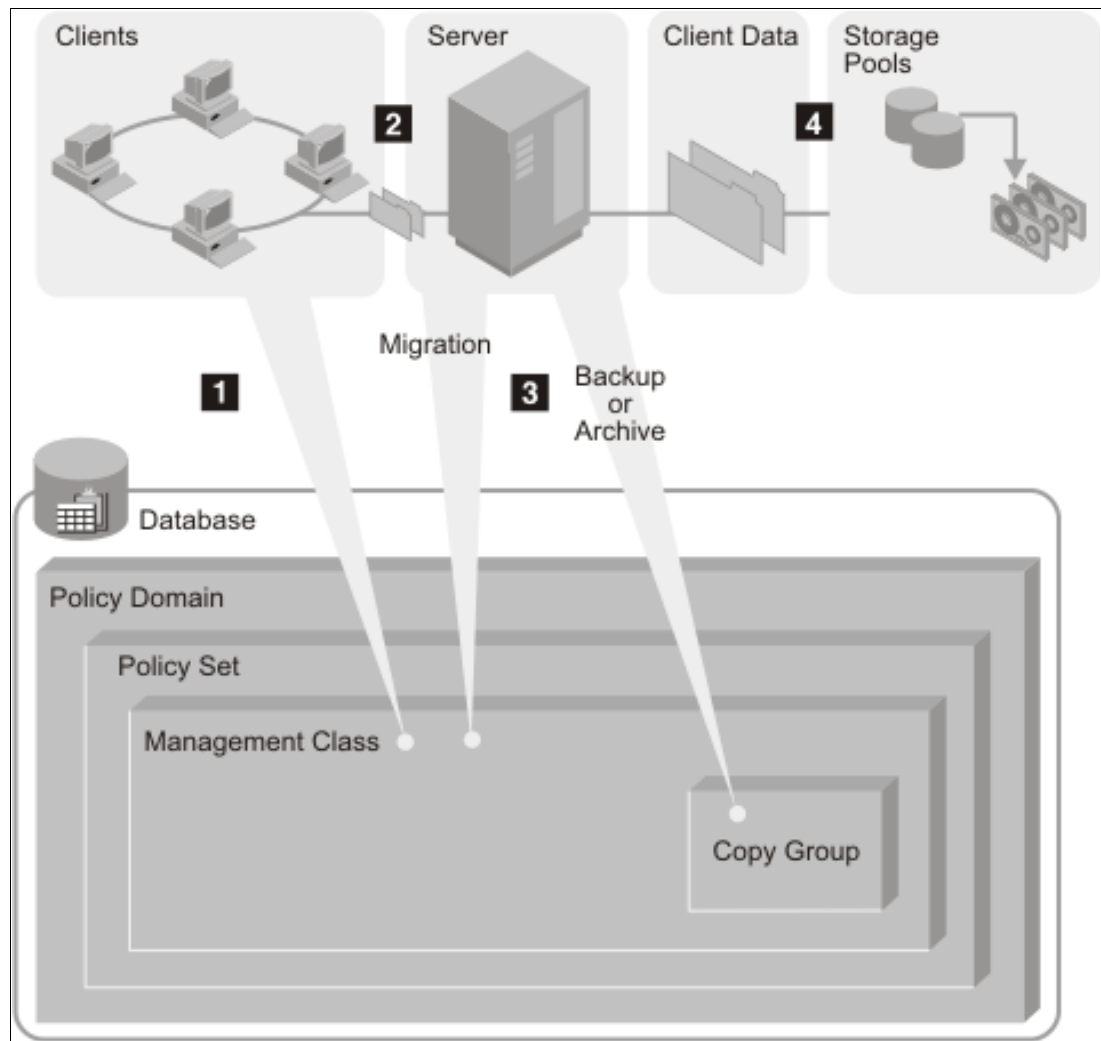


Figure 17-6 How IBM Tivoli Storage Manager Controls Backup, Archive, and Migration Processes

The steps in the process are as follows:

1. A client initiates a backup, archive, or migration operation. The file involved in the operation is bound to a management class. The management class is either the default or one specified for the file in client options (the client's include-exclude list).
2. If the file is a candidate for backup, archive, or migration based on information in the management class, the client sends the file and file information to the server.
3. The server checks the management class that is bound to the file to determine the destination, the name of the Tivoli Storage Manager storage pool where the server initially stores the file. For backed-up and archived files, destinations are assigned in the backup and archive copy groups, which are within management classes. For space-managed files, destinations are assigned in the management class itself.
4. The server stores the file in the storage pool that is identified as the storage destination. The Tivoli Storage Manager server saves information in its database about each file that it backs up, archives, or migrates. If you set up server storage in a hierarchy, Tivoli Storage Manager can later migrate the file to a storage pool other than from the one where the file was initially stored. For example, you might set up server storage so that Tivoli Storage Manager migrates files from a disk storage pool to tape volumes in a tape storage pool.

Files remain in server storage until they expire and expiration processing occurs, or until they are deleted from server storage. A file expires because of criteria that are set in policy. For example, the criteria include the number of versions allowed for a file and the number of days that have elapsed since a file was deleted from the client's file system. If data retention protection is activated, an archive object cannot be inadvertently deleted.

17.3.3 How the server manages storage

Through the server, you manage the devices and media used to store client data. The server integrates the management of storage with the policies that you define for managing client data. For more information, see the documentation listed in 17.10, “More information” on page 471.

Device support

With Tivoli Storage Manager, you can use of a variety of devices for server storage. Tivoli Storage Manager can use direct-attached storage devices as well as network-attached storage devices.

See the current list on the IBM Tivoli Storage Manager website:

<http://www.ibm.com/software/sysmgmt/products/support/IBMTivoliStorageManager.html>

Tivoli Storage Manager represents physical storage devices and media with the following administrator-defined objects:

- ▶ **Library:**
A library is one or more drives (and possibly robotic devices) with similar media mounting requirements.
- ▶ **Drive:**
Each drive represents a drive mechanism in a tape or optical device.
- ▶ **Data mover:**
A data mover represents a device that accepts requests from Tivoli Storage Manager to transfer data on behalf of the server. Data movers transfer data between storage devices.
- ▶ **Path:**
A path represents how a source accesses a destination. For example, the source can be a server, and the destination can be a tape drive. A path defines the one-to-one relationship between a source and a destination. Data might flow from the source to the destination, and back.
- ▶ **Device class:**
Each device is associated with a device class that specifies the device type and how the device manages its media.
- ▶ **Storage pools and volumes:**
A storage pool is a named collection of volumes that have the same media type. A storage pool is associated with a device class. A storage pool volume is associated with a specific storage pool.

Data migration through the storage hierarchy

You can organize the server's storage pools into one or more hierarchical structures. This storage hierarchy allows flexibility in a number of ways. For example, you can set policy to have clients send their backup data to disks for faster backup operations, then later have the server automatically migrate the data to tape.

Storage pool hierarchies

You can arrange storage pools in a storage hierarchies, which consist of at least one primary storage pool to which a client node backs up, archives, or migrates data. Typically, data is stored initially in a disk storage pool for fast client restores, and then moved to a tape-based storage pool, which is slower to access but which has greater capacity. The location of all data objects is automatically tracked within the server database.

You can set up your devices so that the server automatically moves data from one device to another, or one media type to another. The selection can be based on characteristics such as file size or storage capacity. A typical implementation might have a disk storage pool with a subordinate tape storage pool. When a client backs up a file, the server might initially store the file on disk according to the policy for that file. Later, the server might move the file to tape when the disk becomes full. This action by the server is called *migration*. You can also place a size limit on files that are stored on disk, so that large files are stored initially on tape instead of on disk.

For example, your fastest devices are disks, but you do not have enough space on these devices to store all data that needs to be backed up over the long term. You have tape drives, which are slower to access, but have much greater capacity. You define a hierarchy so that files are initially stored on the fast disk volumes in one storage pool. This provides clients with quick response to backup requests and certain recall requests. As the disk storage pool becomes full, the server migrates, or moves, data to volumes in the tape storage pool.

Another option to consider for your storage pool hierarchy is IBM 3592 tape cartridges and drives, which can be configured for an optimal combination of access time and storage capacity.

Migration of files from disk to sequential storage pool volumes is particularly useful because the server migrates all the files for a group of nodes or a single node together. This gives you partial collocation for clients. Migration of files is especially helpful if you decide not to enable collocation for sequential storage pools.

Keeping client files together using collocation

With collocation enabled, the server attempts to keep files belonging to a group of client nodes, a single client node, or client file space on a minimal number of sequential-access storage volumes. Collocation reduces the number of volume mounts required when users restore, retrieve, or recall a large number of files from the storage pool. Collocation thus reduces the amount of time required for these operations.

You can set collocation for each sequential-access storage pool when you define or update the pool.

Figure 17-7 shows an example of collocation by client node with three clients, each having a separate volume containing that client's data.

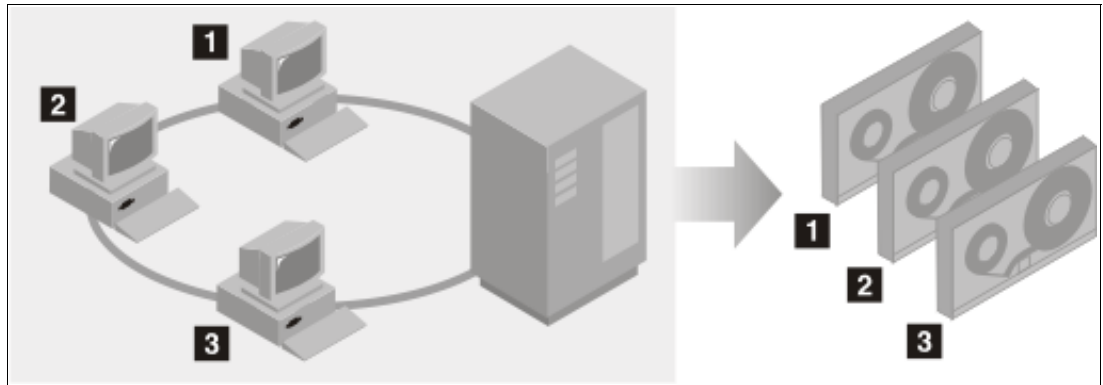


Figure 17-7 Example of collocation enabled (by node)

Figure 17-8 shows an example of collocation by group of client nodes. Three groups have been defined, and the data for each group is stored on separate volumes.

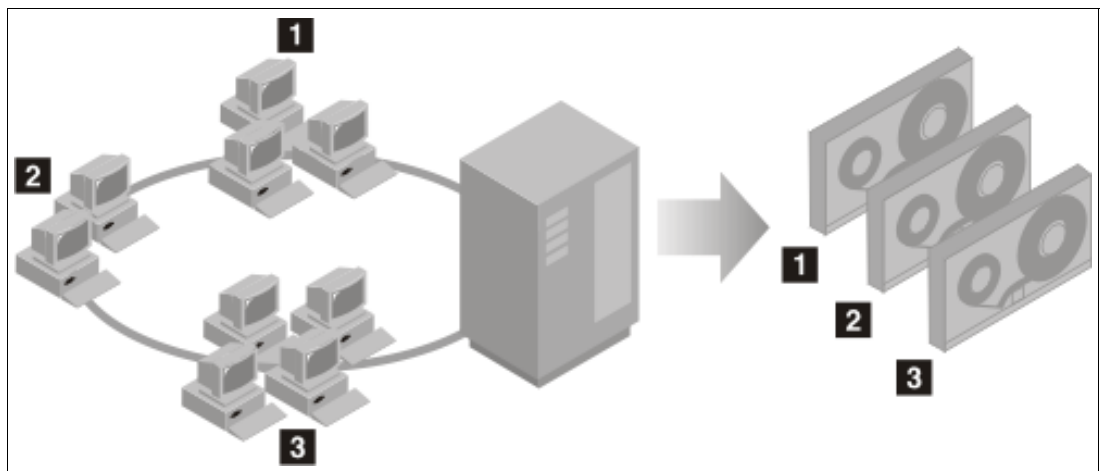


Figure 17-8 Example of collocation enabled (by group)

When collocation is disabled, the server attempts to use all available space on each volume before selecting a new volume. While this process provides better utilization of individual volumes, user files can become scattered across many volumes. Figure 17-9 shows an example of collocation disabled, with three clients sharing space on single volume.

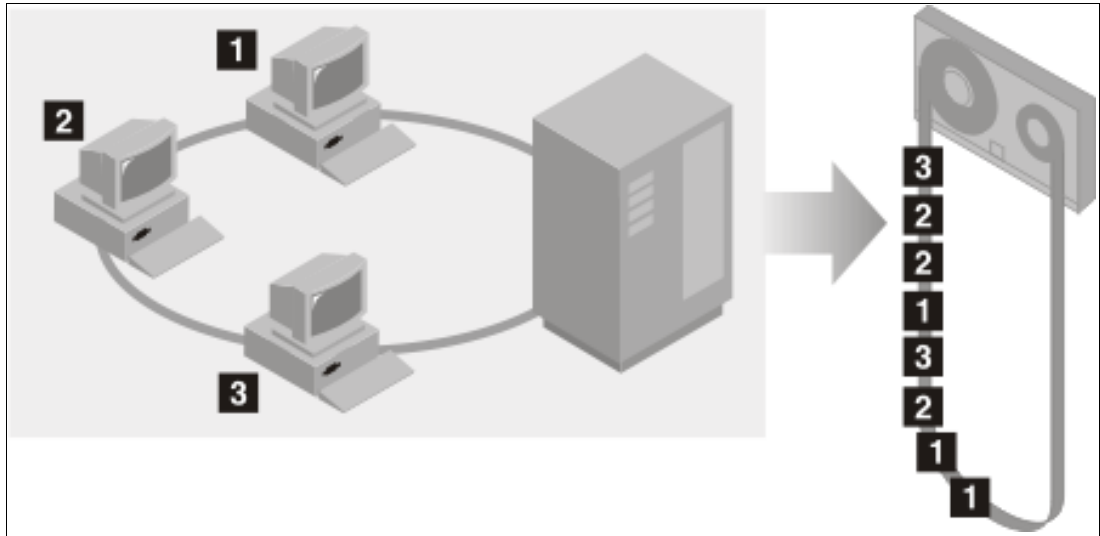


Figure 17-9 Example of collocation disabled

With collocation disabled, more media mount operations might be required to mount volumes when users restore, retrieve, or recall a large number of files.

Collocation by group is the Tivoli Storage Manager system default for primary sequential-access storage pools. The default for copy storage pools and active-data pools is *no collocation*.

Removal of expired data

A policy that you define controls when client data automatically expires from the Tivoli Storage Manager server. The expiration process is how the server implements the policy.

For example, you have a backup policy that specifies that three versions of a file be kept. File A is created on the client, and backed up. Over time, the user changes file A, and three versions of the file are backed up to the server. Then the user changes file A again. When the next incremental backup occurs, a fourth version of file A is stored, and the oldest of the four versions is eligible for expiration.

To remove data that is eligible for expiration, a server expiration process marks data as expired and deletes metadata for the expired data from the database. The space occupied by the expired data is then available for new data.

You control the frequency of the expiration process by using a server option, or you can start the expiration processing by command or scheduled command.

Media reuse by reclamation

As server policies automatically expire data, the media where the data is stored accumulates unused space. The Tivoli Storage Manager server implements a process, called *reclamation*, which allows you to reuse media without traditional tape rotation.

Reclamation is a server process that automatically defragments media by consolidating unexpired data onto other media when the free space on media reaches a defined level. The reclaimed media can then be used again by the server. Reclaiming media allows the automated circulation of media through the storage management process. Use of reclamation can help minimize the number of media that you need to have available.

17.3.4 Tivoli Storage Manager backup and archive concepts

Backup, in Tivoli Storage Manager terms, means creating a copy of a data object to be used for recovery. A Tivoli Storage Manager data object can be a file, a part of a file, a volume image, a directory or a user-defined data object like a database table. The backup version of this data object is stored separately in the Tivoli Storage Manager server storage hierarchy. Tivoli Storage Manager policy tools allow great flexibility for the way data is managed for each client. Backup frequency, retention, and copy policies are easily implemented on the Tivoli Storage Manager client.

In addition to data backup, *archive* copies of data can also be created using Tivoli Storage Manager. Archive creates an additional copy of data and stores it for a specific amount of time, known as the retention period. Tivoli Storage Manager archives are not expired until the retention period is past, even if the original files are deleted from the client system.

Therefore, the difference between *backup* and *archive* is that backup creates and controls multiple backup versions, whereas archive creates an additional file that is retained for a specific period of time.

For details, see the documentation listed in 17.10, “More information” on page 471.

17.3.5 Progressive incremental backups

One of the key differentiators between Tivoli Storage Manager and other data protection products is the *progressive incremental backup methodology*. Tivoli Storage manager only backs up new or changed files. It tracks all of the backups at a file level. It has no concept of a full backup with dependent incrementals or differentials. Because of Tivoli Storage Manager’s powerful relational database, it does not require periodic full backups. Incremental backup by date is also available. This methodology reduces network and storage resource consumption and lowers the overall cost of storage management. Tivoli Storage Manager’s file level progressive backup methodology is far superior to other traditional backup methods such as Full+Incremental or Full+Differential, because progressive incremental backups are never redundant.

For more information, see the documentation listed in 17.10, “More information” on page 471.

17.3.6 Tivoli Storage Manager reporting and monitoring

The IBM Tivoli Storage Manager reporting and monitoring feature uses a combination of reporting and monitoring components to offer you historical reports and real-time monitoring information for the IBM Tivoli Storage Manager servers and clients.

You can view the historical reports to see if there are any issues that need attention, such as uncontrolled growth over time. You can also view workspaces that are being monitored to see the Tivoli Storage Manager server IDs, database size, agent status, client node status, scheduled events, and so on.

The reporting component, sometimes referred to as Tivoli Common Reporting, reports on the retrieved historical data. IBM Tivoli Monitoring acts as a monitoring application that provides workspaces for you to monitor real-time information.

The Tivoli Storage Manager monitoring agent communicates with the Tivoli Storage Manager reporting and monitoring server to retrieve data from its database and return this data to the Tivoli Monitoring server. The Tivoli Storage Manager monitoring agent communicates with the Tivoli Storage Manager reporting and monitoring server to retrieve data from its database and return this data to the Tivoli Monitoring server.

Important: You can view reporting and monitoring information for Solaris and HP-UX systems, but only on a Windows, AIX, or Linux IBM Tivoli Monitoring server.

The monitoring server stores this data in the Tivoli Data Warehouse.

The Tivoli Storage Manager reporting and monitoring feature uses the following components:

► Tivoli Monitoring:

Consists of a number of components that accumulate and monitor historical data for reporting:

- IBM Tivoli Enterprise Portal
- IBM Tivoli Enterprise Management Server
- Tivoli Data Warehouse.

► IBM DB2:

Stores historical data that is obtained from Tivoli Storage Manager servers that are monitored using IBM Tivoli Monitoring.

► Tivoli Storage Manager monitoring agent:

Queries and formats data to be presented to you in the following ways:

- As workspaces using the Tivoli Enterprise Portal
- As reports using the Tivoli Data Warehouse and the reporting portion of the Tivoli Storage Manager reporting and monitoring feature

The agent is installed on the Tivoli Storage Manager server or the Tivoli Monitoring server, and is a multi-instance data collection agent.

► Tivoli Storage Manager reporting and monitoring reporting infrastructure:

Reports on the Tivoli Storage Manager server activities from data that is collected using the Tivoli Storage Manager monitoring agent. The monitoring feature uses the Tivoli Enterprise Portal to view the current status of the Tivoli Storage Manager server.

Figure 17-10 shows how the data flows between the various components.

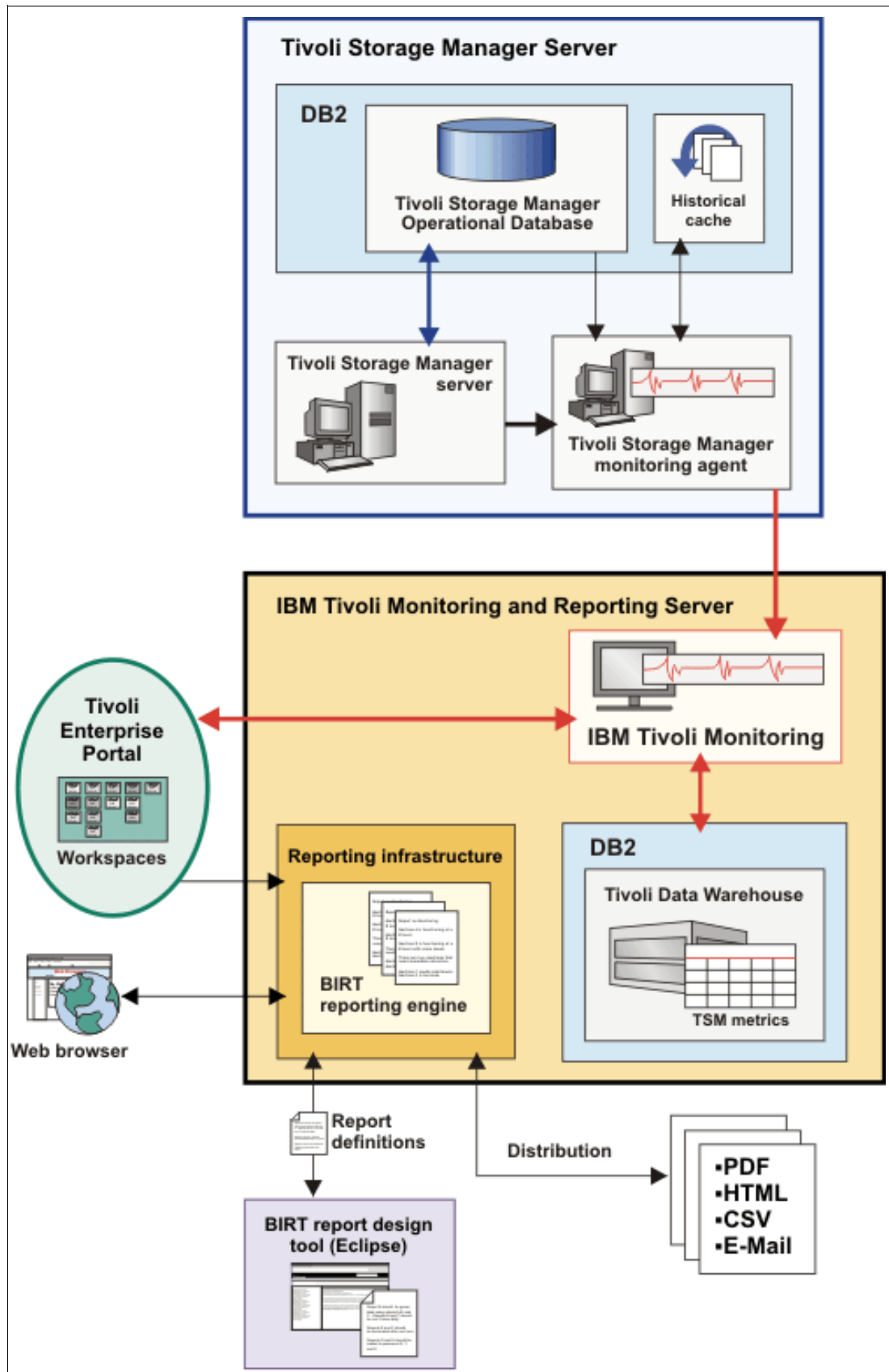


Figure 17-10 Tivoli Storage Manager reporting and monitoring feature infrastructure and data flow

If you are interested in creating your own custom reports, you are required to install the Business Intelligence and Reporting Tools (BIRT) software. The Tivoli Storage Manager reporting and monitoring feature does not have an option to create custom reports using Tivoli Storage Manager.

For more information, see the documentation listed in 17.10, “More information” on page 471.

17.3.7 Tivoli Storage Manager Administration Center

The Tivoli Storage Manager Administration Center is a Web-based interface for centrally configuring and managing IBM Tivoli Storage Manager Version server V5.3 and later, including V6.1. This topic provides general information about Tivoli Storage Manager Administration Center.

For details, see the documentation listed in 17.10, “More information” on page 471.

The Administration Center is a task-oriented interface that replaces the previous administrative web interface. The Administration Center provides wizards to help guide you through common configuration tasks. Properties notebooks allow you to modify settings and perform advanced management tasks.

Important: The Tivoli Storage Manager V5.3 server and Administration Center are out of regular service. To obtain technical support for the V5.3 server and Administration Center, you need an extended-support contract.

To administer a V6.1 server using the Administration Center, the Administration Center must be V6.1.

Key features of the Administration Center

Here we list the key features of the Administration Center:

- ▶ You only need to log in once to access multiple Tivoli Storage Manager servers from a single interface.
- ▶ You can easily monitor the health of your storage environment. Regular status updates are provided for:
 - Scheduled events
 - The server database and recovery log in server versions 5.3, 5.4 and 5.5 (using rules based on best practices)
 - The database manager in server version 6.1
 - Storage devices, including information about off-line drives and paths, and mounted volumes.
- ▶ You can filter and sort storage objects, such as client nodes and library volumes.
- ▶ You can use wizards to more easily perform complex tasks, such as:
 - Creating schedules to perform client node and administrative operations.
 - Creating a server maintenance script to perform database and storage pool backup, migration, expiration, and reclamation.
 - Configuring storage devices. A comprehensive wizard helps you create a library, add drives, check in media volumes, and create storage pools.
 - Configuring a version 6.1 server on the local or a remote UNIX system.

IBM Integrated Solutions Console infrastructure

The Integrated Solutions Console (ISC), is a component framework that allows you to install components provided by multiple IBM applications, and access them from a single web interface. The Administration Center is installed as an Integrated Solutions Console component.

The Administration Center provides several other features that can help you monitor and manage your storage management environment:

- ▶ To centrally manage multiple Tivoli Storage Manager servers from a single server
- ▶ To monitor Tivoli Storage Manager server status
- ▶ To generate usage and security reports for your Tivoli Storage Manager servers.

Attention: The V6.1 Administration Center can only be used with V5.4, V5.5, and V6.1 servers. The Tivoli Storage Manager Version 5.3 server is out of regular service. To obtain technical support for the V5.3 server, you need an extended-support contract.

17.3.8 Tivoli Storage Manager API

The Tivoli Storage Manager application program interface (API) enables an application client to use storage management functions.

The API includes function calls you can use in an application to perform these operations:

- ▶ Start or end a session
- ▶ Assign management classes to objects before they are stored on a server
- ▶ Back up or archive objects to a server
- ▶ Restore or retrieve objects from a server
- ▶ Query the server for information about stored objects
- ▶ Manage file spaces
- ▶ Send retention events.

For information about installing the API, see the documentation listed in 17.10, “More information” on page 471.

17.3.9 Tivoli Storage Manager security

Security is a vital aspect for enterprise storage management. Data must be protected, available, and secure. From the moment data is backed up from the client, IBM Tivoli Storage Manager provides a secure storage management environment. Tivoli Storage Manager is the only interface to your backup and archive data.

Before a communication session between the Tivoli Storage Manager client and the Tivoli Storage Manager server begins, an authentication handshaking process occurs with authentication tickets and a mutual suspicion algorithm. The Tivoli Storage Manager security protocol is modeled after the Kerberos network authentication protocol, which is a highly respected method for secure sign-on cryptography. The client uses its password as part of an encryption key, and does not send the password over the network. Each session key is unique, so replaying a session stream will not result in a sign-on to the Tivoli Storage Manager server. This significantly lowers the chance of a Tivoli Storage Manager session being hijacked by an outside user.

To heighten security for Tivoli Storage Manager sessions, data sent to the Tivoli Storage Manager server during backup and archive operations can be encrypted with standard DES 56-bit encryption. For WAN implementations of Tivoli Storage Manager across public networks, data encryption complements and completes data security for Tivoli Storage Manager.

IBM Tivoli Storage Manager has enhanced support for environments with firewalls in which communication originating from outside the firewall is to be restricted. Clients normally contact the server but with the new firewall support, you can choose to restrict session initiation to the server (sessioninitiation=serveronly). Scheduled backup-archive client operations can be restricted to server-initiated sessions.

For details, see the documentation listed in 17.10, “More information” on page 471.

17.4 IBM Tivoli Storage Manager family of products

This topic provides information about Tivoli Storage Manager family of products. For more information about Tivoli Storage Manager family of products, see the documentation listed in 17.10, “More information” on page 471.

17.4.1 IBM Tivoli Storage Manager Basic or Standard Edition

With the advent of Version 6.1, there is no longer an Express Edition, so there is no longer a Basic or Standard Edition. The standard version of IBM Tivoli Storage Manager contains a rich set of features and provides the core functions of backup, recovery, and archive management:

- ▶ **Progressive backup methodology:**
Saves time and storage space by backing up only new files and modified files. The progressive backup feature uses Tivoli Storage Manager's own relational database to track data wherever it is stored, delivering a direct one-step file restore. Progressive backup eliminates the requirement for traditional full-plus-incremental or full-plus-differential backup and restore procedures, commonly used by other storage management products.
- ▶ **Tape resource sharing:**
Enables multiple Tivoli Storage Manager servers to use the same tape library and drives, optimizing tape hardware asset utilization.
- ▶ **Network-free rapid recovery:**
Supports high-speed client data recovery directly from tape or optical devices. Recovery time is minimized by eliminating the use of network and central server resources.
- ▶ **Dynamic multithreaded transfer:**
Permits multiple clients to simultaneously transfer data to and from the same Tivoli Storage Manager server. Performance is boosted to more than three times the rate of a single-threaded session. The higher speed is achieved by transparently optimizing the number of data transfer sessions, based on available system resources.
- ▶ **Adaptive differencing technology:**
Changes the way data is backed up from the client. Using adaptive differencing, data is transferred to the server either by byte, block, or file level, based on the size of the file being backed up, and the portion of the file that has changed since the last backup. Adaptive differencing technology supports all connectivity strategies, including LANs, WANs, SANs, internet, and dial-up connections. Adaptive differencing was initially designed with mobile computer users in mind. However, other users with a requirement to minimize data transmitted over the network can also benefit from the technology.

- ▶ **Enterprise administration:**

Simplifies centralized control across multiple Tivoli Storage Manager implementations without sacrificing network performance. Tivoli Storage Manager 6.1 employs the Integrated Solutions Console (ISC), which provides a task-based GUI interface to Tivoli Storage Manager administrative tasks.

- ▶ **Clustering:**

Tivoli Storage Manager includes enhanced support for IBM High Availability Cluster Multi-Processing (HACMP™), Microsoft Cluster Server (MSCS), Novell Cluster Services (NCS), and VERITAS Cluster Services (VCS) on Windows.

Tivoli Storage Manager 6.1 has improved the support for Small Computer System Interface (SCSI) and fibre-attached tape device failover on Windows and UNIX, and support for Storage Agents, Library Managers, and Library Clients as cluster members.

- ▶ **LAN-free data transfer:**

An optional module for Tivoli Storage Manager effectively exploits SAN environments by moving data transfers from the communication network to a SAN. Communication bandwidth availability is therefore improved, increasing service levels for users and customers.

- ▶ **Hierarchical Storage Management:**

An optional module for Tivoli Storage Manager automatically and transparently moves unused data files from online disk storage to offline tape storage. In the event that a file is accessed after it has been moved to offline storage, Tivoli Storage Manager transparently recalls the file.

- ▶ **Library and device support:**

Tivoli Storage Manager supports libraries with up to three tape drives and up to 40 cartridge capacity. Larger libraries can be accommodated, but with only three devices and 40 slots enabled.

You can find more information about IBM Tivoli Storage Manager at the website:

<http://www.ibm.com/software/tivoli/products/storage-mgr/>

17.4.2 IBM Tivoli Storage Manager Extended Edition

The IBM Tivoli Storage Manager Extended Edition expands on the features and possibilities of the standard version described in the previous section.

Tivoli Storage Manager Extended Edition adds disaster recovery planning capability for the server, Network Data Management Protocol (NDMP) control for network-attached storage (NAS) filers, and support for larger capacity tape libraries and more tape drives.

You can find more information at:

<http://www.ibm.com/software/tivoli/products/storage-mgr-extended/>

Disaster Recovery Manager

The Disaster Recovery Manager (DRM) component of Tivoli Storage Manager Extended Edition provides disaster recovery for the Tivoli Storage Manager server and assists with disaster recovery for clients.

DRM offers various options to configure, control, and automatically generate a disaster recovery plan (DRP) file. The plan contains the information, scripts, and procedures required to automate restoration and help ensure quick recovery of data after a disaster. The scripts contain the commands necessary to rebuild the Tivoli Storage Manager server.

One of the key features of Tivoli Storage Manager and DRM is the ability to track media in all possible states, such as on-site, in transit, or in a vault. The media movement features of DRM assist greatly with the daily tasks of sending disaster recovery media off-site, and receiving expired media on-site for reuse. With these features, the system administrator can quickly locate all available copies of data.

DRM functions help maintain business continuity by:

- ▶ Establishing and helping to automate a thorough server DRP; clients can then subsequently restore their data from the server if required, and can continue their daily backup procedures.
- ▶ Ensuring that vital site-specific information is available in the same plan.
- ▶ Automating vital recovery steps to return the Tivoli Storage Manager server and backup environment to normal operation.
- ▶ Managing and identifying off-site media required for recovery.
- ▶ Tracking and reporting destroyed systems in the event of a disaster.
- ▶ Storing client configuration information and assigning client recovery priorities.

With DRM, you can recover at an alternate site, on a replacement system with another hardware configuration, and with people who are not familiar with the applications. The DRP can be periodically tested to certify the recoverability of the server. The DRP can, and must, be recreated easily every day so that it stays up to date.

During a disaster, these are commonly encountered errors:

- ▶ A DRP does not exist
- ▶ The DRP was not tested, or if it was, it is now out of date
- ▶ The testing team's skills were not sufficient to perform and evaluate testing
- ▶ The disk volume definitions for the recovery site are not known
- ▶ The location of recovery tapes is not known
- ▶ It is not known which tapes are to be applied first.

DRM keeps track of all the vital information required to rebuild the Tivoli Storage Manager environment, such as:

- ▶ The current server configuration information and its location
- ▶ The current Tivoli Storage Manager server database volumes (size, location, and number)
- ▶ The recovery sequence
- ▶ The currency of the DRP
- ▶ The server and client machines configurations
- ▶ The people to be contacted in the event of a disaster
- ▶ The location of the recovery media and the organization or persons responsible
- ▶ The point in time (PIT) to which the environment can be restored.

During recovery from a disaster, DRM automates the following procedures to restore the Tivoli Storage Manager servers:

- ▶ Restore Tivoli Storage Manager server's key option files
- ▶ Copy files from alternate locations to production locations
- ▶ Initialize the Tivoli Storage Manager database and log volumes
- ▶ Match sizes and locations of the Tivoli Storage Manager database and log volumes
- ▶ Automatically launch restoration of the Tivoli Storage Manager database
- ▶ Track media required and availability
- ▶ Register installed Tivoli Storage Manager server features and return the server state to a valid license configuration
- ▶ Update Tivoli Storage Manager volume catalog information, including whether volumes have been destroyed during the disaster
- ▶ Rebuild Tivoli Storage Manager hierarchical storage configuration

- Restore destroyed volumes from those available where possible
- Recreate customer backup environment.

A detailed description, recovery scenario, and recovery plan built with DRM can be found in *Disaster Recovery Strategies with Tivoli Storage Management*, SG24-6844. Also, guidelines and examples of using DRM to store client machine information in the DRM plan file for use during a client disaster recovery are given in the same book.

In summary, DRM will systematically rebuild the Tivoli Storage Manager server environment and ensure that current application data for the entire enterprise is available for recovery. This can all be done automatically from a single scripted command.

NDMP support for Network Attached Storage

For NAS devices, Tivoli Storage Manager Extended Edition uses NDMP to perform high-performance, scalable backups and restores. NDMP-based backups and restores minimize network traffic and transfer data outboard of the Tivoli Storage Manager client and server. NDMP enables a full and differential file system image backup and restore of Network Appliance™ file servers with OS Data ONTAP V7.1 or higher, and EMC Celerra systems. Multiple backup and restore operations can be performed simultaneously. General NDMP support also allows other NAS vendors to certify integration with Tivoli Storage Manager.

The NDMP backup and restore features are fully integrated with Tivoli Storage Manager Extended Edition server and client. No extra software is required on the server, client, or NAS appliance. When doing backups and restores, the NAS device and the Tivoli Storage Manager server and client all have specific roles, as shown in Figure 17-11.

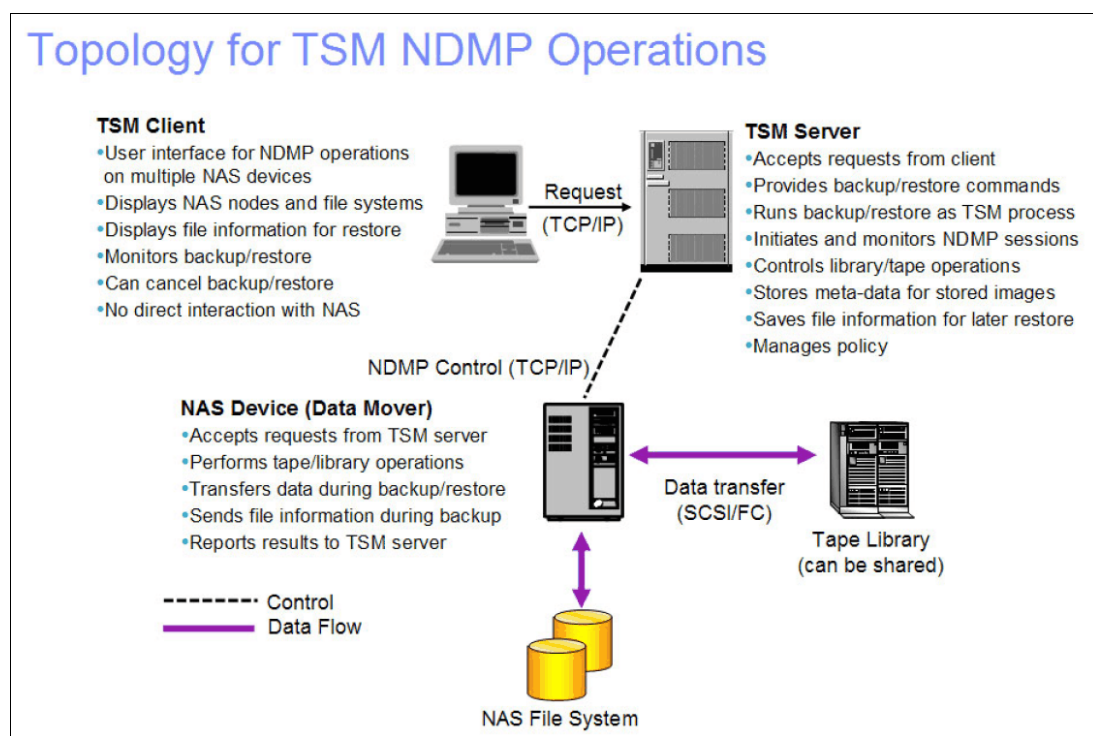


Figure 17-11 Topology for NDMP using IBM Tivoli Storage Manager

Tivoli Storage Manager Extended Edition offers the ability to do file-level and full/differential file system image backups and restore of servers that support the NDMP protocol. You can back up directly to the Tivoli Storage Manager hierarchy and also implement DRM, as it now supports NAS storage. Multiple backup and restore operations can be performed in parallel.

During backup and restore operations, data flows directly between the tape drive and the NAS appliance. NDMP for NAS backup uses either an SCSI-attached tape device local to the NAS appliance, or a SAN-attached SCSI or Automated Cartridge System Library Software (ACSL) device that can be shared with the Tivoli Storage Manager server. Library robotics can be controlled directly by the Tivoli Storage Manager server or by passing SCSI commands through an NAS file server.

Drives must be supported by both the NAS appliance and the NAS OS. Drives can be dedicated to NDMP operations from a single NAS file server or can be shared. Multiple NAS appliances can share SAN-attached shared tape resources if backups are performed through the same Tivoli Storage Manager server. Drives can be also shared with LAN-free backup/restore operations, provided that the library is controlled directly by the Tivoli Storage Manager server.

Extended library and drive support

Tivoli Storage Manager Extended Edition supports larger tape libraries, thus removing the 48-cartridge limit for library capacity, and allowing more than four tape drives within a single library.

Database and application online protection

This feature protects a wide range of application data through the protection of the underlying databases and application management systems holding that data. This module automates data protection tasks and allows database and application servers to continue running their primary applications while they back up and restore data to and from offline storage.

Data retention

IBM System Storage Archive Manager, previously known as Tivoli Storage Manager for Data Retention, helps manage and simplify the retrieval of the ever-increasing amount of data that organizations must retain for strict records retention regulations. Many regulations demand archiving of records, emails, design documents, and other data for many years, in addition to requiring that the data is not changed or deleted.

Bare Machine Recovery

This feature backs up and automatically restores the operating system structures required to rebuild the operating system, as well as data files. It schedules regular operating system backups so that a recovery brings back the latest information.

17.5 IBM Tivoli Storage Manager complementary products

IBM Tivoli Storage Manager can be integrated with several optional applications that together form a powerful integrated storage management solution:

- ▶ IBM Tivoli Storage Manager for Mail
- ▶ IBM Tivoli Storage Manager for Databases
- ▶ IBM Tivoli Storage Manager for Microsoft SharePoint
- ▶ IBM Tivoli Storage Manager HSM for Windows
- ▶ IBM Tivoli Storage Manager for Advanced Copy Services
- ▶ IBM Tivoli Storage Manager for Copy Services
- ▶ IBM Tivoli Storage Manager for Enterprise Resource Planning
- ▶ IBM Tivoli Storage Manager for Space Management
- ▶ IBM Tivoli Storage Manager for Storage Area Networks
- ▶ IBM Tivoli Storage Manager for System Backup and Recovery
- ▶ IBM Tivoli Storage Manager FastBack

- ▶ IBM Tivoli Storage Manager FastBack for Microsoft Exchange
- ▶ IBM Tivoli Storage Manager FastBack for Bare Machine Recovery
- ▶ IBM Tivoli Continuous Data Protection for Files.

17.5.1 IBM Tivoli Storage Manager for Mail

IBM Tivoli Storage Manager for Mail is a software module for IBM Tivoli Storage Manager that automates the data protection of email servers running either Lotus Domino or Microsoft Exchange. This module utilizes the application programming interfaces (APIs) provided by email application vendors to perform online backups without shutting down the email server and improve data-restore performance. As a result, it can help protect the growing amount of new and changing data that must be securely backed up to help maintain mail server application availability 24x7, 365 days a year. The two products available are:

- ▶ IBM Tivoli Storage Manager for Mail: Data Protection for Microsoft Exchange Server 6.1
- ▶ IBM Tivoli Storage Manager for Mail: Data Protection for Lotus Domino 5.5.

For more information, see *Tivoli Storage Manager for Mail - Data Protection for Microsoft Exchange Server, Installation and User's Guide*, SC23-9796-00.

17.5.2 IBM Tivoli Storage Manager for Databases

IBM Tivoli Storage Manager for Databases is a software module designed to work with IBM Tivoli Storage Manager to protect a wide range of application data through the protection of the underlying database management systems holding that data. IBM Tivoli Storage Manager for Databases exploits the various backup-certified utilities and interfaces provided for Oracle using RMAN functionality and Microsoft SQL Server (with or without VSS.)

This same functionality is included in the IBM DB2 Universal Database™ package and Informix® Dynamic Server, using ONBAR enabling them to work directly with IBM Tivoli Storage Manager without the need to buy any additional modules.

For more information, see the website:

<http://www.ibm.com/software/tivoli/products/storage-mgr-db/>

17.5.3 IBM Tivoli Storage Manager for Microsoft SharePoint

IBM Tivoli Storage Manager 6 for Microsoft SharePoint is a policy-based backup and recovery solution. It allows you to restore of your Microsoft SharePoint business data and content after almost any kind of business interruption.

It provides the following functionality:

- ▶ Performs backup and recovery of Microsoft SharePoint Portal 2003 and Microsoft Office SharePoint Server 2007 environments
- ▶ Restores portals, top level sites, subsites and individual document libraries, attachments, lists, folders, areas, and sub areas
- ▶ Schedules full, incremental, or differential backup at the site-level, subsite-level and item-level
- ▶ Preserves all meta-data versions
- ▶ Integrates with the Tivoli Storage Manager server so that you can create synchronous or asynchronous copies of SharePoint data for off-site protection
- ▶ Includes an easy-to-use browser-based graphical user interface (GUI).

For more information, see the website:

<http://www.ibm.com/software/tivoli/products/storage-mgr-sharepoint/>

17.5.4 IBM Tivoli Storage Manager HSM for Windows

IBM Tivoli Storage Manager for HSM for Windows provides Hierarchical Storage Management functionality to the Windows platform. As with IBM Tivoli Storage Manager for Space Management, HSM for Windows automatically migrates rarely-accessed files to alternative storage, without disrupting the most frequently used files in local Windows file systems. Similarly, migrated files are automatically and transparently recalled to their original location when required by applications or users.

HSM for Windows allows various levels of granularity for migration of files. Files can be migrated individually, and file systems can be partially or fully migrated, based on a comprehensive set of policy options.

IBM Tivoli Storage Manager for Space Management complements both IBM Tivoli Storage Manager and IBM Tivoli Storage Manager Extended Edition.

For more information, see the website:

<http://www.ibm.com/software/tivoli/products/storage-mgr-hsm/>

17.5.5 IBM Tivoli Storage Manager for Advanced Copy Services

IBM Tivoli Storage Manager for Advanced Copy Services (formerly known as IBM Tivoli Storage Manager for Hardware) is an optional software module for AIX that integrates with Tivoli Storage Manager Extended Edition. Tivoli Storage Manager for Advanced Copy Services protects mission-critical data that must be available 24x7, and integrates hardware- and software-based snapshot capabilities with Tivoli Storage Manager and its Data Protection components for DB2 UDB, Oracle, and mySAP.

Tivoli Storage Manager for Advanced Copy Services supports a wide range of hardware:

- ▶ IBM Enterprise Storage Server (ESS)
- ▶ IBM DS6000
- ▶ IBM DS8000
- ▶ SAN Volume Controller (SVC) and all IBM and non-IBM devices supported by the SVC.

For a complete list, see the website:

<http://www.ibm.com/systems/storage/software/virtualization/svc/interop.html>

Tivoli Storage Manager for Advanced Copy Services also provides the following functionality:

- ▶ IBM FlashCopy support for ESS for Oracle
- ▶ FlashCopy support for ESS for DB2
- ▶ FlashCopy support for ESS for mySAP on DB2 UDB
- ▶ FlashCopy support for ESS for mySAP on Oracle
- ▶ Snapshot support for DS8000, DS6000, and SVC for DB2 UDB
- ▶ Snapshot support for DS8000, DS6000, and SVC for Oracle
- ▶ Snapshot support for DS8000, DS6000, and SVC for mySAP on DB2 UDB
- ▶ Snapshot support for DS8000, DS6000, and SVC for mySAP on Oracle
- ▶ Multiple snapshot versions managed by Tivoli Storage Manager policy
- ▶ Coordinated FlashCopy backup of multi-partition DB2 UDB databases distributed across multiple host systems.

Support of FlashCopy and snapshot functionality allows for “Zero Impact” backups and instant recovery. Data transfer to the Tivoli Storage Manager server is handled from a separate storage server, allowing the primary production data to remain online and undisturbed.

For more information, see *Tivoli Storage Manager for Advanced Copy Services - Data Protection for Snapshot Devices Installation and User's Guide*, SC33-8331-00.

17.5.6 IBM Tivoli Storage Manager for Copy Services

IBM Tivoli Storage Manager for Copy Services is a new optional module for Windows that integrates with Tivoli Storage Manager or Tivoli Storage Manager Extended Edition. It is designed to utilize Microsoft's Volume Snapshot Services (VSS) on Windows 2003. Tivoli Storage Manager for Copy Services provides similar functionality to Tivoli Storage Manager for Advanced Copy Services, but supports Windows VSS and Microsoft Exchange Server 2003 only.

Tivoli Storage Manager for Copy Services includes the following features:

- ▶ Single command-line interface (CLI) for performing earlier and VSS snapshot backup, restore, and query operations
- ▶ Single GUI for performing earlier and VSS snapshot backup, restore, and query operations
- ▶ Support for both hardware and software VSS providers that strictly adhere to Microsoft VSS provider requirements
- ▶ Support for a clustered Exchange environment.

Full and Copy backup types are supported, with granularity at the Exchange Storage Group level. Backups are managed by Tivoli Storage Manager policies and can be stored on the Tivoli Storage Manager server, local disks or both. Different policies can be assigned for the various storage locations and backup types (Full or Copy). As with Tivoli Storage Manager for Advanced Copy Services, zero impact backups and instant recovery allow the primary production data to remain online and undisturbed. Data movement to Tivoli Storage Manager storage can be off-loaded to a secondary machine through a VSS hardware provider that supports transportable shadow copy volumes.

For more information, see the website:

<http://www-01.ibm.com/software/tivoli/products/storage-mgr-copy-services/>

17.5.7 IBM Tivoli Storage Manager for SAP V6.1

Previously known as IBM Tivoli Storage Manager for Enterprise Resource Planning, Data Protection for SAP for V6.1 is a software module that works with IBM Tivoli Storage Manager to better protect infrastructure and application data and improve the availability of SAP R/3 servers.

For more details, see the following publications:

- ▶ *Tivoli Storage Manager for Enterprise Resource Planning - Data Protection for SAP, Installation and User's Guide for DB2*, SC33-6341-10
- ▶ *Tivoli Storage Manager for Enterprise Resource Planning - Data Protection for SAP, Installation and User's Guide for Oracle*, SC33-6340-10.

17.5.8 IBM Tivoli Storage Manager for Space Management

IBM Tivoli Storage Manager for Space Management provides Hierarchical Storage Management (HSM) to automatically migrate rarely-accessed files to alternative storage, without disrupting the most frequently used files in local storage. Migrated files are automatically and transparently recalled to primary storage when required by applications or users. Administrators and users are freed from manual file system maintenance tasks, and more online disk space is available for more important active data. Tivoli Storage Manager for Space Management can also help defer the requirement to purchase additional disk storage for clients, by making optimal use of available client storage.

Tivoli Storage Manager for Space Management offers increased scalability and performance through parallel migrations, improved candidate search, and optimized synchronization between the IBM Tivoli Storage Manager server and the HSM client.

IBM Tivoli Storage Manager for Space Management complements both IBM Tivoli Storage Manager and IBM Tivoli Storage Manager Extended Edition, and is supported on AIX, HP/UX, Solaris, and Linux.

For more information, see *Tivoli Storage Manager for Space Management for UNIX and Linux - User's Guide*, SC23-9794-00.

17.5.9 IBM Tivoli Storage Manager for Storage Area Networks

IBM Tivoli Storage Manager for Storage Area Networks enables your SAN-connected Tivoli Storage Manager servers and client computers to make maximum use of their direct network connection to storage. This software extension enables both servers and client computers to make the bulk of their backup/restore and archive/retrieve data transfers over the SAN instead of the LAN, either directly to tape or to the Tivoli Storage Manager disk storage pool. This ability greatly reduces the impact of data protection on the LAN while also reducing CPU utilization on both client and server.

Certain SAN configurations allow specific SAN devices to perform data movements directly to and from certain tape devices, further reducing client and server CPU utilization.

Tivoli Storage Manager for Storage Area Networks complements and coexists with the standard library-sharing functionality of both standard version and Extended editions of the Tivoli Storage Manager server.

The core functions of IBM Tivoli Storage Manager for Storage Area Network are:

- ▶ LAN-free backup/restore
- ▶ SAN-connected tape library.

This Tivoli Storage Manager component is also commonly referred to as a Storage Agent.

For more information, see the following publications:

- ▶ *Tivoli Storage Manager for SAN for AIX - Storage Agent User's Guide*, SC23-9797-00
- ▶ *Tivoli Storage Manager for SAN for HP-UX - Storage Agent User's Guide*, SC23-9798-00
- ▶ *Tivoli Storage Manager for SAN for Linux - Storage Agent User's Guide*, SC23-9799-00
- ▶ *Tivoli Storage Manager for SAN for Sun Solaris - Storage Agent User's Guide*, SC23-9800-00
- ▶ *Tivoli Storage Manager for SAN for Windows - Storage Agent User's Guide*, SC23-9553-00.

17.5.10 IBM Tivoli Storage Manager for System Backup and Recovery

IBM Tivoli Storage Manager for System Backup and Recovery (SysBack®) provides a flexible backup method for AIX systems to help protect data and provide bare machine recovery capabilities. It offers a comprehensive system backup, restore, and reinstallation tool. SysBack is a simple-to-use and highly effective tool. Any feature might be executed from either the AIX command line or by using the SMIT menu interface.

For Windows platforms, bare machine recovery can be achieved with the Tivoli Storage Manager Backup/Archive client's Automated System Recovery capability.

In addition, Windows, Sun, and Linux bare machine recovery can be done with Cristie Bare Machine Recovery. This integrates directly with Tivoli Storage Manager to provide operating system recovery for these platforms.

Tivoli Storage Manager for FastBack products also provide similar recovery capabilities for the Windows environment only.

For more information, see *Tivoli Storage Manager Backup-Archive Clients - Installation and User's Guide*.

17.5.11 IBM Tivoli Storage Manager FastBack

IBM Tivoli Storage Manager FastBack is recovery software, a specific kind of storage management that ensures that applications and users are back up and running within minutes following any data loss while performing full data recovery in the background:

- ▶ Protect and recover data for critical Windows applications
- ▶ Reduce the need for traditional backup windows with storage management software that captures data changes at the block level while providing extremely low systems overhead
- ▶ Schedule automated data transfers based on flexible, policy-based settings, helping administrators meet data protection and retention requirements on a per-application basis
- ▶ Enable data asset recovery from any Windows application, including Microsoft Exchange, Microsoft SQL Server, Oracle, IBM DB2 and SAP
- ▶ Make the most effective use of available bandwidth with strategies such as multi-threading, bundling of small files and industry-standard compression
- ▶ Operating systems supported: Windows.

For more information about Tivoli Storage Manager FastBack, see the website:

<http://www.ibm.com/software/tivoli/products/storage-mgr-fastback/>

17.5.12 IBM Tivoli Storage Manager FastBack for Microsoft Exchange

IBM Tivoli Storage Manager FastBack for Microsoft Exchange provides the ability to recover Microsoft Exchange data such as email, attachments, calendar entries, contacts and tasks.

It helps you to accomplish the following tasks:

- ▶ Minimize business risk with the ability to recover Microsoft Exchange data objects from virtually any Microsoft Exchange database, even corrupt databases
- ▶ Optimize Microsoft Exchange recovery by applying it at a granular level to any individual data object or group of objects, such as individual email messages, contact lists, tasks, or calendar entries

- ▶ Increase operational efficiency and user productivity by helping reduce recovery time from hours or days to minutes
- ▶ Improve service levels by helping to minimize the downtime associated with data recovery
- ▶ Enable recovery of objects that were previously considered unrecoverable, such as deleted email messages, or address books lost due to synchronization errors
- ▶ Restore objects directly to an Exchange Server or to send objects to a user-defined destination by SMTP
- ▶ Integrate with Active Directory and Exchange Server security to help limit unauthorized access to backup and restore systems.

For more information, see the website:

<http://www.ibm.com/software/tivoli/products/storage-mgr-fastback-exchange/>

17.5.13 IBM Tivoli Storage Manager FastBack for Bare Machine Recovery

IBM Tivoli Storage Manager FastBack for Bare Machine Recovery provides recovery following a disaster or catastrophic server failure, restoring systems within an hour:

- ▶ Enables systems recovery following a disaster or catastrophic server failure
- ▶ Provides the flexibility of recovering to comparable hardware, to dissimilar hardware or to a virtual machine using VMware or Microsoft Virtual Server
- ▶ Helps protect remote or branch offices with a cost-effective disaster recovery and business resiliency strategy that requires a minimum of standby hardware
- ▶ Utilizes IBM Tivoli Storage Manager FastBack to provide near-instant access to applications and data while full recovery takes place in the background
- ▶ Facilitates the migration of workloads to new hardware platforms, making it fast and easy to move workloads from old hardware or stand-alone servers to new hardware or blade servers
- ▶ Enables organizations to perform bare machine recovery in a local office, in a data center or in a central recovery site
- ▶ Operating systems supported: Windows.

For more information, see the website:

<http://www.ibm.com/software/tivoli/products/storage-mgr-fastback-bmr/>

17.5.14 IBM Tivoli Continuous Data Protection for Files

According to industry surveys, almost 70% of corporate data exists on notebooks (mobile computers) or desktop machines, and less than 8% of it is backed up regularly. For notebook, desktop, and file server machines that contain important, critical, or sensitive data that is constantly being updated, a typical 24-hour backup cycle might not be sufficient to provide adequate data protection. The addition of IBM Tivoli Continuous Data Protection for Files provides a client machine with the capability of being able, transparently in real time, to back up a file to a Tivoli Storage Manager server as soon as the file is saved. Files that are backed up by this method are managed in the same way as other corporate data by the Tivoli Storage Manager server.

Tivoli Continuous Data Protection for Files was developed with notebook (mobile computer) and desktop users in mind, but can be applied to any client with a high rate of change of data on its file systems.

Tivoli Continuous Data Protection for Files provides clients with true point-in-time recoverability. It is supported on AIX, Solaris, Linux, and Windows platforms.

For more information, see the website:

<http://www.ibm.com/software/tivoli/products/continuous-data-protection/>

17.6 Client backup and restore operations

It is important to understand Tivoli Storage Manager options for client backup and restore operations. It is also important to understand the characteristics of each of these operations because each method might have an effect on backup and restore efficiency, retention periods, portability, CPU utilization, connection time, and network utilization. The standard backup method that Tivoli Storage Manager uses is called progressive incremental backup. It is a unique and efficient method for backup.

For details, see the documentation listed in 17.10, “More information” on page 471.

17.6.1 Traditional LAN and WAN backup topology

In a traditional LAN and WAN environment the Tivoli Storage Manager backup and archive client or application reads data from locally attached disks and sends it over the LAN to the Tivoli Storage Manager backup server. The server receives the data, and then writes it out to its storage pool (tape, disk, or optical media) based on predefined policies and server configuration. Data is read and written by both the Tivoli Storage Manager Client and Tivoli Storage Manager Server machines. In addition, control information is also sent over the LAN to the Tivoli Storage Manager Server.

Traditional LAN and WAN backup flow is shown in Figure 17-12.

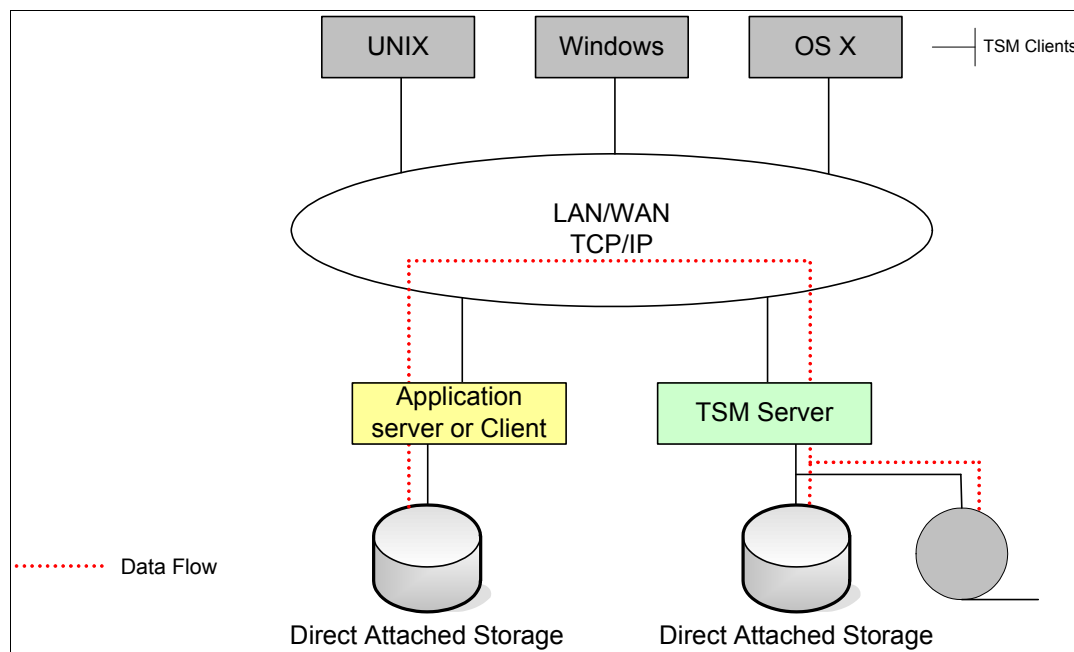


Figure 17-12 Tivoli Storage Manager LAN and WAN backup flow

17.6.2 SAN backup topology: LAN-free

SAN technology provides an alternative path for data movement between the Tivoli Storage Manager client and the server. Shared storage resources (disk, tape) are accessible to both the client and the server through the Storage Area Network. Data movement is off-loaded from the LAN and from the server processor and allows for greater scalability. LAN-free backups decrease the load on the LAN by introducing a Storage Agent.

The Storage Agent can be thought of as a small Tivoli Storage Manager server (without a database or recovery log), which is installed and run on the Tivoli Storage Manager client machine. The Storage Agent handles the communication with the Tivoli Storage Manager server over the LAN but sends the data directly to SAN attached tape devices or disks, relieving the Tivoli Storage Manager server from the actual I/O transfer.

For more information, see the following publications:

- ▶ *Tivoli Storage Manager for SAN for AIX - Storage Agent User's Guide*, SC23-9797-00
- ▶ *Tivoli Storage Manager for SAN for HP-UX - Storage Agent User's Guide*, SC23-9798-00
- ▶ *Tivoli Storage Manager for SAN for Linux - Storage Agent User's Guide*, SC23-9799-00
- ▶ *Tivoli Storage Manager for SAN for Sun Solaris - Storage Agent User's Guide*, SC23-9800-00
- ▶ *Tivoli Storage Manager for SAN for Windows - Storage Agent User's Guide*, SC23-9553-00.

17.6.3 Split-mirror/point-in-time copy backup using SAN

A split-mirror/point-in-time backup occurs when a copy volume generated by Operating System mirroring or a hardware assisted instant copy function (as found on many of today's high-end storage systems) is backed up to a Tivoli Storage Manager server. Such a backup method virtually eliminates the backup-related performance impact on the production host. This approach is facilitated and automated with the Tivoli Storage Manager for Advanced Copy Services product, which integrates IBM FlashCopy on disk systems with Tivoli Storage Manager and its database protection capabilities for Oracle, SAP, and DB2. This *Copy-Backup* procedure adds value to storage and backup procedures, because it helps ensure that essential applications can continue to run 24x7 with minimal backup-related impact.

Figure 17-13 shown split-mirror/point-in-time backup flow.

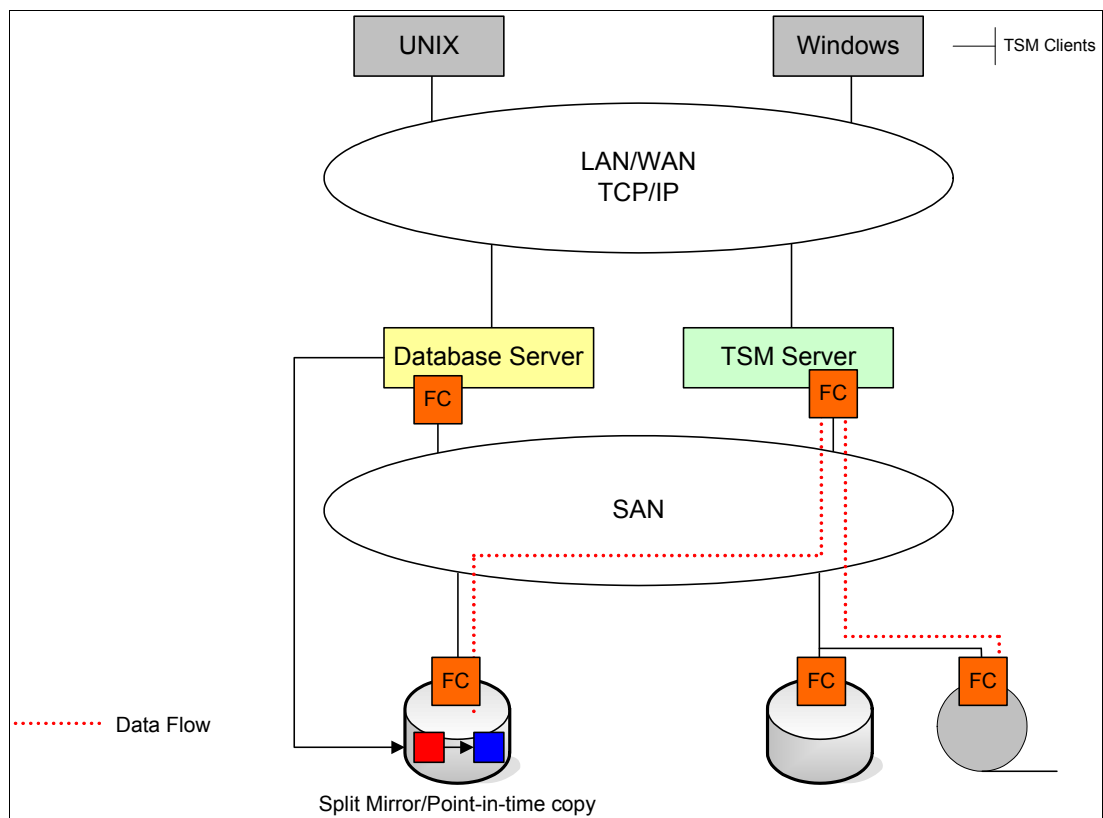


Figure 17-13 Tivoli Storage Manager split-mirror/point-in-time copy backup flow

17.6.4 Network-attached storage backup and restore

Network-attached storage (NAS) file servers are dedicated storage machines whose operating systems are optimized for file-serving functions. NAS file servers typically do not run third-party software. Instead, they interact with programs like Tivoli Storage Manager through industry-standard network protocols, such as network data management protocol (NDMP).

Tivoli Storage Manager provides two basic types of configurations that use NDMP for backing up and managing NAS file servers. In one type of configuration, Tivoli Storage Manager uses NDMP to back up a NAS file server to a library device directly attached to the NAS file server. The NAS file server, which can be distant from the Tivoli Storage Manager server, transfers backup data directly to a drive in a SCSI-attached tape library. Data is stored in special, NDMP-formatted storage pools, which can be backed up to storage media that can be moved offsite for protection in case of an on-site disaster.

See Figure 17-14 for details.

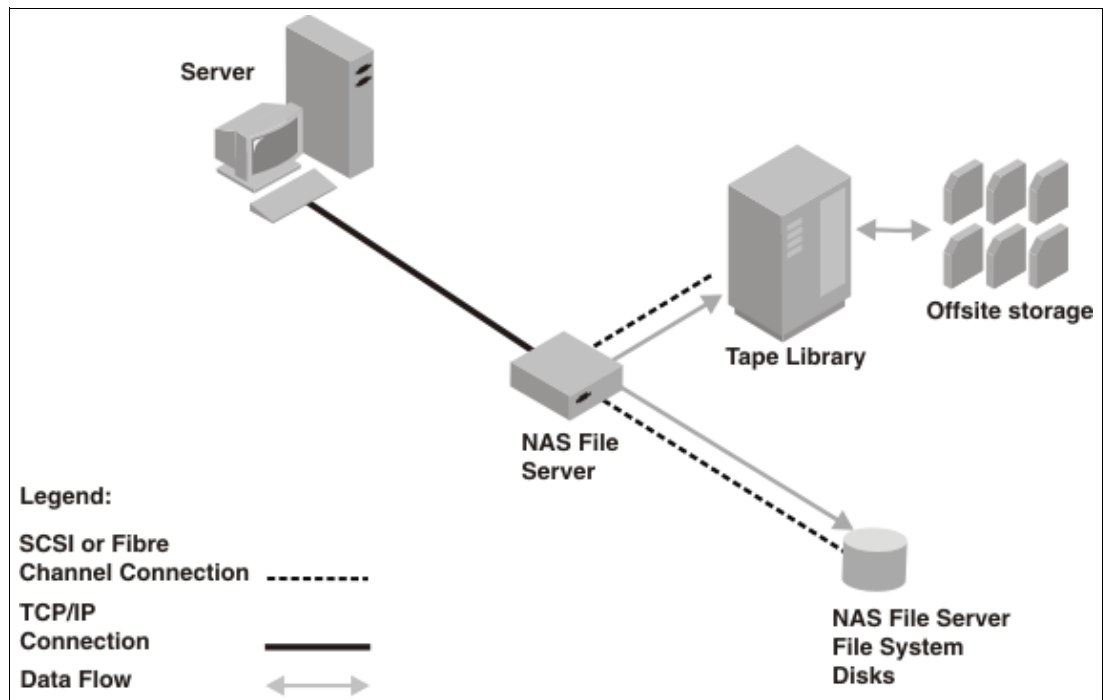


Figure 17-14 Library device directly attached to a NAS file server

In the other type of NDMP-based configuration, Tivoli Storage Manager uses NDMP to back up a NAS file server to a Tivoli Storage Manager storage-pool hierarchy. With this type of configuration, you can store NAS data directly to disk (either random access or sequential access) and then migrate the data to tape. Data can also be backed up to storage media that can then be moved offsite. The advantage of this type of configuration is that it gives you all the back-end data management features associated with a conventional Tivoli Storage Manager storage pool hierarchy, including migration and reclamation.

Figure 17-15 shows the NAS file server to Tivoli Storage Manager storage-pool hierarchy.

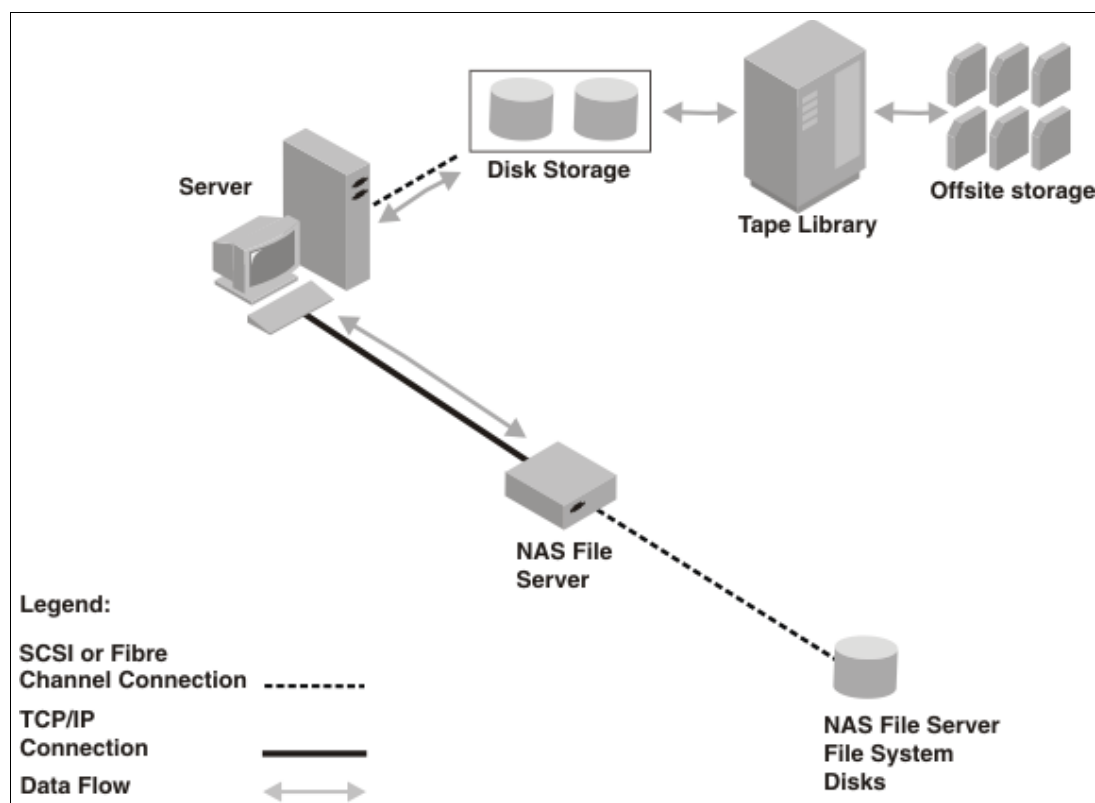


Figure 17-15 NAS file server to Tivoli Storage Manager storage-pool hierarchy

In both types of configurations, Tivoli Storage Manager tracks file-system image backups and has the capability to perform NDMP file-level restores.

For more information, see the documentation listed in 17.10, “More information” on page 471.

17.6.5 Image backup

An image backup is a block-by-block copy, single object backup of a volume (typically a UNIX file system or raw logical volume, or Windows drive) on a Tivoli Storage Manager client. Being able to restore an entire volume as one object can lead to faster recoveries. Image backup is available at the time of writing on AIX, HP, Sun, Linux, and Windows client platforms. Review the specific image backup requirements for each platform in their associated installation guides found at:

<http://publib.boulder.ibm.com/infocenter/tsminfo/v6/index.jsp>

With image backup, the Tivoli Storage Manager server does not track individual files in the file system image. File system images are tracked as individual objects and the management class policy are applied to the file system image as a whole.

An image backup provides the following benefits:

- Backs up file systems containing a large number of files faster than a full file system incremental back up
- Improves the speed with which Tivoli Storage Manager restores file systems containing many small files

- ▶ Conserves resources on the server during backups because only one entry is required for the image
- ▶ Provides a point-in-time picture of your logical volume, which might be useful if your enterprise needs to recall that information
- ▶ Restores a corrupt file system or raw logical volume. Data is restored to the same state it was when the last logical volume backup was performed.

To restore an image backup of a volume, the Tivoli Storage Manager client must be able to obtain an exclusive lock on the volume being restored.

Important: Do not use the Tivoli Storage Manager dynamic image backups for file systems, because the file system's design can provide inconsistent data even when there is no write activity. Dynamic image backup might result in a fuzzy image, and is not guaranteed to be valid or complete when restored.

On Windows client platforms a Logical Volume Storage Agent (LVSA) is used, which can take snapshot of the volume while it is online. Optionally only occupied blocks can be copied. If the snapshot option is used (rather than static) then any blocks which change during the backup process are first kept unaltered in an Original Block File. In this way the client can send a consistent image of the volume as it was at the start of the snapshot process to the Tivoli Storage Manager server.

Attention: File systems managed by Tivoli Storage Manager for space management are not enabled for image backup.

17.6.6 VMware Consolidated Backup

VMware Consolidated Backup is a VMware backup solution for the ESX 3.0 server and above in a storage area network (SAN) and non-SAN environment.

By offloading the backup away from the virtual machines and the ESX host, you can reduce the load on the production VMware ESX host for backup operations. It also allows backup in a LAN-free environment.

17.7 Tivoli Storage Manager server-to-server communications

With a solution that includes multiple Tivoli Storage Manager servers, you can use server-to-server communications and virtual volumes to enhance management and improve disaster recoverability. Tivoli Storage Manager server-to-server communications provide the capability to do the following tasks:

- ▶ Configure and manage multiple servers with enterprise administration
- ▶ Distribute a consistent configuration for Tivoli Storage Manager servers through a configuration manager to managed servers. By having consistent configurations, you can simplify the management of a large number of servers and clients
- ▶ Perform tasks simultaneously on multiple servers by using command routing, enterprise logon and enterprise console
- ▶ Send server and client events to another server for logging
- ▶ Monitor events of many servers and clients from a single server
- ▶ Store data on another server using virtual volumes.

For details, see the documentation listed in 17.10, "More information" on page 471.

17.7.1 Enterprise configuration

The Tivoli Storage Manager enterprise configuration functions make it easier to consistently set up and manage a network of Tivoli Storage Manager servers. You can set up configurations on one server and distribute the configurations to other servers. You can make changes to configurations and have the changes automatically distributed.

To use enterprise configuration, select the Tivoli Storage Manager server that is to act as the *configuration manager*. You might want to dedicate a new server for this purpose.

Figure 17-16 illustrates a simple configuration.

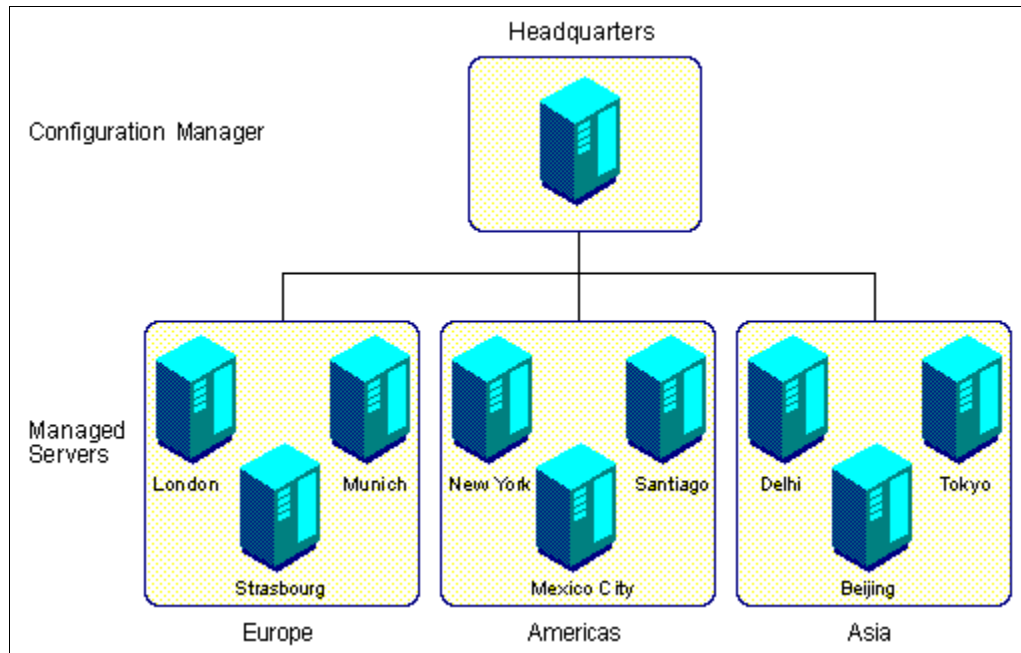


Figure 17-16 A scenario for implementing enterprise configuration

At the configuration manager, you can define the details of the server configurations that you want to distribute, for example:

- ▶ Set up backup and archive policies and client option sets
- ▶ Designate one or more administrators to have access to the servers, and control their authority levels
- ▶ Define the servers that you want the configuration manager to manage or communicate with, and you set up communications among the servers

In one or more *profiles*, point to the definitions of the configuration information that you want to use to manage other servers.

On each server that is to receive the configuration information, identify the server as a *managed server* by defining a *subscription* to one or more profiles owned by the configuration manager. All the definitions associated with the profiles are then copied into the managed server's database. Things defined to the managed server in this way are managed objects that cannot be changed by the managed server. From then on, the managed server gets any changes to the managed objects from the configuration manager by the profiles. Managed servers receive changes to configuration information at time intervals set by the servers, or by command. For details, see Figure 17-17.

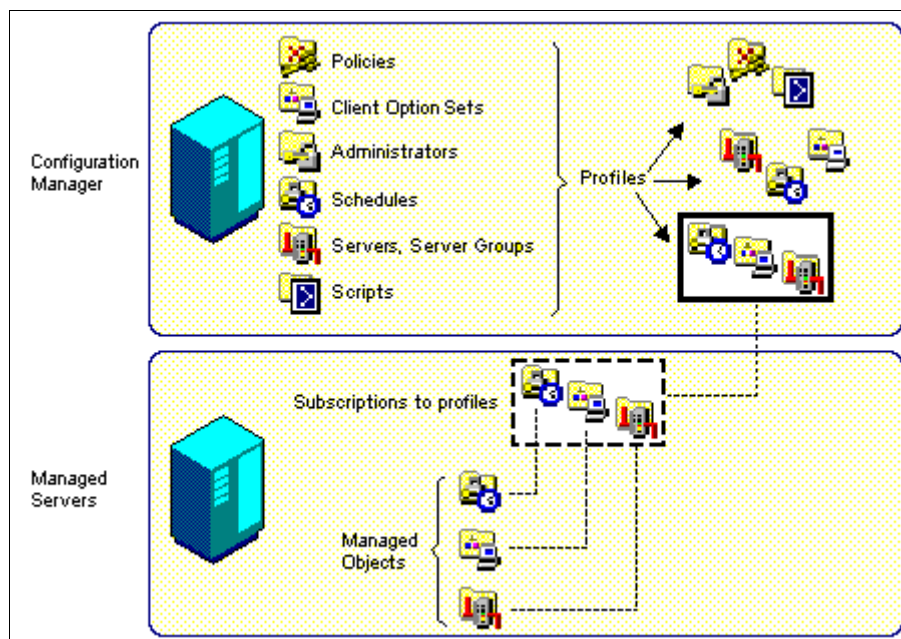


Figure 17-17 Enterprise configuration

For more information about enterprise configuration, see the documentation listed in 17.10, “More information” on page 471.

17.7.2 Using server-to-server virtual volumes for disaster recovery

Server-to-server virtual volumes can be used as a strategy for site disaster protection and disaster recovery.

For site disaster protection, server-to-server virtual volumes (with Tivoli Disaster Recovery Manager) can be used to back up storage pools from one Tivoli Storage Manager server to another and back up the Tivoli Storage Manager database of one server to another. This form of virtual vaulting can occur on a peer-to-peer basis, between any number of servers, and at separate data centers or separate sites. For example, two servers can back up each other in an equal level relationship.

For disaster recovery, server-to-server virtual volumes can be used to store the Disaster Recovery Plan file remotely. In this strategy, the source server creates the Disaster Recovery Manager plan files, then stores the files on a remote target server. You can display information about recovery plan files from the server that created the files (the source server) or from the server on which the files are stored (the target server). You can easily display a list of all recovery plan files that have been saved on a target server.

For details, see the documentation listed in 17.10, “More information” on page 471.

17.8 Tivoli Storage Manager and high availability clustering

High availability clustering can be used to configure either or both a highly available Tivoli Storage Manager server and/or clients. A highly available Tivoli Storage Manager server provides highly available services to its clients without significant disruptions; and a highly available Tivoli Storage Manager client provides highly available backup or restore services to an application and data.

For a highly available Tivoli Storage Manager server, critical resources such as the database and storage pools must be shared between the clustered nodes. A highly available Tivoli Storage Manager client will be able to back up file systems that are commonly available to all nodes in the cluster.

Tivoli Storage Manager supports clustering using supporting software such as IBM PowerHA™, Microsoft Cluster Server (MSCS), VERITAS Storage Foundation, and Tivoli System Automation.

This section provides overview information about a few of these environments. For more information, see the documentation listed in 17.10, “More information” on page 471.

17.8.1 IBM High Availability Cluster Multi-Processing

A Tivoli Storage Manager server can use IBM High Availability Cluster Multi-Processing (HACMP) software for high availability. HACMP is an AIX-based clustering solution, which allows automatic system recovery on system failure detection. Using HACMP together with Tivoli Storage Manager ensures server availability. HACMP offers local or campus disaster survivability with real-time automated failover and reintegration within distance limitations. In an HACMP environment, TCP/IP is the communications method used to support the checking of status and availability of the production and failover server, also commonly referred to as the *heartbeat* connection.

HACMP detects system failures and manages failover to a recovery processor with a minimal end-user downtime. You can set up a 2-way active-passive HACMP clustered configuration so that, if the production Tivoli Storage Manager server fails, the server will be brought up on the standby system in the cluster, as shown in Figure 17-18. Scripts are provided with Tivoli Storage Manager to automate the failover and failback operations. The Tivoli Storage Manager database, recovery log, and storage pools are stored on shared storage, available to both servers in the cluster.

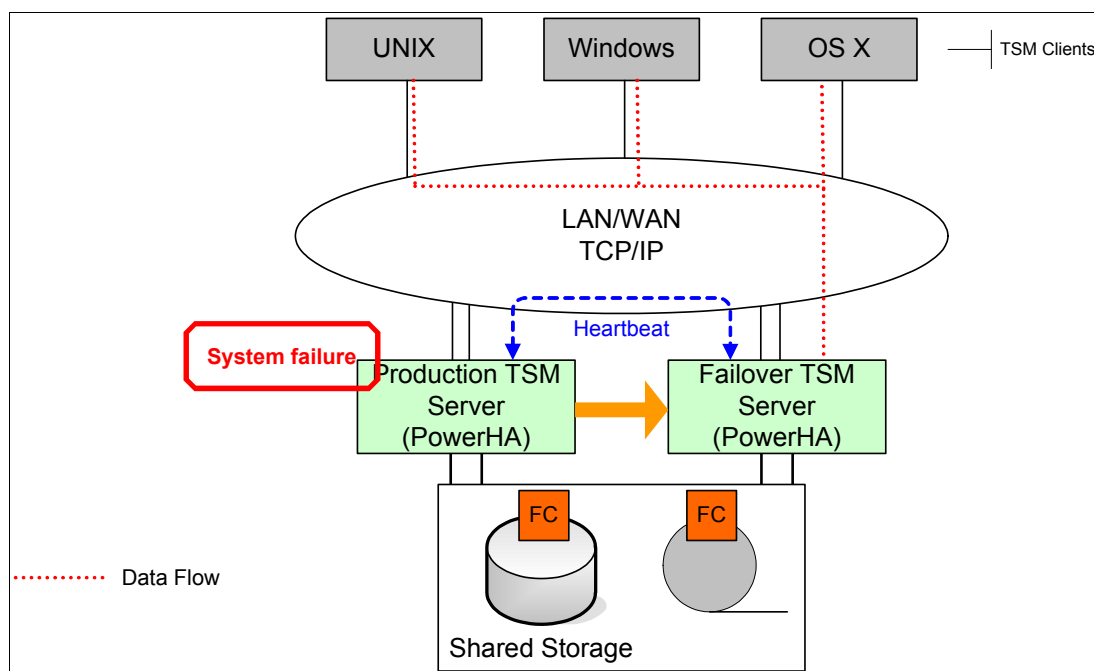


Figure 17-18 PowerHA and Tivoli Storage Manager server high availability configuration

Both failover and fallback act as though a Tivoli Storage Manager server has crashed or halted and was then restarted - the fact that the server is running on other hardware is transparent to the clients. Any transactions that were in-flight at the time are rolled back, and all completed committed transactions are still complete. Tivoli Storage Manager clients see this as a communications failure and try to re-establish the connection. The backup-archive clients can usually restart from the last committed transaction.

Note: For a detailed discussion of supported environments, prerequisites, install, setup, and testing of a PowerHA and Tivoli Storage Manager server failover environment, see *Tivoli Storage Manager for AIX - Administrator's Guide*, SC23-9769-00, and *PowerHA for AIX Cookbook*, SG24-7739-00.

17.8.2 Backup-archive client support with HACMP

The Tivoli Storage Manager backup-archive client itself (including the administrator, backup-archive, HSM, and API pieces) can also be clustered using PowerHA to provide highly available backup-restore operations. This configuration allows scheduled Tivoli Storage Manager client operations to continue processing in the event of a system failure on a redundant clustered failover server. See Figure 17-19 for an illustration of how this works.

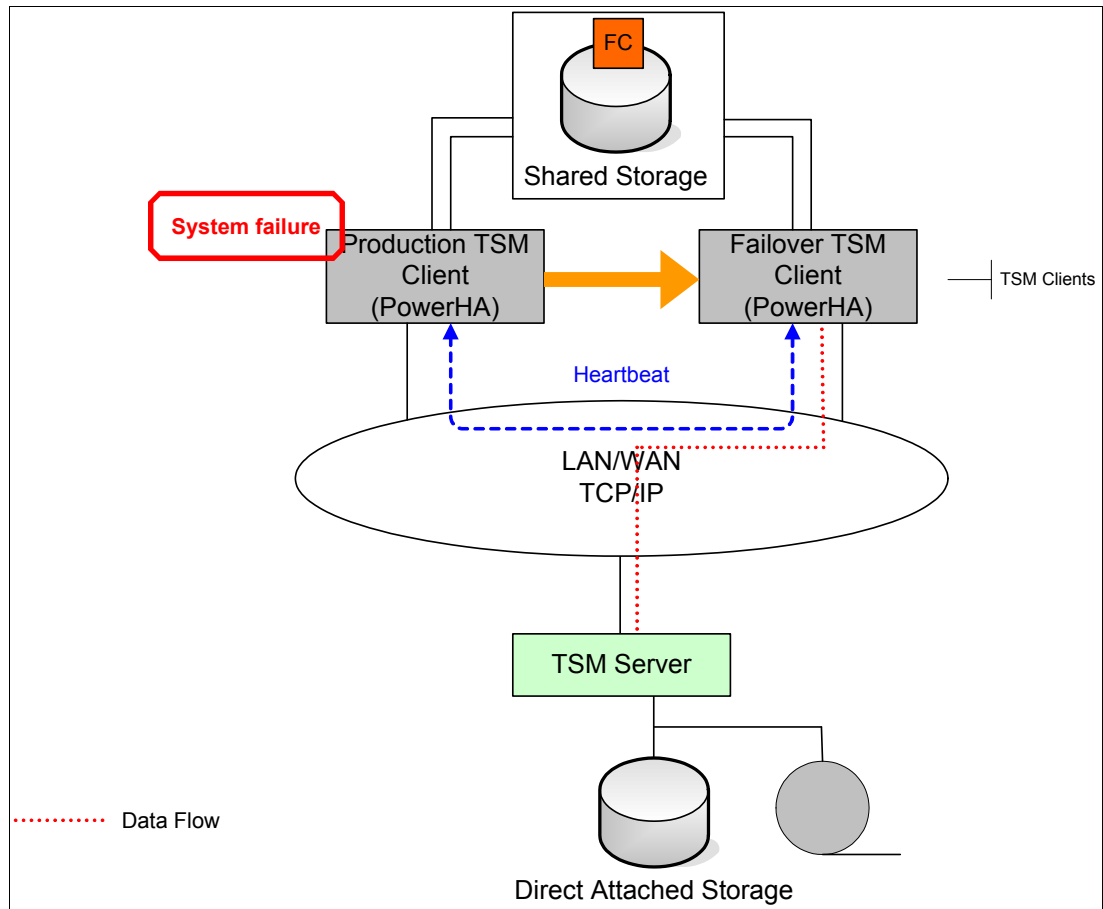


Figure 17-19 PowerHA and Tivoli Storage Manager client high availability configuration

In this configuration, the data to be backed up is stored on shared disk which is available to either of the clustered client nodes.

If a scheduled incremental backup of a clustered volume is running on machine-a and a system failure causes a failover to machine-b, machine-b then reconnects to the server. If the re-connection occurs within the start window for that event, the scheduled command is restarted. This scheduled incremental backup will reexamine files sent to the server before the failover. The backup will then *catch up* to where it terminated before the failover situation.

If a failover occurs during a user initiated (that is, non-scheduled) client session, the Tivoli Storage Manager client starts on the node that is handling the takeover. This allows it to process scheduled events and provide web client access. You can install the Tivoli Storage Manager client locally on each node of an PowerHA environment. You can also install and configure the Tivoli Storage Manager Scheduler service for each cluster node to manage all local disks and each cluster group containing physical disk resources.

PowerHA supports for Hierarchical Storage Management (HSM) clients on AIX provides support for PowerHA failover on AIX so that HSM managed file systems can continue to operate in the case of an PowerHA node failover and fallback.

17.8.3 Tivoli Storage Manager server and Microsoft Cluster Server (MSCS)

Tivoli Storage Manager is a cluster-aware application and can be configured in a MSCS high availability environment. Clustered systems are connected to the same disk subsystem and provide a high-availability solution that minimizes or eliminates many potential sources of downtime. Microsoft Cluster Server (MSCS) is software that helps configure, monitor, and control applications and hardware components that are deployed on a Windows cluster. When you use cluster configurations, you enhance the availability of your servers.

With a Microsoft Cluster Server, you can place Tivoli Storage Manager server cluster resources into a virtual server. A *virtual server* is an MSCS cluster group that looks like a Windows server. The virtual server has a network name, an IP address, one or more physical disks, and a service. A Tivoli Storage Manager server can be one of the virtual services provided by an MSCS virtual server.

The virtual server name is independent of the name of the physical node on which the virtual server runs. The virtual server name and address migrate from node to node with the virtual server. Clients connect to a Tivoli Storage Manager server using the virtual server name, rather than the Windows server name. The virtual server name is implemented as a cluster network name resource and maps to a primary or backup node. The mapping is dependent on where the virtual server currently resides. Any client that uses WINS or directory services to locate servers can automatically track the virtual server as it moves between nodes. Automatically tracking the virtual server does not require client modification or reconfiguration.

As mentioned earlier, each virtual server has its own disk as part of a cluster resource group. Therefore, they cannot share data. Each Tivoli Storage Manager server that has been implemented as a virtual server has its database, recovery log, and set of storage pool volumes on a separate disk owned by that virtual server.

Because the server's location is transparent to client applications, this affords Tivoli Storage Manager the maximum ease of failover and fallback, while minimizing the impact on the Tivoli Storage Manager clients.

Note: MSCS only supports an IP Address as a resource. This means that any Tivoli Storage Manager server running on a cluster must limit its supported communication method to just TCP/IP. Any client *not* using TCP/IP as a communication method will not be able to reach the virtual server if it fails over to the other cluster node.

The following example demonstrates the way the MSCS virtual server concept works.

Assume that a clustered Tivoli Storage Manager server called TSMSEVER1 is running on node A and a clustered Tivoli Storage Manager server called TSMSEVER2 is running on node B. Clients connect to the Tivoli Storage Manager server TSMSEVER1 and the Tivoli Storage Manager server TSMSEVER2 without knowing which node currently hosts their server. The MSCS concept of a *virtual server* ensures that the server's location is transparent to client applications. To the client, it appears that the Tivoli Storage Manager server is running on a virtual server called TSMSEVER1.

Figure 17-20 illustrates the MSCS virtual server concept.

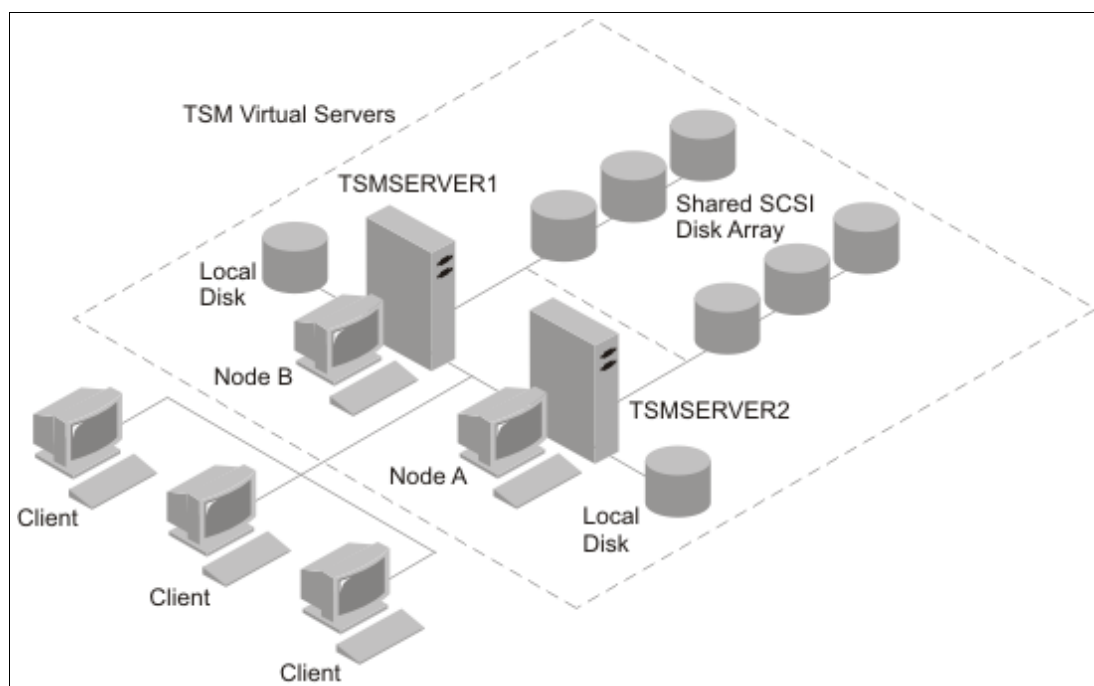


Figure 17-20 Clustering with TMSERVER1 as node a and TMSERVER2 as node b

When one of the software or hardware resources fails, failover occurs. Resources (for example: applications, disks, or an IP address) migrate from the failed node to the remaining node. The remaining node takes over the Tivoli Storage Manager server resource group, restarts the Tivoli Storage Manager service, and provides access to administrators and clients.

If node A fails, node B assumes the role of running TMSERVER1. To a client, it is exactly as if node A were turned off and immediately turned back on again. Clients experience the loss of all connections to TMSERVER1 and all active transactions are rolled back to the client. Clients must reconnect to TMSERVER1 after this occurs. The location of TMSERVER1 is transparent to the client.

Figure 17-21 demonstrates clustering with TSMSERVER2 as node b and assuming the role of TSMSERVER1 as node a.

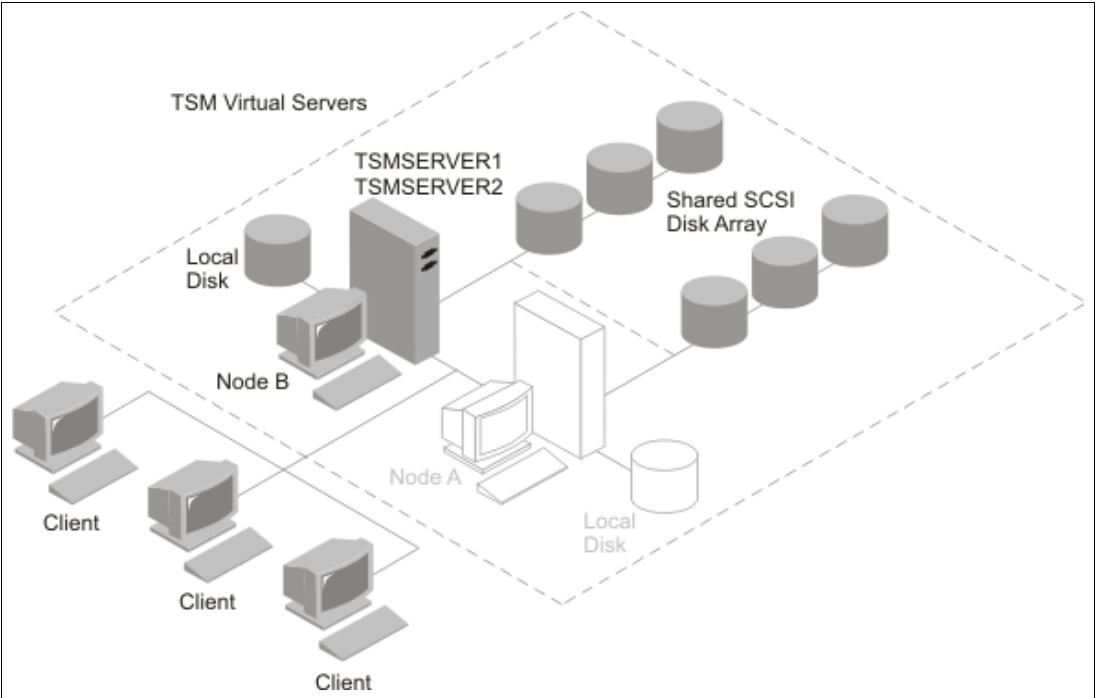


Figure 17-21 Clustering with TSMSERVER2 as node b and assuming the role of TSMSERVER1 as node a

SCSI tape failover

With MSCS and VCS, you cannot use failover for tape devices. However, Tivoli Storage Manager can handle this type of a failover pattern.

Tivoli Storage Manager uses a shared SCSI bus for the tape devices. Each node (two only) involved in the tape failover must contain an additional SCSI adapter card. The tape devices (library and drives) are connected to the shared bus. When failover occurs, the Tivoli Storage Manager server issues a SCSI bus reset during initialization. In a failover situation, the bus reset is expected to clear any SCSI bus reserves held on the tape devices. This allows the Tivoli Storage Manager server to acquire the devices after the failover.

Table 17-1 describes the hardware and software tested for use with Tivoli Storage Manager SCSI tape failover.

Table 17-1 Hardware and software supported for SCSI tape failover

Operating System	Windows Advanced Server or Datacenter Server
SCSI Adapter	Adaptec AHA-2944UW PCI SCSI Controller
SCSI Tape Library	IBM 3590-B11 or IBM 3570-C12

Fibre tape failover

With MSCS and VCS, you cannot use failover for tape and library devices. However, Tivoli Storage Manager can handle the failover of Fibre Channel direct attached tape and library devices on the Windows 2003 or Windows 2008 server without any additional hardware.

Fibre Channel devices have to be directly attached and not connected through a gateway (for example, Passlight 5000) in order for failover to be successful. The following section describes a specific hardware and software configuration for Tivoli Storage Manager Fibre tape failover support. Table 17-2 on page 465 describes the hardware and software tested for use with Tivoli Storage Manager Fibre tape failover.

Table 17-2 Hardware and software supported for fibre tape failover

Operating system	Windows 2003 or Windows 2008 32-bit server
Fibre Channel Adapter	QLogic HBA 2340 with Storport driver
Fibre Channel Tape Library and Drives	IBM and other vendor Fibre Channel directly attached tape and library devices

For a detailed discussion on supported environments, prerequisites, install, setup, and testing of a MSCS and Tivoli Storage Manager server failover environment, see *Tivoli Storage Manager for Windows - Administrator's Guide*, SC23-9773-00.

17.8.4 Tivoli Storage Manager backup-archive client support with MSCS

The Tivoli Storage Manager client is supported in an MSCS cluster environment. This configuration allows scheduled Tivoli Storage Manager client operations to continue processing in the event of a system failure on a redundant clustered failover server as shown in Figure 17-22.

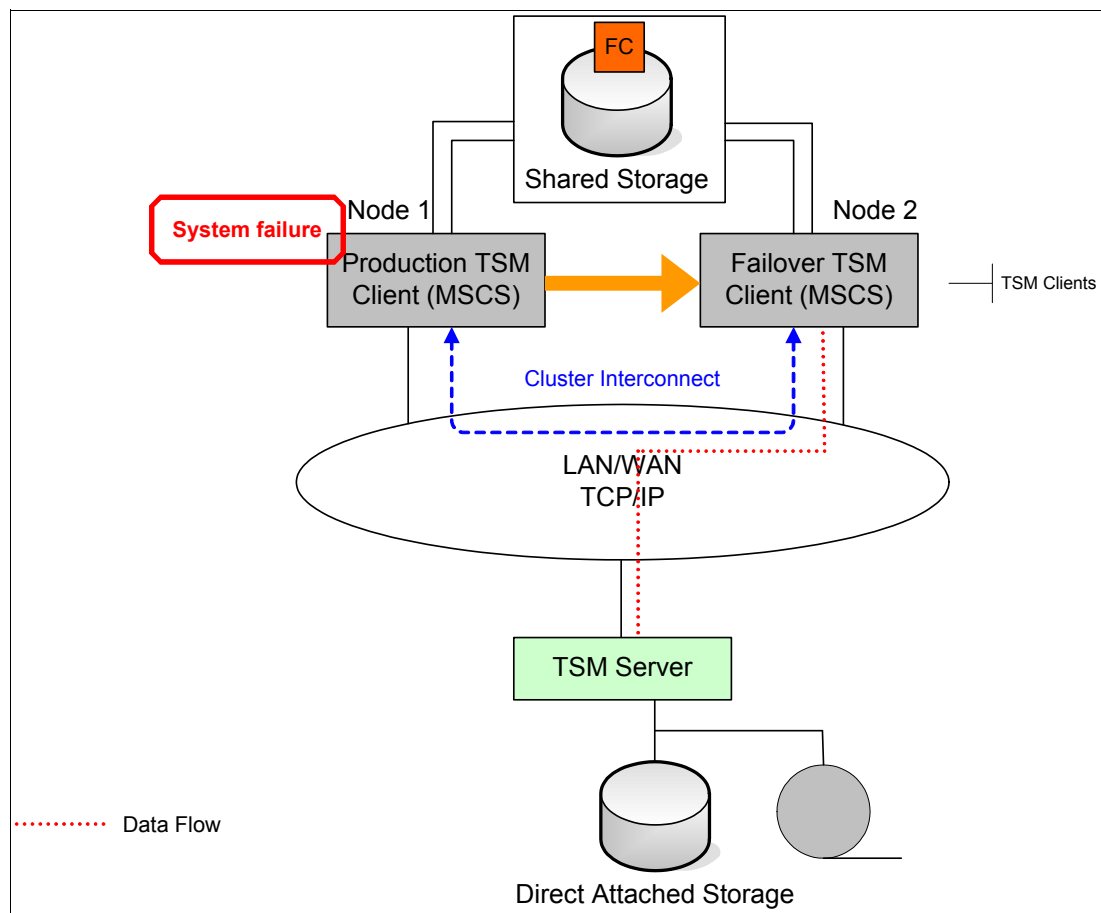


Figure 17-22 MSCS and Tivoli Storage Manager client high availability configuration

In this example, the cluster contains two nodes: Node 1 and Node 2, and two cluster groups containing physical disk resources. In this case, an instance of the Tivoli Storage Manager Backup-Archive Scheduler Service is installed for each node: Node 1 and Node 2. This ensures that proper resources are available to the Backup-Archive client when disks move (or fail) between cluster nodes. The CLUSTERNODE option in the client option file ensures that Tivoli Storage Manager manages backup data logically, regardless of which cluster node backs up a cluster disk resource.

17.9 Tivoli Storage Manager and tape vaulting

Traditionally, disaster recovery plans include daily off-site tape backups that are picked up from the local site and transported by a courier to a secure facility, which is often a tape vaulting service provider. Vaulting of tapes at off-site locations can provide a secure means to protect data in the event of a disaster at the primary site. To recover from a disaster, you must know the location of off-site recovery media. Tivoli Storage Manager DRM helps determine which volumes to move off-site and back on-site and tracks the location of the volumes.

With tape vaulting, you can back up primary storage pools to a copy storage pool and then send the copy storage pool volumes off-site. You can track these copy storage pool volumes by changing their access mode to off-site, and updating the volume history to identify their location. If an off-site volume becomes expired, the server does not immediately return the volume to the scratch pool. The delay prevents the empty volumes from being deleted from the database, making it easier to determine which volumes must be returned to the on-site location. DRM handles all of this automatically.

17.9.1 Electronic tape vaulting

Using Tivoli Storage Manager with electronic tape vaulting provides additional data protection capabilities, with backups made to remote tape drives over communication links. Electronic tape vaulting can enable shorter recovery times and reduced data loss if the server is damaged. An electronic tape vaulting solution combined with Tivoli Storage Manager is fundamental to achieving Tier 3 and above RPO and RTOs, that is, less than 24 hours.

With electronic tape vaulting, the Tivoli Storage Manager server will have an alternate location to store primary and copy storage pools as though they are directly attached. The Tivoli Storage Manager server can first write a copy of disk storage pool data to tape pools at the remote site (Datacenter #2), then the data can be migrated to the tape storage pools at the primary site (Datacenter #1). See Figure 17-23.

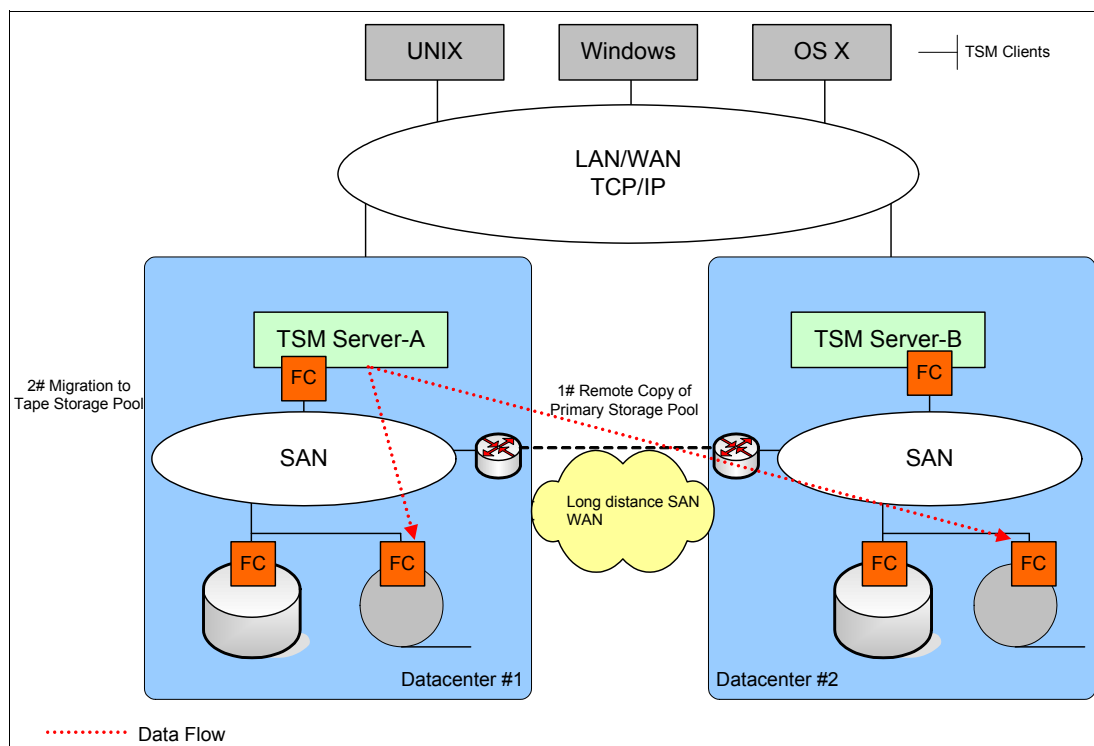


Figure 17-23 Tivoli Storage Manager and electronic tape vaulting

Depending on your configuration (and whether or not remote disk replication is being used in conjunction with electronic tape vaulting) you might choose to backup the Tivoli Storage Manager database and configuration files by this method. This ensures a copy of the data is stored at both sites and that the Tivoli Storage Manager server can rapidly recover at the remote site. If remote disk replication is used for mirroring of the Tivoli Storage Manager database and storage pools, the Tivoli Storage Manager server can be recovered very quickly without any loss of client data. A peer-to-peer configuration can be used to balance the load of Tivoli Storage Manager services in the enterprise and provide data protection and rapid recovery for a failure at either site.

Using electronic tape vaulting with Tivoli Storage Manager offers these advantages:

- ▶ Critical data can be frequently and rapidly vaulted to remote off-site locations.
- ▶ In the event of a disaster, up-to-date data can be restored at a hotsite, therefore improving recovery time and recovery point objectives.
- ▶ It eliminates physical tape handling that can result in damaged tapes, lost tapes, tapes delayed in transit, or data that is sabotaged. Increased reliability.
- ▶ It eliminates costs associated with couriers and off-site vaulting vendors.
- ▶ Government off-site vaulting regulations are satisfied.
- ▶ It lowers the cost of downtime and storage management.
- ▶ It increases company control and data security.
- ▶ Peer solutions eliminate or reduce costs associated with hot-site service providers.

17.9.2 Remote disk mirroring and tape vaulting solutions

A variety of technologies exist for remote disk mirroring and electronic tape vaulting:

- ▶ Long distance SANs
- ▶ Dense Wavelength Division Multiplexing (DWDM)
- ▶ Fibre extenders
- ▶ WAN based channel extension using telco and IP protocols
- ▶ Newly emerging NAS and iSCSI gateways.

Table 17-3 summarizes several of these technologies. The use of such technologies also might depend on a particular vendor's replication or vaulting solution. For example, the IBM System Storage DS6000 and DS8000 use PPRC to achieve data replication. Metro Mirror is supported by ESCON links, which can be further extended by DWDM or WAN channel extension.

Table 17-3 Extended electronic vaulting technologies

Electronic vaulting technology	Commonly supported distances between sites	Common product vendors	Relative technology costs
Long Distance SAN (shortwave/longwave Fibre Channel)	up to 30km	Brocade (OEM) McDATA (OEM)	Low
CWDM, DWDM and fibre extenders	up to 300km	Cisco Nortel Ciena Finisar CNT	Medium
WAN, IP Based Routers and Channel Extension	up to 1000's km	CNT Cisco (OEM)	Low to High (depending on solution)

SANs were designed to overcome the distance limitations of other storage channel protocols, such as SCSI. Longwave laser GBICs available on most SAN hubs, switches and directors enable a transmission distance of up to 10 kilometers (11 kilometers when including switch to host connections) when used with 9 micron diameter single-mode optical fibre. Shortwave GBICs use multi-mode fibre and is the ideal choice for shorter distance (less than 500 meters from transmitter to receiver or vice versa).

CWDM or DWDM is a way to open up the conventional optical fibre bandwidth by breaking it up into many channels, each at a particular optical wavelength (a unique color of light). Each wavelength can carry a signal at any bit rate less than an upper limit defined by the electronics, typically up to several gigabits per second. DWDMs are implemented in areas that have dark fibre available through telcos and service providers. The DWDM is deployed as part of the physical layer. It is therefore independent of protocol, simply passing signal information in the format it is received. Examples of the protocols it can support are ATM, Gigabit Ethernet, ESCON, FICON and Fibre Channel.

WAN and IP based channel extenders typically use telecommunication lines for data transfer and therefore enable application and recovery sites to be located longer distances apart. The use of WAN and IP channel extenders provides the separation for disaster recovery purposes and avoids various barriers imposed when customers do not have a "right of way" to lay their fibre cable. WAN and IP channel extenders generally compress the data before sending it over the transport network, however the compression ratio needs to be determined based on the application characteristics and the distance.

Network attached storage (NAS) and iSCSI solutions are beginning to offer low cost IP based storage. Copies of Tivoli Storage Manager storage pools and the Tivoli Storage Manager database can be stored at a remote site using IP based storage to offer a low cost implementation while utilizing existing infrastructure. Configurations can include Tivoli Storage Manager clients attached to iSCSI based data backing up to a Tivoli Storage Manager server or Tivoli Storage Manager servers using iSCSI based storage as storage pools.

For a detailed overview of technologies, products, costs, and best practices with distance solutions, review *Introduction to Storage Area Networks*, SG24-5470-03.

17.9.3 Collocation considerations for off-site vaulting

With collocation, large numbers of files belonging to a client node can be restored, retrieved and recalled more quickly. However, using collocation on copy storage pools requires special consideration. Primary and copy storage pools perform various recovery roles. Normally you use primary storage pools to recover data to clients directly. You use copy storage pools to recover data to the primary storage pools. In a disaster where both clients and the server are lost, the copy storage pool volumes will probably be used directly to recover clients. The types of recovery scenarios that concern you the most will help you to determine whether to use collocation on your copy storage pools.

You might also want to consider that collocation on copy storage pools will result in more partially filled volumes and probably increased off-site reclamation activity. Collocation typically results in a partially filled sequential volume for each client or client file space. This might be acceptable for primary storage pools because these partially filled volumes remain available and can be filled during the next migration process. However, for copy storage pools this might be unacceptable because the storage pool backups are usually made to be taken off-site immediately. If you use collocation for copy storage pools, you will have to decide between:

- ▶ Taking more partially filled volumes off-site, thereby increasing the reclamation activity when the reclamation threshold is lowered or reached
- ▶ Leaving these partially filled volumes on-site until they fill and risk not having an off-site copy of the data on these volumes.

With collocation disabled for a copy storage pool, typically there will be only a few partially filled volumes after storage pool backups to the copy storage pool are complete. Consider carefully before using collocation for copy storage pools. Even if you use collocation for your primary storage pools, you might want to disable collocation for copy storage pools. Or, you might want to restrict collocation on copy storage pools to certain critical clients, as identified by the Business Impact Analysis.

17.9.4 Reclamation considerations for off-site vaulting

Space on a sequential volume becomes reclaimable as files expire or are deleted from the volume. For example, files become obsolete because of aging or limits on the number of versions of a file. In reclamation processing, the Tivoli Storage Manager server rewrites files on the volume being reclaimed to other volumes in the storage pool, making the reclaimed volume available for reuse.

When an off-site volume is reclaimed, the files on the volume are rewritten to another copy storage pool volume that is on-site. The Tivoli Storage Manager server copies valid files contained on the off-site volumes being reclaimed, from the original files in the primary storage pools. In this way, the server can reclaim off-site copy storage pool volumes without having to recall and mount these volumes. Logically, these files are moved back to the on-site location. The new volume must be moved off-site as soon as possible. However, the files have not been physically deleted from the original off-site volume. In the event of a disaster occurring before the newly written copy storage pool volume has been taken off-site, these files can still be recovered from the off-site volume, provided that it has not already been reused and the database backup that you use for recovery references the files on the off-site volume.

The server reclaims an off-site volume which has reached the reclamation threshold as follows:

1. The server determines which files on the volume are still valid
2. The server obtains these valid files from a primary storage pool, or if necessary, from an on-site volume of a copy storage pool
3. The server writes the files to one or more volumes in the copy storage pool and updates the database. If a file is an aggregate file with unused space, the unused space is removed during this process
4. A message is issued indicating that the off-site volume was reclaimed
5. The newly written volumes are then marked to be sent off-site, and after this has occurred, the reclaimed volume can be returned to an on-site scratch pool.

Volumes with the access value of off-site are eligible for reclamation if the amount of empty space on a volume exceeds the reclamation threshold for the copy storage pool. The default reclamation threshold for copy storage pools is 100%, which means that reclamation is not performed.

If you plan to make daily storage pool backups to a copy storage pool, then mark all new volumes in the copy storage pool as off-site and send them to the off-site storage location. This strategy works well with one consideration; if you are using automatic reclamation (the reclamation threshold is less than 100%). Each day's storage pool backups will create a number of new copy storage pool volumes, the last one being only partially filled. If the percentage of empty space on this partially filled volume is higher than the reclaim percentage, this volume becomes eligible for reclamation as soon as you mark it off-site. The reclamation process causes a new volume to be created with the same files on it. The volume you take off-site is then empty according to the Tivoli Storage Manager database. If you do not recognize what is happening, you can perpetuate this process by marking the new partially filled volume off-site.

If you send copy storage pool volumes off-site, it is best that you control copy storage pool reclamation by using the default value of 100. This turns reclamation off for the copy storage pool. You can start reclamation processing at desired times by changing the reclamation threshold for the storage pool.

Depending on your data expiration patterns, you might not need to do reclamation of off-site volumes each day. You might choose to perform off-site reclamation on a less frequent basis. For example, suppose you send copy storage pool volumes to and from your off-site storage location once a week. You can run reclamation for the copy storage pool weekly, so that as off-site volumes become empty they are sent back for reuse.

When you perform reclamation for off-site volumes, use the following sequence:

1. Back up your primary storage pools to copy storage pools.
2. Turn on reclamation for copy storage pools by lowering the reclamation threshold below 100%.
3. When reclamation processing completes, turn off reclamation for copy storage pools by raising the reclamation threshold to 100%.
4. Mark any newly created copy storage pool volumes as off-site and then move them to the off-site location.

This sequence ensures that the files on the new copy storage pool volumes are sent off-site, and are not inadvertently kept on-site because of reclamation.

Attention: If collocation is enabled and reclamation occurs, the server tries to reclaim the files for each client node or client file space onto a minimal number of volumes.

17.10 More information

- ▶ More information about Tivoli Storage Manager can be found on the following website:
<http://www.ibm.com/software/tivoli/products/storage-mgr/>
- ▶ Product manuals for API and Backup-Archive Clients:
 - *Using the Application Programming Interface*, SC23-9793-00
 - *Tivoli Storage Manager for UNIX and Linux Backup-Archive Clients Version 6.1 - Installation and User's Guide*, SC23-9791-00
 - *Tivoli Storage Manager for Windows Backup-Archive Clients*, SC23-9792-00
- ▶ Product manuals for Storage Manager for AIX:
 - *Tivoli Storage Manager for AIX - Administrator's Guide*, SC23-9769-00
 - *Tivoli Storage Manager for AIX - Administrator's Reference*, SC23-9775-00
 - *Tivoli Storage Manager for AIX - Installation Guide*, GC23-9781-00
 - *Tivoli Storage Manager for SAN for AIX - Storage Agent User's Guide*, SC23-9797-00
- ▶ Product manuals for Storage Manager for HP-UX:
 - *Tivoli Storage Manager for HP-UX - Administrator's Guide*, SC23-9770-00
 - *Tivoli Storage Manager for HP-UX - Administrator's Reference*, SC23-9776-00
 - *Tivoli Storage Manager for HP-UX - Installation Guide*, GC23-9782-00
 - *Tivoli Storage Manager for SAN for HP-UX - Storage Agent User's Guide*, SC23-9798-00
- ▶ Product manuals for Storage Manager for Linux:
 - *Tivoli Storage Manager for Linux - Administrator's Guide*, SC23-9771-00
 - *Tivoli Storage Manager for Linux - Administrator's Reference*, SC23-9777-00
 - *Tivoli Storage Manager for Linux - Installation Guide*, GC23-9783-00
 - *Tivoli Storage Manager for SAN for Linux - Storage Agent User's Guide*, SC23-9799-00
- ▶ Product manuals for Storage Manager for Sun Solaris:
 - *Tivoli Storage Manager for Sun Solaris - Administrator's Guide*, SC23-9772-00
 - *Tivoli Storage Manager for Sun Solaris - Administrator's Reference*, SC23-9778-00
 - *Tivoli Storage Manager for Sun Solaris - Installation Guide*, GC23-9784-00
 - *Tivoli Storage Manager for SAN for Sun Solaris - Storage Agent User's Guide*, SC23-9800-00

- ▶ Product manuals for Storage Manager for Windows:
 - *Tivoli Storage Manager for Windows - Administrator's Guide*, SC23-9773-00
 - *Tivoli Storage Manager for Windows - Administrator's Reference*, SC23-9779-00
 - *Tivoli Storage Manager for Windows - Installation Guide*, GC23-9785-00
 - *Tivoli Storage Manager for SAN for Windows - Storage Agent User's Guide*, SC23-9553-00
- ▶ Product manual for Space Management:

Tivoli Storage Manager for Space Management for UNIX and Linux - User's Guide, SC23-9794-00
- ▶ Product manual for ACS:

Tivoli Storage Manager for Advanced Copy Services - Data Protection for Snapshot Devices Installation and User's Guide, SC33-8331-00
- ▶ Product manuals for SAP:
 - *Tivoli Storage Manager for Enterprise Resource Planning - Data Protection for SAP, Installation and User's Guide for DB2*, SC33-6341-10
 - *Tivoli Storage Manager for Enterprise Resource Planning - Data Protection for SAP, Installation and User's Guide for Oracle*, SC33-6340-10
- ▶ Product manual for Mail:

Tivoli Storage Manager for Mail - Data Protection for Microsoft Exchange Server, Installation and User's Guide, SC23-9796-00.



IBM Tivoli Storage Productivity Center

IBM Tivoli Storage Productivity Center (TPC) is an integrated suite that includes a single user interface to manage capacity utilization of storage systems, file systems, and databases, and to automate file system capacity provisioning in both physical and virtual environments.

IBM Tivoli Storage Productivity Center components can be ordered as a bundled suite or ordered separately based on specific needs.

These individual components are included:

- ▶ IBM Tivoli Storage Productivity Center Basic Edition
- ▶ IBM Tivoli Storage Productivity Center for Data
- ▶ IBM Tivoli Storage Productivity Center for Disk
- ▶ IBM Tivoli Storage Productivity Center for Replication.

These integrated packages are included:

- ▶ IBM Tivoli Storage Productivity Center Standard Edition
- ▶ IBM Tivoli Storage Productivity Center with Advanced Provisioning
- ▶ IBM System Storage Productivity Center.

18.1 New with IBM Tivoli Storage Productivity Center V4.1.1

You can use this information to learn about new features and enhancements in IBM Tivoli Storage Productivity Center version 4.1.1. This section highlights the changes since IBM Tivoli Storage Productivity Center 4.1 and IBM TotalStorage Productivity Center 3.3.2.

For more information about each of the features, go to the Tivoli Storage Productivity Center Information Center and search for Planning for the IBM Tivoli Storage Productivity Center family. For information about how to use the features, see the *IBM Tivoli Storage Productivity Center User's Guide*, SC27-2338-01.

Tivoli Storage Productivity Center 4.1.1 adds the following new features, functions, and enhancements.

18.1.1 Name and licensing changes

Here we list the changes in name and licensing.

Name change

IBM Tivoli Storage Productivity Center V4.1.1 has been renamed from IBM TotalStorage Productivity Center. All user interfaces, documentation, online help, and messages have also been changed to reflect the name change.

Licensing changes

These are the licenses available for IBM Tivoli Storage Productivity Center:

- ▶ IBM Tivoli Storage Productivity Center Basic Edition
- ▶ IBM Tivoli Storage Productivity Center for Data
- ▶ IBM Tivoli Storage Productivity Center for Disk
- ▶ IBM Tivoli Storage Productivity Center Standard Edition.

If you have an IBM TotalStorage Productivity Center for Fabric license only, you can upgrade to IBM Tivoli Storage Productivity Center Standard Edition.

If you have an IBM TotalStorage Productivity Center for Basic Edition license only, you can upgrade to IBM Tivoli Storage Productivity Center Basic Edition, IBM Tivoli Storage Productivity Center for Disk, IBM Tivoli Storage Productivity Center for Data, or IBM Tivoli Storage Productivity Center Standard Edition.

If you have an IBM TotalStorage Productivity Center for Data license only, you can upgrade to IBM Tivoli Storage Productivity Center for Data or IBM Tivoli Storage Productivity Center Standard Edition.

If you have an IBM TotalStorage Productivity Center for Disk license only, you can upgrade to IBM Tivoli Storage Productivity Center for Disk, IBM Tivoli Storage Productivity Center for Data (Disk plus Data), or IBM Tivoli Storage Productivity Center Standard Edition.

If you have an IBM TotalStorage Productivity Center Standard Edition license, you can upgrade to IBM Tivoli Storage Productivity Center Standard Edition.

18.1.2 Integration features

Tivoli Storage Productivity Center provides these integration features:

Integration of Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication

Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication, previously separate products, are now integrated. You can start the IBM Tivoli Storage Productivity Center for Replication user interface from within the Tivoli Storage Productivity Center user interface.

The IBM Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication Installation and Configuration Guide SC27-2337-01 also includes the installation, upgrade, and uninstallation information for IBM Tivoli Storage Productivity Center for Replication.

This integration enables you to perform the following tasks:

- ▶ Start the IBM Tivoli Storage Productivity Center for Replication user interface from within the Tivoli Storage Productivity Center user interface
- ▶ Use the Tivoli Storage Productivity Center GUI to set up IBM Tivoli Storage Productivity Center for Replication SNMP alerts and IBM Tivoli Enterprise Console® events
- ▶ Provide a Tivoli Storage Productivity Center superuser role that has authority over all Tivoli Storage Productivity Center commands. IBM Tivoli Storage Productivity Center for Replication includes a replication administrator role that has authority to all IBM Tivoli Storage Productivity Center for Replication commands. IBM Tivoli Storage Productivity Center for Replication will honor the Tivoli Storage Productivity Center superuser role giving the superuser role authority over all Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication commands.

Integration of Tivoli Storage Productivity Center and IBM Tivoli Integrated Portal

Tivoli Integrated Portal is a standards-based architecture for web administration. Tivoli Integrated Portal enables developers to build administrative interfaces for IBM and independent software products as individual plug-ins to a common console network. The installation of Tivoli Integrated Portal is required to enable single sign-on for Tivoli Storage Productivity Center.

Single sign-on is an authentication process that enables you to enter one user ID and password to access multiple applications. Single sign-on integrates with the launch in context feature to enable you to move smoothly from one application to a specific location in a second application.

Launch in context feature

The launch in context feature enables you to access external applications from the Tivoli Storage Productivity Center GUI. Element managers are the most prevalent external applications that use the launch in context feature. An element manager is usually the vendor-specific software that is used to administer a particular storage device. The launch in context feature provides starting points in the Tivoli Storage Productivity Center GUI so you can click a button or select a menu item to start an element manager.

When you install Tivoli Storage Productivity Center, Tivoli Integrated Portal, and Tivoli Storage Productivity Center for Replication, the components are automatically configured to use launch in context. You can access Tivoli Storage Productivity Center and Tivoli Storage Productivity Center for Replication from the Tivoli Integrated Portal GUI and you can access Tivoli Storage Productivity Center for Replication from the Tivoli Storage Productivity Center GUI.

Single sign-on

Single sign-on is an authentication process that enables you to enter one user ID and password to access multiple applications. Single sign-on enables you to access:

- ▶ Tivoli Storage Productivity Center and Tivoli Storage Productivity Center for Replication from the Tivoli Integrated Portal GUI
- ▶ Tivoli Storage Productivity Center for Replication from the Tivoli Storage Productivity Center GUI
- ▶ External applications such as element managers from the Tivoli Storage Productivity Center GUI

The single sign-on feature requires a centralized user and group repository, such as an LDAP-compliant directory, that all participating applications can access.

Tivoli Storage Productivity Center uses Lightweight Third Party Authentication (LTPA) tokens to pass the user information between applications. To use LTPA tokens for single sign-on, each participating application must possess the same set of keys to encode and decode the user information contained in the token. As an additional security feature, the LTPA tokens expire after a determined amount of time. When the tokens expire, you must re-enter your user ID and password information.

If you select operating system authentication, then the use of the single sign-on feature is limited. Operating system authentication does not support single sign-on for element managers, even when the element manager is installed on the same machine as Tivoli Storage Productivity Center.

18.1.3 Storage Resource agents

Tivoli Storage Productivity Center now supports Storage Resource agents on Microsoft Windows, AIX, and Linux. The Storage Resource agent probe is equivalent to the information that is collected by probes using the Data agent.

The Storage Resource agents do not require the Agent Manager and can be deployed to other systems using the Tivoli Storage Productivity Center GUI on the server system.

You can use the following functions:

- ▶ Asset reports (including HBA)
- ▶ Capacity reports
- ▶ Subsystem to host storage correlation including multipathing information
- ▶ Topology and Data Path explorer functions.

This support does not include file system scans, NAS discovery or topology, zoning and zone control functions or subsystem device driver configuration. You can still use the Data agent and Fabric agent for this information.

18.1.4 SQL access to Tivoli Storage Productivity Center database

Tivoli Storage Productivity Center will provide a set of DB2 views that represent key information that has been collected by monitoring jobs and stored in the database repository. A view is a way of describing data that exists in one or more tables within the database repository. It does not contain data but, instead, is a stored set of SQL commands that define a subset of rows and columns in the base tables.

You can use the Structured Query Language (SQL) to retrieve the information from the views and create reports using your own tools, such as Business Intelligence and Reporting Tools

(BIRT) or Microsoft Excel. Other applications can also use these views to gather and import information that is collected by Tivoli Storage Productivity Center.

The following categories of views will contain information collected by Tivoli Storage Productivity Center:

Storage entity views

These views include information about the properties of the entity. For example, the name, capacity, freespace, and so forth for a storage subsystem.

Entities defined by Tivoli Storage Productivity Center

These entities include Data agents, Fabric agents, alert log, Tivoli Storage Productivity Center server, computer groups, storage subsystem groups, file system groups, storage resource groups, and so forth.

Aggregated views

These views provide summary information for the database history, data in a database instance, and the Data agent file system.

Reporting views

These views combine several various entities in one view for a report.

Rollup views

These views include rollup report information from the master and subordinate Tivoli Storage Productivity Center servers, Data agents and Fabric agents, host cluster data, computer group, host, database computer groups, fabric SAN assets, switch assets, storage subsystem group, storage subsystems, and Tivoli Storage Productivity Center for Databases.

18.1.5 Performance management improvements

Tivoli Storage Productivity Center release 4.1.1 provides performance management improvements for metrics, thresholds, and alert definitions. For more information, search for **improvements to performance metrics** in following Information Center:
<http://www-03.ibm.com/systems/storage/software/center/index.html>

18.1.6 Storage Optimizer

The Storage Optimizer is a tool to help you analyze your storage networks to identify hot spots or bottlenecks, plan for storage growth, improve performance, and help develop storage migration or storage consolidation plans. Using the data in the Tivoli Storage Productivity Center database, the Storage Optimizer enables you to create an analysis report and an optimization report. The analysis report analyzes your data storage environment and proposes changes to improve your environment. Based on the analysis report, the optimization report includes storage migration or storage consolidation guidelines.

This feature requires an IBM Tivoli Storage Productivity Center Standard Edition license.

18.1.7 Advanced Brocade Discovery

The Advanced Brocade Discovery function has been removed from Tivoli Storage Productivity Center. This function uses the Brocade API. If you have a fabric that is managed by the Brocade API and do not have a Brocade SMI-S agent, install and configure the

Brocade SMI-S agent to manage the fabric before or after you upgrade Tivoli Storage Productivity Center. You must install and configure the Brocade SMI-S agent to view zoning data or perform zone control through Tivoli Storage Productivity Center. When you upgrade Tivoli Storage Productivity Center, you receive a warning message if you are using the Advanced Brocade API but no Brocade SMI-S agent is configured to manage the fabric.

18.1.8 Discovery Library Adapter

Tivoli Storage Productivity Center provides a Discovery Library Adapter which can be used in a multiple IBM Tivoli product environments to exchange data with other Tivoli products. The data gathered by Tivoli Storage Productivity Center is put in a common data model format and written to a file using the Identification Markup Language. This file can be loaded into IBM Tivoli Change and Configuration Management database (CCMDB) so that other Tivoli products such as IBM Tivoli Application Dependency Discovery Manager or IBM Tivoli Business Service Manager can use that data.

18.1.9 Storage resource groups

Storage resource groups are new objects provided to help storage administrators plan, monitor, and report on the managed environment.

A storage resource group is a set of entities managed by Tivoli Storage Productivity Center. These entities can be servers, switches, storage subsystems, fabrics, storage pools, and storage volumes. Storage resource groups can be a group of heterogeneous objects and can also contain other storage resource groups without any connectivity.

Policies for provisioning (volume creation and selection, workload profiles, zoning and multipathing configuration) can be specified and associated with storage resource groups. These policies are used by the SAN Planner to populate default settings.

Storage resource groups are used primarily for planning functions but is also available with the Tivoli Storage Productivity Center Basic Edition license. With the basic license, you can create and view storage resource groups in the topology. With the Standard Edition license, the planner function is enabled and you can use storage resource groups as input.

Storage resource groups also work with the following profiles.

Workload profiles

Describes the requirements that define the performance characteristics of newly provisioned capacity.

Provisioning profiles

Describes the requirements such as total capacity, number of volumes, Redundant Array of Independent Disks (RAID) level, volume name prefix, multipathing options, zoning options, and so forth.

18.1.10 IBM Support Assistant

Tivoli Storage Productivity Center Standard Edition contains a plug-in for IBM Support Assistant (ISA). IBM Support Assistant is an application provided at no charge to help discover solutions, and fix issues with IBM software. For more information about ISA, see the website:

<http://www-01.ibm.com/software/support/isa/>

18.1.11 IBM General Parallel File System

Tivoli Storage Productivity Center supports the monitoring of the IBM General Parallel File System (GPFS) 3.2 on AIX. GPFS provides access to critical file data. GPFS also provides concurrent high-speed file access to applications that are running on multiple nodes of an AIX cluster, a Linux cluster, or a heterogeneous cluster of AIX and Linux nodes. In addition to providing file storage capabilities, GPFS provides storage management, information life-cycle tools, centralized administration and allows for shared access to file systems from remote GPFS clusters.

For more information about GPFS, see the *General Parallel File System - Concepts, Planning, and Installation Guide*, GA76-0413-02.

18.1.12 Installation changes

This section highlights the installation changes since IBM TotalStorage Productivity Center 3.3.2.

IBM Tivoli Storage Productivity Center for Replication

IBM Tivoli Storage Productivity Center for Replication is now installed with IBM Tivoli Storage Productivity Center. The *IBM Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication Installation and Configuration Guide* SC27-2337-01 also includes the installation, upgrade, and uninstallation information for IBM Tivoli Storage Productivity Center for Replication.

IBM DB2 Database for Linux, UNIX, and Windows

Tivoli Storage Productivity Center now supports DB2 9.5. You will be able to migrate your Tivoli Storage Productivity Center databases from DB2 9.1 or DB2 8.2 to DB2 9.5. DB2 9.5 is optional. Tivoli Storage Productivity Center still supports DB2 9.1.

Installation of IBM Tivoli Integrated Portal

Tivoli Storage Productivity Center now installs IBM Tivoli Integrated Portal along with Tivoli Storage Productivity Center.

Embedded WebSphere 6.1 and JRE 1.5

The Device server is upgraded to run under Embedded WebSphere 6.1 (from Embedded WebSphere 6.0.2). The Data server, GUI, and CLI is upgraded to use JRE version 1.5. InstallShield uses JRE 1.5 during the installation and uninstallation process when Tivoli Storage Productivity Center is installed using the disk1 image. The image to perform local agent installations uses JRE version 1.4.2.

Silent installation

Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication do not support silent installation except for the Data agents and Fabric agents.

18.1.13 New device and application support

You can use this information to learn about new device and application support in IBM Tivoli Storage Productivity Center version 4.1.1.

IBM System Storage DS8000 4.3

This release supports DS8000 4.3 with these additional items:

- ▶ Storage pool striping
- ▶ Dynamic volume expansion
- ▶ Internet Protocol Version 6 (IPv6)
- ▶ Redundant Array of Independent Disks (RAID 6)
- ▶ Variable logically partitioned mode (LPARs)
- ▶ Space-efficient FlashCopy
- ▶ Data encryption
- ▶ Solid-state drives
- ▶ 1 TB SATA drives.

IBM System Storage SAN Volume Controller 4.3.1

This release supports SAN Volume Controller 4.3.1 and above with these additional items:

- ▶ Embedded CIM agent
- ▶ 64-bit logical block address (LBA) for the back end array
- ▶ 2 TB virtual disks (VDisks) and managed disks (MDisks).

Microsoft SQL Server 2005 and Microsoft SQL Server 2008 databases

Tivoli Storage Productivity Center can now monitor the Microsoft SQL Server 2005 and Microsoft SQL Server 2008 databases. You must configure Microsoft SQL Server before you can monitor the database. For information about configuration, see the Information Center. Search for **Configuring Microsoft SQL Server 2005 or 2008**.

EMC PowerPath

With Tivoli Storage Productivity Center, you can now use EMC PowerPath storage systems like CLARiiON and Symmetrix. Using these storage systems, you can discover host volume information and display detailed information for the volume for capacity planning purposes. Connection reports can show the connectivity from the host to the storage subsystems.

EMC PowerPath version 4.0 or later is supported.

Network Appliance (NetApp)

With Tivoli Storage Productivity Center, you can use the Network Appliance SMI-S agent to support block storage devices. The SMI-S agent supports the SMI-S 1.2 array profile.

IBM XIV Storage System

XIV Storage System will have an embedded CIM agent that Tivoli Storage Productivity Center will use to run discovery and probe jobs.

You will be able to start the XIV Storage System GUI from within Tivoli Storage Productivity Center if the GUI is installed on the same system as the Tivoli Storage Productivity Center GUI. The XIV Storage System GUI will be supported on Windows and Linux.

Both the Data agent and Storage Resource agent will support XIV Storage System.

For the IBM XIV Storage System software release 10.2 and later, the XIV Storage System CIM agent must use user authentication credentials that are defined for the XIV Storage System system. XIV Storage System offers the option of using a Lightweight Directory Access Protocol (LDAP) repository for user authentication and authorization. For information about user authentication for XIV Storage System, see the *IBM XIV Storage System User Manual*.

Important: The XIV Storage System information provided in the Tivoli Storage Productivity Center 4.1.1 documentation is only for planning purposes until the supported XIV Storage System software is available. Tivoli Storage Productivity Center support is targeted for a future XIV Storage System software release. A flash will be issued when Tivoli Storage Productivity Center support for XIV Storage System is available.

Multipath subsystem device drivers

Tivoli Storage Productivity Center supports these subsystem device drivers (SDD):

- ▶ AIX SDD
- ▶ Windows SDD
- ▶ Windows SDD DSM
- ▶ Linux SDD
- ▶ HP SDD
- ▶ Solaris SDD
- ▶ Novell SDD (reporting only)
- ▶ AIX SDD PCM
- ▶ Linux DM_Multipath.

IBM System Storage N Series Gateway servers

IBM Tivoli Storage Productivity Center supports IBM System Storage N Series Gateway servers as Other NAS. This support allows you to monitor and report on file systems through the Windows CIFS or UNIX NFS shares that are accessible to the scan or probe jobs for the Data agent. No back-end storage information such as controllers, disks, and logical volumes is collected or reported.

High-Availability Cluster Multi-Processing

This release provides additional support for High-Availability Cluster Multi-Processing version 5.5.

Tivoli Enterprise Portal

A Universal Agent for Tivoli Storage Productivity Center that utilizes a set of Tivoli Storage Productivity Center web services calls to gather information and provide results files that will display enhanced information such as job status and Tivoli Storage Productivity Center status in the IBM Tivoli Integrated Portal.

Terminology

The Tivoli Storage Productivity Center documentation uses the term *storage subsystem* and the Tivoli Storage Productivity Center for Replication documentation uses the term *storage system*. Both terms mean the devices used for storage management.

18.1.14 Features supported in a future release

The following features will be supported in a future release of Tivoli Storage Productivity Center. Information pertaining to these features is provided in this topic for planning purposes only. A flash will be issued when Tivoli Storage Productivity Center support for these features is available.

LSI SMI-S Provider

Support for the LSI SMI-S Provider for IBM System Storage DS4000 and IBM System Storage DS5000 storage systems is targeted for a future release of Tivoli Storage Productivity Center. The LSI SMI-S Provider supports the Storage Management Initiative Specification (SMI-S) 1.3 specifications. The LSI SMI-S Provider also provides performance metrics.

DS5000 series

Support for the following DS5000 features is targeted for a future release of Tivoli Storage Productivity Center. These features are included in the following DS5000 series systems:

- ▶ DS5100 and DS5300 - full disk encryption, solid-state drives, 1 TB SATA drives
- ▶ DS5020 - full disk encryption.

SAN Volume Controller

Tivoli Storage Productivity Center 4.1.1 will support SAN Volume Controller 5.1.0/5.1.1, which is targeted for release after the 4.1.1 release. SAN Volume Controller 5.1.0/5.1.1 will include the following features:

- ▶ Maximum managed disk size is 2 TB
- ▶ Maximum virtual disk size is 256 TB
- ▶ The SAN Volume Controller hardware platform runs in a 64-bit mode rather than 32-bit.

Solid-state drives can be in a SAN Volume Controller node. Solid-state drives are an optional feature for SAN Volume Controller.

18.2 Tivoli Storage Productivity Center overview

IBM Tivoli Storage Productivity Center provides a set of policy-driven automated tools for managing storage capacity, availability, events, performance and assets in your enterprise environment, including NetWare, NAS, Tivoli Storage Enterprise Storage Server, and Microsoft Cluster Server technologies, as well as RDBMSs such as Oracle, Sybase SQL Server, Microsoft SQL Server, and DB2 UDB. Tivoli Storage Productivity Center provides storage management from the host and application to the target storage device. It provides disk and tape subsystem configuration and management, performance management, SAN fabric management and configuration, and usage reporting and monitoring.

Tivoli Storage Productivity Center can help you identify, evaluate, control and predict your enterprise storage management assets. Because it is policy-based, it can detect potential problems and automatically make adjustments based on the policies and actions that you define. For example, it can notify you when your system is running out of disk space or warn you of impending storage hardware failure. By alerting you to these and other issues related to your stored data, it enables you to prevent unnecessary system and application downtime.

Tivoli Storage Productivity Center is designed to:

- ▶ Simplifies the management of storage infrastructures
- ▶ Manages, configures, and provisions SAN-attached storage
- ▶ Monitors and tracks performance of SAN-attached devices
- ▶ Monitors, manages, and controls (through zones) SAN fabric components
- ▶ Manages the capacity utilization and availability of file systems and databases.

Figure 18-1 shows the main window, which is displayed when you log into IBM Tivoli Storage Productivity Center.

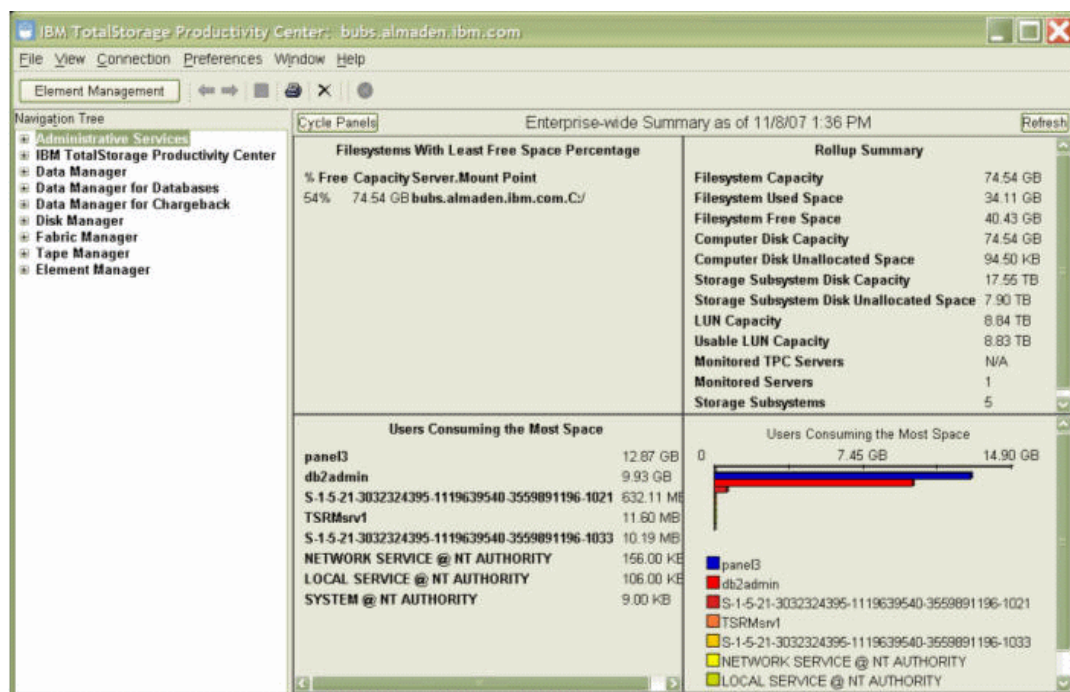


Figure 18-1 IBM Tivoli Storage Productivity Center main window

18.3 Tivoli Storage Productivity Center architecture

The IBM Tivoli Storage Productivity Center (TPC) consists of several key components. This topic identifies these components and shows how they are related.

For more information about the TPC architecture, see the *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

18.3.1 Tivoli Storage Productivity Center components

Here we discuss the components of the Tivoli Storage Productivity Center.

Data server

This component is the control point for product scheduling functions, configuration, event information, reporting, and GUI support. It coordinates communication with and data collection from agents that scan file systems and databases to gather storage demographics and populate the database with results. Automated actions can be defined to perform file system extension, data deletion, and Tivoli Storage Productivity Center backup or archiving, or event reporting when defined thresholds are encountered. The Data server is the primary contact point for GUI user interface functions. It also includes functions that schedule data collection and discovery for the Device server.

Device server

This component discovers, gathers information from, analyzes performance of, and controls storage subsystems and SAN fabrics. It coordinates communication with and data collection from agents that scan SAN fabrics and storage devices.

Database

A single database instance serves as the repository for all Tivoli Storage Productivity Center components.

Agents

Data agents, Storage Resource agents, and fabric agents gather host, application, and SAN fabric information and send this information to the Data Server or Device server.

GUI

The graphical user interface (GUI) lets you enter or receive information for all Tivoli Storage Productivity Center components.

CLI

The command-line interface (CLI) lets you issue commands for major Tivoli Storage Productivity Center functions.

Tivoli Integrated Portal

The IBM Tivoli Storage Productivity Center installation program includes IBM Tivoli Integrated Portal. Tivoli Integrated Portal is a standards-based architecture for web administration. Tivoli Integrated Portal enables developers to build administrative interfaces for IBM and independent software products as individual plug-ins to a common console network.

The installation of Tivoli Integrated Portal is required to enable single sign-on for Tivoli Storage Productivity Center. Single sign-on is an authentication process that enables you to enter one user ID and password to access multiple applications. Single sign-on integrates with the launch in context feature to enable you to move smoothly from one application to a functionally-related location in a second application.

Note: The single sign-on feature is not available for the following programs:

- ▶ IBM XIV Storage System element manager
- ▶ IBM System Storage SAN Volume Controller
- ▶ IBM System Storage DS4000.

Figure 18-2 shows a diagram of the Tivoli Storage Productivity Center components.

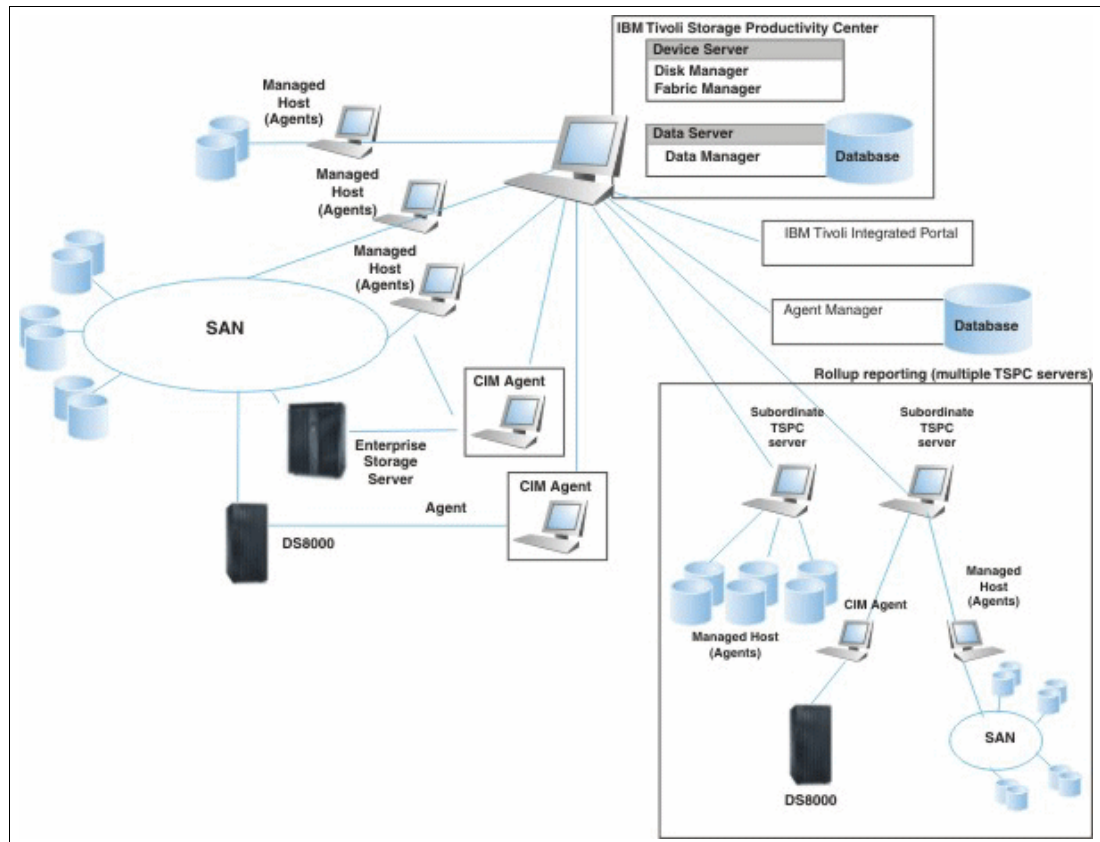


Figure 18-2 Tivoli Storage Productivity Center components

18.3.2 Tivoli Storage Productivity Center structure

In this section we look at the Tivoli Storage Productivity Center structure from the logical and physical view.

Logical structure

The logical structure of Tivoli Storage Productivity Center V4.1.1 has three layers, as shown in Figure 18-3.

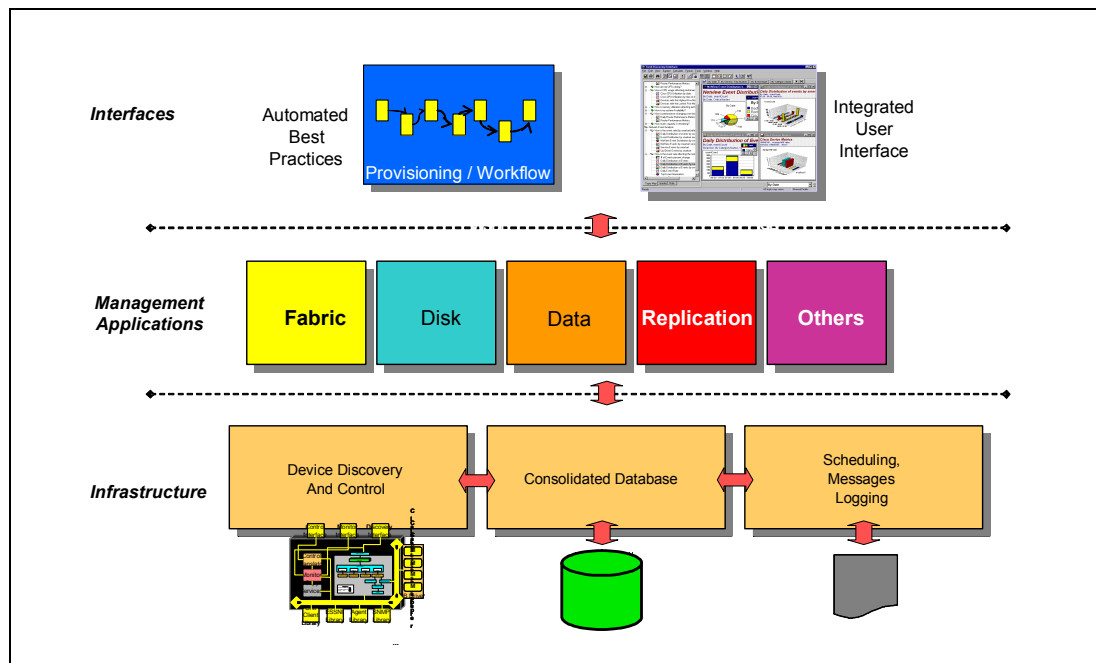


Figure 18-3 IBM TPC v4.1.1 logical structure

The infrastructure layer consists of basic functions such as messaging, scheduling, logging, device discovery, and a consolidated database shared by all components of Tivoli Storage Productivity to ensure consistent operation and performance.

The application layer consists of core Tivoli Storage Productivity Center management functions, that rely on the common base infrastructure to provide various options of storage or data management. These application components are most often associated with the product components that make up the product suite, such as fabric management, disk management, replication management, and data management.

The interface layer presents integration points for the products that make up the suite. The integrated graphical user interface (GUI) brings together product and component functions into a single representation that seamlessly interacts with the components to centralize the tasks for planning, monitoring, configuring, reporting, topology viewing, and problem resolving.

Physical structure

IBM Tivoli Storage Productivity Center is comprised of the following components:

- ▶ A data component: Data Manager
- ▶ A disk component: Disk Manager
- ▶ A fabric component: Fabric Manager
- ▶ A tape component: Tape Manager.

Figure 18-4 shows the Tivoli Storage Productivity Center V4.1.1 physical structure.

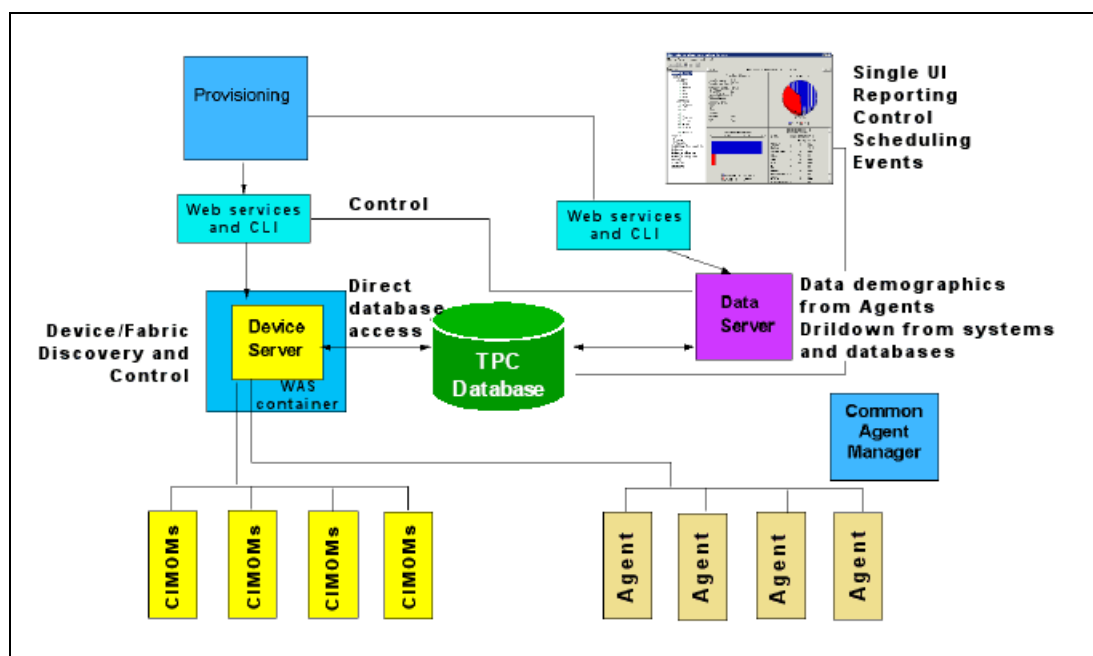


Figure 18-4 TPC V4.1.1 physical structure

The Data server is the control point for product scheduling functions, configuration, event information, reporting, and GUI support. It coordinates communication with agents and data collection from agents that scan file systems and databases to gather storage demographics and populate the database with results. Automated actions can be defined to perform file system extension, data deletion, and backup or archiving or event reporting when defined thresholds are encountered. The data server is the primary contact point for GUI user interface functions. It also includes functions that schedule data collection and discovery for the device server.

The device server component discovers, gathers information from, analyzes performance of, and controls storage subsystems and SAN fabrics. It coordinates communication with agents and data collection from agents that scan SAN fabrics.

The single database instance serves as the repository for all Tivoli Storage Productivity Center components.

The Data agents and Fabric agents gather host, application, and SAN fabric information and send this information to the data server or device server.

The GUI allows you to enter information or receive information for all Tivoli Storage Productivity Center components.

The command-line interface (CLI) allows you to issue commands for major Tivoli Storage Productivity Center functions.

For more details, see 18.5, “IBM Tivoli Storage Productivity Center management applications” on page 490.

18.4 Standards and protocols used in IBM Tivoli Storage Productivity Center

TPC was built upon storage industry standards. This section presents an overview the standards used within the various TPC components.

For more information about standards and protocols used in the IBM Tivoli Storage Productivity Center architecture, see the *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

18.4.1 Common Information Model and Web-Based Enterprise Management

Web-Based Enterprise Management (WBEM) is an initiative of the Distributed Management Task Force (DMTF) with the objective to enable the management of complex IT environments. It defines a set of management and Internet standard technologies to unify the management of complex IT environments.

The three main conceptual elements of the WBEM initiative are:

- ▶ Common Information Model (CIM): A formal object-oriented modeling language that is used to describe the management aspects of systems.
- ▶ xmlCIM: The syntax used to describe CIM declarations and messages used by the CIM protocol.
- ▶ Hypertext Transfer Protocol (HTTP)
- ▶ Hypertext Transfer Protocol over Secure Socket Layer (HTTPS)

HTTP and HTTPS are used as a way to enable communication between a management application and a device that both use CIM.

The CIM Agent provides a means by which a device can be managed by common building blocks rather than proprietary software. If a device is CIM-compliant, software that is also CIM-compliant can manage the device. Using CIM, you can perform tasks in a consistent manner across devices and vendors.

The CIM/WBEM architecture defines the following elements:

- ▶ Agent code or CIM Agent:
An open-systems standard that interprets CIM requests and responses as they transfer between the client application and the device. The Agent is normally embedded into a device, which can be hardware or software. When not embedded (which is the case for devices that are not CIM-ready such as the DS4000/DS5000), a device provider (usually provided by the device manufacturer) is required.
- ▶ CIM Object Manager (CIMOM):
The common conceptual framework for data management that receives, validates, and authenticates the CIM requests from the client application (such as TPC for disk). It then directs the requests to the appropriate component or a device provider.
- ▶ Client application or CIM Client:
A storage management program, such as Tivoli Storage Productivity Center, that initiates CIM requests to the CIM Agent for the device. A CIM Client can reside anywhere in the network, because it uses HTTP to talk to CIM Object Managers and Agents.

- ▶ Device or CIM Managed Object:

A Managed Object is a hardware or software component that can be managed by a management application by using CIM.

- ▶ Device provider:

A device-specific handler that serves as a plug-in for the CIMOM. That is, the CIMOM uses the handler to interface with the device.

Note: The terms *CIM Agent* and *CIMOM* are often used interchangeably. At this time, few devices come with an integrated CIM Agent. Most devices need a external CIMOM for CIM to enable management applications (CIM Clients) to talk to the device.

For more information, see the website:

<http://www.dmtf.org/standards/wbem/>

18.4.2 Storage Management Initiative - Specification

The Storage Networking Industry Association (SNIA) has fully adopted and enhanced the CIM for Storage Management in its Storage Management Initiative - Specification (SMI-S). SMI-S was launched in mid-2002 to create and develop a universal open interface for managing storage devices, including storage networks.

The idea behind SMI-S is to standardize the management interfaces so that management applications can use these and provide cross-device management. This means that a newly introduced device can be immediately managed as it conforms to the standards. TPC for disk uses that standard.

The API provided by SMI-S makes it easier for management applications to collect data from devices. The specification also provides consistent definitions of the data fields in the API, which in theory will give consistent views and comparisons between devices. In theory, because vendors are still free to interpret the fields as they see fit. So, it is important to know your devices and understand just what is meant by the performance metrics returned by the device.

TPC uses SMI-S for:

- ▶ Storage Array discovery and reporting
- ▶ Storage Array provisioning
- ▶ Storage Array performance monitoring and reporting
- ▶ Fabric (switch port) performance monitoring and reporting
- ▶ Tape discovery and reporting
- ▶ SVC discovery and reporting (IBM only).

Note: For a list of IBM and non-IBM SMI-S Compliant devices that TPC supports, see the website:

<http://www.snia.org/ctp/conformingproviders/>

18.4.3 Service Location Protocol

The Service Location Protocol (SLP) is an IETF standard that SLP provides as a scalable framework for the discovery and selection of network services.

SLP enables the discovery and selection of generic services, which can range in function from hardware services such as those for printers or fax machines, to software services such

as those for file servers, email servers, web servers, databases, or any other possible services that are accessible through an IP network.

Traditionally, to use a particular service, an end-user or client application needs to supply the host name or network IP address of that service. With SLP, however, the user or client no longer needs to know individual host names or IP addresses (for the most part). Instead, the user or client can search the network for the desired service type and an optional set of qualifying attributes.

The SLP architecture includes three major components:

- ▶ **Service agent (SA):**
A process working on the behalf of one or more network services to broadcast the services.
- ▶ **User agent (UA):**
A process working on the behalf of the user to establish contact with a network service. The UA retrieves network service information from the service agents or directory agents.
- ▶ **Directory agent (DA):**
A process that collects network service broadcasts.

The SA and UA are required components in an SLP environment, where the SLP DA is optional.

The SMI-S specification introduces SLP as the method for the management applications (the CIM clients) to locate managed objects. In SLP, an SA is used to report to UAs that a service that has been registered with the SA is available.

18.5 IBM Tivoli Storage Productivity Center management applications

Tivoli Storage Productivity Center provides four general types of support: data management, disk management, fabric management, and tape management. The following table summarizes the support provided by each function. To obtain each function, you must install the appropriate package.

The Tivoli Storage Productivity Center functions are shown in Table 18-1.

Table 18-1 Tivoli Storage Productivity Center functions

Data management	Disk management	Fabric management	Tape management
Host-centric: ▶ Discovery ▶ Monitoring ▶ File system extension ▶ Enterprise- wide reporting Application-centric: ▶ Monitor DB2, Oracle, SQL Server, Sybase ▶ Discovery ▶ Monitoring ▶ Chargeback	For storage subsystems: ▶ Discovery ▶ Monitoring ▶ Configuration (for example, creating volumes) ▶ Performance management	For fabrics: ▶ Discovery ▶ Monitoring ▶ Configuration (for example, zoning) ▶ Performance management	For tape libraries: ▶ Discovery ▶ Monitoring

Each license determines the functions that are accessible in the user interface.

For more information about IBM Tivoli Storage Productivity Center management applications, see these publications:

- ▶ *IBM Tivoli Storage Productivity Center Workflow User's Guide*, SC27-2341-01
- ▶ *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

18.5.1 Data Manager

This topic introduces the Data Manager application and discusses its architecture and key features.

For details, see *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

Data Manager Overview

Data Manager is a comprehensive file and capacity management solution for heterogeneous storage environments. Data Manager includes enterprise-wide reporting and monitoring, policy-based management and automated capacity provisioning for Direct Attached Storage (DAS), network attached storage (NAS), and SAN environments.

Data Manager helps you improve storage utilization, plan for future capacity, and ensure availability by providing storage on demand for file systems. Use Data Manager to perform the following functions:

- ▶ Discover and monitor disks, partitions, shared directories, and servers
- ▶ Monitor and report on capacity and utilization across platforms to help you to identify trends and prevent problems
- ▶ Provides a wide variety of standardized reports about file systems and storage infrastructure to track usage and availability
- ▶ Provide file analysis across platforms to help you to identify and reclaim space used by non-essential files
- ▶ Provide policy-based management and automated capacity provisioning for file systems when user-defined thresholds are reached.

Using these functions, Data Manager helps you lower storage costs by:

- ▶ Improving storage utilization
- ▶ Enabling intelligent capacity planning
- ▶ Helping you manage more storage with the same staff
- ▶ Supporting application availability through computer uptime reporting and application database monitoring.

Information collected by Data Manager helps you understand what is really going on with data on your servers and in your storage environment. View when files are created, accessed, and modified, and by what group or user. This type of information helps system administrators map storage resources to the consumers of the resource. The ability to map storage consumption to storage hardware has become increasingly important as the size of open systems environments has increased.

In addition to understanding the current consumption and usage of data within the enterprise, Data Manager tracks the information over time. Not only does this historical view of storage consumption and utilization show usage trends over time, the system administrator can also see a projected use of storage in the future. System administrators can prepare for the need

to purchase additional capacity in a planned proactive manner rather than reacting to out-of-space emergencies.

Use Data Manager policy functions to help you evaluate and control the usage and status of your enterprise storage management assets. Because Data Manager is policy-based, it has autonomic self-healing capabilities that can detect potential problems and automatically make adjustments based on the policies and actions you have established. Use the capability to provision storage based upon storage policies to expand a file system, and to allocate storage to a volume.

For example, Data Manager can notify you when your system is running out of disk space or warn you of impending storage hardware failure. By alerting you to these and other issues related to your stored data, Data Manager enables you to prevent unnecessary system and application downtime.

Data management

Every storage administrator likes to have an unlimited amount of storage. Unfortunately in the real world, this is not possible. The best you can hope for is to manage effectively the storage which is available, and accurately predict what is needed for the future. An important tool to assist this is knowing how much storage is actually needed, and to do that, you need to know what you actually need to store. Data stored electronically essentially can be divided into the categories shown in Figure 18-5 on page 492

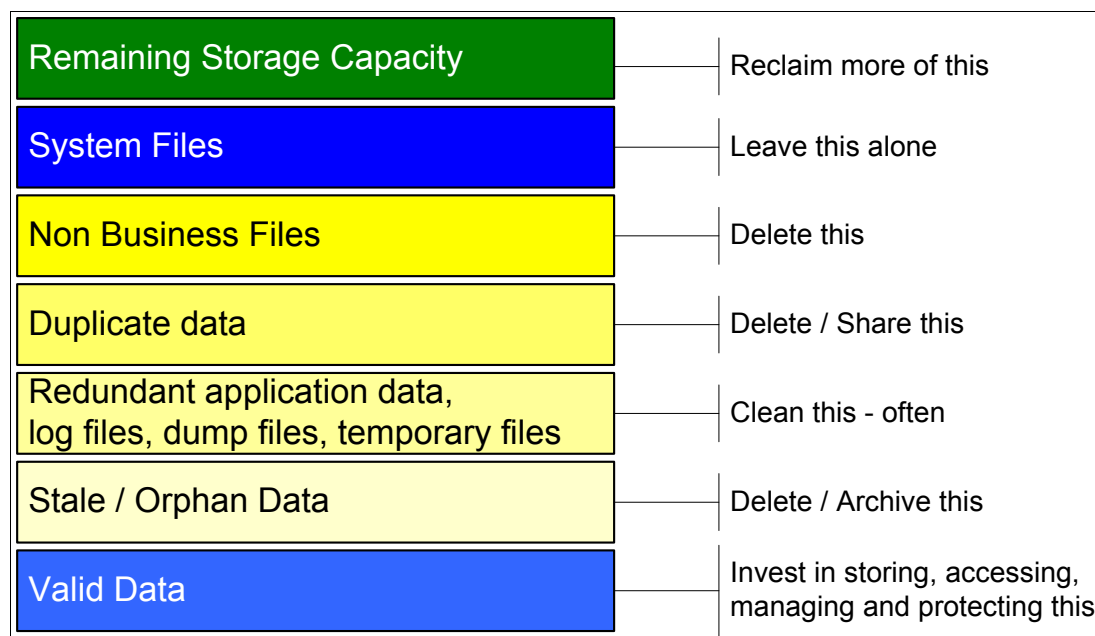


Figure 18-5 Types of data typically found in an enterprise and what do with them

Data Manager helps you identify how much of each type of data that you have, and automate actions to deal with each type.

Data Manager Functions

Data Manager provides you with a number of functions that enable you to collect information about and perform tasks against the storage resources within your enterprise. This section provides a brief overview of the main functions that can help you manage and automate capacity utilization of your file systems and databases.

Monitoring

Data Manager uses three methods to discover information about the assets in the storage environment:

► Pings:

A ping is a standard ICMP (Internet Control Message Protocol) ping which collects information about the availability of the storage assets in your storage environment. By running pings and viewing the results of those pings in Availability and System reports, you can monitor and report on the availability of your storage, both from a network point of view and from a computer uptime perspective. This allows you to see what percentage of the time the storage in your enterprise (or on particular server or group of servers), is off-network due to network problems or perhaps is unavailable due to system downtime.

You can define any number of pings to check the availability of various storage resources at various times. If an agent does not respond to a ping (or a predefined number of pings), you can set up an alert to take action. The actions can be one, any or all of SNMP Trap, TEC Event, Login Notification, Windows Event Log, UNIX Syslog, running a script, and email.

► Scans:

Use scans to collect statistics about the usage and trending of storage consumption within your environment. You can view the information collected by scans in Capacity, Usage, Usage Violations, Backup, and System reports, as well as use that information as part of quota analysis.

Scans are always directed against a Data agent and deliver very detailed information about the file systems, files, and RDBMS objects (instances, databases, devices, table spaces, tables, indexes, data files, containers) of computers. Use the statistics gathered by scans to:

- View information about when storage resources are created, accessed, and modified and by what group or user
- Map the actual storage resource to the consumers of that resource
- Generate a historical view of storage consumption to help determine usage trends over time.

You can define any number of scans to gather information about various storage resources at various times.

► Probes:

Probes are used to gather information about the assets and system resources of monitored servers, such as processor count and speed, memory size, disk count and size, file systems, and so on. Probes also gather information about the files, instances, logs and objects that makeup the monitored databases. Data collected by probes is used in Asset, Capacity, Storage Subsystems, and System reports.

Alerting

An alert defines an action to be performed if a particular event occurs or condition is found. Alerts in Data Manager can be set on computers, file systems, and directories. Alerts in Data Manager for Databases can be set on instances, database-table spaces, and tables. Alerts can tell you, for instance, if a file system or database is approaching capacity.

An Alert will register in the Alert log, plus you can also define one, any or all of the following actions to be performed in addition:

- Send an SNMP trap
- Send an alert to the Tivoli Enterprise Console (TEC)
- Generate a login notification next time a specified user logs in to TPC for Data

- ▶ Write out alert messages to the OS log (Windows Event Log or UNIX Syslog)
- ▶ Run a script in response to an alert (you can execute any third-party tools for actions, such as archiving, backup and recovery, or provisioning)
- ▶ Send an email or a page indicating the nature of the alert.

Reporting

Reporting in Data Manager is very rich, with over 300 predefined views, and the capability to customize those standard views, save the custom report, and add it to your menu for scheduled or ad hoc reports. You can also create your own individual reports according to particular needs and set them to run as needed, or in batch (regularly). Reports can be produced in table format or a variety of charting (graph) views. You can export reports to a PDF file, an HTML file, a CSV file or a formatted text file for external usage.

Reports are generated against data already in the repository. A common practice is to schedule scans and probes just before running reports.

Reporting can be done at almost any level in the system, from the enterprise down to a specific entity and any level in between.

Data Manager allows you to group information about similar entities (disk, file systems, and so on) from various servers or business units into a Summary Report so that business and technology administrators can manage an enterprise infrastructure. Or, you can summarize information from a specific server. The flexibility and choice of configuration is entirely up to the administrator.

Major reporting categories for file systems and databases are:

- ▶ Assets Reporting uses the data collected by probes to build a hardware inventory of the storage assets. You can then navigate through a hierarchical view of the assets by drilling down through computers, controllers, disks, file systems, directories, and exports. For databases, information about instances, databases, tables, and data files is presented for reporting.
- ▶ Availability Reporting shows responses to ping jobs, as well as computer uptime.
- ▶ Capacity Reporting shows how much storage capacity is installed, how much of the installed capacity is being used and how much is available for future growth. Reporting is done by disk and file system, and for databases, by database.
- ▶ Usage Reporting shows the usage and growth of storage consumption, grouped by file system, and computers, individual users, or enterprise-wide.
- ▶ Usage Violation Reporting shows violations to the corporate storage usage policies, as defined through Data Manager. Violations are either of Quota (defining how much storage a user or group of users is allowed) or Constraint (defining which file types, owners and file sizes are allowed on a computer or storage entity). You can define what action is to be taken when a violation is detected; for example, SNMPT trap, email, or running a user-written script.
- ▶ Backup Reporting identifies the most at risk files in your enterprise that are not backed up properly.

Policy Management

Data Manager can enable you to define and enforce storage policies through user-defined alerts, quotas, and constraints, notifying the user by email, pager, or the event log, or a systems management console for events such as when a quota has been exceeded or a constraint violated.

However, finding a problem is not enough. You need a way to fix problems, or potential problems as they are discovered. Data Manager can provide automated solutions through event management.

Policy Management function provides the following features:

- ▶ **Setting storage usage quotas:**

Use quotas to set limits on the amount of storage space that a user or group of users can consume. You can use Data Manager and Data Manager for Databases to set quotas at the file system, computer, network, database, table space, and instance levels.

- ▶ **Storage usage for Network Appliance quotas:**

Use Network Appliance quotas to import quotas from Network Appliance file servers and determine how and when you will be alerted to the hard limits defined in those quotas.

- ▶ **Setting file constraints:**

Use constraints to define the acceptable and unacceptable file types, file sizes, and file owners for a NAS filer or a set of NAS filers in your environment. You can also use constraints to request an IBM Tivoli Storage Manager archive and backup of the largest violating files identified by a constraint.

- ▶ **Using file system extensions:**

Use file system extensions to create additional space in the local file systems of managed hosts. You can extend file systems manually, or set up policies to do it automatically.

File system extension is supported for JFS file systems running on AIX 5.1 and VxFS file systems running on Sun Solaris 2.8. The AIX JFS2 file system is not supported.

File system extension/LUN provisioning jobs are not supported under Solaris 9.

- ▶ **LUN provisioning:**

By default, IBM Tivoli Storage Productivity Center will attempt to extend a file system in its local volume group. If there is not enough space in the local volume group and an TotalStorage Enterprise Storage Server subsystem is available, additional LUNs can be provisioned for file system extension. Provisioning consists of creating and formatting a new LUN according to the parameters specified in the file system extension policy, and assigning the new LUN to the managed host. You can enable and configure LUN provisioning using the Provisioning tab in the File System Extension window.

- ▶ **Scheduling script-based actions:**

Use scheduled actions to run user-defined scripts against selected computers and computer groups.

You can access the scheduled actions facility using either Data Manager or Data Manager for Chargeback. The actions that you schedule to run are based on user-defined scripts, allowing you to use any third-party tools for actions, such as recovery or provisioning. You can schedule scripts to run:

- Immediately
- Once at a specified time and date
- Repeatedly according to a schedule you specify.

- ▶ **Archive and backup functions:**

Data Manager provides a method within the alerting function of a constraint by which you can automatically invoke a IBM Tivoli Storage Manager archive or backup to run against the files that violate that constraint.

Let us say that, as a result of running a report, you find a large number of files that have not been accessed in over two years. Looking at the cumulative sizes of these files, you

determine that they are taking up space that is better utilized elsewhere. You can archive and delete these files to reclaim that space. If, at a time in the future, those files are needed, they can be retrieved from the archive. The result of the archive job is more free space without data loss.

Data Manager for Databases

Data Manager for Databases provides a set of policy-driven automated tools for managing storage capacity, availability, events, performance and assets in your relational databases. It can help you identify, evaluate, control and predict the storage needs of Relational Database Management Systems (RDBMSs), which include Oracle, Sybase SQL Server, Microsoft SQL Server, and UDB/DB2.

Data Manager for Databases is policy-based and through its autonomic self-healing capabilities it can detect potential problems and automatically make adjustments based on the policies and actions you have established. For example, it can notify you when your database tables are running out of storage space or warn you of a dropped table space. By alerting you to these and other issues related to your stored data, it enables you to prevent unnecessary system and application downtime.

Table 18-2 lists the RDBMS objects whose storage usage is monitored by Data Manager for Databases.

Table 18-2 RDBMS objects

Oracle	SQL Server	Sybase	DB2/UDB
Instances	Instances	Instances	Instances
Databases	Databases	Devices	Databases
Table spaces	Data files	Databases	Table spaces
Data files	Tables	Fragments	Containers
Tables	Indexes	Tables	Tables
Indexes	Log Files	Indexes	Indexes
Extents			
Segments			
Redo Log files			
Control Files			
Archive Log Directories			

Data Manager for Databases collects information about RDBMS storage assets and performs tasks against the RDBMS storage assets. Methods for managing and automating capacity utilization of your databases can help you:

- ▶ Monitor storage assets associated with enterprise-wide databases and notification of potential problems before they occur
- ▶ Make intelligent capacity management decisions based on current and trended historical data
- ▶ See how much storage is being consumed by users, groups of users and OS's within the database

- ▶ Create policy-based management for databases when user-defined threshold is reached
- ▶ View storage utilization management from a database and application perspective.

Attention: Data Manager for Databases does not currently support the monitoring of clustered database applications.

Data Manager for Chargeback

Data Manager for Chargeback uses the storage usage information gathered by Data Manager and Data Manager for Databases to generate invoices that chargeback for storage usage.

With Data Manager for Chargeback you can collect usage information about a departmental, group, or user level. You can allocate costs by storage usage by user, disk capacity by computer, table space, or file system/physical device. You can create cost centers by creating user, computer, or table space groups, allowing organization to chargeback individuals or business units for their storage usage. By understanding the costs associated with existing and future storage usage, you can improve the use of that storage and reduce the costs associated with its maintenances and upgrades.

Using Data Manager, you can run monitoring jobs that gather detailed statistics about the usage and trending of the storage consumed by the users and user groups within your organization. Using Data Manager for Databases, you can run monitoring jobs that gather detailed statistics about the usage and trending of the storage consumed within RDBMS Instances by the users and user groups within your organization.

In addition to providing invoicing for storage usage, Data Manager for Chargeback also integrates with the chargeback systems already implemented in your environment. It provides you with a higher level, application-specific CIMS output format which can be imported into CIMS applications. With Data Manager for Chargeback, you can export chargeback data for direct import into CIMS, giving you the ability to integrate your storage usage data with other enterprise chargeback information and processes.

18.5.2 Disk Manager

This topic introduces the Disk Manager application and discusses its architecture and key features.

For details, see *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

18.5.3 Disk Manager Overview

As a component of the IBM Tivoli Storage Productivity Center, Disk Manager helps you manage SANs and heterogeneous storage from a single console.

With Disk Manager, you can manage network storage components based on SMI-S, such as:

- ▶ IBM System Storage SAN Volume Controller
- ▶ IBM TotalStorage Enterprise Storage Server (TotalStorage Enterprise Storage Server)
- ▶ Tivoli Storage Productivity Center disk systems (DS4000, DS5000, DS6000, and DS8000 series)
- ▶ IBM XIV Storage System
- ▶ Other storage subsystems that support the SMI-S standards.

Note: Certain disk management features are not supported for XIV Storage System as described in the planning for XIV Storage System information in the *IBM Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication Installation and Configuration Guide*.

Disk Manager can:

- ▶ Collect and store performance data and provides alerts
- ▶ Provide graphical performance reports
- ▶ Help optimize storage allocation
- ▶ Provide volume contention analysis.

Through the use of data collection, setting of thresholds and use of performance reports, performance can be monitored for the TotalStorage Enterprise Storage Server, DS4000, DS5000, DS6000, DS8000, SAN Volume Controller, and any other storage subsystem that supports the SMI-S block server performance subprofile. The performance function starts with the data collection task, which captures performance statistics for the devices and stores the data in the database.

After data collection is done, you can set thresholds to identify exceptions for certain performance metrics, depending on the type of device. Threshold checking is performed during data collection, and when performance is outside the specified bounds, alerts can be generated.

After performance data has been collected, you can configure Disk Manager to present graphical or text reports on the historical performance behavior of specified devices, or of specific internal components of those devices. The performance reports provide information about past period performance metrics or current performance in graphical form.

Device discovery is done using the Service Location Protocol (SLP), as specified by SMI-S. Configuration of the discovered devices is possible in conjunction with CIM agents associated with those devices, using the standard mechanisms defined in SMI-S. Disk Manager also gathers events, and can launch an element manager specific to each device.

Note: Tivoli Storage Productivity Center supports discovery of and reporting on both FB and CKD volumes (and pools and arrays), but can create only FB volumes. Tivoli Storage Productivity Center does not support the creation of CKD volumes.

Disk Manager functions

Disk Manager provides you with a number of functions that enable you to manage SANs and heterogeneous storage from a single console. This section provides a brief overview of the main functions that management application provides.

▶ Monitoring:

Monitoring functions run regularly scheduled or as one-time data collection jobs. These jobs gather detailed statistics about the storage subsystem performance. This data is later used in reports, probes, and so forth.

The rules of monitoring are:

- Threshold violations (exceptions) are always logged and saved in the TPC database
- The user can view the exceptions that were detected in the TPC alert log, or can generate an exception report. The exception report allows the user to drill-down into the details of individual exceptions.

► Alerting:

TPC for Disk allows configured alerts such as email, SNMP trap, SMTP, windows Event Log, run a script for all devices managed from TPC. Use the Alerting feature of Disk Manager to set alerts that notify you of storage-related events that occur within your environment. After you have defined the events or conditions for which you want to be alerted, you can let Disk Manager monitor your storage so that you do not have to.

► Reporting:

Disk Manager can provide performance monitoring, customization of thresholds based on your storage environment, and generation of events if thresholds are exceeded. In addition, Disk Manager can help the IT administrators select LUNs for better performance.

The reporting data can be displayed in table format, giving the actual values of the various collected metrics, or can be displayed in a graphical format, primarily as line-graphs. Reports can be displayed on recent data, historical data, or on exception conditions. Exception reports allow the user to view a list of any threshold violations in a given time range. They are usually displayed as bar graphs, showing the total number of threshold violations.

Disk Manager has defined the following performance reports:

- Storage Subsystem Performance
- Storage Subsystem Controller Performance
- Storage Subsystem Controller Cache Performance
- I/O Group Performance
- Array Performance
- Managed Disk Group Performance
- Port Performance
- Top 25 Volumes I/O Rate
- Top 25 Volumes Data Rate
- Top 25 Volumes Cache Hit
- Top 25 Volumes Response Time
- Top 25 Volumes Disk
- SAN Switch Report
- Switch Port Error Report
- Top 25 Switch Ports Ops Rate Report
- Top 25 Switch Ports Data Rate Report

However, the user can define his own reports:

- The By Controller report (user-defined) has all possible metrics available, and the user can select whatever they want. And they can save their customizations.
- They can make it act like a system-defined report (it's a "My Report") if they want.
- "Customizing" here means: selecting which components are in the report, which columns/metrics to include, what order to show them in. Reports can also be scheduled to run in a "batch" on a schedule and generate, say, HTML files that can be viewed later.

Performance management capabilities

Disk Manager has the following performance management capabilities:

- "By Volume", "By Array", "By Port", "By Subsystem" Reports
- By Volume, they are viewing a (textual) report of volume-level performance statistics across multiple volumes

- ▶ A set of default thresholds are in effect for all devices of a particular device type. In addition, thresholds can be user-defined for individual devices, overriding the default thresholds for those individual devices
- ▶ Thresholds can be enabled or disabled, and the upper (stress) and lower (idle) boundaries of the threshold can be set
- ▶ Performance data collection must be active for thresholds to be monitored
- ▶ All thresholds defined by PM are bi-level thresholds, meaning that two unique boundary conditions can be defined, one to indicate a warning condition, and another to indicate a critical condition
- ▶ Each threshold can allow a total of 4 boundary values to be defined, the critical stress, warning stress, warning idle, and critical idle boundaries, if so desired.

Configuring performance thresholds

You can use the Disk Manager to set performance thresholds for each device type. Setting thresholds for certain criteria enables Disk Manager to notify you when a certain threshold has been exceeded, so that you to take action before a critical event occurs.

You can specify what action must be taken when a threshold-exceeded condition occurs. The action might be to log the occurrence or to trigger an event. The threshold settings can vary by individual device.

Managing TagmaStore CIM agents with TPC for Disk

The TagmaStore CIM agents are provided by Hitachi Data Systems for the TagmaStore storage subsystem. The TagmaStore CIM agent collects information from the TagmaStore storage subsystem.

IBM Tivoli Storage Productivity Center now supports the Hitachi Data Systems TagmaStore CIM Agent 5.8. This version of the CIM Agent supports only the Array Profile and not the Storage Virtualizer Profile. However, Tivoli Storage Productivity Center will support the TagmaStore as a Storage Virtualizer. Tivoli Storage Productivity Center will be able to display information for virtual disks and local disks.

Tivoli Storage Productivity Center cannot provide correlation information between the virtual storage used by TagmaStore and the volumes created on the storage due to an existing limitation of the CIM agent 5.8 from Hitachi Data Systems. However, Tivoli Storage Productivity Center reports correctly display the correlation between volumes created on the local storage and the local disks. This limitation has no impact on the topology but it does affect several reports that show the correlation:

- ▶ **Data Manager → Reporting → Asset → By Storage Subsystem → HDS_device → Disks → Volumes** (will not show the relation between the disk and Volume or volume in the tree)
- ▶ **Disk Manager → Reporting → Storage Subsystem → Volume to Back-end Volume Assignment**
- ▶ **Disk Manager → Reporting → Storage Subsystem → Computer Views** (because Tivoli Storage Productivity Center cannot populate the information for volumes created on the virtual disks)

All volumes are created from a storage pool that is allocated from a primordial storage pool and an imported primordial pool. A volume cannot be created over both local and virtual extents.

For volume correlation, the host machine must have the Data agent installed and the TagmaStore device must be in the same SAN fabric. There must also be a zone configured in the active zoneset between the ports of the host machine and the ports of the TagmaStore device. The Fabric agent needs to be configured for the fabric to which the host is connected.

For back-end correlation, the TagmaStore device ports and back-end subsystem ports must be in the same zone and the back-end subsystem has assigned storage volumes to all ports of the TagmaStore device.

18.5.4 Fabric Manager

This topic introduces the Fabric Manager application and discusses its architecture and key features.

For details, see *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

Fabric Manager overview

Fabric Manager helps you manage the SAN fabric that connects the host systems and applications to the storage devices. It is a comprehensive management solution for multi-vendor SANs and includes automatic resource and topology discovery, monitoring and alerts, and zone control.

Fabric Manager is an enterprise scalable solution architected to ANSI SAN standards, allowing you to choose the products right for your environment. It helps you:

- ▶ Simplify the task of SAN management and configuration
- ▶ Ensure SAN availability
- ▶ Improve SAN return on investment.

Fabric Manager is able to do SAN fabric performance and capacity management reporting and monitoring. Zoning is one of many SAN security options. Zoning is the only configuration supported for this release. Along with subsystem based volume masking, it is almost universally employed to ensure that only systems that are authorized can access selected data. Zoning is provided by the switch at the port level, so that, for example, a host on port X can access the subsystem connected by port Y. When business needs change, often zones must change also.

Zone control is enabled from a centralized location. Existing zones can be discovered and the members that make up the zone can be viewed and modified by adding or deleting them. New zones can be created and existing zones can be deleted. Fabric Manager makes working with zones easy by providing a GUI that discovers SAN devices and makes them available to add or remove from zones as individual devices.

The ability to do switch performance and capacity management reporting and monitoring can help you to determine if more bandwidth is needed. Wide-area and local-area IP networks, and storage area networks (SANs) move data from one place to another. The management of the bandwidth is needed to continually monitor for link utilization and errors. Data needs to be gathered to tune resources, balance workloads, and do capacity planning.

With Fabric Manager, the comprehensive management of multi-vendor SANs can help simplify the IT infrastructure by providing automatic resource and topology discovery, monitoring, and alerts, and zone control. It brings all the sources of information about SAN topology and configuration into a single place, and creates topology mapping of the SANs. This topology offers both host-centric and device-centric views that can be displayed graphically. The SAN topology display tracks all topology and configuration changes through in-band, out-of-band, and SMI-S monitoring. Without this type of centralized topology,

information from a number of sources, such as Element Management tools, device logs, and SNMP traps, have to be continually monitored and manually correlated to determine the current SAN configuration and topology.

Fabric Manager provides you with the ability to view events happening in your SAN environment and records state changes. The events are displayed in a color-coded fashion and can be further customized to reflect organizational priorities. It will forward events signaling topology changes or updates to the IBM Tivoli Enterprise Console, another SNMP manager, or both. Source side filtering by the Fabric Manager enables you to control what events get sent to your centralized console. Filtering helps ensure that the administrators get the information they need and are not overwhelmed by extraneous information.

Fabric Manager functions

Fabric Manager provides you with a number of functions that enable you to discover and manage SAN components and devices, and the topology of the storage area network environment. You can monitor storage utilization on the SAN and determine the availability of SAN components.

This section provides a brief overview of the main functions that management application provides.

Discovery

The process of finding resources within an enterprise, including detection of network topology, is called *discovery*. Fabric Manager uses the following two methods to discover the network:

- ▶ **In-band:**

Fabrics are automatically discovered and the Fabric agents can provide information about the fabrics or switches to which the computers are attached. When Fabric agents (in-band agents) are installed on the computers attached to the fabric and these agents are registered with IBM Tivoli Storage Productivity Center, you do not have to perform a specific discovery step for these agents.

For example, if you have McDATA and QLogic switches and have Fabric agents installed on the computers that are attached to these switches, the switches are automatically discovered when you install the Fabric agents.

If you have a McDATA switch, ensure that the switch has the Open Systems Management Server (OSMS) feature enabled. OSMS is required for in-band discovery. The host control option needs to be enabled if you want to do zone control and if no CIM Agent for the fabric is present. See your McDATA documentation for information about these features and how to enable them.

In-band discovery provides this following information:

- Topology information for switches, connections, nodes, and ports of the devices connected to the fabric, zones, zone sets
- Host and device information (collected from in-band agents only)
- Basic information to identify the host and devices in the fabric.

- ▶ **Out-of-band:**

Monitors the switches, managed hosts, and storage over an IP network by standardized Management Information Base (MIB) files, which typically are loaded only onto the managed switches. Fabric Manager gathers SNMP-collected information from out-of-band agents.

Managing fabrics

A Storage Area Network *zone* is a grouping of multiple ports to form a virtual private storage network. A *zone set* is a collection of zones that belong to a single SAN and can be activated or deactivated as a single entity across all switches in the fabric. A zone set can contain one or more zones, and a zone might be a member of more than one zone set. A zone might contain one or more *zone members* (aliases), and an alias might belong to more than one zone. Ports that are members of a group or zone can communicate with each other, but are isolated from ports in other zones. Devices, however, can belong to more than one zone.

Using zoning, you can automatically or dynamically arrange fabric-connected devices into logical groups, or zones, across the physical fabric. Support for aliases is also provided. An alias is a name assigned to a device so that the name can be meaningful and familiar. Also, an alias can be a group of devices that can be managed together to make the zoning process easier.

To configure zones, you can use Fabric Manager, or the management application for the devices. Advantages to zoning through Fabric Manager are that with Fabric Manager you can perform zoning tasks from a single interface. Fabric Manager supports industry zone management standards, which means that you can work with devices from multiple vendors using consistent, familiar methods.

Topology viewer console lets you view zones in SANs. Fabric Manager lets you create, update, and delete zones and zone sets within a SAN. In addition, you can:

- ▶ Assign zones to zone sets
- ▶ Activate and deactivate zone sets
- ▶ View zone membership
- ▶ Add and remove zone members.

Topology views

The Topology Viewer provides an extended graphical topology view; a graphical representation of the physical and logical resources (for example, computers, fabrics, and storage subsystems) that have been discovered in the storage environment. In addition, the Topology Viewer depicts the relationships among resources (for example, the disks comprising a particular storage subsystem). Detailed, tabular information (for example, attributes of a disk) is also provided. With all the information that topology viewer provides, you can easily and more quickly monitor and troubleshoot your storage environment.

The Topology Viewer provides a central location to view a storage environment, quickly monitor and troubleshoot problems, and gain access to additional tasks and function within the TPC user interface without users losing their orientation to the environment.

Monitoring fabrics

With monitoring feature, you can monitor fabrics by following divisions:

- ▶ Groups:

A group represents a user-defined set of related objects against which you want to run monitoring jobs. An object represents a storage resource, such as a computer, fabric, storage subsystem, table space, and so on. For example, if you want to scan all of the SAN switches in your Payroll department, create a Fabric Group named Payroll and populate that group with the individual SAN switch from the Payroll department. Then, whenever you want to collect information about those switches, you can simply select the Fabric Group named Payroll when defining a monitoring job. As switches are added or removed from the Payroll department, you can simply update the Payroll Fabric Group to reflect those changes.

► Switch Performance Monitors:

You can monitor the performance of selected fabric switches. Fabric Manager can collect performance data for the switches and generate reports from the results of those monitors.

Report data is collected on data rates, operations rates, switch port errors, and other performance statistics.

Alerting

The alerting feature lets you define alerts for events that occur in your environment. You can define alerts as part of monitoring jobs, or use the alerting nodes under the Fabric Manager component. For example, you can define an alert that will send you an email when a switch goes offline or comes online. You can also define an alert to be triggered if one of your monitoring jobs fails. There are three type of Fabric Manager alerts:

► Fabric alerts:

You can use these alerts to monitor fabrics for changes such as these:

- Fabric Discovered: A fabric is discovered
- Fabric State Change: A fabric is missing or is rediscovered
- Fabric Status Change: Offline A fabric goes offline
- Fabric Status Change: Online A fabric comes online
- Fabric to Switch Change: A switch to fabric association is discovered, rediscovered, or missing
- Zone State Change: A zone is missing or is rediscovered
- Zone Discovered: A zone is discovered
- Zone Set State Change: A zone set in a fabric was activated, deactivated, or has gone missing
- Zone Set Discovered: A zone set is discovered
- Zone to Zone Alias Change: A zone to zone alias association is discovered, missing, or rediscovered
- Zone to Zone Member Change: A zone to zone member association is discovered, is missing, or is rediscovered
- Zone Set to Zone Change: A zone set to zone association is discovered, is missing, or is rediscovered
- New Connection: A new connection is detected
- Connection State Change: A connection is missing or rediscovered
- Zone Alias to Member Change: A zone alias to member association is discovered, rediscovered, or is missing

► Switch alerts:

You can use these alerts to monitor switches for changes in configuration or status such as these:

- Switch Discovered: A switch is discovered
- Switch State Change: A switch is missing or is rediscovered
- Switch Property Changes: A switch port is discovered, is missing, or is rediscovered
- Switch Status Change Offline: The status of the switch has changed from OK to a lower status. This does not necessarily mean that the switch has gone missing

- Switch Status Change Online: The status of the switch status of the switch has been upgraded to OK
- Switch Version Change: The version of a switch changes
- Switch to Port Change: A switch to switch port association is discovered, is missing, or is rediscovered
- Switch Blade Change: A switch module is discovered, is missing, or is rediscovered
- Switch Blade Change Offline: A switch module goes offline
- Switch Blade Change Online: A switch module comes online
- Total Port Data Rate Threshold: A total port data rate threshold is reached
- Port Receive Bandwidth Percentage Threshold: The threshold on the average port bandwidth utilization percentage for receive operations. The Port Send Utilization Percentage metric is checked against the threshold boundaries for each collection interval. This threshold uses the following default values: 85,75,-1,-1
- Port Send Bandwidth Percentage Threshold: The threshold on the average port bandwidth utilization percentage for send operations. The Port Send Utilization Percentage metric is checked against the threshold boundaries for each collection interval. This threshold uses the following default values: 85,75,-1,-1
- Link Failure Rate Threshold: A link failure rate threshold is reached
- Error Frame Rate Threshold: An error frame rate threshold is reached
- Total Port Packet Rate Threshold: A total port packet rate threshold is reached.
- Endpoint device alerts:

You can use these alerts to monitor endpoint devices for following conditions:

 - Endpoint Discovered: An endpoint device is discovered
 - Endpoint State Change: An endpoint device is missing or is rediscovered
 - Endpoint to Node Change: A peripheral entity to node association is missing, newly discovered, or rediscovered
 - Endpoint Version Change: The version for an endpoint device changes.

Reporting

In addition to performance data collection and performance thresholds, TPC allows the user to display the collected data and the recognized threshold exceptions in a user-friendly format. Data can be displayed in table format, giving the actual values of the various collected metrics, or can be displayed in a graphical format, primarily as line-graphs. Reports can be displayed on recent data, historical data, or on exception conditions. They allow the user to pick a device, a device component, to pick a particular metric, and to specify a time range, and the performance data for the given device component and the given metrics will be plotted over the given time range. Exception reports allow the user to view a list of any threshold violations in a given time range. They are usually displayed as bar graphs, showing the total number of threshold violations. There are predefined reports of Fabric Manager:

- SAN Switch Report
- Switch Port Error Report
- Top 25 Switch Ports Ops Rate Report
- Top 25 Switch Data Rate Report.

18.5.5 Tape Manager

This topic introduces the Tape Manager application and discusses its key features.

For details, see *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

Tape Manager overview

IBM Tivoli Storage Productivity Center Tape Manager helps you manage your SMI-S compliant tape libraries which are based on SMI-S 1.1 profile for tape libraries.

You must have at least Tape Operator authority to access the functions in the Tape Manager node. To view your authority level, open the Role to Group Mappings window in the **Administrative Services** → **Configuration node** in the navigation tree.

With Tape Manager you can:

- ▶ Discover SMI-S compliant tape libraries attached to your SAN fabrics
- ▶ View tape library information
- ▶ Monitor tape libraries by probing for data and viewing alerts
- ▶ Create and view asset and capacity reports using data collected from your tape libraries.

Use the topology viewer to view information about tape libraries.

Tape library support

IBM Tivoli Storage Productivity Center Tape Manager provides support for tape library management. Tape Manager is present and available in the GUI whenever the Device Server is installed. No separate license is required. Using Tape Manager, you can discover tape libraries, group libraries to monitor multiple libraries, view alerts generated by tape libraries, and launch tape library element managers.

Tape Manager supports the following tape libraries:

- ▶ IBM 3584
- ▶ IBM 3494 (limited support).

Managing tape libraries

Tape Manager provides the capability to manage your tape libraries. For this release, tape library management capabilities are limited to support for the IBM 3584 tape library (full support) and the IBM 3494 tape library (partial support).

The tape library management functions allow you to:

- ▶ Discover IBM 3584 and IBM 3494 tape libraries attached to your SAN fabrics
- ▶ Create groups of tape libraries with similar attributes, making it easier to perform management tasks on multiple tape libraries at one time
- ▶ Create and view asset reports using data collected from your tape libraries; these reports can include such information as the library identifier and the status of the library
- ▶ Create and view capacity reports using data collected from your tape libraries; these reports can include such information as the number of drives in a library, the number of tape slots, the number of occupied tape slots
- ▶ View alerts generated by tape libraries; these alerts include read or write errors, media changer errors, and maintenance requests
- ▶ Launch the element manager from several places in the product user interface.

18.6 Performance analysis with Storage Optimizer

This section provides brief description of the Storage Optimizer and how to use it through Tivoli Storage Productivity Center to analyze storage subsystems to identify performance bottlenecks.

For more information about performance analysis with Storage Optimizer, see the *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

18.6.1 Storage Optimizer overview

Analyzing large amounts of performance data and interpreting the results is a time consuming exercise and requires a certain set of skills. The Storage Optimizer changes all this by automating the task of interpreting the vast amounts of configuration and performance data, which allows even the most junior storage administrator the ability to identify areas for improvement within their storage networks.

The Storage Optimizer uses data in the IBM Tivoli Storage Productivity Center database to analyze your storage subsystems to identify performance bottlenecks, and propose changes to improve performance. This topic lists the supported subsystems and describes the general steps for using Storage Optimizer, as well as a look into how the Storage Optimizer functions.

The Storage Optimizer also helps you develop storage migration or storage consolidation plans, and helps you plan for the growth of your storage infrastructure, based on the various performance behaviors.

Two main aspects to the Storage Optimizer

There are two main aspects to the Storage Optimizer that we discuss in more detail later on in this topic. The first portion of work is called the Analysis, and the second portion is the actual Optimization itself. At a high level, the Analysis is responsible for collecting, aggregating and predicting the utilizations of the storage subsystem infrastructure. The Optimization is responsible for utilizing the results of the Analysis to help put together a plan for potential migrations and consolidations, to improve overall utilization of the infrastructure.

Note: The Storage Optimizer does not actually perform any migrations or make any modifications to subsystem configurations. It is the Storage Optimizer's primary purpose to provide you with a performance analysis and optimization report that you can choose to implement at your own discretion.

Who ultimately uses the Storage Optimizer

The Storage Optimizer is used by the following staff members:

- ▶ Junior storage administrators that do not have much experience with interpreting existing performance reports
 - ▶ Senior storage administrators looking to validate a hunch or what other reports indicate to be the case
 - ▶ All storage administrators looking for a plan to consolidate/retire storage hardware
- This is useful when looking at retiring obsolete subsystems, or alternatively, when looking at consolidating onto a subset of subsystems to improve their density, thus leaving various subsystems shut down (saving money on maintenance, depreciation, and energy costs) until such a time that their data growth actually catches up to their hardware capacities.
- ▶ Anyone with a *Standard Edition* licence.

The Storage Optimizer uses a unique “Heat Map” style of display that makes narrowing in on “Hot Spots” (bottlenecks) much easier than having to analyze huge amounts of raw performance data to identify problem areas.

The two primary use-cases are:

- ▶ Migration Scenario: Provides guidelines to neutralize the “hot spots”

- Retirement Scenario: Provides guidelines to retire selected pools (move all volumes away from the selected pools).

Important: To use the Storage Optimizer, you must have a IBM Tivoli Storage Productivity Center Standard Edition license.

18.6.2 How the Storage Optimizer works

The first process of the Storage Optimizer is the *analysis*. The analysis process uses the Disk Magic¹ tool to produce *predicted* utilizations based on actual, measured performance metrics that are collected for that device by Tivoli Storage Productivity Center. These predicted utilizations are provided on a *per storage pool* basis.

What do we mean by “utilizations”? The analysis reports on four unique aspects of the internals of a storage subsystem and how these are specific to a particular pool, and produces a percentage that represents how much of the total capacity of that component is being used by the measured workload and configuration.

The analysis predicts these utilizations:

- HDD (hard disk drive):

This is the estimated overall “saturation” of the physical disks in a given subsystem. This number is impacted by the RPM speed, capacity and workload being measured. This does not apply to situations where the subsystem is a virtualizer with no internal disk (such as the SAN Volume Controller). This utilization is closely related to the “Disk Utilization” metric in the *By Array* performance reports, but not exactly the same, because this number comes from Disk Magic.

- HA (host adapter):

Sometimes known as a fibre adapter or fibre card, this is the estimated “saturation” of the ports on the subsystem. This number is impacted by the number of total ports, maximum hardware bandwidth, and workload being measured. You can see raw performance metrics related to this utilization in the *By Port* performance reports.

- DA (device adapter):

Sometimes known as a Redundant Array of Independent Disks (RAID) controller, this is the estimated “saturation” of the RAID controllers of the subsystem. This number is impacted by the workload being measured and logical configuration of disks to pools.

- SMP (controller):

This is the estimated “saturation” of the processor complex of the subsystem. This number is impacted by the model of subsystem and the workload being measured.

Note: If you are analyzing large numbers of subsystems, it is best to schedule a time to run the Storage Optimizer when processor demand is at a minimum.

¹ Disk Magic is an IBM licensed component from a company called IntelliMagic.

Additionally, there are two utilization numbers provided by the Storage Optimizer that are somewhat unique in their computation:

► Utilization:

This overall “utilization” number is computed as MAX(HDD, HA, DA, SMP), meaning it represents the largest value of the four other components – it does not represent a separate component.

► Space:

This utilization number is unique in that it is not *predicted*, it is measured. This is provided to give an indication of how full a particular pool has become, in terms of allocated-capacity versus total-capacity.

After the analysis is completed, the optimization capabilities can be used to migrate or consolidate the storage infrastructure. *Migration* is the ability to alleviate hot spots by relocating “hot” volumes in the “hot” pools to “colder” pools and subsystems. *Consolidation* is the ability to identify possibilities for reclaiming poorly utilized space.

18.6.3 Important performance metrics

The Storage Optimizer looks at eleven raw metrics to feed into the Disk Magic models:

- Read I/O Rate (normal)
- Read I/O Rate (sequential)
- Read I/O Rate (overall)
- Write I/O Rate (normal)
- Write I/O Rate (sequential)
- Write I/O Rate (overall)
- Read Cache Hits Percentage (overall)
- Write Cache Hits Percentage (overall)
- Cache to Disk Transfer Rate
- Read Transfer Size
- Write Transfer Size.

18.6.4 Supported subsystems or applications

The following IBM storage subsystems or applications are supported, see Table 18-3

Table 18-3 Support Matrix

Support Matrix	FC Drives	SATA Drives	SAS Drives	Solid State Drives
DS3000	No	No	No	No
DS4000	Yes	No	No	No
DS5000	No	No	No	No
DS6000	Yes	No	No	No
DS8000	Yes	No	No	No
SVC	Yes	No	No	No
XIV	No	No	No	No

No special firmware or CIMOM level requirements are introduced by this feature and the supported TPC levels will be sufficient.

Note: Non-IBM subsystems are not supported by the Storage Optimizer.

You can analyze the front-end of an SVC that has non-IBM back-end storage, however no guidelines will be offered for these non-IBM back-end subsystems because there are no Disk Magic models for those devices. The analysis will be useful when looking at the “Hot Spots” and determining where the bottlenecks lie on the SVC.

18.7 IBM Tivoli Provisioning Manager

This section provides brief description of IBM Tivoli Provisioning Manager.

For more information about performance analysis with Storage Optimizer, see the *IBM Tivoli Storage Productivity Center User's Guide*, SC27-2338-01.

18.7.1 Tivoli Provisioning Manager overview

IBM Tivoli Provisioning Manager, built on a Service Oriented Architecture (SOA)², enhances usability for executing changes while keeping server and desktop software compliant. Tivoli Provisioning Manager helps organizations with provisioning, configuration and maintenance of servers and virtual servers, operating systems, middleware, applications, storage and network devices acting as routers, switches, firewalls, and load balancers. With this release of the Tivoli Provisioning Manager, users can benefit from significant ease of use enhancements including a new graphical user interface, better reporting and streamlined use scenarios that provide more out of the box value.

Key functional areas of enhancement include:

- ▶ Improved discovery and reporting of datacenter resources
- ▶ Operating system imaging and bare metal provisioning
- ▶ Software distribution and patch management over a scalable secure infrastructure
- ▶ Automated deployment of servers through software templates
- ▶ Simplification and automation of datacenter tasks
- ▶ Compliance reporting and remediation.

These capabilities allow companies to manage the complete lifecycle of datacenter and distributed resources from initial provisioning to patching and configuration maintenance to resource repurposing or end of life. Through the automation of tasks, IT administrators can manage more resources with fewer people, allowing key skilled people to focus on activities that grow the business rather than serve to maintain it.

IBM Tivoli Provisioning Manager also contains a number of other new features designed to simplify use and improve time-to-value. These features include a graphical user interface with role based views designed to simplify change execution tasks and pre-canned reports that can be easily modified to fit user needs.

² SOA is an architectural approach to building applications and enables rapid development and deployment of new services to meet business needs.

IBM Tivoli Provisioning Manager core functions:

- ▶ Discovery of Resources
- ▶ Operating system provisioning
- ▶ Software deployment and configuration
- ▶ Automating Tasks through best practice automation workflows
- ▶ Auditing, Compliance and Remediation
- ▶ Common management Infrastructure for datacenter or distributed software management
- ▶ Storage Provisioning
- ▶ Change and Configuration Management.

18.7.2 Tivoli Provisioning Manager workflow

The next two figures show how IBM Tivoli Provisioning Manager makes it easier to provide storage. Figure 18-6 shows the workflow that is necessary to provide a new disk to a server without using IBM Tivoli Provisioning Manager.

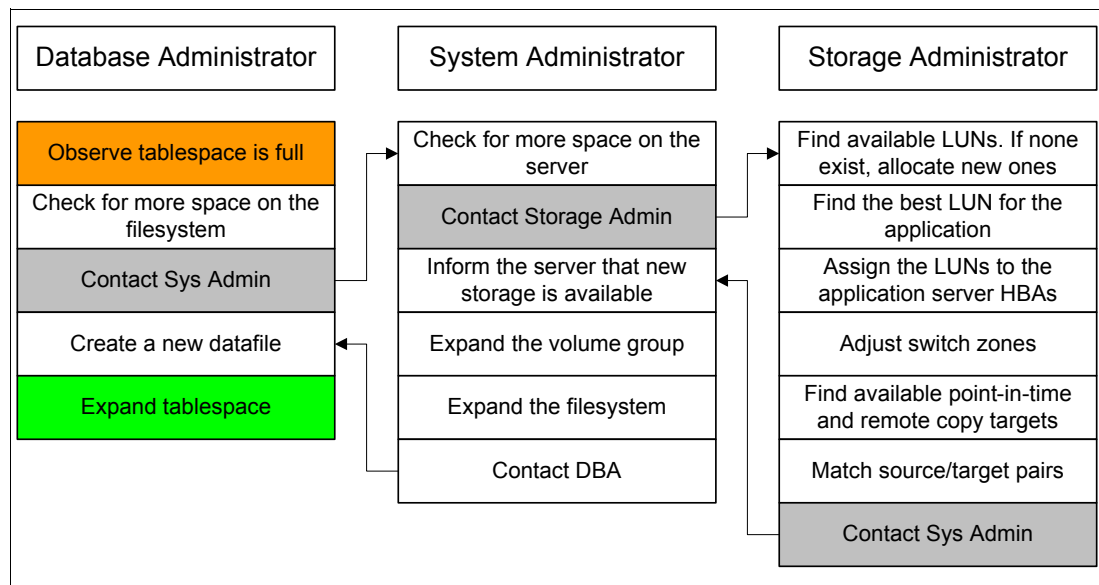


Figure 18-6 Storage provisioning: Manually

Manual Storage provisioning includes the following steps:

1. The database administrator (DBA) will find out there is a need for more space on the file system because a table space is going to be full. The DBA will have to contact the system administrator who will have to contact the storage administrator, if there is no more space available on the server.
2. The storage administrator will have to allocate a new LUN, if none exists, assign the LUN, adjust the zoning, prepare copy targets, and match source and target pairs, because the database requires mirrored data. Finally, the storage administrator will have to contact the system administrator again when the LUN is ready.
3. The system administrator will have to add the new LUN to the server, expand the volume group, and expand the file system. Finally, the system administrator will have to contact the database administrator.
4. The database administrator now can create a new data file and expand the table space.

Figure 18-7 shows the workflow to do the same thing if you have Tivoli Provisioning Manager supporting you in this task. After defining the tasks within Tivoli Provisioning Manager once, the storage administrator just has to initiate the process.

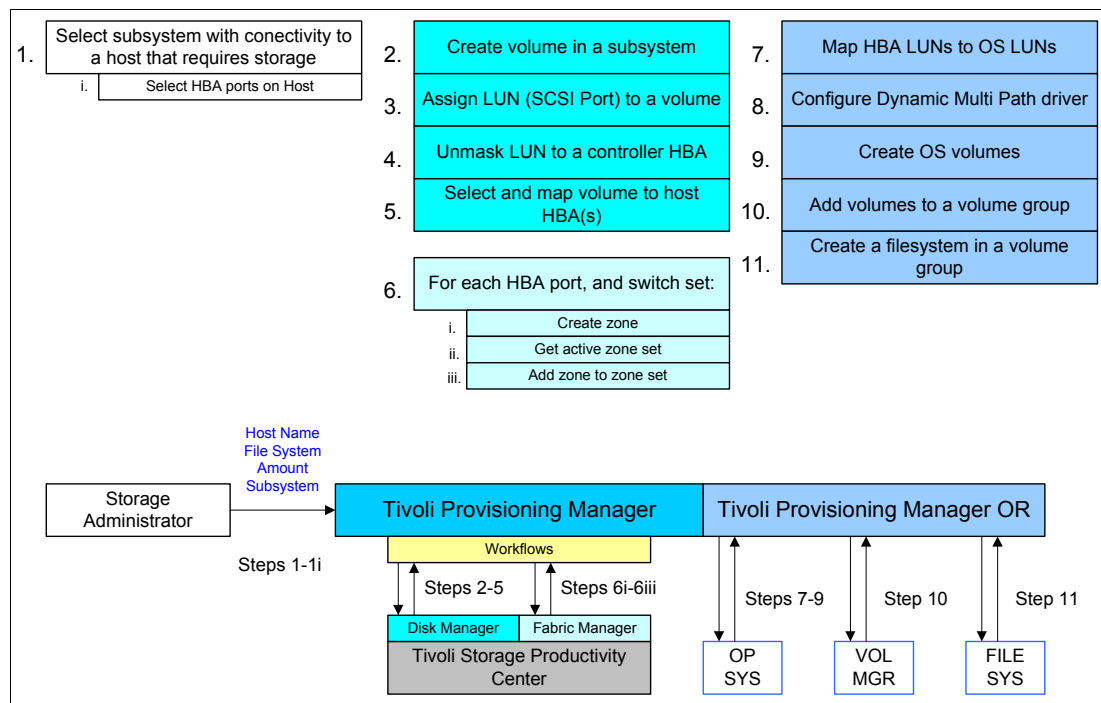


Figure 18-7 Storage provisioning supported by Tivoli Provisioning Manager

18.8 Tivoli Storage Productivity Center components and packages

This topic provides information about the IBM Tivoli Storage Productivity Center components and packages. For more information about TPC individual components and packages, see the *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

Use Table 18-4 to compare specific functions and data sources that are available in each license.

Table 18-4 TPC package comparison

Data source	Functionality	Required TPC license
Hypervisor		
	Probe	Data Edition
Storage Resource Agent		
	Host Probe	Basic Edition
Storage Resource Agent / Fabric Agent		
	Fabric Probe	Basic Edition
	Fabric Provisioning	Basic Edition
	Fabric Health Monitoring	Basic Edition

Data source	Functionality	Required TPC license
Data Agent		
	NetApp/NAS Probe	Data Edition
	Database Probe	Data Edition
	Deep Filesystem Scans	Data Edition
	Batch Reporting	Data Edition
Storage Resource Agent / Data Agent		
	Planning	Standard Edition
	Analytics	Standard Edition
Fibre Channel Switch		
	Probe	Basic Edition
	Provisioning	Basic Edition
	Health Monitoring	Basic Edition
	Performance Monitoring	Standard Edition
	Planning	Standard Edition
	Analytics	Standard Edition
Storage Subsystem		
	Probe	Basic Edition
	Provisioning	Basic Edition
	Health Monitoring	Basic Edition
	Performance Monitoring	Disk Edition
	Planning	Standard Edition
	Analytics	Standard Edition

18.8.1 IBM Tivoli Storage Productivity Center Basic Edition

IBM Tivoli Storage Productivity Center Basic Edition is focused on providing basic device management services for IBM System Storage DS3000, DS4000, DS5000, DS6000, and DS8000 products, IBM SAN Volume Controller and heterogeneous storage environments. This is a management option available with IBM Tivoli Storage hardware acquisitions. This tool provides storage administrators a simple way to conduct device management for multiple storage arrays and SAN fabric components from a single integrated console that also is the base of operations for the IBM Tivoli Storage Productivity Center suite. Productivity Center Basic Edition also does discovery and asset management of tape libraries, specifically IBM 3494 and 3584 Tape Libraries.

Key features include:

- ▶ Discovery and configuration of supported devices
- ▶ Event gathering, error logging, and launch of device element managers
- ▶ Provisioning support, including LUN allocation and assignment
- ▶ Basic Asset and Capacity Reporting
- ▶ Topology viewer: End to End Storage Topology View and Health Console
- ▶ Simple upgrade path to IBM Tivoli Storage Productivity Center for Disk, Data or Fabric

IBM Tivoli Storage Productivity Center Basic Edition enables device configuration and management of SAN-attached devices from a single console. Basic Edition simplifies the complexity of managing multiple SAN-attached storage devices. It enables you to manage SANs and heterogeneous storage from a single console.

Basic Edition enables you to manage network storage components based on SMI-S, such as:

- ▶ IBM System Storage SAN Volume Controller
- ▶ IBM System Storage disk systems (DS3000, DS4000, DS5000, DS6000, and DS8000 series)
- ▶ Other storage systems that support SMI-S standards.

Device discovery is performed by Service Location Protocol (SLP), as specified by SMI-S. Configuration of the discovered devices is possible in conjunction with CIM agents associated with those devices, using the standard mechanisms defined in SMI-S. Tivoli Storage Productivity Center gathers events and can start an element manager specific to each discovered device.

For SAN fabrics, Tivoli Storage Productivity Center simplifies the management and improves the availability of the SAN environment. You can monitor and report on SAN resources and use a single location for zone control. Tivoli Storage Productivity Center discovers existing zones and zone members and allows you to modify or delete them. In addition, you can create new zones. Support for aliases is also provided.

Tivoli Storage Productivity Center gives you the ability to view events happening in your SAN environment and record state changes. Events related to topology changes or updates can be forwarded to the IBM Tivoli Enterprise Console, to another manager that uses Simple Network Manager Protocol (SNMP), or to both.

For the SAN fabric, Tivoli Storage Productivity Center supports host bus adapters (HBAs), disk systems, tape systems, SAN switches, routers, and gateways.

Tip: IBM Tivoli Storage Productivity Center Basic Edition includes basic disk, fabric, tape, and data management functions, but does not include chargeback, database monitoring, and performance monitoring functions.

18.8.2 IBM Tivoli Storage Productivity Center for Data

TPC for Data is an advanced Storage Resource Management (SRM) tool for storage environments that provides a set of policy-driven automated features for Data for managing storage capacity, availability, events, data assets, and green initiatives including DAS, NAS, and SAN technologies.

Key features include the following capabilities:

- ▶ Utilize and optimize existing storage resources and perform storage management with a high level of control
- ▶ Improve storage utilization
- ▶ Enable intelligent capacity planning
- ▶ Manage more storage with the same staff
- ▶ Support high application availability
- ▶ Detect failed backup and archive tasks
- ▶ Provide enterprise-wide reports on capacity by role-based management
- ▶ Assist customers in support of data classification, ILM assessments, and ITIL® storage practices.

Tip: IBM Tivoli Storage Productivity Center for Data includes data management, basic tape, disk, and fabric management, database monitoring, and chargeback functions, but does not include performance monitoring functions.

18.8.3 IBM Tivoli Storage Productivity Center for Disk

IBM Tivoli Storage Productivity Center for Disk is designed to enable device configuration and management of SAN-attached devices from a single console. In addition, it also includes performance capabilities to help monitor and manage the performance of the devices' disks. IBM Tivoli Storage Productivity Center for Disk is designed to help simplify the complexity of managing multiple SAN-attached storage devices.

Productivity Manager for Disk is designed to enable the IT administrator to:

- ▶ Enable proactive performance management by provide an IT administrator with a single, integrated console for the performance management of IBM System Storage devices.
- ▶ Monitor performance metrics across multiple storage subsystems from a single console.
- ▶ Designed to allow administrators to monitor metrics, such as I/O rates and cache utilization, and support optimization of storage through the identification of the best LUNs across multiple storage subsystems.
- ▶ Monitor and analyze performance statistics for storage systems to measure services levels by storing received performance statistics into database tables for later use, and analyze and generate reports on monitored devices for display in central administrative console.
- ▶ Receive timely alerts to enable event action based on customer policies by setting performance thresholds for the devices based on performance metrics and the generation of alerts when those thresholds are exceeded.
- ▶ Provides a Performance Optimization Engine that can help reduce service times of resource-constrained application by an average of 48% and up to a maximum of 90% using a heat map that provides storage IO utilization reports that can be used to determine hotspots in the environment and generate plans for migration to distribute workload.

Key benefits include:

- ▶ Designed to help you potentially improve storage return on investment by keeping SANs operational reliably and dependably
- ▶ Designed to help reduce storage administration costs by simplifying the management of complex SANs
- ▶ Designed to offer continuous real-time monitoring and fault identification to improve SAN availability.

Tip: IBM Tivoli Storage Productivity Center for Disk includes basic disk, fabric, tape, and data management functions and storage system performance monitoring, but does not include fabric performance monitoring, chargeback, and database monitoring functions.

18.8.4 IBM Tivoli Storage Productivity Center For Replication

This section provides an overview of the IBM Tivoli Storage Productivity Center (TPC) for Replication, describes the key concepts necessary to use the product and its components, and contains several scenarios that illustrate how to perform specific types of replication.

For more information about the TPC for Replication, see *IBM Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication Installation and Configuration Guide*, SC27-2337-01.

TPC for Replication overview

IBM Tivoli Storage Productivity Center For Replication can help simplify and automate the configuration of your replication environment. The basic functions of Tivoli Storage Productivity Center (TPC) for Replication are designed to provide management of the advanced copy services: IBM FlashCopy, Metro Mirror and Global Mirror capabilities for the IBM System Storage DS8000, IBM System Storage DS6000 and the IBM System Storage SAN Volume Controller (SVC).

It is designed to:

- ▶ Automate the configuration of IBM DS8000, DS6000, and the IBM SAN Volume Controller advanced copy services features
- ▶ Monitor the progress of the copy services so you can verify the amount of replication that has been done as well as the amount of time needed to complete the replication
- ▶ Manage and coordinate the copy operations to ensure successful completion from your source volumes to your disaster recovery volumes:
 - Flash Copy
 - Metro Mirror
 - Global Mirror
 - Metro Global Mirror.
- ▶ Execute automated failover to keep your critical data online and available to your users even if your primary site fails. When the primary site comes back on, the software manages failback to the default configuration as well.

Attention: With Tivoli Storage Productivity Center V4.1.1, Tivoli Storage Productivity Center and Tivoli Storage Productivity Center for Replication, previously separate products, are now integrated. You can start the IBM Tivoli Storage Productivity Center for Replication user interface from within the Tivoli Storage Productivity Center user interface.

IBM Tivoli Storage Productivity Center for Replication is still a separate, stand-alone product on z/OS.

The IBM Tivoli Storage Productivity Center (TPC) for Replication V4 includes new capabilities and enhancements:

- ▶ Metro Global Mirror with Practice Volumes
- ▶ Metro Global Mirror with HyperSwap®
- ▶ Global Mirror with Practice Volume (Both Directions)
- ▶ Global Copy, to minimize application impact when initializing Metro Mirror sessions
- ▶ Replicating session progress indicators.

TPC for Replication architecture

The architecture of TPC for Replication is shown in Figure 18-8.

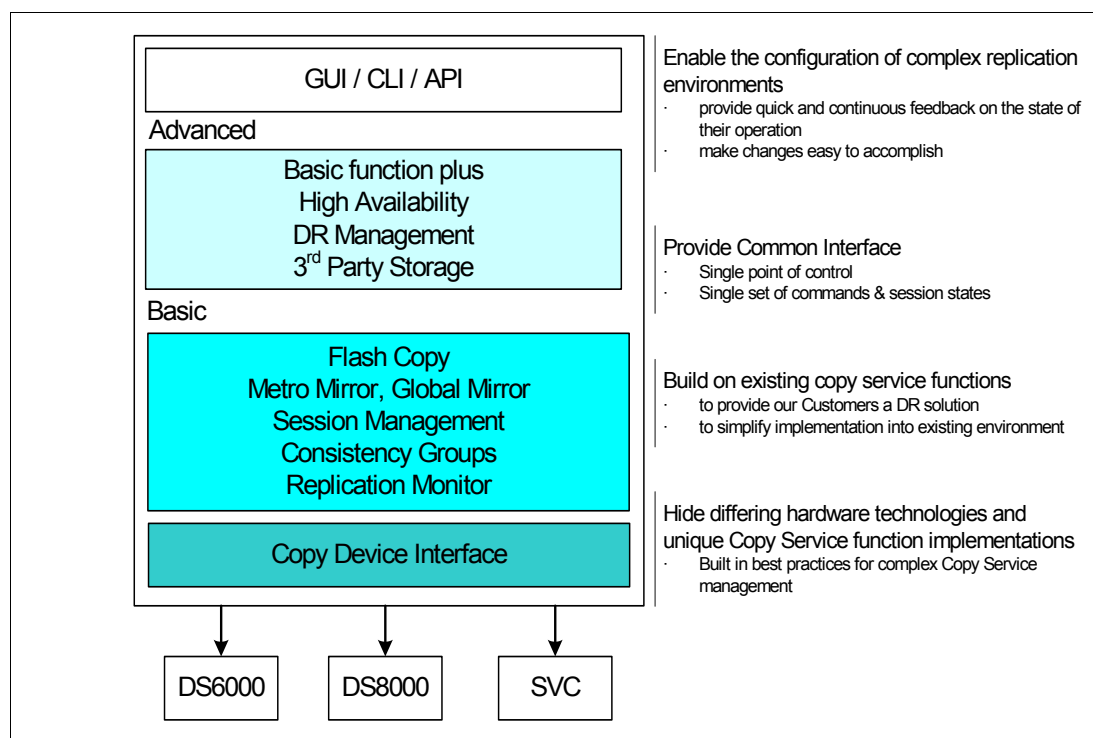


Figure 18-8 TPC for Replication architecture

Consistency management

To maintain integrity of data in any operating system, the sequence in which updates are being written is crucial. If that sequence is changed, data corruption will occur. The correct sequence must be maintained within a volume, across volumes, and across multiple storage systems.

In remote disk mirroring environments, the order of dependant writes across volumes and across multiple storage systems must be maintained at the remote location. During a real disaster (fire, flooding, earthquake), all the components in your environment will almost certainly not fail simultaneously. Failures will more likely be intermittent and gradual, and the disaster time frame will take many seconds or even minutes. This is known as a rolling disaster.

Advanced Copy Services integrated with automation software such as TPC for Replication can protect data from being a mirror of a “dying scenario” - that is, where multiple failures are occurring during a rolling disaster.

TPC for Replication FREEZES affected sessions at a known point instead of mirroring literally hundreds of time-offset failures in a short amount of time, as shown in Figure 18-9.

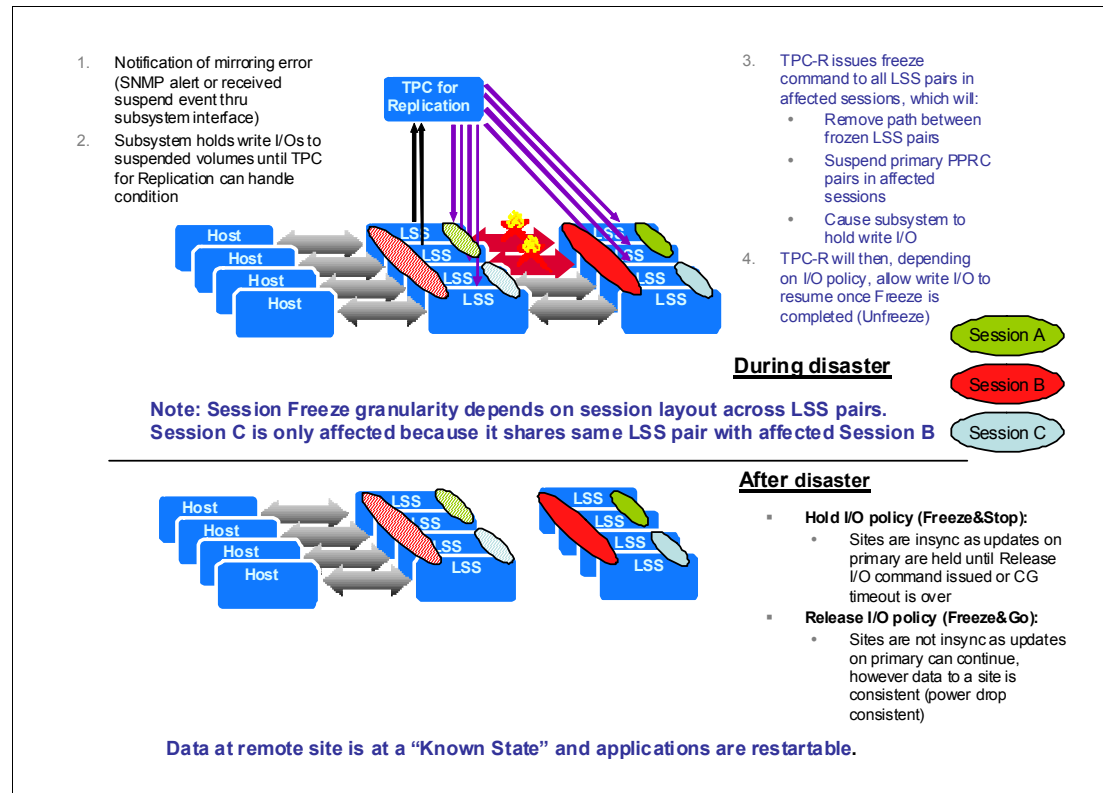


Figure 18-9 TPC for Replication monitoring and freeze capability

The heartbeat functionality on managed LSSes (logical subsystems) is used to avoid losing the alerting and freeze functionality of isolated subsystems (when the TCP/IP connection is broken). See Figure 18-10.

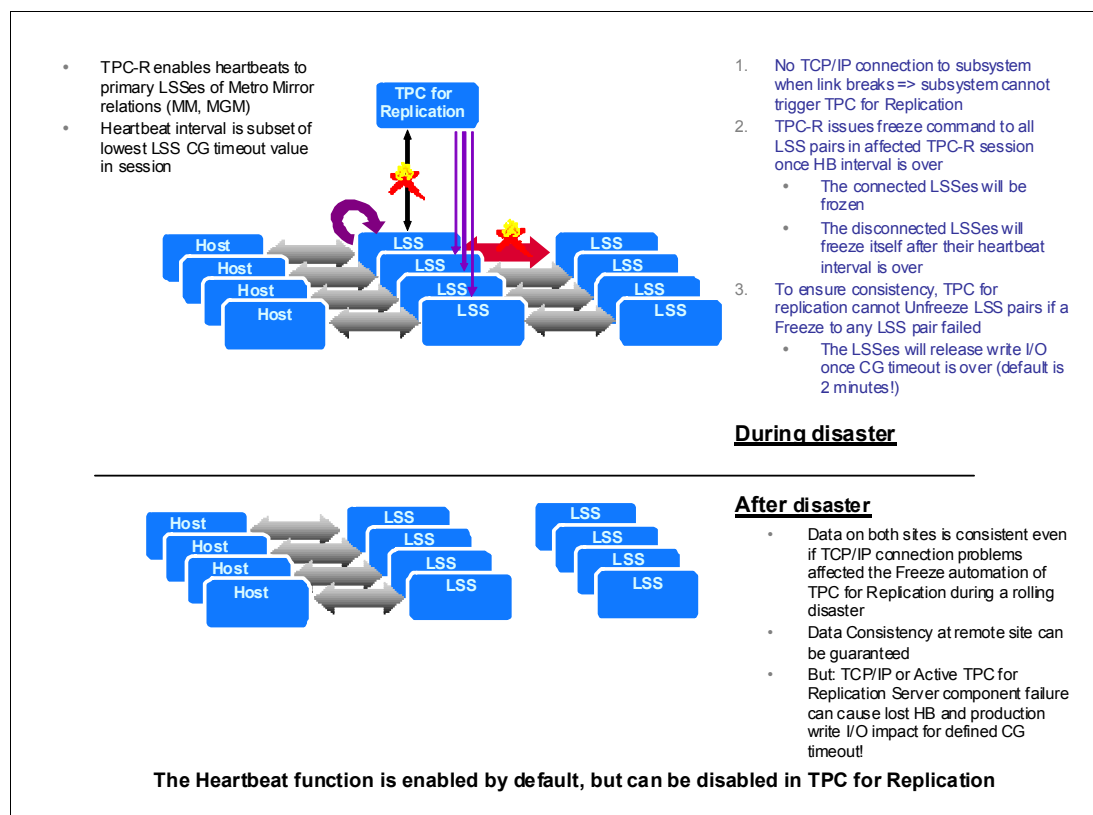


Figure 18-10 TPC for Replication Heartbeat on DS6000/DS8000

Replication management

TPC for Replication uses a simple graphical user interface (GUI) as well as a command line interface (CLI) to configure automation, manage ongoing activities and monitor progress of all key tasks. Different types of storage systems can be managed by a single integrated tool.

The functions are:

- ▶ Manage the entire Copy Services environment:
 - Add, modify and remove subsystems
 - Add and remove Paths
 - Add, modify and remove Sessions
 - Add and remove Copy Sets in Sessions
 - Manage Sessions with simple commands depending on Session type and state.
- ▶ Monitor the entire Copy Services environment:
 - Health Check screen
 - All sessions overview screen
 - Individual session views with Copy sets overview and details
 - Progress Bars.
- ▶ Diagnostics:
 - Error overview
 - Detail view of each error
 - Error messaging and history console.

Session and copy sets

A copy set is a set of volumes that contain copies of the same data, shown in Figure 18-11. All the volumes in a copy set are the same format (count key data (CKD) or fixed block) and size. In a replication session, the number of volumes in a copy set and the role that each volume in the copy set plays are determined by the copy type.

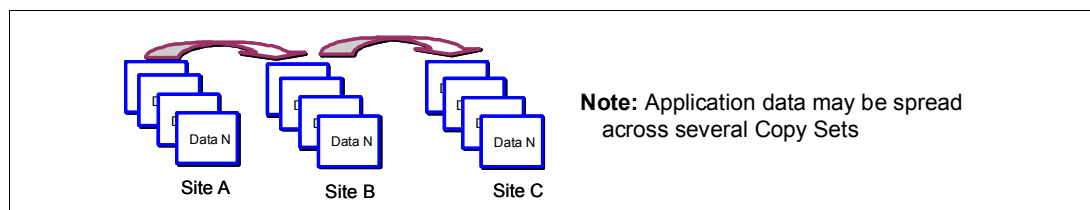


Figure 18-11 TPC for Replication copy set

A session is a container of multiple copy sets making up a consistency group, shown in Figure 18-12.

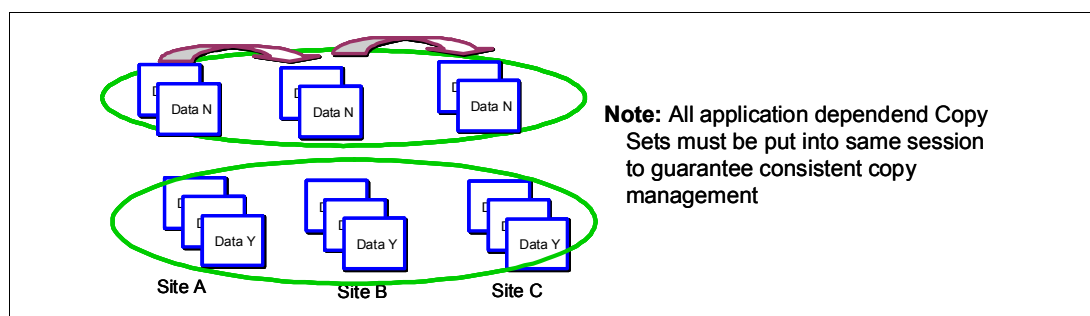


Figure 18-12 TPC for Replication sessions

Three types of sessions can be created, corresponding to the type of the copy service required:

- FlashCopy:

A FlashCopy is a point-in-time local copy. FlashCopy replication creates the target volume with a copy of the data that represents the data at the time the relationship was established. Both source and target can be referenced and updated independently.

- Metro Mirror:

A Metro Mirror session is synchronous remote replication. In this context, remote means the source is in one physical subsystem and the target is in a separate physical subsystem. Synchronous means that when a write is issued to change the source, the change is propagated to the target before the write is completely posted. This method of replication maintains identical data in both the source and target (unlike FlashCopy where the copy is static). The advantage of Metro Mirror is that when a disaster occurs, there is no data loss at the recovery site because both writes must complete before signaling completion of a write to the source application.

- **Global Mirror:**

A Global Mirror session is continuous asynchronous remote replication. As with Metro Mirror, remote means the source is in one physical subsystem and the target is in a separate subsystem. Asynchronous means that when a write is issued to the source copy, the change is not synchronously propagated and applied to the target. The updates to the target occur after control is given back to the application. Asynchronous replication greatly reduced (and might eliminate) performance impact of the copy on the application and is preferred for copying data at distances of greater than 300 km. However, because the write is not applied synchronously, there is a chance for data loss in the case of a disaster.

After you have decided which method of replication you want to use (copy type), you set up the copy sets. In a replication session, the number of volumes in a copy set and the role that each volume in the copy set plays is determined by the copy type.

The following terms are used in describing sessions:

- **Host volume:**

A host volume is a volume that an application, such as a database, reads data from and writes data to. A host volume can be the source for a copy function. A host volume can also be the target of the copy function, in which case you cannot write to the volume directly - it is only written by the copy service.

- **Journal volume:**

A journal volume holds a consistent copy of the data until a new consistent copy is formed. The journal volume restores the last consistent point during a recovery.

- **Role:**

A volume's role is the function it assumes in the copy set, and is composed of the intended use and, for Global Mirror and Metro Mirror, the volume's site location. Every volume in a copy set is assigned a role. A role can assume the functions of a host volume, journal volume, or target volume.

- **Role pair:**

A role pair is the association of two roles in a session. For example, in a Metro Mirror session, the role pair can be the association between volume roles Host 1 and Host 2.

- **Target volume:**

A target volume is an intermediate volume that receives data from a source. Depending on the session type, that data might or might not be consistent.

Copy sets as well as sessions can be created by using TPC for Replication wizards.

IBM Tivoli Storage Productivity Center for Replication Two Site Business Continuity (AIX, Linux, Windows, z/OS)

IBM TPC for Replication Two Site Business Continuity helps to obtain continuous availability and disaster recovery solutions by using point-in-time replication, which includes FlashCopy; and continuous replication, which includes Metro Mirror and Global Mirror.

Use IBM Tivoli Storage Productivity Center for Replication Two Site Business Continuity for the following tasks:

- ▶ Perform a FlashCopy for DS6000, and DS8000
- ▶ Perform a FlashCopy for SAN Volume Controller
- ▶ Perform a Global Mirror Single Direction for DS6000, and DS8000
- ▶ Perform a Global Mirror Either Direction with Two-Site Practice for DS6000, and DS8000
- ▶ Perform a Global Mirror Single Direction for SAN Volume Controller
- ▶ Perform a Global Mirror Failover/Failback for DS6000, and DS8000
- ▶ Perform a Global Mirror Failover/Failback for SAN Volume Controller
- ▶ Perform a Global Mirror Failover/Failback with Practice for DS6000, and DS8000
- ▶ Perform a Global Mirror Failover/Failback with Practice for SAN Volume Controller
- ▶ Perform a Metro Mirror Single Direction for SAN Volume Controller
- ▶ Perform a Metro Mirror Failover/Failback for SAN Volume Controller
- ▶ Perform a Metro Mirror Failover/Failback with Practice for SAN Volume Controller
- ▶ Perform a Metro Mirror or Global Copy for DS6000, and DS8000
- ▶ Perform a Metro Mirror Failover/Failback or Global Copy for DS6000, and DS8000
- ▶ Perform a Metro Mirror Failover/Failback with HyperSwap for DS6000, and DS8000
- ▶ Practice recovery procedures while maintaining disaster recover capability.

IBM Tivoli Storage Productivity Center for Replication Three Site Business Continuity (AIX, Linux, Windows, z/OS)

IBM TPC for Replication Three Site Business Continuity helps to obtain continuous availability and disaster recovery solutions by using point-in-time replication, which includes FlashCopy, and continuous replication, which includes Metro Mirror, Global Mirror, and Metro Global Mirror to secondary and tertiary sites.

Use IBM Tivoli Storage Productivity Center for Replication Three Site Business Continuity for the following tasks:

- ▶ Perform a FlashCopy for DS6000, and DS8000
- ▶ Perform a FlashCopy for SAN Volume Controller
- ▶ Perform a Global Mirror Single Direction for DS6000, and DS8000
- ▶ Perform a Global Mirror Either Direction with Two-Site Practice for DS6000, and DS8000
- ▶ Perform a Global Mirror Single Direction for SAN Volume Controller
- ▶ Perform a Global Mirror Failover/Failback for DS6000, and DS8000
- ▶ Perform a Global Mirror Failover/Failback for SAN Volume Controller
- ▶ Perform a Global Mirror Failover/Failback with Practice for DS6000, and DS8000
- ▶ Perform a Global Mirror Failover/Failback with Practice for SAN Volume Controller
- ▶ Perform a Metro Mirror Single Direction for SAN Volume Controller
- ▶ Perform a Metro Mirror Failover/Failback for SAN Volume Controller
- ▶ Perform a Metro Mirror Failover/Failback with Practice for SAN Volume Controller
- ▶ Perform a Metro Mirror or Global Copy for DS6000, and DS8000
- ▶ Perform a Metro Mirror Failover/Failback or Global Copy for DS6000, and DS8000
- ▶ Perform a Metro Mirror Failover/Failback with HyperSwap for DS6000, and DS8000
- ▶ Perform a Metro Global Mirror for DS8000 (with failover and failback)
- ▶ Perform a Metro Global Mirror for DS8000 (with failover and failback)
- ▶ Perform a Metro Global Mirror with Practice for DS8000 (with failover and failback)
- ▶ Perform a Global Metro Mirror with HyperSwap for DS8000
- ▶ Practice recovery procedures while maintaining disaster recover capability.

Important: IBM Tivoli Storage Productivity Center for Replication Three Site Business Continuity requires a BC license key.

IBM Tivoli Storage Productivity Center for Replication Basic Edition for System z (z/OS)

IBM TPC for Replication Basic Edition for System z provides a disaster-recovery capable solution, and helps protect from storage system failures.

Use IBM Tivoli Storage Productivity Center for Replication Basic Edition for System z with Basic HyperSwap for the following tasks:

- ▶ Monitoring for events that indicate a storage device has failed
- ▶ Determining if the failing storage device is part of a peer-to-peer remote copy (PPRC) pair
- ▶ Determining from policy, the action to be taken
- ▶ Ensuring that data consistency is not violated
- ▶ Swapping the I/O between the primary logical devices in the consistency group with the secondary logical devices in the consistency group (performing a HyperSwap for DS6000, and DS8000)
- ▶ Allowing only z/OS attached count key data (CKD) volumes to be added to the HyperSwap session.

IBM Tivoli Storage Productivity Center for Replication Basic Edition for System z provides only the HyperSwap session and not the functionality of the other IBM Tivoli Storage Productivity Center for Replication products.

Note: No license is required for IBM Tivoli Storage Productivity Center for Replication Basic Edition for System z. If you want to use IBM Tivoli Storage Productivity Center for Replication or IBM Tivoli Storage Productivity Center for Replication Three Site Business Continuity, a separate license for each of these products is required.

The z/OS HyperSwap license is required for running Basic HyperSwap.

IBM Tivoli Storage Productivity Center for Replication for System z (z/OS)

IBM Tivoli Storage Productivity Center for Replication for System z provides the functions as the IBM Tivoli Storage Productivity Center for Replication open systems products. It runs on System z servers, using a mixture of FICON and TCP/IP communications, to provide replication management of DS8000, and DS6000, regardless of type of data on them (ECKD™ or FBA).

IBM Tivoli Storage Productivity Center for Replication for System z provides:

- ▶ Volume protection to exclude any volumes from being used for disaster-protection copy operations
- ▶ Command prompting to confirm storage administrator actions before running the copy services commands
- ▶ User roles for administrative levels of access
- ▶ Site awareness to indicate site locations of the storage volumes and help assure copies are done correctly
- ▶ Metro Global Mirror support for the DS8000, providing failover and failback support, fast reestablishment of three-site mirroring, quick resynchronization of mirrored sites using incremental changes only, and data currency at the remote site.

Note: If you want to use IBM Tivoli Storage Productivity Center for Replication or IBM Tivoli Storage Productivity Center for Replication Three Site Business Continuity, a separate license for each of these features is required.

18.8.5 IBM Tivoli Storage Productivity Center Standard Edition

IBM Tivoli Storage Productivity Center Standard Edition is one of the industry's most comprehensive storage resource management solutions by combining the consolidated benefits of the following three components as one bundle:

- ▶ Tivoli Storage Productivity Center for Data
- ▶ Tivoli Storage Productivity Center for Disk
- ▶ Tivoli Storage Productivity Center Basic Edition.

In addition to the benefits and features of Data, Disk and Basic Edition, Productivity Center Standard Edition offers additional management, control and performance reporting for the Fibre Channel SAN infrastructure. These features include the following benefits:

- ▶ Provide automated device discovery, topology rendering, error detection and fault isolation, SAN error predictor, zone control, real-time monitoring and alerts, and event management for heterogeneous enterprise SAN environments.
- ▶ Help simplify the management and improve the availability of the SAN environment
- ▶ Provide automatic device discovery and allow multiple SAN views, including physical, logical and zone views.
- ▶ View and analyze multiple aspects of the storage environment, including capacity, utilization, assets and availability.
- ▶ Detect storage events and generate the appropriate alerts to the administrator.
- ▶ Provide a number of SAN availability features, including SAN error protector. An autonomic computing capability, SAN error predictor is designed to help predict SAN network problems before they become severe and impact data and application availability. This functionality incorporates predictive failure analysis into the storage network environment, designed to allow administrators to be proactive in managing SAN availability.
- ▶ Provide basic diagnostic capabilities to show which resources are impacted by an availability or performance issue in the SAN. Today, it can monitor performance at the port and switch level, and display this information in the common topology viewer for IBM Tivoli Storage Productivity Center.

Tip: IBM Tivoli Storage Productivity Center Standard Edition contains all the functions for data management, disk management, fabric management, and tape management, plus:

- ▶ Analytical functions
- ▶ Configuration Analysis
- ▶ Configuration History
- ▶ SAN Planner

18.8.6 IBM Tivoli Storage Productivity Center with Advanced Provisioning

IBM Tivoli Storage Productivity Center with Advanced Provisioning is an integrated storage capacity provisioning solution designed to simplify and automate complex cross-discipline tasks for provisioning storage capacity in the enterprise environment and is designed to move you from just-in-case provisioning to intelligent, automated on demand provisioning.

The IBM Tivoli Storage Productivity Center with Advanced Provisioning is comprised of:

- ▶ IBM Tivoli Provisioning Manager
- ▶ IBM Tivoli Storage Productivity Center
 - IBM Tivoli Storage Productivity Center for Data
 - IBM Tivoli Storage Productivity Center for Disk
 - IBM Tivoli Storage Productivity Center for Replication

Tivoli Storage Productivity Center with Advanced Provisioning is designed to allow you to automate and execute the numerous steps necessary to provision storage in a consistent manner, helping reduce or eliminate human error. Provisioning capabilities are provided by the Tivoli Storage Productivity Center, while Tivoli Provisioning Manager facilitates the automation of the provisioning tasks through the use of storage workflows.

Tivoli Storage Productivity Center with Advanced Provisioning includes workflow building blocks designed to help you easily build and customize your own solution workflows with minimal effort. The simplicity and flexibility of workflows helps enable you to automate at your own pace. Through the use of automated storage capacity provisioning workflows, helps allow you to:

- ▶ Provision storage capacity with consistency in a repeatable and auditable manner
- ▶ Automate storage capacity provisioning at your own pace with minimal effort
- ▶ Enable staff to manage cross discipline IT projects more effectively and efficiently
- ▶ Improve responsiveness to the storage needs of applications

18.9 IBM Tivoli Storage Productivity Center licenses

The following product licenses are available for IBM Tivoli Storage Productivity Center: Basic, Data, Disk, and Standard. Each license determines the functions that are accessible in the user interface.

For more information about licensing in the TPC, see the *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725-00.

Table 18-5 provides the list of functions that are included with each license to help you determine which one best meets your storage management needs.

Table 18-5 Tivoli Storage Productivity Center packages

Package/License	Function
IBM Tivoli Storage Productivity Center for Data	Includes data management, basic tape, disk, and fabric management, database monitoring, and chargeback functions, but does not include performance monitoring functions.
IBM Tivoli Storage Productivity Center for Disk	Includes basic disk, fabric, tape, and data management functions and storage system performance monitoring, but does not include fabric performance monitoring, chargeback, and database monitoring functions.
IBM Tivoli Storage Productivity Center for Replication (Two Site Business Continuity license and Three Site Business Continuity license)	IBM Tivoli Storage Productivity Center for Replication functions. When you install IBM Tivoli Storage Productivity Center for Replication, no licenses will be installed. You must install the Two Site or Three Site Business Continuity license after you install IBM Tivoli Storage Productivity Center for Replication.
IBM Tivoli Storage Productivity Center Basic Edition	Includes basic disk, fabric, tape, and data management functions, but does not include chargeback, database monitoring, and performance monitoring functions.
IBM Tivoli Storage Productivity Center Standard Edition	Contains all the functions for data management, disk management, fabric management, and tape management, plus: ▶ Analytical functions ▶ Data path explorer view (or context-sensitive performance analysis) ▶ Configuration Analysis ▶ Configuration History ▶ SAN Planner

Important: Each marketing package contains the complete Tivoli Storage Productivity Center product, and all functions are installed. However the various licenses that are shipped with each marketing package will limit the functionality available at the graphical user interface level.

Attention: IBM Tivoli Storage Productivity Center for Replication is still a separate, stand-alone product on z/OS.

18.10 IBM System Storage Productivity Center

This topic describes the features and functionalities of IBM System Storage Productivity Center (SSPC).

For more information about SSPC, see *IBM System Storage Productivity Center V1.4: Introduction and Planning Guide*, SC23-8824-06.

18.10.1 New with IBM System Storage Productivity Center Version 1.4

You can use this information to learn about the new features and enhancements that are included in IBM System Storage Productivity Center Version 1.4. SSPC 1.4 adds the following new features, functions, and enhancements.

Name change

IBM TotalStorage Productivity Center has been renamed to Tivoli Storage Productivity Center V4.1.1. All user interfaces, documentation, online help, and messages reflect the new name.

IBM Tivoli Storage Productivity Center

IBM Tivoli Storage Productivity Center Basic Edition 4.1.1 is preinstalled on the IBM System Storage Productivity Center server.

New Model 2805-MC4

System Storage Productivity Center Model 2805-MC4 includes a hard drive that is 10 GB larger than the previous model. System Storage Productivity Center Models 2805-MC2 and 2805-MC3 are no longer available for purchase.

IBM System Storage DS8000 R4.3 is supported

You can monitor System Storage DS8000 4.3 with System Storage Productivity Center. (You can also monitor previous releases.)

IBM System Storage SAN Volume Controller

IBM SAN Volume Controller Console 5.1.0 is preinstalled on the System Storage Productivity Center server. Because this level of the console no longer requires a CIM agent to communicate with the SAN Volume Controller, a CIM agent is not installed with the console. Instead, you can use the CIM agent that is embedded in the SAN Volume Controller hardware. To manage older levels of the SAN Volume Controller, install the corresponding CIM agent on the SSPC server. PuTTY remains installed on the System Storage Productivity Center and available for key generation.

IBM System Storage DS Storage Manager

A new version of the System Storage DS Storage Manager user interface is available for you to optionally install on the System Storage Productivity Center server or on a remote server. The DS Storage Manager 10.60 can manage the IBM DS3000, IBM DS4000, and IBM DS5000. With DS Storage Manager 10.60, when you use Tivoli Storage Productivity Center to add and discover a DS CIM agent you can launch the DS Storage Manager from the topology viewer, the Configuration Utility, or the Disk Manager of the Tivoli Storage Productivity Center.

IBM Java 1.5 is preinstalled

IBM Java is preinstalled and supports DS Storage Manager 10.60. You do not need to download Java from Sun Microsystems.

DS CIM agent management commands

The DS CIM agent management commands (DSCIMCLI) for 5.4.3 are preinstalled on the System Storage Productivity Center system.

Enhancements to Password Tool

The Password Tool that you use to change the DB2 user, WebSphere user, and DB2 service passwords now generates fewer messages and command windows.

18.10.2 System Storage Productivity Center overview

The IBM System Storage Productivity Center (SSPC) is an integrated hardware and software solution that will help you improve and centralize the management of your storage environment through the integration of products. It provides a single point from which to manage your storage systems.

The SSPC combines the power of a customized IBM System x server with preinstalled storage software that represents a significant point of centralized management. SSPC enhances several rudimentary device utilities for easier, more intuitive, context based administration and on the whole lowers resource overhead.

System Storage Productivity Center simplifies storage management by:

- ▶ Centralizing the management of storage network resources with IBM storage management software
- ▶ Providing greater synergy between storage management software and IBM storage devices
- ▶ Reducing the number of servers that are required to manage your software infrastructure
- ▶ Providing a simple migration path from basic device management to using storage management applications that provide higher-level functions.

The System Storage Productivity Center is a 1U, rack-mounted hardware appliance that has IBM Tivoli Storage Productivity Center Basic Edition and IBM Tivoli Storage Productivity Center for Replication installed on it.

Figure 18-13 shows an overview of how the SSPC, TPC, TPC for Replication, and the hardware and software for IBM System Storage DS8000, IBM System Storage SAN Volume Controller, IBM System Storage DS3000, IBM System Storage DS4000, and IBM System Storage D5000 all relate with each other.

Note: The IBM Tivoli Storage Productivity Center was formerly the IBM TotalStorage Productivity Center.

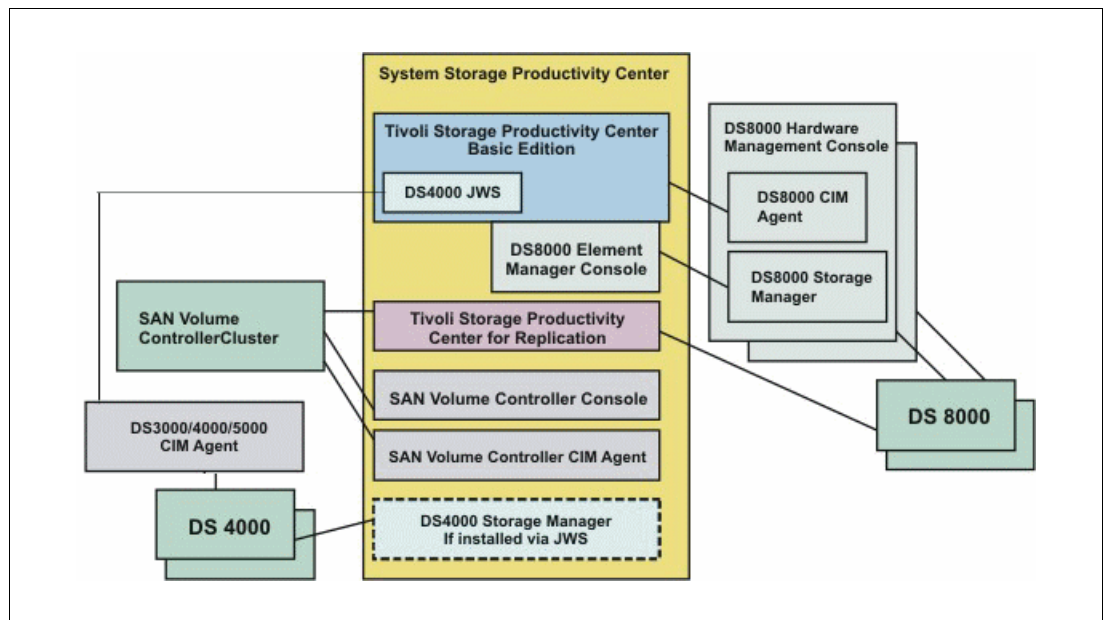


Figure 18-13 Overview of System Storage Productivity Center

The System Storage Productivity Center provides integration of storage management applications by providing support for the following products:

- ▶ **IBM Tivoli Storage Productivity Center Basic Edition:** IBM Tivoli Storage Productivity Center Basic Edition enables device configuration and management of SAN-attached devices from a single console. Basic Edition simplifies the complexity of managing multiple SAN-attached storage devices. It enables you to manage SANs and heterogeneous storage from a single console.
- ▶ **IBM System Storage DS8000 Storage Manager:** The IBM System Storage DS8000 Storage Manager user interface is accessible from the Tivoli Storage Productivity Center user interface. Tivoli Storage Productivity Center provides a DS8000 element manager page, which allows you to add and manage multiple DS8000 element managers from one console. You can use the DS8000 element manager page to:
 - Access the IBM DS8000 Storage Manager interface that is embedded in the DS8000 Hardware Management Console (HMC)
 - Use the DS8000 Storage Manager interface to configure and administer its associated IBM DS8000 storage arrays.
- ▶ **IBM System Storage SAN Volume Controller:** The IBM System Storage SAN Volume Controller Console is installed with System Storage Productivity Center. The SAN Volume Controller Console is a separate icon from the Tivoli Storage Productivity Center user interface icon on your desktop.
- ▶ **IBM Tivoli Storage Productivity Center for Replication:** Beginning with Release 4.1, IBM Tivoli Storage Productivity Center for Replication and Tivoli Storage Productivity Center, previously separate products, are integrated. This means that you can start the IBM Tivoli Storage Productivity Center for Replication user interface from within the Tivoli Storage Productivity Center user interface.
- ▶ **IBM System Storage DS3000, DS4000, and DS5000 Storage Manager:** You can optionally install the DS3000, DS4000, and DS5000 Storage Manager user interface on the System Storage Productivity Center server, or a remote server so that you can manage the storage systems from a single console.

To use Tivoli Storage Productivity Center to view reports and topology from the DS3000, DS4000, or DS5000, you must first install the Engenio provider (CIM agent) for the storage system (on a separate server; not the System Storage Productivity Center server). For information about the Engenio provider, contact LSI at the website:

http://www.lsi.com/storage_home/products_home/external_raid/management_software/smis_provider/index.html

- ▶ **DS CIM agent management commands:** The DS CIM agent management commands (DSCIMCLI) are installed on the System Storage Productivity Center system. The DSCIMCLI can manage the CIM agents for DS8000, and DS6000.

For information about the CIM agent management commands, see *DS Open Application Programming Interface Reference* at the website:

<http://www.ibm.com/servers/storage/support/software/cimdsoapi/>

Figure 18-14 illustrates the steps that the IBM service representative and you must perform to install and configure the IBM System Storage Productivity Center.

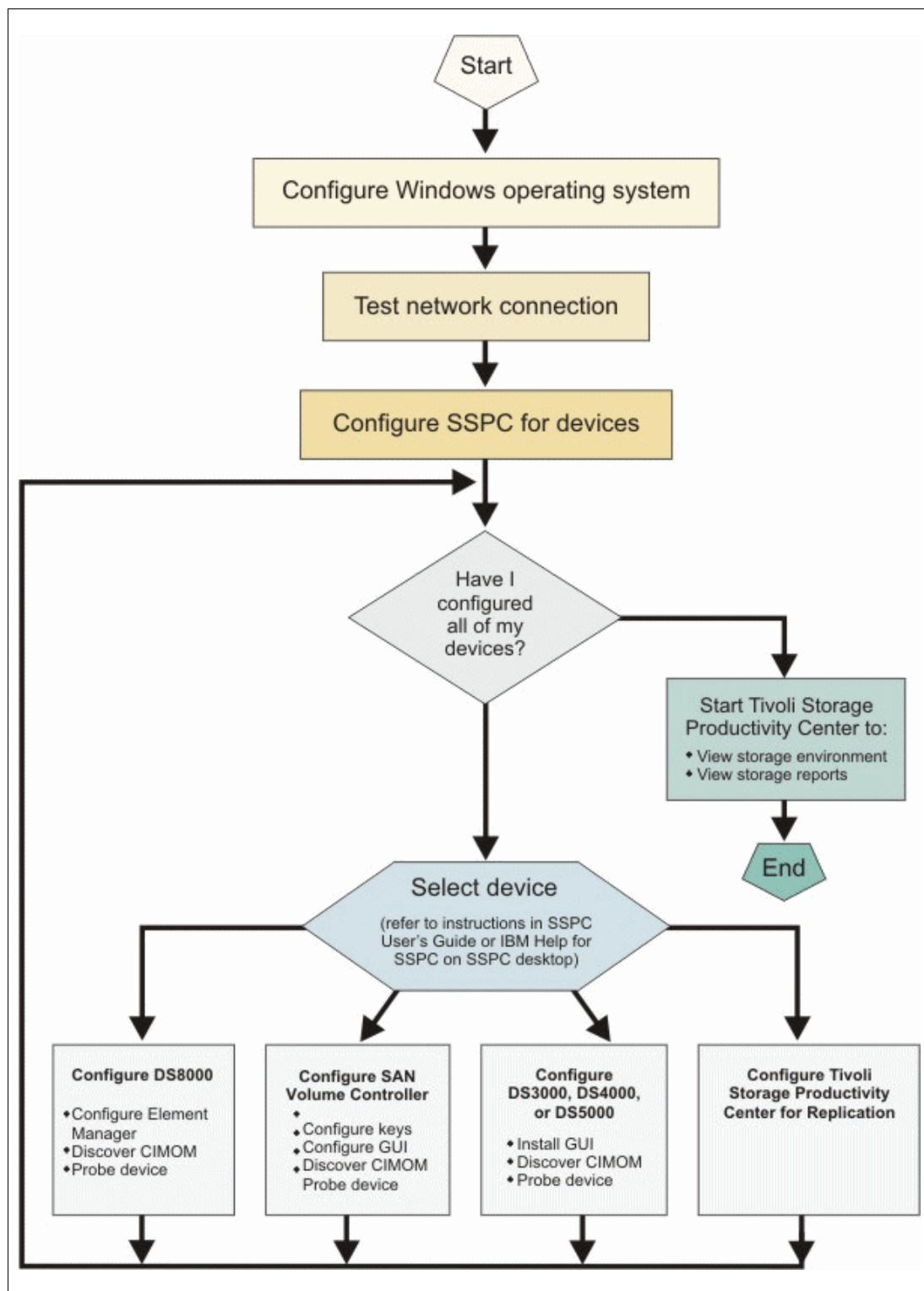


Figure 18-14 Overview of installing and configuring the System Storage Productivity Center

18.10.3 System Storage Productivity Center hardware components

The SSPC consists of an IBM System x server running the Microsoft Windows Server 2008 Enterprise Edition operating system. This topic describes the standard SSPC server configuration.

The SSPC Model 2805-MC4 server is based on the IBM System x3550 M2 Type 7946 server, with IBM Tivoli Storage Productivity Center, DS CIM agent management commands (DSCIMCLI), the SAN Volume Controller Console, and IBM Tivoli Storage Productivity Center for Replication all preinstalled. The System x3550 server is a rack-model server for high-volume network-transaction processing. This quad-core server is suited for networking environments that require microprocessor performance, input/output (I/O) flexibility, and manageability. For detailed information, see the *IBM System x3550 M2 Type 7946 Installation and User's Guide*. This publication is available on the *IBM System Storage Productivity Center Documentation CD*, which comes with the server.

The System Storage Productivity Center Model 2805-MC4 server includes the following characteristics:

- ▶ One Intel Xeon quad-core central processing unit, with speed of 2.4 GHz, cache of 8 MB, and power consumption of 80 W
- ▶ 8 GB of RAM (eight 1-inch dual inline memory modules of double-data-rate three memory, with a data rate of 1333 MHz)
- ▶ Two 146-GB hard disk drives, each with a speed of 15 K
- ▶ One Broadcom 6708 Ethernet card
- ▶ One CD/DVD bay with read and write-read capability

Optionally, you can add a keyboard, video, and mouse (KVM) unit, or a redundant power supply. The SSPC Model 2805-MC4 server supports six 2.5-inch hot-swap Serial Attached SCSI (SAS) hard disk drives.

Note: You can create a server that is equivalent to the SSPC by installing the SSPC software on any supported server. However, the hardware support included with the SSPC is available only if the software is installed on a Model 2805-MC4 server. For more information, see the section about installing SSPC components in the IBM System Storage Productivity Center User's Guide.

See this publication for more information about IBM System Storage Productivity Center: *IBM System Storage Productivity Center V1.4: Introduction and Planning Guide*, SC23-8824-06.

18.10.4 System Storage Productivity Center licensing

License entitlement for preinstalled and optional software must be purchased separately from the IBM System Storage Productivity Center server. Ensure that the necessary licenses have been purchased before the IBM service representative arrives to install the System Storage Productivity Center server.

Preinstalled software

The following software is preinstalled on the System Storage Productivity Center server:

- ▶ Microsoft Windows Server 2008 Enterprise Edition
- ▶ IBM Tivoli Storage Productivity Center 4.1.1 Basic Edition
- ▶ IBM Tivoli Storage Productivity Center for Replication 4.1.1
- ▶ IBM DB2 Enterprise Server Edition 9.5 with Fix Pack 3
- ▶ IBM System Storage SAN Volume Controller Console version 5.1.0
- ▶ PuTTY

The Basic Edition of Tivoli Storage Productivity Center can be upgraded to the Standard Edition, which includes additional data analysis and performance reporting features. For information about upgrading to these packages, see the topic about upgrading in the IBM System Storage Productivity Center Information Center at:

<http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp>

Also see the *IBM System Storage Productivity Center User's Guide*.

Licensing details

The IBM Tivoli Storage Productivity Center Basic Edition software is licensed per System Storage Productivity Center server, regardless of the number of processors or amount of data managed. Licensing includes the first year of software support. After the first year, support is provided on a subscription basis.

Tivoli Storage Productivity Center for Replication is installed with local OS authentication. This program is non-licensed. You are required to purchase and install the 2-site or 3-site license to use the functions of Tivoli Storage Productivity Center for Replication.

Licensing for the IBM DB2 Enterprise Server is included in the Tivoli Storage Productivity Center Basic Edition license agreement.

Attention: As a best practice, install antivirus software on the SSPC server but not other software (except for the upgrades to the software listed in this topic). While antivirus applications can offer protection to servers that might be in an unprotected environment, installing other software can cause port conflicts or performance issues. Other precautions that you can take to protect the SSPC server include enabling security updates to the operating system and establishing a backup procedure for the SSPC.

18.10.5 System Storage Productivity Center support matrix

This topic describes IBM System Storage Productivity Center software support for each supported hardware system.

Table 18-6 provides details for support matrix.

Table 18-6 System Storage Productivity Center supported hardware

Supported hardware (see note)	Supported software				
	IBM Tivoli Storage Productivity Center	IBM Tivoli Storage Productivity Center for Replication	IBM System Storage SAN Volume Controller	DSCIMCLI	IBM DS Storage Manager
DS3000	X				X
DS4000	X				X
DS5000	X				X
DS6000	X	X		X	
DS8000	X	X		X	
SAN Volume Controller	X	X	X		
TS3310 Tape Library	X				
TS3500 Tape Library	X				

Note: SSPC might also support other hardware because that hardware is supported by Tivoli Storage Productivity Center. For a complete list of supported devices, go to the following website:

<http://www.ibm.com/software/sysmgmt/products/support/IBMTotalStorageProductivityCenterStandardEdition.html>

Then select: → **Plan** → **[View all Plan documents]**.

18.11 Interoperability

IBM Tivoli Storage Productivity Center supports:

- Storage Systems:
 - EMC
 - Fujitsu
 - Hewlett-Packard
 - Hitachi Data Systems
 - IBM
 - LSI
 - NEC
 - NetApp
 - SGI
 - Sun Microsystems.

- ▶ Operating Systems:
 - AIX
 - HP-UX
 - Solaris
 - Red Hat Linux
 - SUSE Linux
 - Microsoft Windows.
- ▶ Host Business Adapters:
 - IBM
 - Brocade
 - QLogic
 - Emulex
 - LSI
 - Hewlett-Packard.
- ▶ Switches and Directors:
 - Brocade
 - Cisco
 - QLogic
 - McData.

18.12 More information

See these publications for more information about IBM Tivoli Storage Productivity Center:

- ▶ *IBM Tivoli Storage Productivity Center V4.1: Release Guide*, SG24-7725
- ▶ *IBM Tivoli Storage Productivity Center User's Guide*, SC27-2338-01
- ▶ *IBM Tivoli Storage Productivity Center Workflow User's Guide*, SC27-2341-01
- ▶ *IBM System Storage Productivity Center V1.4: Introduction and Planning Guide*, SC23-8824-06.

For the latest information about IBM Tivoli Storage Productivity Center, see following link:

<http://www-03.ibm.com/systems/storage/software/center/index.html>

The latest documentation is available from the TPC Infocenter at:

http://publib.boulder.ibm.com/infocenter/tivihelp/v4r1/index.jsp?topic=/com.ibm.itpc.doc_2.1/tpc_infocenter_home.htm



IBM Grid Access Manager

IBM System Storage Multilevel Grid Access Manager Software (Grid Access Manager Software) is built on an open, high-performance grid architecture that delivers data protection, information life-cycle management, simplified storage management and multi-site data access based on open standards. The potential benefits derived from these features can help deliver important cost savings and operational efficiencies, including:

- ▶ Simplified management and improved storage utilization, with excellent performance
- ▶ Data protection and improved business continuity
- ▶ Support for global access, multi-site operation

19.1 GAM overview

Grid Access Manager (GAM) Software enables customers with single or multiple sites and with fixed content or reference data storage requirements to improve storage utilization and investment across sites by way of an enterprise-wide, fault-tolerant storage grid with real-time failover capabilities. Grid Access Manager Software can help protect enterprise data through automated replication, life-cycle management and digital signature functionality.

The need to retain massive volumes of business-critical fixed-content data for long periods of time presents new data and storage management challenges for IT. Users continue to demand fast performance and higher and broader availability of enterprise data. Massive data volumes combined with long retention periods require storage administrators to deliver a cost-effective storage strategy that meets the user's needs, protects valuable data, scales on demand, simplifies data migration and automates recovery from both planned and unplanned downtime.

Grid Access Manager Software (Figure 19-1) can help customers address these critical challenges by delivering an automated and virtualized storage and data management layer that can be deployed locally or across multiple facilities and even on heterogeneous storage hardware. Grid Access Manager Software achieves this by presenting itself to enterprise applications as a single storage system that can aggregate underlying storage silos into a single enterprise storage pool, even if these silos are spread across multiple facilities and consist of heterogeneous storage media. This highly automated system enables IT administrators to focus on storage planning and infrastructure improvements and eliminate many manual administration tasks.

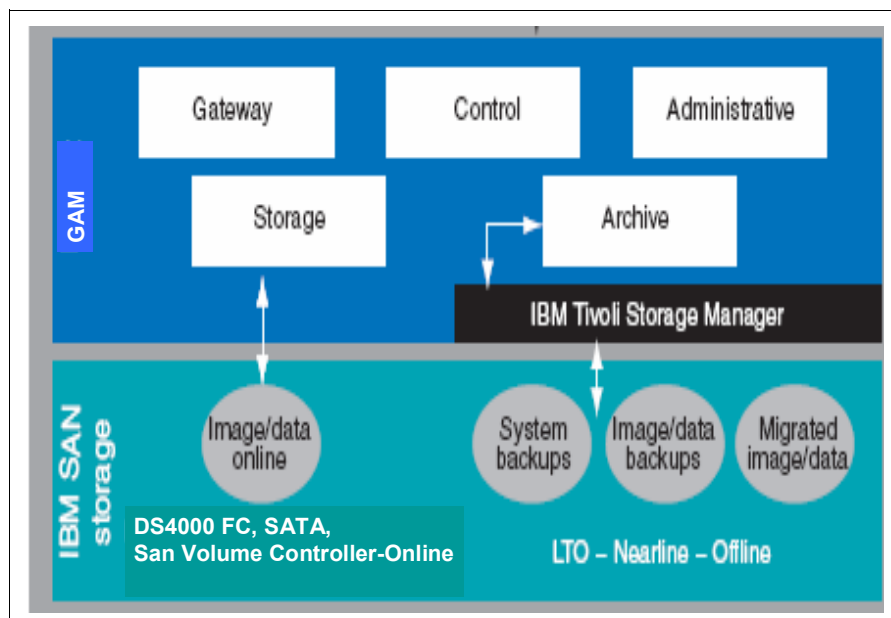


Figure 19-1 IBM GAM

19.2 Performance and availability

Grid Access Manager Software employs data streaming, caching and the power of grid architecture to deliver a fast, responsive storage system. Grid Access Manager Software implements a fast data-streaming mechanism to deliver data. To improve performance, Grid Access Manager Software load-balances requests based on storage tier, CPU and network bandwidth.

The optional Distributed Gateway based on IBM GPFS technology increases the scalability and performance of the file-system front end for the Grid, allowing NAS capability for transactional as well as archive data.

19.3 Data protection and business continuity

Grid Access Manager Software supports business continuity by creating multiple copies of data and automatically replicating them geographically or to various storage tiers based on customer rules so they are no longer vulnerable to potential issues. By supporting continuous operations even in the event of failure, Grid Access Manager Software provides automated recovery to help maintain business resiliency. The solution also supports continuous operation during hardware refreshes, capacity upgrades and data migration.

19.3.1 Data integrity

Grid Access Manager Software helps protect the integrity and authenticity of stored data using digital fingerprints. Data integrity is proactively monitored and verified when data is stored, replicated, restored and retrieved. Integrity checks can be performed on demand and if errors are encountered, Grid Access Manager Software can automatically create another replica from a known good copy.

Grid Access Manager Software provides a fault-tolerant, resilient and self-healing solution. Storage services can remain available even after hardware, storage and network failures occur or where entire sites are completely destroyed.

In the event of site, storage or system failure, Grid Access Manager Software automates data recovery to fully restore the storage system to the original state. During this period, storage services continue to be available.

19.3.2 Multi site operation

Fast, enterprise-wide availability can be a challenge for traditional storage systems when transporting petabytes of data across sites that are linked with varying bandwidth networks. Grid Access Manager Software enables formation of archives that can scale to petabytes of data across hundreds of sites. Key enabling features include:

- Object store with global name space:

Grid Access Manager Software provides an object-based storage system that can scale to petabytes (capacity) and billions of objects (object count) in a single unified system that can span multiple sites and tiers. A global name space gives access to stored objects using globally unique object identifiers, which can be accessed from any grid resource, regardless of location or storage tier. This eliminates storage silos to enable efficient and effective multi-site access.

- WAN-optimized replication:

Grid Access Manager Software provides policy-based N-way data replication among sites and heterogeneous storage hardware. The replication semantics are highly optimized for WAN operation and designed to recover from congestion or link failure.

19.4 Simplified management and optimized storage utilization

Grid Access Manager Software helps customers implement a virtualized, shared storage pool for reference data and fixed content. By operating across locations, multiple storage tiers and storage hardware brands, Grid Access Manager Software offers storage flexibility and easy extensibility to customers, and supports tiered storage architectures so organizations can have greater flexibility to implement the most cost-effective storage according to their defined policies while continuing to use existing investments. Key features include:

- ▶ Intelligent information life-cycle management
- ▶ Operation across facilities and tiers
- ▶ Flexible Hardware Configurations
- ▶ Automated data migration
- ▶ Open, standards-based access
- ▶ Proactive monitoring and management

19.4.1 Intelligent information life-cycle management

Grid Access Manager Software enables the creation of information life-cycle management (ILM) policies that govern the geographic location and placement of data in a multi-tier, multi-facility storage system. Intelligent ILM can help deploy the right storage in the right place, on the right media, at the right time based on its relevance and value to the organization.

19.4.2 Operation across facilities and tiers

Capacity, performance and geographic scope of Grid Access Manager Software might be increased by adding new nodes. As nodes are added typically in a non-disruptive manner, Grid Access Manager Software can continue to present itself as one large scalable computing system. Grid Access Manager Software automatically reconfigures the environment to take advantage of the new resources.

19.4.3 Flexible hardware configurations

Grid Access Manager Software utilizes industry-standard servers and storage, empowering administrators to deliver a “best of breed” enterprise storage strategy with the flexibility to adapt over time. A Grid Access Manager Software solution can be composed of multiple storage tiers utilizing online and nearline storage media (such as SCSI, SATA and tape).

19.4.4 Automated data migration

Grid Access Manager Software can automate the movement of data from obsolete resources to new resources thus reducing or eliminating manual, disruptive data-migration projects.

19.4.5 Open, standards-based access

With support for industry standard interfaces (including CIFS and NFS) and protocols (such as HTTP), Grid Access Manager Software enables organizations to implement an enterprise fixed-content-storage strategy.

19.4.6 Proactive monitoring and management

Grid Access Manager Software is designed to reduce the cost of management by continuously monitoring the health of the system, including storage, servers and network, and proactively alerts administrators if issues arise. Administrators can monitor and manage the grid from across the network using the single-pane, Web-based management console. Administrators can track CPU, network, storage resources and digital data assets using the 120 customizable alerts and more than 500 on demand reports.

19.5 Summary of the key features and benefits

Table 19-1 summarizes the key features and benefits of those features provided by IBM Grid Access Manager software.

Table 19-1 GAM key features

Feature	Benefit
Simplified management and optimized storage utilization	Intelligent information life-cycle management policies that govern the geographic location and placement of data in a multi-tier, multi-facility storage system.
Operation across facilities and tiers	Capacity, performance and geographic scope of Grid Access Manager Software might be increased by adding new nodes in a non-disruptive manner.
Hardware agnostic	Heterogeneous storage platforms, no vendor tie-in.
Automated data migration	Automate the movement of data from obsolete resources to new resources, Place on the 'correct' resource automatically.
Open, standards-based access	GAMS enables organizations to implement an enterprise fixed-content-storage strategy.

19.6 More information

For more information about GAMS see

- <http://www.ibm.com/systems/storage/software/gam/index.html>
- *Introducing the IBM Grid Access Manager, SG24-7612*



IBM GPFS and SoFS

IBM General Parallel File System (GPFS) is a scalable, parallel, clustered file system. In addition to high-speed parallel file access, GPFS provides fault tolerance, including automatic recovery from disk and node failures. IBM General Parallel File System (GPFS) currently powers many of the world's largest scientific supercomputers and commercial applications requiring high-speed access to large volumes of data.

IBM scale out file services (SOFS) is a flexible storage virtualization service that will help alleviate data storage challenges by enabling quick implementation of highly scalable, global, clustered network attached storage systems.

These virtualization products are part of an IBM commitment to the open standards adopted by the Storage Networking Industry Association (SNIA). They implement standard common information model (CIM) based APIs to allow management applications from IBM, and other vendors to administer and monitor their activities.

Storage software solutions are a key element in the IBM overall storage strategy. This strategy addresses several of the most pressing needs currently facing Chief Information Officers (CIO) and IT managers. As part of its strategy, IBM intends to deliver industry-leading technologies that will help dramatically reduce total cost of ownership (TCO) for storage.

We first discuss the IBM GPFS and then give you an overview of the IBM SoFS.

20.1 IBM General Parallel File System

The IBM General Parallel File System (GPFS) is more than just a file system; it is a file management infrastructure. It provides unmatched performance and reliability with scalable access to critical file data. GPFS distinguishes itself from other cluster file systems by providing concurrent high-speed file access to applications executing on multiple nodes of an AIX cluster, a Linux cluster, or a heterogeneous cluster of AIX and Linux nodes. New with GPFS 3.2.1.5 you can add windows client nodes to an AIX, Linux or mixed cluster. In addition to providing file storage capabilities, GPFS provides storage management, information life cycle management tools, centralized administration and allows for shared access to file systems from remote GPFS clusters.

GPFS provides file system services to parallel and serial applications. GPFS allows parallel applications simultaneous access to the same files, or other files, from any node which has the GPFS file system mounted while managing a high level of control over all file system operations.

GPFS is particularly appropriate in an environment where the aggregate peak need for data bandwidth exceeds the capability of a distributed file system server. GPFS allows users shared file access within a single GPFS cluster and across multiple GPFS clusters.

GPFS currently powers many of the world's largest scientific supercomputers and commercial applications requiring high-speed access to large volumes of data such as:

- ▶ Engineering design
- ▶ Digital Media
- ▶ Business intelligence
- ▶ Financial Analytics
- ▶ Seismic data processing
- ▶ Geographic information systems
- ▶ Scalable file serving

GPFS provides online storage management, scalable access, and integrated information lifecycle management tools capable of managing petabytes of data and billions of files. Virtualizing your file storage space and allowing multiple systems and applications to share common pools of storage provides you the flexibility to transparently administer the infrastructure without disrupting applications, improving cost and energy efficiency, while reducing management overhead.

20.2 GPFS overview

A GPFS cluster consists of:

- ▶ AIX nodes, Linux nodes, Windows nodes, or a combination thereof

A node can be:

- An individual operating system image on a single computer within a cluster.
 - A system partition that contains an operating system. Certain IBM System p5 and IBM System p machines allow multiple system partitions, each of which is considered to be a node within a GPFS cluster.
- ▶ Network shared disks (NSDs) created and maintained by the NSD component of GPFS
 - All disks used by GPFS must first be given a globally-accessible NSD name.
 - The GPFS NSD component provides a method for cluster-wide disk naming and access.

- On Linux machines running GPFS, you might give an NSD name to:
 - Physical disks
 - Logical partitions of a disk
 - Representations of physical disks (such as LUNs)
- On AIX machines running GPFS, you might give an NSD name to:
 - Physical disks
 - Virtual shared disks
- ▶ A shared network for GPFS communications allowing a single network view of the configuration. A single network, a LAN or a switch, is used for GPFS communication, including the NSD communication.

GPFS allows users shared access to files in either the cluster where the file system was created or other GPFS clusters. Each site in the network is managed as a separate cluster, while allowing shared file system access. When multiple clusters are configured to access the same GPFS file system, Open Secure Sockets Layer (OpenSSL) is used to authenticate and check authorization for all network connections.

20.3 Benefits and highlights of GPFS

GPFS provides a wide range of beneficial features, among which are high levels of scalability and performance, data availability, flexibility, and simplified storage management.

20.3.1 Scalable high-performance shared disk file systems

GPFS is not a network file system like NFS. There is no single-server bottleneck or protocol overhead for data transfer. GPFS takes file management beyond a single system by providing scalable access from multiple systems to a global namespace. This is done by:

- ▶ Allowing multiple processes or applications on all nodes in the cluster simultaneous access to the same file using standard file system calls.
- ▶ Increasing aggregate bandwidth of your file system by spreading reads and writes across multiple disks.
- ▶ Balancing the load evenly across all disks to maximize their combined throughput. One disk is no more active than another.
- ▶ Supporting very large file and file system sizes.
- ▶ Allowing concurrent reads and writes from multiple nodes.
- ▶ Allowing for distributed token (lock) management.
- ▶ Independent paths to the same file by the same name from anywhere in the cluster
- ▶ Allowing for the specification of various networks for GPFS daemon communication and for GPFS administration command usage within your cluster.

Achieving high throughput to a single, large file requires striping data across multiple disks and multiple disk controllers. Rather than relying on striping in a separate volume manager layer, GPFS implements striping in the file system. Managing its own striping affords GPFS the control it needs to achieve fault tolerance and to balance load across adapters, storage controllers, and disks. Large files in GPFS are divided into equal sized blocks, and consecutive blocks are placed on separate disks in a round-robin fashion.

To exploit disk parallelism when reading a large file from a single-threaded application, whenever it can recognize a pattern, GPFS prefetches data into its buffer pool, issuing I/O requests in parallel to as many disks as necessary to achieve the bandwidth of which the switching fabric is capable. GPFS recognizes sequential, reverse sequential, and various forms of striped access patterns.

GPFS uses a sophisticated token management system to provide data consistency while allowing multiple independent paths to the same file by the same name from anywhere in the cluster

20.3.2 Increased data availability

GPFS failover support allows you to organize your hardware into failure groups. A failure group is a set of disks that share a common point of failure that can cause them all to become simultaneously unavailable. When used in conjunction with the replication feature of GPFS, the creation of multiple failure groups provides for increased file availability if a group of disks fails. GPFS maintains each instance of replicated data and metadata on disks in various failure groups. If a set of disks becomes unavailable, GPFS fails over to the replicated copies in another failure group.

Disks can be added or deleted while the file system is mounted. When the time is right and system demand is low, you can then rebalance the file system across all currently configured disks. In addition, you can also add or delete nodes without having to stop and restart the GPFS daemon on all nodes.

In a SAN configuration where you have also defined NSD servers, if the physical connection to the disk is broken, GPFS dynamically switches disk access to the servers nodes and continues to provide data through NSD server nodes. GPFS falls back to local disk access when it has discovered the path has been repaired.

20.3.3 Simplified storage management

GPFS provides storage management based on the definition and use of:

- ▶ Storage pools
- ▶ Policies
- ▶ Filesets

Storage pools

A storage pool is a collection of disks or RAID configurations with similar properties that are managed together as a group. Storage pools provide a method to partition storage on the file system, which assists with:

- ▶ Improved price-performance, by matching the cost of storage to the value of the data.
- ▶ Improved performance, by reducing the contention for premium storage, and reducing the impact of slower devices
- ▶ Improved reliability, by providing for replication based on need and better failure containment

Policies

Files are assigned to a storage pool based on defined policies. Policies provide for:

- ▶ Placing files in a specific storage pool when the files are created
- ▶ Migrating files from one storage pool to another
- ▶ File deletion based on file characteristics
- ▶ Snapshot metadata scans and file list creation

Filesets

Filesets provide a method for partitioning a file system and allow administrative operations at a finer granularity than the entire file system. For example, filesets allow you to:

- ▶ Define data block and inode quotas at the fileset level
- ▶ Apply policy rules to specific filesets

Highlights of the comprehensive features and benefits list of the GPFS file system can be seen in Table 20-1.

20.3.4 Simplified administration: Features and benefits of GPFS

GPFS offers many of the standard UNIX file system interfaces allowing most applications to execute without modification or recompiling. UNIX file system utilities are also supported by GPFS. That is, users can continue to use the UNIX commands they have always used for ordinary file operations

The only unique commands are those for administering the GPFS file system. GPFS administration commands are similar in name and function to UNIX file system commands, with one important difference: the GPFS commands operate on multiple nodes. A single GPFS command performs a file system function across the entire cluster.

GPFS commands save configuration and file system information in one or more files, collectively known as GPFS cluster configuration data files. The GPFS administration commands are designed to keep these files synchronized between each other and with the GPFS system files on each node in the cluster, thereby providing for accurate configuration data.

Table 20-1 summarizes key features and the benefits that they provide.

Table 20-1 GPFS Features and Benefits.

Feature	Benefit
Highly Scalable File System supports thousands of nodes and thousands of disks	GPFS enables the scale out expansion of the file system using a building block approach with commodity components as service demands and data volumes increase.
Data Protection Both user data and metadata can be replicated (mirrored) in the GPFS installation.	Improves availability by ensuring reliable access to data and metadata.
Parallel File System. Divides individual files into multiple blocks and stores these blocks across multiple disks or disk arrays in parallel.	Higher performance and reliability by eliminating the bottlenecks that typically arise when an entire file resides on a single node.
Storage Pools. User-defined grouping of storage resources.	Storage can be grouped or tiered based on factors such as location, performance and reliability.
External Storage Pools. Policy driven interactions with external applications for near-line data storage.	External pools utilize the high performance policy engine to expand your data storage options. For example, an external pool data can be stored on a tape library.
Policy-Driven Automation	Simplifies data management and allows matching the cost of storage to the value/importance of the data. For example, critical data can be stored on the fastest storage.

Feature	Benefit
Highly Scalable File System supports thousands of nodes and thousands of disks	GPFS enables the scale out expansion of the file system using a building block approach with commodity components as service demands and data volumes increase.
Data Protection Both user data and metadata can be replicated (mirrored) in the GPFS installation.	Improves availability by ensuring reliable access to data and metadata.
Parallel File System. Divides individual files into multiple blocks and stores these blocks across multiple disks or disk arrays in parallel.	Higher performance and reliability by eliminating the bottlenecks that typically arise when an entire file resides on a single node.
Storage Pools. User-defined grouping of storage resources.	Storage can be grouped or tiered based on factors such as location, performance and reliability.
External Storage Pools. Policy driven interactions with external applications for near-line data storage.	External pools utilize the high performance policy engine to expand your data storage options. For example, an external pool data can be stored on a tape library.
Concurrent File Access and Block-Level Locking	Enables multiple users, programs or nodes to access a single file simultaneously – accelerates processing for parallel applications that share data and eliminates the need for multiple copies of data, reducing data storage and management costs.
High Availability	Automatically recovers from events that can normally interrupt data availability. Applications and users can continue without interruption in the event of a node, path or storage system failure.
Clustered NFS. Tools to deploy and manage a clustered NFS file serving solution.	Provides an integrated solution enabling scalable NFS file serving from a cluster of Linux servers.
Large Blocksize Options	Data blocks of 16K, to 4MB can be used to optimize throughput for various applications. Large blocks allow more data to be written or read in a single operation. The wide range of block sizes enables more efficient use of disk bandwidth and available storage space.

20.4 SoFS overview

IBM scale out file services (SOFS) is a combination of hardware and software components combined to form a system, or solution, which provides highly scalable and highly available virtualized shared file system access, as shown in Figure 20-1. Virtualizing the file sharing system (NFS, CIFS etc) using a single, global, namespace means that an end user sees one device rather than individual servers and storage devices that comprise this virtual NAS device.

At the heart of the SoFS system is the IBM General Parallel File System (GPFS). GPFS is one of the most scalable commercial file systems available, installed in thousands of nodes and storing petabytes of data in discrete installations. Scale out file services enable the virtualization of the NFS service, transparently distributing service requests across multiple servers, which in turn enables the creation of a scale-out NAS farm installation

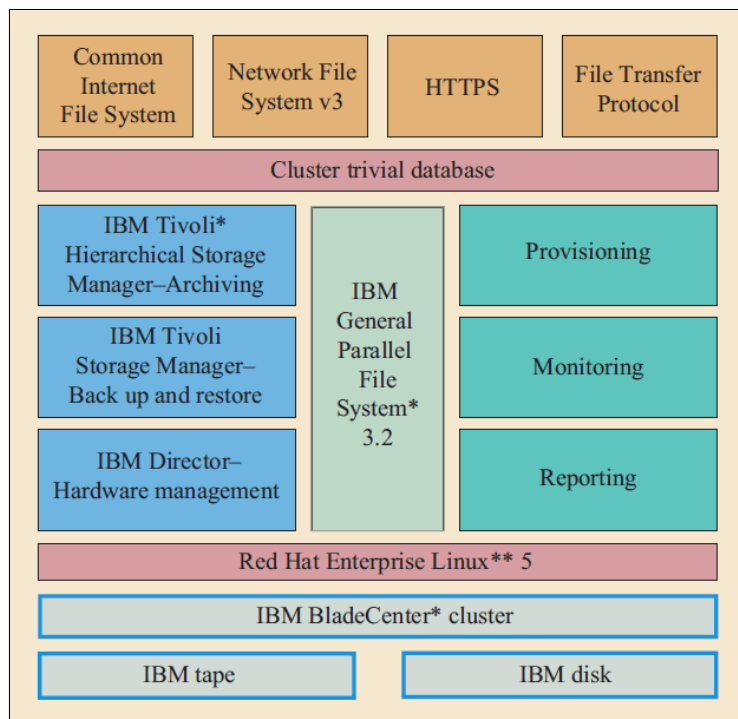


Figure 20-1 SoFS components

The SoFS solution is extremely resilient. It is designed so there is no single point of failure in the system. It also allows for a disaster recovery configuration that uses the global file system. It also supports snapshots and mirroring for high availability and smaller backup windows. The SoFS solution also supports data lifecycle management. Data can be stored in ways relative to the data value and performance requirements. And SoFS also supports LTO-4 tape backup with data encryption. This offers protection for archived data.

20.5 How SoFS works

The SoFS system provides file services by implementing a *scale-out* approach, hence the IBM Scale out File Services product name. SoFS uses the wide-striping capability of IBM GPFS to spread the blocks of a single file across as many nodes as possible to combine the streaming performance of many midrange back-end storage devices. By doing so, a bandwidth of multiple gigabytes per second is possible because the I/O performance of many disks can be combined instead of storing data only to a small RAID device as filers do today.

By fully exploiting the capabilities of GPFS, the integration of a policy-based life-cycle management interface and the possibility to place files in a directory that spans multiple independent storage classes creates a single file system with multiple storage classes. Each has its own performance, availability, and cost criteria, on a file basis, not on a filer basis, as has been done in the past.

SoFS utilizes and expands the IBM Tivoli Storage Manager and Tivoli Hierarchical Storage Manager products to integrate tape as an additional storage class inside the scale-out NAS. This allows transparent migration of data between disk and tape storage and provides end-to-end life-cycle management of the data.

The physical implementation of SoFS utilizes the IBM BladeCenter for highly efficient and scalable processing with various IBM storage devices. The BladeCenter provides a scalable platform for adding storage controllers into the SoFS pool to achieve the right performance balance. Processing blades share power, cooling, and I/O connectivity within the BladeCenter chassis and achieve very efficient performance versus stand-alone servers.

To provide a single global namespace to the user, SoFS uses virtualization and redirection technologies available within GPFS clusters. Each SoFS cluster node has internal access to all data blocks simultaneously and is able to assume the responsibility for each of the other nodes in a cluster. This provides the means for a transparent, nondisruptive protocol failover for client applications. As none of the existing cluster technologies provides a transparent failover of network-based file systems, SoFS introduces an extremely scalable cluster suite dealing with the specifics of file systems to provide a high degree of semantic correctness while being able to scale to very high bandwidths.

20.6 Why use SoFS

The efficient management of data is an ongoing struggle between access and scalability. Providing access to file-level data (that is, those files associated with individual documents, multimedia content, databases, and other applications) becomes more difficult as the number of users with access, and the amount of data stored, both grow.

20.6.1 The scale-up approach

Achieving the scalability needed to respond to the growth of data volume typically results in higher hardware and software costs and greater management challenges. Network-attached storage (NAS) solutions provide simplicity, manageability, and access, but until now, they lacked a single capability that has kept them from playing a role beyond the departmental level: scalability.

If a traditional filer is reaching a scalability or capacity limit, users upgrade to the next higher filer level until they reach the highest available system; this is the *scale-up* approach.

Instead of scaling up in a single high-end enterprise-class server node or building failover clusters to provide high availability for the filers, the SoFS system utilizes software functions to build a global namespace and exposes to the end user *a single virtual cluster node* with very high scalability.

20.6.2 Difference between SoFS and other clustered file systems

While other clustered File systems are just that, SoFS is more of a 'storage controller in a grid' meaning that the ownership of data is handled dynamically. In many clustered file system implementations, there is one node that is the default owner of a portion of the data. To maintain this structure requires high volumes of internode communication and in case the owning node goes down, the IP and data must be taken over by another node. While IP takeover is not a big problem, the takeover of data often proves difficult. In SoFS, this problem is completely non-existent, as any node can see any data at any time.

20.7 Components within an SoFS solution

In this section, we discuss the various hardware and software components of the SoFS solution.

20.7.1 SoFS hardware components

The hardware side of a SoFS solution consists of servers, SAN switches, and disk storage. Tape storage for archiving data is an optional part of any SoFS solution.

Storage server

The servers used in a SoFS solution are clusters of BladeCenter HS21 or HS21 XM blades or rack mounted X3650M2 servers. Production servers might be connected by Gigabit or 10 Gigabit Ethernet. Connectivity to the switches and storage units is by Fibre Channel only for the blades, and Fibre Channel or SAS attachment for the rack mounted servers.

As a minimum three System x servers are required. All servers in a configuration will be the same type. A separate management console might be added or that function may be located in one of the System x servers. The BladeCenter HS21 or HS21 XM blades might be housed in a BladeCenter H with Gigabit Ethernet or a BladeCenter H with 10 Gigabit Ethernet.

SAN hardware

The SAN switch hardware currently used for SoFS systems can be seen in Table 20-2. The base configuration of these switches activates less than the maximum number of Fibre Channel ports on the switch but for SoFS configurations all ports will be activated.

Table 20-2 SoFS SAN switch hardware

Switch Model	Description
2005-B16	SAN16B-2, 16-port 4Gbps Fibre Channel Switch
2498-24E	SAN24B-4, 24-port 8Gbps Fibre Channel Switch
2498-40E	SAN40B-4, 40-port 8Gbps Fibre Channel Switch

Storage units

Data is stored on one or more of the IBM storage system units described in

Table 20-3 SoFS Storage Units

Storage Unit	Description
XIV (2810-A14)	This large solution provides a capacity of 180TB and uses only 1 TB SATA disk drives. The system has up to 120 GB cache and 24 FC host ports.
DS8000	SoFS will support the use of DS8100, DS8300 and DS8700 systems, but a special request must be made. Configurations will be customized to the particular solution. This large solution provides up to 512 TB using FATA disks or 307.2 TB using FC disks, up to 32 GB cache, up to 128 FC host ports
IBM SVC	SoFS will support attachment of IBM SVC managed devices, this means any devices not on the supported hardware list can be used for SoFS storage devices if they are supported by SVC.
DCS9950	SoFS will support the use of the DCS9950, but a special request must be made. Configurations will be customized to the particular solution. This mid size solution provides up to 960TB using SATA disks or 448TB with FC, RAID-6, 8 FC host ports, 3 GB/sec read/write speed for sequential I/O
DS4800	This mid size solution provides up to 168 TB using SATA disks or 67.2 TB using FC disks, up to 16 GB cache, 8 FC host ports
DS4200	This mid size solution provides up to 84 TB using SATA disks, 2 GB cache, 4 FC host ports.
DS3400	This dual controller model of the DS3400 is the only one supported in SoFS configurations. Any number of DS3400 systems that can fit into the SAN network are supported in SoFS configurations. The preferred LUN configuration with 4 enclosures is RAID 5 in a 3+P arrangement so that each drive is in a separate enclosure. As a result, the system can survive the loss of a complete enclosure without losing a LUN.
DS3200	This entry level solution provides up 14.4 TB using SAS disks, up to 1 GB cache, and 6 SAS host ports. See the section and figure below on SAS configurations.

Note: Supported hardware device information is supplied by Techline and current as of 1 October 2009. Consult your IBM or business partner sales team for updated information.

20.7.2 SoFS software components

IBM GPFS 3.2 product on servers running Red Hat Enterprise Linux 5 is the heart of the SoFS system; this provides the SoFS cluster nodes with a single-system view of the file system. See Figure 20-1 on page 549 for an overview of the software components of SoFS system.

Included within SoFS is a CIFS component which is based on the Samba open source project and a newly developed, high speed, cluster-aware version of the TDB metadata store database used by Samba. This Cluster Trivial Database (CTDB) allows Samba to scale well across multiple cluster nodes. A more detailed explanation of this is available in the following paper: IBM Scale out File Services: Reinventing network-attached storage:

<http://www-935.ibm.com/services/us/its/pdf/sofs-am-journal-final-070108.pdf>

Other software components included and their function are shown in Table 20-4.

Table 20-4 SoFS software components

Component	SoFS Function
IBM Director	Hardware management
IBM Tivoli Storage Manager	Backup and restore
IBM Tivoli Hierarchical Storage manager	Archiving
DS Storage Manager SMclient	Storage device management
SoFS GUI	User interface
SoFS Toolset	Management of SoFS system

20.8 Data protection and disaster recovery

There are a number of choices the operator can make in the area of data protection and disaster recovery. SoFS allows for appropriate choices for each type of data, giving maximum flexibility to enable the right choice for each tier of data.

20.8.1 Data protection

The data protection choices match the choices available on the storage system used in the SoFS solution. These typically include RAID 5, RAID 6, RAID 10, and the unique XIV data protection algorithm.

20.8.2 Simple SoFS single site solution: No disaster recovery

A simple SoFS solution as shown in Figure 20-2 allows users to access data locally using a LAN, and remotely using a WAN. While the high availability is built into a single system, there is no disaster recovery built into this solution.

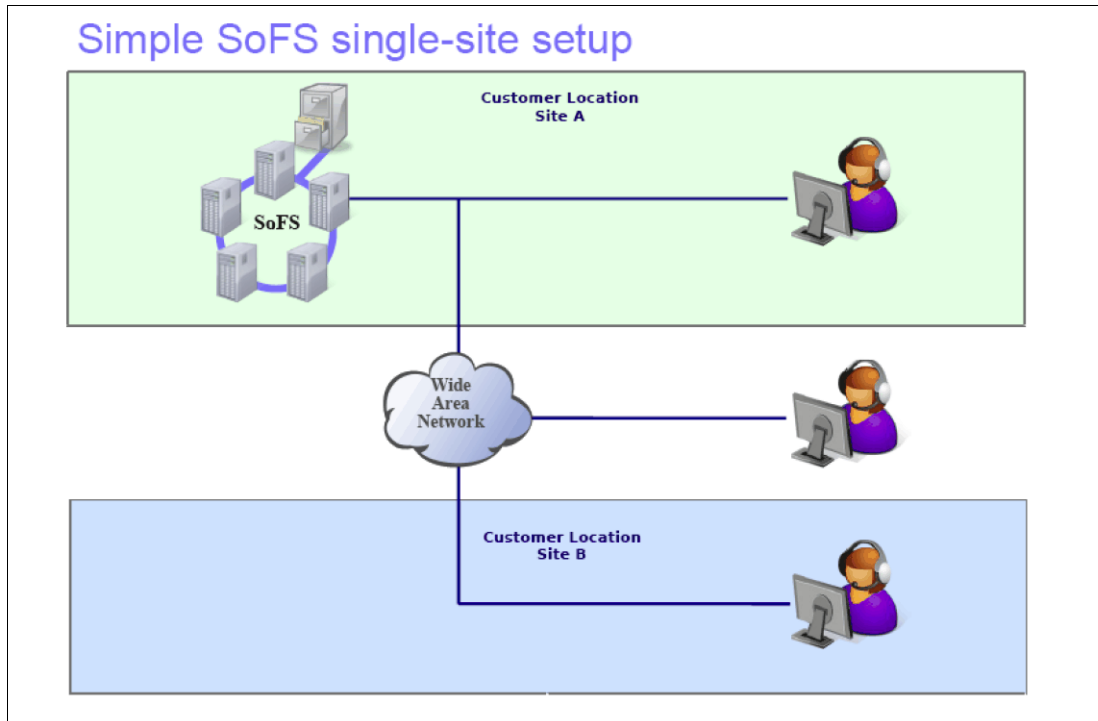


Figure 20-2 Single-site setup

20.8.3 SoFS with synchronous replication on-site

The operator can set up synchronous replication of part or all of the data to other nodes within the same SoFS site as shown in Figure 20-3.

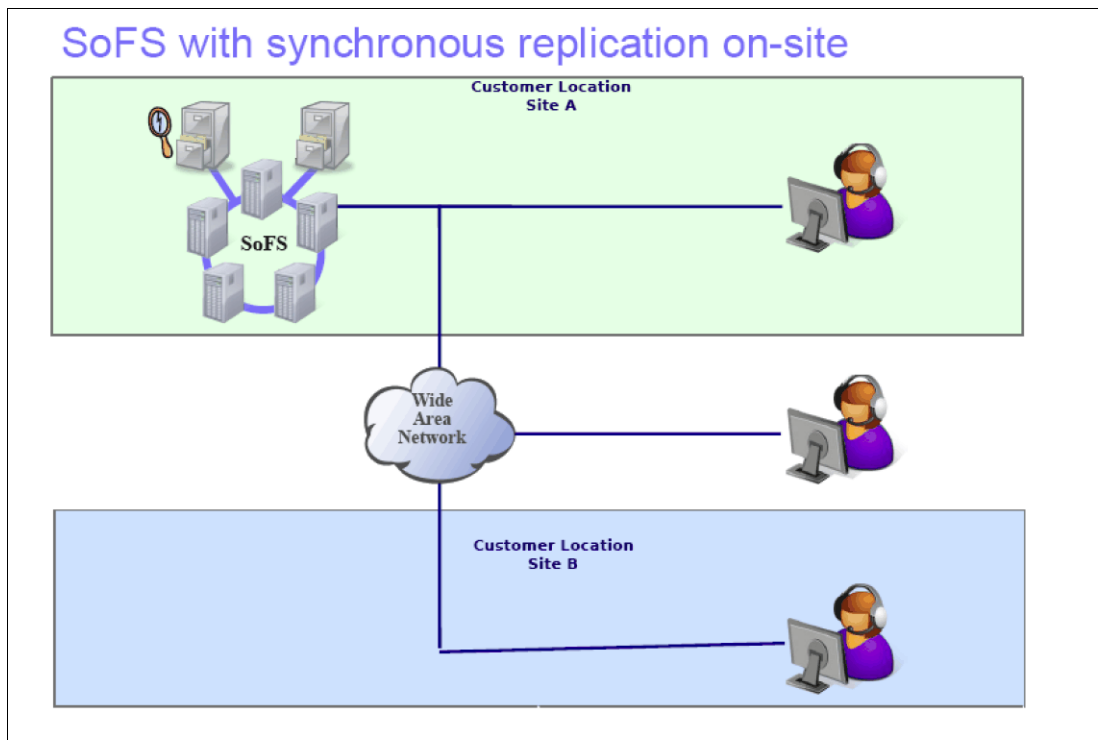


Figure 20-3 Sync replication on a single site

20.8.4 SoFS with cross site replication

The operator can set up replication of data between SoFS clusters on various sites as shown in Figure 20-4. The replication can be synchronous if the distance can be traversed by Fibre Channel cables. It is asynchronous if the distance is longer.

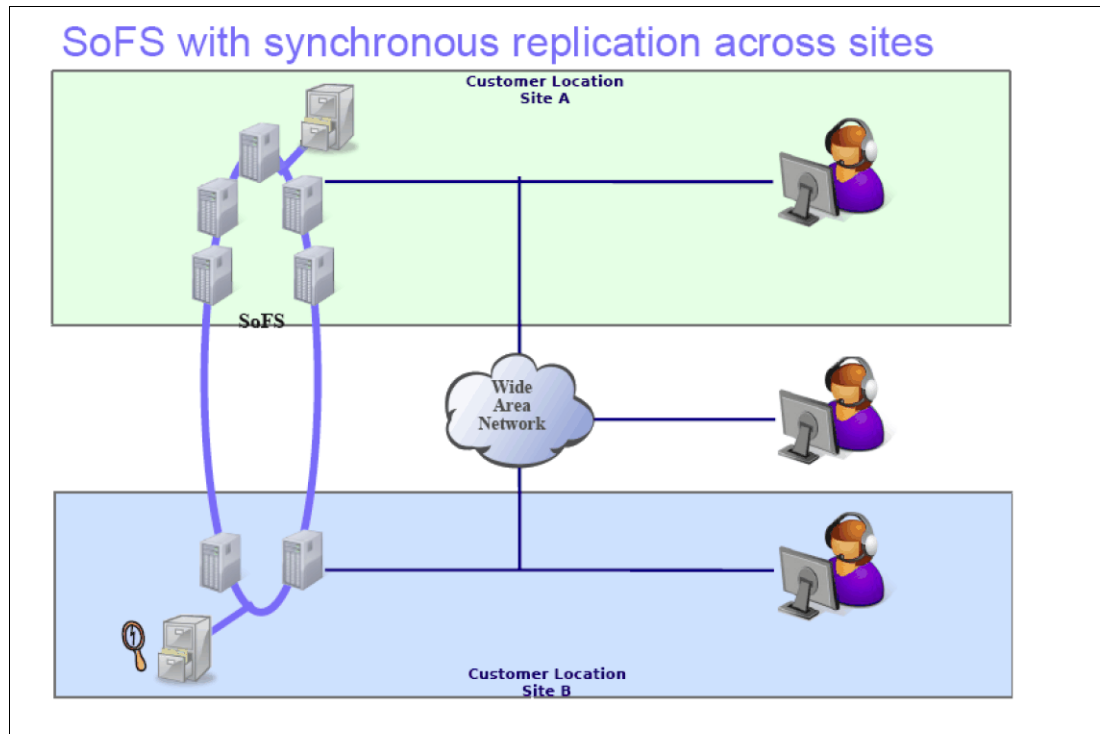


Figure 20-4 Cross site replication

20.8.5 SoFS with a cross cluster mount

Two SoFS sites can be connected and each file system mounted on the other to provide a solution that works as one large SoFS that covers two sites as shown in Figure 20-5. Within that complex, mirrors of data can be set up on various sites within the single SoFS solution. The cross cluster mount requires a TCP/IP connection between the clusters. This IP connection can be used for data and metadata traffic. As an option, there is also the ability to send data traffic over a Fibre Channel (SAN) connection between the clusters.

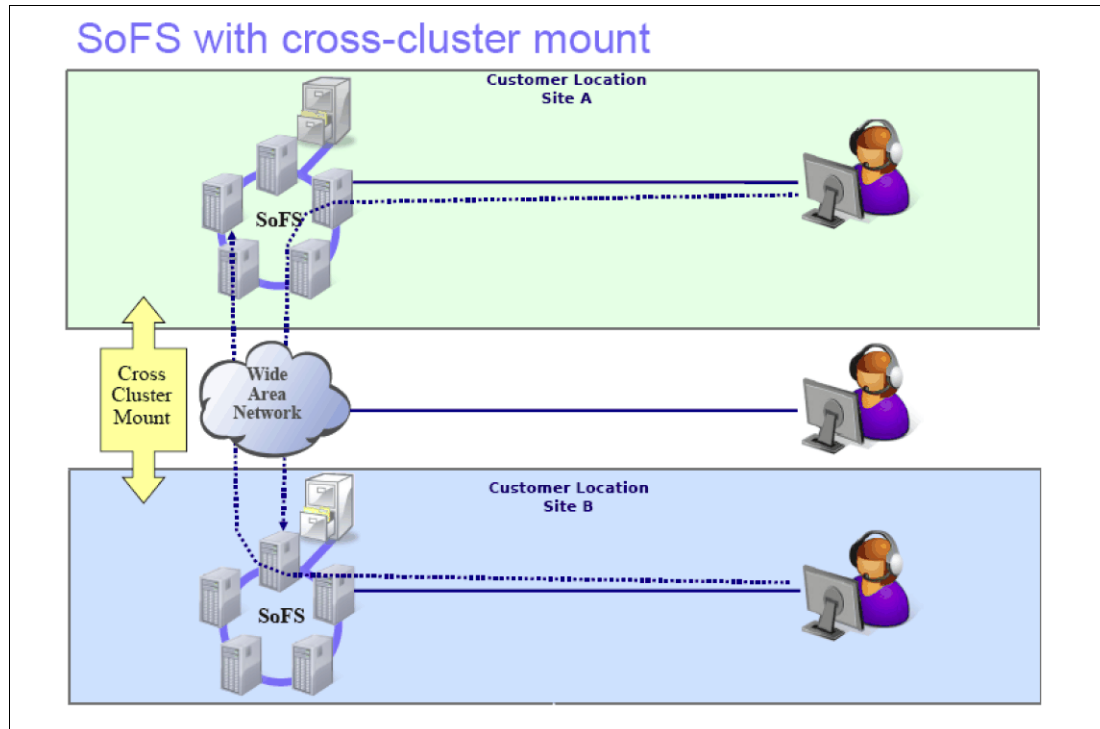


Figure 20-5 Cross cluster mounts

20.8.6 Snapshot technology

A snapshot of an entire GPFS file system might be created to preserve the file system's contents at a single point in time. A snapshot contains a copy of only the file system data that has been changed since it was created, using a copy on write technique. Snapshots of a file system are read-only; changes might only be made to the active (that is, normal, non-snapshot) files and directories.

The snapshot function allows a backup or mirror program to run concurrently with user updates and still obtain a consistent copy of the file system as of the time that the snapshot was created. Snapshots also provide an online backup capability that allows easy recovery from common problems such as accidental deletion of a file, and comparison with older versions of a file. Because snapshots are not copies of the entire file system, they must not be used as protection against media failures.

20.9 More information

Further information about SoFS can be found here:

<http://www-935.ibm.com/services/us/its/html/sofs-landing.html>

A more detailed overview of SoFS is available in this white paper:

<http://www-935.ibm.com/services/us/its/pdf/sofs-am-journal-final-070108.pdf>

An ATS white paper provides an overview and sample configuration information:

<http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101565>

Many thanks: Much of the information in this chapter was taken from the two documents referenced previously and is reproduced with the kind permission of the original authors.

For more information about GPFS, see these links:

- ▶ GPFS product home page:

<http://www.ibm.com/systems/clusters/software/gpfs/index.html>

- ▶ GPFS 3.2 information flyer:

<ftp://ftp.software.ibm.com/common/ssi/pm/fy/n/clf03001usen/CLF03001USEN.PDF>

- ▶ Introduction to GPFS version 3.2.1 (current version at the time of writing):

<ftp://ftp.software.ibm.com/common/ssi/sa/wh/n/clw03005usen/CLW03005USEN.PDF>



Part 5

z/OS storage management

In Part 5 we describe the following products for z/OS storage management:

- ▶ DFSMS
- ▶ DFSORT
- ▶ z/OS storage management tools



Data Management with DFSMS and Data Sorting with DFSORT

Data Facility Storage Management Subsystem (DFSMS) is a software suite that automatically manages data from creation to expiration. DFSMS provides allocation control for availability and performance, backup/recovery and disaster recovery services, space management, tape management, and reporting and simulation for performance and configuration tuning. It offers the following benefits:

- ▶ Integrated, comprehensive data management solution
- ▶ Delivers managed, predictable service to meet Service Level Agreements for Storage
- ▶ Up to 50% increased effective storage capacity with reduced overall storage cost.

On the other hand Data Facility Sort (DFSORT) is the IBM high-performance sort, merge, copy, analysis, and reporting product. DFSORT is an optional feature of z/OS.

21.1 DFSMS components

All data has a lifecycle, whether it is a temporary work file or customer information that regulatory laws require to be kept for several years. There is more to managing data than simply creating and deleting a data set, and the DFSMS group of products supports all the stages involved.

DFSMS is made up of the following components:

- ▶ DFSMSdfp (Data Facility Product), a base element of z/OS: Provides storage, data, program, and device management functions and DFSMS Copy Services capabilities.
- ▶ DFSMSdss (Data Set Services): An optional feature of z/OS providing data movement, copy, backup, and space management functions.
- ▶ DFSMShsm (Hierarchical Storage Manager): An optional feature of z/OS providing backup, recovery, migration, and space management functions.
- ▶ DFSMSrmm (Removable Media Manager): An optional feature of z/OS providing management functions for removable media.
- ▶ DFSMStvs (Transactional VSAM Services): An optional feature of z/OS that enables batch jobs and CICS® online transactions to update shared VSAM data sets concurrently.

21.2 DFSMSdfp (Data Facility Product)

Data Facility Product (DFSMSdfp) is the heart of the storage management subsystem; it provides the logical and physical input and output for z/OS storage, it keeps track of all data and programs managed within z/OS, and it provides data access both for native z/OS applications and other platforms such as AIX/UNIX, the Windows family, AS/400®, or OS/2.

21.2.1 DFSMS SDM Copy Services

The IBM Data Facility Storage Management Subsystem (DFSMS) SDM (System Data Mover) Copy Services are enterprise-level data duplication and disaster recovery capabilities that include:

- ▶ z/OS Global Mirror, also known as Extended Remote Copy (XRC)
- ▶ Peer-to-Peer Remote Copy (PPRC) including:
 - Metro Mirror, also known as synchronous PPRC
 - Global Copy, also known as PPRC-XD
 - Global Mirror
- ▶ 3-site solutions:
 - Metro/Global Copy
 - Metro/Global Mirror (M/GM)
 - Metro/zGlobal Mirror (M/zGM)
- ▶ FlashCopy
- ▶ Concurrent Copy
- ▶ Virtual Concurrent Copy
- ▶ SnapShot

21.2.2 Features

Here are the main features provided:

- ▶ Provide a comprehensive disaster recovery solution for the z/OS environment.
- ▶ Deliver a range of data duplication solutions that solve the customer requirement for an immediate point-in-time copy of data.
- ▶ SDM provides TSO commands and the ANTRQST API to enable disk hardware copy services functions.

Note: To understand more about Flash Copy, PPRC, and other features listed previously, see the website:

<http://www.ibm.com/systems/storage/software/sms/sdm/index.html>

21.2.3 z/OS NFS

z/OS NFS (Network File System) provides a way to optimize performance and efficiency in a distributed network while still capitalizing on the capacity, security, and integrity of zOS.

21.2.4 XRC Performance Monitor

Extended Remote Copy (XRC), also known as z/OS Global Mirror is IBM's premier disaster recovery option for zSeries storage systems.

The XRC Performance Monitor provides additional information to better manage an XRC environment and help XRC achieve the desired performance (typically measured as the latency in updating the mirrored system) at a low cost. The monitor provides information in realtime and also maintains historic records. From a very high level or from a detailed perspective, administrators can view the parameters that contribute to system performance.

Several parts of an IBM zSeries environment affect the performance of XRC. The XRC Performance Monitor gathers information from important sources. This functionality not only helps determine the optimum XRC configuration, but also helps quickly identify performance bottlenecks. The XRC Performance Monitor is also useful in spotting trends that might lead to future performance problems; thus it can also assist with capacity planning.

Note: For more details on XRC, see the website:

<http://www.ibm.com/systems/storage/software/xrc/index.html>

21.2.5 IBM TotalStorage Expert

The IBM TotalStorage ETL Expert is a program whose purpose is to help administrators manage Enterprise Tape Library (ETL) products using a web browser interface. TotalStorage ETL Expert helps with the following tasks:

- ▶ Performance management
- ▶ Asset management
- ▶ Capacity management

Performance management for ETL

System administrators might need to analyze performance information when an applications performance had been degraded. The IBM TotalStorage ETL Expert can gather performance information for the IBM TotalStorage 3494 Tape Library, IBM TotalStorage Virtual Tape Servers (VTS) or peer-to-peer VTS or any combination of the three. TotalStorage ETL Expert allows administrators to manage ETL's performance no matter what kind of host platform is used.

Asset management for ETL

An IBM ETL can contain up to three libraries. From a host system's viewpoint, all of these can be seen as a "tape library," but the physical configuration varies extensively for each. For example, a VTS is comprised of VTS controller, Tape Volume Cache (TVC) and its back end native tape library. Administrators might want to review which VTS uses which native tape library or might want to review how many physical tape drives in a native library are allocated for a VTS's use. The TotalStorage ETL Expert shows the relationship between libraries and the configuration is made plain and simple.

Capacity management for VTS/ETL

The TotalStorage Expert can show the capacity information for VTS, for example, how much data is active in back-end tape cartridges, and how many empty tape cartridges are available for stacked logical volumes. This information is important not only from a capacity planning perspective, but also from the VTS's performance viewpoint, because the tape cartridge utilization is tightly aligned with the VTS's back-end operation.

Note: For more details on TotalStorage Expert, see the website:
<http://www.ibm.com/systems/storage/software/expert/index.html>

21.3 DFSMSdss (Data Set Services)

Data Set Services (DFSMSdss), a component of the Data Facility Storage Management Subsystem (DFSMS), offers the following features:

- ▶ DFSMSdss offers powerful, user-friendly functions that let you move or copy data between volumes of like and unlike device types. It can also copy data that has been backed up.
- ▶ DFSMSdss can increase performance by reducing or eliminating DASD free-space fragmentation.
- ▶ DFSMSdss provides you with host system backup and recovery functions at both the data set and volume levels. It also includes an enhanced Stand-alone Restore Program that lets you restore vital system packs during disaster recovery, without a host operating system.
- ▶ DFSMSdss can convert your data sets and volumes to system-managed storage, or return your data to a non-system-managed state as part of a recovery procedure, all without data movement.

21.3.1 Concurrent Copy

Concurrent Copy (CC) is both an extended function in cached IBM Storage Controls and a component of DFSMSdss. CC enables you to copy or dump data while applications are updating the data. Both IMS/ESA® and DB2 can use CC for their backups.

21.3.2 FlashCopy

FlashCopy provides a fast data duplication capability. This option helps eliminate the need to stop applications for extended periods of time in order to perform backups and restores. FlashCopy is automatically initiated by DFSMSdss on IBM Enterprise Storage Servers (ESS) with the FlashCopy feature.

21.3.3 DFSMSdss Stand-Alone Services

DFSMSdss Stand-Alone Services was previously introduced as an APAR enhancement to all releases of DFSMS. The stand-alone restore function is a single-purpose program designed to allow the system programmer to restore vital system packs during disaster recovery without needing to rely on an OS/390 or z/OS environment.

Note: For more information of DFSMSdss, see the website:

<http://www.ibm.com/systems/storage/software/sms/dss/index.html>

21.4 DFSMSHsm (Hierarchical Storage Manager)

DFSMSHsm is an automated hierarchical storage manager and productivity tool designed to support Availability Management, Space Management, and Disaster Recovery.

21.4.1 The storage hierarchy

A storage device hierarchy consists of a set of storage devices that have differing costs for storing data, differing amounts of data stored, and differing speeds for accessing the data. Storage devices at the lower level typically provide lower cost per byte storage and might have slower access time. These devices might be disk or tape.

21.4.2 Availability management

Availability management is the process that helps maintain a backup copy of data sets which can be easily recovered if the original copy is damaged or accidentally deleted.

DFSMSHsm can automatically make backup copies of changed data sets and copies complete volumes (dumps) according to a definable schedule. DFSMSHsm can also create a point-in-time copy of a set of volumes using the FlashCopy process and automatically move those copies to tape.

21.4.3 Space management

Space management is a process that monitors usage to maintain enough space on user volumes for new and active data.

- DFSMSHsm uses a three-level storage device hierarchy for space management:
 - Level 0 contains data directly accessible to the end user or program.
 - Level 1 is disk containing data sets that DFSMSHsm has moved from level 0 volumes.
 - Level 2 is tape containing data sets that DFSMSHsm has moved from level 1 or level 0 volumes.

- ▶ DFSMSHsm can automatically move (migrates and recalls) data sets, which have not been referenced within a specified time, to lower cost devices. A migrated data set is one that has been moved to a Level 1 or Level 2 volume, DASD or tape, in a special space saving form on the volume.
- ▶ DFSMSHsm can automatically delete data sets that have passed their expiration dates.

21.4.4 Disaster recovery

Disaster recovery is provided by ABARS (Aggregate Backup And Recovery Support).

Aggregate backup is a command-driven function that can back up a user-defined group (called an aggregate group) of data sets for recovery at another computer site or at the same site. Aggregate recovery is the command-driven function that recovers the data sets that were previously backed up by aggregate backup.

21.4.5 List, report, query, audit

DFSMSHsm can manage backed up, dumped, and migrated volumes and data sets. It can maintain an inventory and provides features so users might find information about their own volumes and data sets through LIST, REPORT, QUERY, and AUDIT functions.

Note: IBM recently announced new, higher capacity cartridges for the IBM System Storage TS1120 tape subsystem. These new cartridges have a physical capacity of around 700 GB and a logical capacity of around 2 terabytes when compressed at 3:1. DFSMSHsm can use these high capacity tapes for all tape related functions. For more details on DFSMSHsm, see the website:

<http://www.ibm.com/systems/storage/software/sms/hsm/index.html>

21.5 DFSMSrmm (Removable Media Manager)

DFSMSrmm is the functional component of DFSMS for managing all removable media resources including automatic, manual, and virtual tape libraries (such as IBM's TS7700 series), for the z/OS environment.

21.5.1 Management of the tape library

DFSMSrmm is an integral part of DFSMS and is shipped as part of DFSMS with the z/OS operating system. DFSMSrmm cooperates with and exploits numerous system components that are required for the complete security, management, and well being of your tape library.

21.5.2 DFSMSrmm volume retention and movement policies

DFSMSrmm volume retention and movement policies, can optionally be specified interactively with ISPF panels. This allows authorized application owners to alter existing values without contacting the tape librarian.

21.5.3 Report generator

DFSMSrmm includes the report generator which is an ISPF based tool for easily building and customizing numerous storage management reports. Sample report types and reports enable reporting based on DCOLLECT data, DFSMSHsm functional statistics, and DFSMSrmm created data.

21.5.4 Full function removable media management system

DFSMSrmm is a full function removable media management system for your enterprise. DFSMSrmm supports tape volumes which are used anywhere in your enterprise, including z/VM, and those managed by the IBM Integrated Removable Media Manager for the enterprise on System z.

21.5.5 Compatible with existing tape management systems

DFSMSrmm is functionally compatible with existing tape management systems and runs in parallel during conversion. The easy to use conversion tools and related documentation and Redbooks publications enable the easy migration of your existing tape environment to management by DFSMSrmm. The special expiration date formats of 99nnn and 98nnn are supported by DFSMSrmm.

21.6 DFSMStvs

DFSMS Transactional VSAM Services (DFSMStvs) is an optional z/OS feature that enables batch jobs and CICS online transactions to update shared VSAM data sets concurrently. Multiple batch jobs and online transactions can be run against the same VSAM data sets. DFSMStvs helps ensure data integrity for concurrent batch updates, while CICS ensures it for online updates of these features:

- ▶ Contributes to the reduction or elimination of the batch window for CICS applications and other VSAM applications by allowing concurrent batch and online updating of VSAM recoverable data sets.
- ▶ Provides the ability to share VSAM data sets at record-level with integrity and commit and rollback functions for non-CICS applications.
- ▶ Offers Backup-while-open to be taken using DFSMSdss and DFSMSHsm.
- ▶ Enables batch applications to use the same forward recovery logs.
- ▶ Delivers increased system availability with simpler, faster, and more reliable recovery operations for VSAM storage structures.
- ▶ Simplifies scheduling batch jobs because multiple batch jobs that access the same files can be run concurrently on one or more z/OS images in a Geographically Dispersed Parallel Sysplex™ (GDPS®) instead of serially on one image.

21.7 More information

For more information about DFSMS, see the website:

<http://www.ibm.com/systems/storage/software/sms/index.html>

Also see the following Redbooks publications:

- ▶ *DFSMSHsm ABARS and Mainstar Solutions*, SG24-5089
- ▶ *IBM TotalStorage Virtual Tape Server: Planning, Implementing, and Monitoring*, SG24-2229
- ▶ *z/OS V1R3 DFSMS Technical Guide*, SG24-6569
- ▶ *Z/OS V1R3 and V1R5 DFSMS Technical Guide*, SG24-6979

21.8 DFSORT

Many organizations have a “batch window,” the time during which the overnight batch jobs have to perform processing related to the previous day’s activity. As system capabilities and usage have increased, so too, the number of reports required and the amount of data to be processed by the batch jobs has increased. As the batch window cannot be extended, ways have to be found to process the data as quickly and efficiently as possible. As many of the batch jobs require data to be sorted, particularly those producing reports, improved sorting can make a contribution to the goal of processing more data in the same period of time.

21.8.1 Product highlights

Here we list highlights for DFSORT:

- ▶ DFSORT, together with DFSMS and RACF®, form the strategic product base for the evolving system-managed storage environment.
- ▶ DFSORT adds the ability to do faster and easier sorting, merging, copying, reporting and analysis of your business information, as well as versatile data handling at the record, fixed position/length or variable position/length field, and bit level.
- ▶ DFSORT is designed to optimize the efficiency and speed with which operations are completed through synergy with processor, device, and system features (for example, memory objects, Hyperspace, data space, striping, compression, extended addressing, DASD and tape device architecture, processor memory, processor cache, and so on) and other products (for example, The SAS System, COBOL, PL/I, IDCAMS BLDINDEX, and so on).
- ▶ DFSORT includes the high-performance ICEGENER facility, the versatile ICETOOL utility, Symbols, and multiple output capability with the powerful OUTFIL feature.
- ▶ Sort, merge or copy data sets while including, excluding or reformatting records.
- ▶ Analyze data and produce detailed reports using the ICETOOL utility or the OUTFIL function. In addition, OUTFIL allows you to create various views of the data and various reports with a single pass over the data.
- ▶ Use symbols for fields and constants in DFSORT and ICETOOL statements.
- ▶ Join and match records from separate data sets in various ways.
- ▶ Convert FB data sets to VB data sets, and VB data sets to FB data sets.
- ▶ Add timestamps to records, include or omit records based on today's date, a past date or a future date, and convert SMF, TOD and ETOD date and time values to readable form.
- ▶ Collate and compare data according to your national and cultural needs using DFSORT's National Language Support.

21.8.2 DFSORT/VSE

DFSORT/VSE is IBM's high-performance sort, merge, copy, analysis and reporting product for VSE/ESA.

21.8.3 Features

Here are the main features provided:

- ▶ Sort, merge, or copy data sets while including, excluding or reformatting records.
- ▶ Collate and compare data according to your national and cultural needs using DFSORT/VSE's National Language Support.
- ▶ Performance: Improvements with DFSORT/VSE help reduce batch window times and increase system throughput.
- ▶ Reduced: Input/output times and work space requirements.
- ▶ Optimize: Efficiency and processing speed increased by using data space, virtual storage, and extended addressing.
- ▶ Analyze data and produce detailed reports using the ICETOOL utility.
- ▶ Sort, merge and transform two-digit year dates correctly using a specified fixed or sliding century window.
- ▶ Create reports quickly and easily using DFSORT/VSE's ICETOOL utility.
- ▶ Sort using virtual disk space rather than DASD.
- ▶ Specify secondary allocation amounts for SAM ESDS work files, which allows more efficient use of work space.

Note: For more details on DFSORT/VSE, see the website:

<http://www.ibm.com/support/docview.wss?rs=0&uid=isg3T7000098>



z/OS Storage Management Tools

The z/OS storage management tools are a collection of products designed to provide a comprehensive menu of solutions in order to meet the storage management requirements of z/OS customers. The tools are positioned to complement and extend the storage management capabilities of DFSMS.

The products allow IBM to provide a more complete suite of storage management solutions:

- ▶ Ability to predict results of alternative actions with minimum risk
- ▶ Overall ability to optimize the responsiveness of storage assets
- ▶ Better response time and throughput by elimination of easily identified bottlenecks
- ▶ Provides storage administrators a more simple, straightforward environment, increasing their productivity while reducing mistakes as well as potential outages
- ▶ An increased system availability from simpler, faster, and more reliable recovery operations for storage structures

The following tools can be used to manage storage on z/OS:

- ▶ IBM Tivoli Omegamon XE
- ▶ IBM DB2 Cloning Tool
- ▶ IBM Tivoli Advanced Catalog Management for z/OS
- ▶ IBM Tivoli Advanced Backup and Recovery
- ▶ IBM Tivoli Advanced Audit for DFSMSHsm
- ▶ IBM Tivoli Automated Tape Allocation Manager
- ▶ IBM Tivoli Storage Optimizer for z/OS
- ▶ IBM Tivoli Tape Optimizer for z/OS
- ▶ DFSMS Optimizer

22.1 IBM Tivoli Omegamon XE

IBM Tivoli OMEGAMON® XE for Storage on z/OS enables you to monitor and manage mainframe-attached storage. It combines comprehensive storage monitoring for z/OS I/O subsystem performance and storage availability.

Features: Here we list the features provided:

- ▶ Enables you to monitor and manage storage (including DASD and tape devices), and the analysis of two important IBM storage software components: Data Facility Systems Managed Storage and Data Facility Hierarchical Storage Manager (DFHSM).
- ▶ Provides the capability to quickly create commands or schedule actions that maintain and administer DASD storage with a new storage toolkit for DFHSM and DFDSS functions.
- ▶ Offers more versatile and granular reporting capabilities using the storage toolkit.
- ▶ Improves problem resolution with Dynamic Workspace Linking (DWL) to OMEGAMON XE on z/OS, which enables you to easily navigate between Tivoli Enterprise Portal workspaces.
- ▶ Adds enhanced problem determination capabilities with new problem-solving workspaces in the Tivoli Enterprise Portal (TEP).
- ▶ V4.1.0 is fully globalized and translated into group one languages

22.2 IBM DB2 Cloning Tool

IBM DB2 Cloning Tool for z/OS is a vendor-independent tool that provides access to data sets on replicated volumes created with point-in-time copies of Fast Data Replication tools (FlashCopy and SnapShot) or Splits of Continuous Mirrors tools (EMC TimeFinder, IBM PPRC, HDS ShadowImage, Softek TDMF®, and Fujitsu Equivalent Copy).

The clones these tools create are not accessible due to an inherent replication problem: the target volume label, internal data, and data set names all reflect the source volume name. DB2 Cloning Tool solves these issues so that the replicated data can be utilized quickly and accessed on the same system.

DB2 Cloning Tool supports the replication of DB2 subsystems with intelligent features that condition the target environment enabling the replicated data to be made available for use within minutes.

DB2 Cloning Tool provides a way for customers to meet the demand for increased online access to databases while continuing to meet business requirements for batch and backup processing. Using the tool, customers can run processes in parallel. In addition, customers can easily create test and QA environments and refresh the data on a regular basis.

Features: Here we list the features provided:

- ▶ IBM DB2 Cloning Tool is an innovative tool that solves the cloned data access dilemma, giving users access to datasets on target volumes created with FlashCopy or Snapshot by renaming and cataloging them.
- ▶ IBM DB2 Cloning Tool resolves catalog conflicts of like-named datasets that are created by cloning a volume. It also solves internal conflicts that are created when copying a volume to another VOLSER.
- ▶ Fixes volume conflicts (VTOC, VTOCIX, and VVDS) and then renames and recatalogs the datasets

- ▶ Rename capability allows users to change any qualifier or add/delete qualifiers
- ▶ Track the behind-the-scenes FlashCopy, enabling a user to know when they can start another cloning operation or withdraw from the current copy
- ▶ Provide automatic pairing of volume characteristics
- ▶ Provide the capability to FlashCopy or Snapshot by Storage Groups

22.3 IBM Tivoli Advanced Catalog Management for z/OS

IBM Tivoli Advanced Catalog Management for z/OS offers vital data protection while helping to increase data availability. With effective auditing tools, secure recovery, and improved day-to-day management, you can access your data, even in the event of problems in the IT environment ranging from simple human errors to natural disasters. Effective catalog management requires the best tools available to ensure that key data assets are protected and can be recovered in any event.

Advanced Catalog Management for z/OS is designed to provide backup and restore facilities that provide high integrity and fast processing.

Infrastructure benefits: Infrastructure benefits for the product are as follows:

- ▶ Backup and restore facilities that provide high integrity and fast processing.
- ▶ A forward recovery facility for BCS and VVDS catalog structures that use SMF data, merged with backup records, to update the BCS and VVDS from the time of the backup to the current time.
- ▶ A full-function backup and restore facility for VSAM files that allows attribute changes during recovery and can back up damaged KSDS data sets.
- ▶ An expanded diagnostic facility that enables frequent health checks of the BCS, VVDS, VTOC, and tape management catalog structures. This facility contains an audit-check and 'fix' capability to re-synchronize entries that are in error.
- ▶ A fast and safe "MERGECAT" facility that enables moving or copying of individual BCS entries, alias groups of BCS entries, or entire catalogs.
- ▶ A fast and safe split/merge facility that enables moving or copying of individual BCS entries, alias groups of BCS entries, or entire catalogs.
- ▶ The "Reorg While Open" and "Repair While Open" features reorganize and repair ICF catalogs while applications are still active. These high availability features can save hours of application downtime over traditional tools.
- ▶ The feature to allow reorganization and repair of, and attribute changes for ICF catalogs while applications are still active. These high availability features can save hours of application downtime over traditional tools.

22.4 IBM Tivoli Advanced Backup and Recovery

IBM Tivoli Advanced Backup and Recovery solution provides powerful function for reliable, accurate backups, and even helps to identify data that is not already backed up. It can reduce costs associated with remote DASD replication.

Advanced Backup and Recovery can centralize and track all backups including ICF catalogs. This enables customer environments to bounce back quickly and fully from unplanned local outages, saving crucial time and money.

Infrastructure benefits: Infrastructure benefits for the product are as follows:

- ▶ ABARS Manager enhances the value of DFSMSHsm ABARS:
 - All backup information and backup status is accessible through one central point
 - Backup status and details can be viewed online
 - Online access to the full activity log and all error messages, as well as condensed status summary, makes ABARS management simple
 - Detail dataset inventory provides recovery capacity requirements, recovery tape VOLSER, and other key planning and operational details
- ▶ Incremental ABARS allows users to combine their disaster recovery back up strategy with an incremental backup capability. The benefits of this are a reduced backup size, improved backup and recovery performance, and reduced resources to vault and maintain the backup.
- ▶ With ABARS Manager, the value of ABARS is expanded to provide a local recovery solution that allows searching and recovering selected data sets from ABARS or DFSMSHsm backups.
- ▶ Automated Critical Data Identification interfaces directly into ABARS Manager to ensure all critical assets are included in the backup.
- ▶ Critical Backup Tracking / Inventory ensures that all critical non-ABARS-managed assets have a backup, tracks and inventories all other backups in the environment and identifies duplicate backups.

22.5 IBM Tivoli Advanced Audit for DFSMSHsm

IBM Tivoli Advanced Audit for DFSMSHsm offers comprehensive and extremely fast audit support in the following areas: DFSMSHsm Migration, Backup, Offline Control DataSets (MCDS, BCDS, and OCDS), and DFSMSHsm owned tape.

Advanced Audit for DFSMSHsm evaluates the Migration, Backup, and Offline Control DataSets (MCDS, BCDS, and OCDS) to resolve structural and logical discrepancies that might have prevented migrated data from being recalled, or backup data from being restored. Advanced Audit for DFSMSHsm checks the ICF catalog relationships to assure that migrated data is cataloged, and that data cataloged as “migrated” is actually migrated. This function duplicates the DFSMSHsm audit and enhanced audit commands with improved performance and flexibility.

The IBM Data Facility System Managed Storage Hierarchical Storage Manager (DFSMSHsm) manages hundreds of thousands to millions of data assets, both critical and non-critical, in licensed IT environments. As businesses grow the sheer quantity of data assets stored on z/OS increases commensurately. This growth is mitigated using DFSMSHsm to manage the life cycle of data assets, from creation to expiration. DFSMSHsm is a complex environment requiring audit and reporting solutions that reduce resource consumption and help bridge the learning curve for new or inexperienced DFSMSHsm administrators.

Advanced Audit for DFSMSHsm enables customers to manage DFSMSHsm environments using less time and fewer resources and provide powerful enhancements over the base DFSMSHsm offering. The solution provides proactive DFSMSHsm management capabilities that permit customers to produce a truly resilient DFSMSHsm environment.

Infrastructure benefits: Advanced Audit for DFSMSHsm provides an easy-to-use, high performance method for auditing the DFSMSHsm environment, including migration and backup tapes. It can easily and automatically identify problems and implement corrective actions to pre-defined scenarios. It makes frequent auditing feasible, and makes it easy to keep the DFSMSHsm environment accurate.

22.6 IBM Tivoli Automated Tape Allocation Manager

IBM Tivoli Automated Tape Allocation Manager for z/OS addresses tape device availability, providing automatic, unattended sharing of tape resources for across multiple images. Tivoli Automated Tape Allocation Manager helps ensure existing tape drives are available on demand for requests from multiple images. Its tape sharing capabilities and commands can help manage tape devices across your System z environment reducing tape-related bottlenecks and job request backlogs

Features: Here we list the features provided:

- ▶ Maximize use of existing tape resources before you invest in additional hardware Minimize batch processing delays and operational overhead by sharing tape storage devices across a multi-system environment without the need for a shared control file.
- ▶ Automate tape resource management and minimize the need for operator intervention.
- ▶ Migrate from alternative vendor products to an IBM solution.
- ▶ New in V3.1 Integration with Tivoli Enterprise Portal (TEP).
- ▶ Configuration T1 parameter can be adjusted to specify particular timing intervals for various device types.
- ▶ A new “E” command shows a map of tape esoteric names and associated devices.
- ▶ Option to use the SAF security interface to help control your ability to manage tape availability via Automated Tape Allocation Manager for z/OS, V3.1.
- ▶ Enhanced ability to employ preferential tape device assignment.

22.7 IBM Tivoli Advanced Allocation Management on z/OS

IBM Tivoli Advanced Allocation Management helps businesses increase productivity and maximize data availability by preventing and recovering from X37 abends, using a single solution.

Features

Here we list the features provided:

- ▶ Proactively intercept, resolve and recover from X37-type abends using a single, integrated solution.
- ▶ Minimize out-of-space conditions that can hamper productivity.
- ▶ Provide comprehensive activity reporting to view total abends and errors avoided or recovered.
- ▶ Centrally control DASD usage both at initial allocation and end-of-volume processing
- ▶ Supports both SMS and non-SMS data sets, using flexible rule definitions to specify criteria.

- ▶ Utilize IBM Tivoli Enterprise Portal as a common user interface for managing information and tasks for a wide variety of solutions.
- ▶ Share information with Tivoli OMEGAMON XE for Storage for more effective problem discovery, investigation and resolution

22.8 IBM Tivoli Storage Optimizer for z/OS

IBM Tivoli Storage Optimizer for z/OS provides a single focal point for the discovery, monitoring and managing of z/OS storage devices across the enterprise. IBM Tivoli Storage Optimizer for z/OS tracks capacity and usage changes over time enabling you to stay ahead of storage demand. FlashCopy

Features: Here we list the features provided:

- ▶ Provides high-availability and fault-tolerant architecture Offers optimized monitoring.
- ▶ Helps you identify root cause for more in-depth analysis and problem resolution.
- ▶ Provides strategic planning that enables you to analyze space, size and growth rates

22.9 IBM Tivoli Tape Optimizer for z/OS

IBM Tivoli Tape Optimizer for z/OS maximizes the efficiency of tape-copy processes and optimize utilization of tapes.

IBM Tivoli Tape Optimizer on z/OS is a tape copy and stacking solution for data residing on tape storage devices that are managed by Data Facility Storage Management Subsystem removable media manager (DFSMSrmm). You can copy a single tape or data set, or potentially your entire tape library

Features: Here we list the features provided:

- ▶ V2.1 can copy tape volumes to other tapes or tape-compatible storage.
- ▶ Rename tape data sets Preserve DFSMSrmm tape library information.
- ▶ Update the system catalog for the copied tapes

22.10 DFSMS Optimizer

The DFSMS Optimizer solution contains two features: The Performance Analyzer uses historical and real time data to provide an overall data usage picture and the Management Class Analyzer provides cost benefit analyzes and what-if simulations of DFSMS Management Class policies. The Optimizer supports both System-Managed Storage (SMS) and non-SMS environments. The Performance Analyzer uses historical and real time data to provide an overall data usage picture and The Management Class Analyzer provides cost benefit analyzes and what-if simulations of DFSMS Management Class policies. The Optimizer supports both System-Managed Storage (SMS) and non-SMS environments.

Maximizing the use of your storage resources while helping minimize your overall storage costs. DFSMS Optimizer includes a Management Class Analyzer that provides cost-benefit analyses and what-if simulations of management class policies. DFSMS Optimizer helps you make informed, timely data management decisions for your storage system environment.

Features: Here we list the features provided:

- ▶ Graphical User Interface running on Windows or OS/2.
- ▶ Includes a Performance Analyzer that uses historical and real-time data to provide an overall data usage picture.
- ▶ Powerful SMS filtering controls.
- ▶ High compression (80-90%) of collected SMF, RMF, CRR, and HSM data.
- ▶ Reporting and analysis for both SMS and non-SMS data.

Note: For more information about all the z/OS storage management in details, visit <http://www.ibm.com/systems/storage/software/toolkit/>



Smart Business Storage Cloud

The IBM smart business storage cloud (SBSC) is a range of flexible storage virtualization solutions and appliances that will help alleviate data storage challenges by enabling quick implementation of highly scalable, global, clustered network attached storage systems.

23.1 Overview

IBM Smart Business Storage Cloud offers a storage virtualization solution designed to support your storage optimization efforts. It can help alleviate your data storage challenges by enabling quick implementation of a scalable, global file storage system with flexibility in deployment and management options. The solution provides virtualized storage to enable storage and server consolidation, a unified management platform to help reduce outages and storage management labor demands and costs, as well as advanced data replication for cost effective business continuity and disaster recovery.

IBM offers several types of cloud solutions and appliances, for storage and other services: Smart Business on the IBM Cloud, Smart Business Cloud services, and Smart Business Systems.

23.1.1 Public cloud versus private cloud

There are two main scenarios for storage clouds that IBM customers can choose to pursue based on their business drivers and technical strategy.

The two scenarios are defined as Public Storage Cloud and Private Storage Cloud, the key differentiators of these scenarios is that the public storage cloud is designed for customers who do not want to own, manage, or maintain the storage environment thus reducing their capital and operational expenditures cost around storage. The public storage cloud provides for variable billing options and shared tenancy of the storage cloud, giving customers the flexibility to manage the usage and growth of their storage needs. This is the industry standard view of a storage cloud offering and is comparable to storage cloud offering by other vendors.

The private cloud has fixed charges and dedicated tenancy, so it is designed for enterprise customers who want flexibility around ownership, management, and maintenance of the storage cloud.

23.1.2 Public cloud

Similar to a rent model, IBM dictates the choice of technology and cloud location, shared infrastructure with variable monthly charges, dynamic physical capacity at customer level and security measures to isolate customer data.

In a public cloud IBM owns the physical assets, facilities, and standard contracts with multiple service level agreements (SLA) to meet specific needs. Public cloud solutions work well for cross industry solutions that are storing from tens of terabytes (TB) to multiple petabytes (PB) of data.

23.1.3 Private cloud

Similar to a purchase or lease model, with a private cloud customers have the choice of technology and location on dedicated infrastructure with fixed monthly charges and physical capacity at the customer level.

Each application can utilize dynamic capacity by sharing the cloud storage among multiple applications. Further the Private Cloud provides built-in security thru platform dedication, choice of asset ownership, and custom service level agreements (SLA) to meet specific needs.

Private clouds also work well for cross industry solutions storing tens of Terabytes (TB) to multiple petabytes (PB) of data.

23.2 Smart Business Storage Cloud components

A Smart Business Storage Cloud is a combination of hardware and software components combined to form a system, or solution, which provides three key features:

- ▶ Global namespace
- ▶ Clustered file system
- ▶ Information Lifecycle Management (ILM)

As shown in Figure 23-1, virtualizing the file sharing system (NFS, CIFS, and so on) using a single, global, namespace means that an end user sees one device rather than individual servers and storage devices that comprise this virtual NAS device.

At the heart of the SBSC system is the IBM General Parallel File System (GPFS). GPFS is one of the most scalable commercial file systems available, installed in thousands of nodes and storing petabytes of data in discrete installations. SBSC enables the virtualization of the NFS service, transparently distributing service requests across multiple servers, which in turn enables the creation of a scale-out NAS farm installation

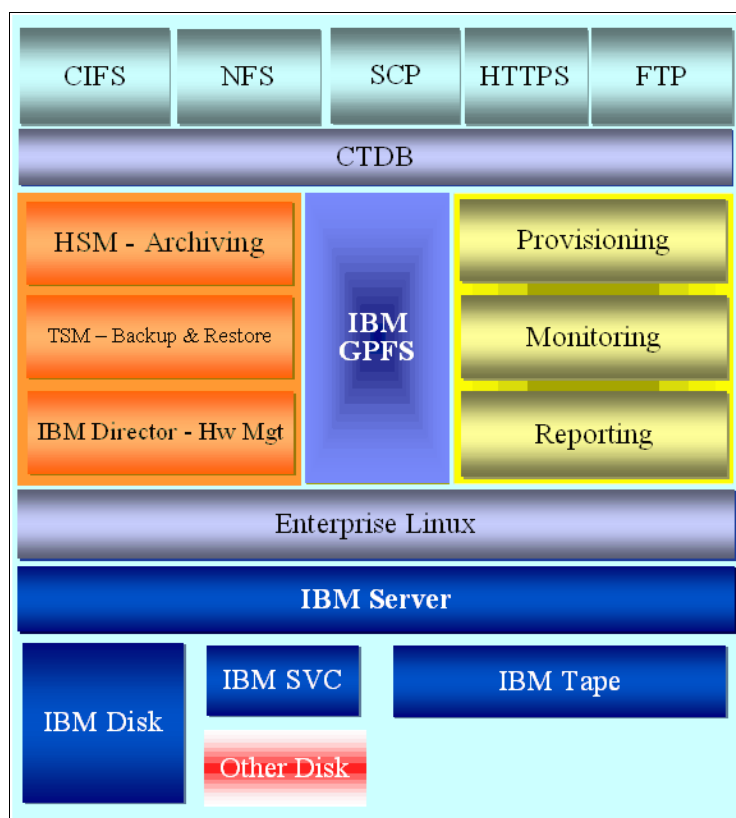


Figure 23-1 SBSC components

The SBSC solution is extremely resilient. It is designed so that there is no single point of failure in the system. It also allows for disaster recovery configurations that use the global file system, these options include synchronous and asynchronous mirroring of data either on site or between sites and cross cluster file system mounts.

SBSC supports snapshots integrated into Windows Explorer using Volume Shadow Copy Services.

The SBSC solution also supports fully configurable data lifecycle management. Data can be stored in ways relative to the data value and performance requirements. File placement can be defined by flexible policies, data can be classified according to service level agreements and placed on appropriate storage tiers automatically at creation time or moved automatically using the integrated ILM functionality. SBSC also supports LTO-4 tape backup with data encryption, this offers protection for archived data.

23.2.1 How it works

The SBSC system provides file services by implementing a *scale-out* approach. SBSC uses the wide-striping capability of IBM GPFS to spread the blocks of a single file across as many nodes as possible to combine the streaming performance of many midrange back-end storage devices. All files are accessed by all nodes, this removes the traditional NAS and earlier global namespace implementation limitations of specific files being 'pinned' to individual filers. By doing so a bandwidth of multiple gigabytes per second is possible because the I/O performance of many disks can be combined instead of storing data only to a small RAID device as filers do today.

By fully exploiting the capabilities of GPFS, the integration of a policy-based life-cycle management interface and the possibility to place files in a directory that spans multiple independent storage classes creates a single file system with multiple storage classes. Each has its own performance, availability, and cost criteria, on a file basis, not on a filer basis, as has been done in the past.

SBSC utilizes and expands the IBM Tivoli Storage Manager and Tivoli Hierarchical Storage Manager products to integrate tape as an additional storage class inside the cloud. This allows transparent migration of data between disk and tape storage and provides end-to-end life-cycle management of the data.

The physical implementation of SBSC utilizes the IBM Bladecenter for highly efficient and scalable processing with various IBM storage devices. The Bladecenter provides a scalable platform for adding storage controllers into the SBSC pool to achieve the right performance balance. Processing blades share power, cooling, and I/O connectivity within the Bladecenter chassis and achieve very efficient performance versus stand-alone servers.

To provide a single global namespace to the user, SBSC uses virtualization and redirection technologies available within GPFS clusters. Each SBSC cluster node has internal access to all data blocks simultaneously and is able to assume the responsibility for each of the other nodes in a cluster. This provides the means for a transparent, nondisruptive protocol failover for client applications. As none of the existing cluster technologies provides a transparent failover of network-based file systems, SBSC introduces an extremely scalable cluster suite dealing with the specifics of file systems to provide a high degree of semantic correctness while being able to scale to very high bandwidths.

23.2.2 The difference between SBSC and other clustered file systems

While other clustered file systems are just that, SBSC is more of a "storage controller in a grid," meaning that the ownership of data is handled dynamically. In many clustered file system implementations, there is one node that is the default owner of a portion of the data. To maintain this structure requires high volumes of internode communication and in case the owning node goes down the IP and data must be taken over by another node. While IP takeover is not a big problem, the takeover of data often proves difficult.

As can be seen in Figure 23-2, with SBSC, this problem is completely overcome, because any node can see any data at any time.

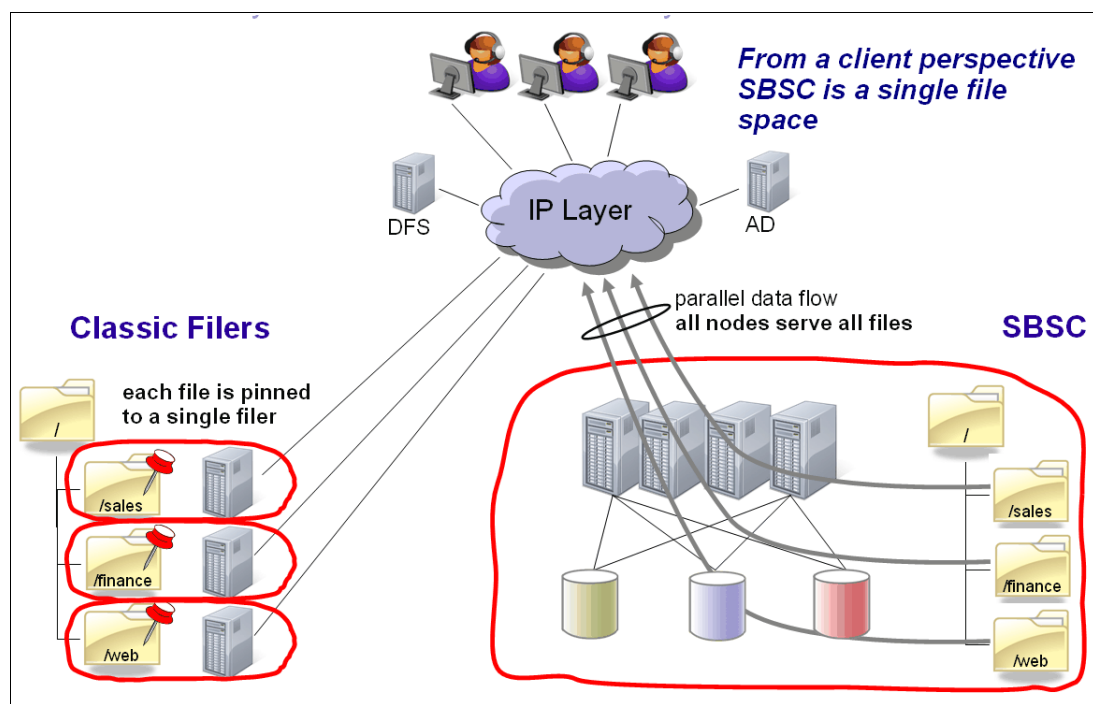


Figure 23-2 The difference between SBSC and traditional approach

23.3 The benefits of using an SBSC

The efficient management of data is an ongoing struggle between access and scalability. Providing access to file-level data (i.e., those files associated with individual documents, multimedia content, databases, and other applications) becomes more difficult as the number of users with access and the amount of data stored both grow.

Achieving the scalability needed to respond to the growth of data volume typically results in higher hardware and software costs and greater management challenges.

Network-attached storage (NAS) solutions provide simplicity, manageability, and access, but until now, they lacked a single capability that has kept them from playing a role beyond the departmental level: scalability.

If a traditional filer is reaching a scalability or capacity limit, users upgrade to the next higher filer level until they reach the highest available system; this is the *scale-up* approach.

Instead of scaling up in a single high-end enterprise-class server node or building failover clusters to provide high availability for the filers, the SBSC system utilizes software functions to build a global namespace and exposes to the end user *a single virtual cluster node* with very high scalability.

23.3.1 Continuous availability

SBSC allows disk subsystems to be added, removed, and migrated non-disruptively as can be seen in Figure 23-3.

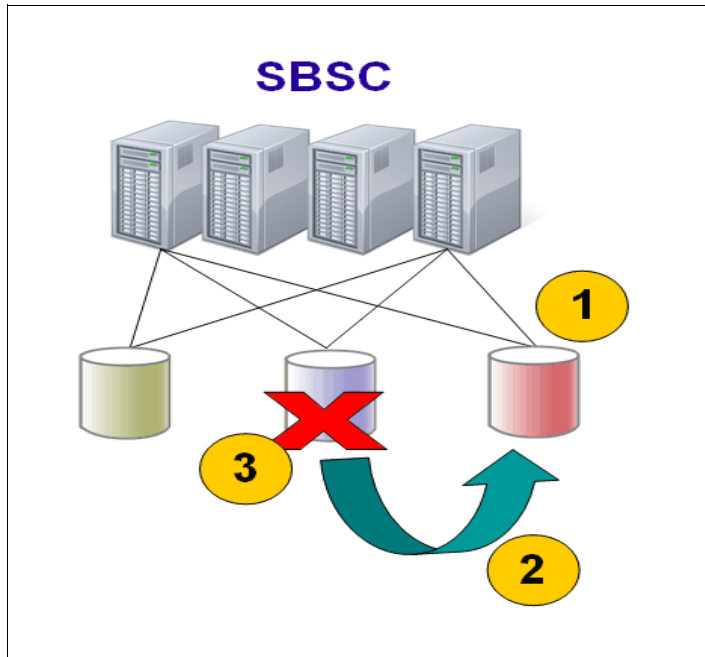


Figure 23-3 non disruptive upgrades

1. Add new disk subsystem
2. Redistribute data excluding storage to be removed
3. Remove old subsystem

23.3.2 Online transparent data migration

Existing NAS storage can be easily integrated into a cloud by allowing existing devices to become part of the global namespace. This enables migration from old storage to be completed automatically and completely transparently as shown in Figure 23-4.

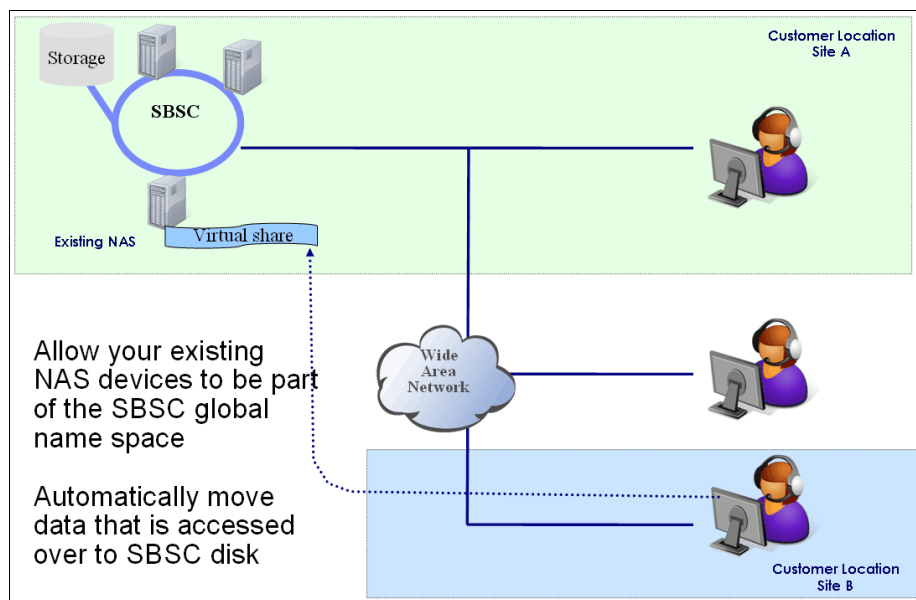


Figure 23-4 online data migration

23.3.3 Summary of benefits

Cloud computing has the potential to make an enormous impact to your business by:

- ▶ Reducing IT labor cost by 50% in configuration, operations, management and monitoring
- ▶ Improving capital utilization by 75%, significantly reducing license costs
- ▶ Reducing provisioning cycle times from weeks to minutes
- ▶ Improving quality, eliminating 30% of software defects
- ▶ Reducing end user IT support costs by up to 40%

For more information, see the website:

<http://www.ibm.com/systems/storage/solutions/cloud/index.html>

Storage services are among the easiest to adapt to the cloud computing model. IT storage teams commonly provide storage “as a service” to other departments, so Best Practices and service level agreements have become standardized over the years.

IBM offers an expanding menu of cloud storage services, available on a pay-per-use basis, delivered from state-of-the-art IBM data centers. IBM also helps clients implement public, private, and hybrid cloud storage services, with expert consulting and workload optimized systems.

The following summary shows business, operational and technical benefits that can be delivered by a Smart Business Storage Cloud.

Business benefits

- ▶ Significant storage cost reduction
- ▶ Faster access to data
- ▶ Geographic access to data
- ▶ Smart/Green adaptability
- ▶ ‘Bottomless’ cloud storage

Operational benefits

- ▶ Simplified Infrastructure
- ▶ Single point of management
- ▶ High data availability
- ▶ Seamless scalability
- ▶ Better storage utilization
- ▶ Improved performance
- ▶ Greater application support

Technical benefits

- ▶ Single global namespace
- ▶ Multi-protocol data access
- ▶ Simplified administration and back-up
- ▶ Automated, policy based placement and migration of files
- ▶ Scalable to billions of files
- ▶ Faster provisioning

23.4 Components within an SBSC solution

The following sections contain a brief overview of the main components of an SBSC system.

23.4.1 Host attachment

The SBSC host attachments shown in Figure 23-5 are primarily provided through clustered CIFS and NFS services, using the new cluster trivial database (CTDB) it is possible to provide transparent, non-disruptive failover of CIFS and NFS between SBSC nodes. This also provides unparalleled aggregate performance scaling, Intelligent load balancing, and simultaneous access to a single file by heterogeneous clients.

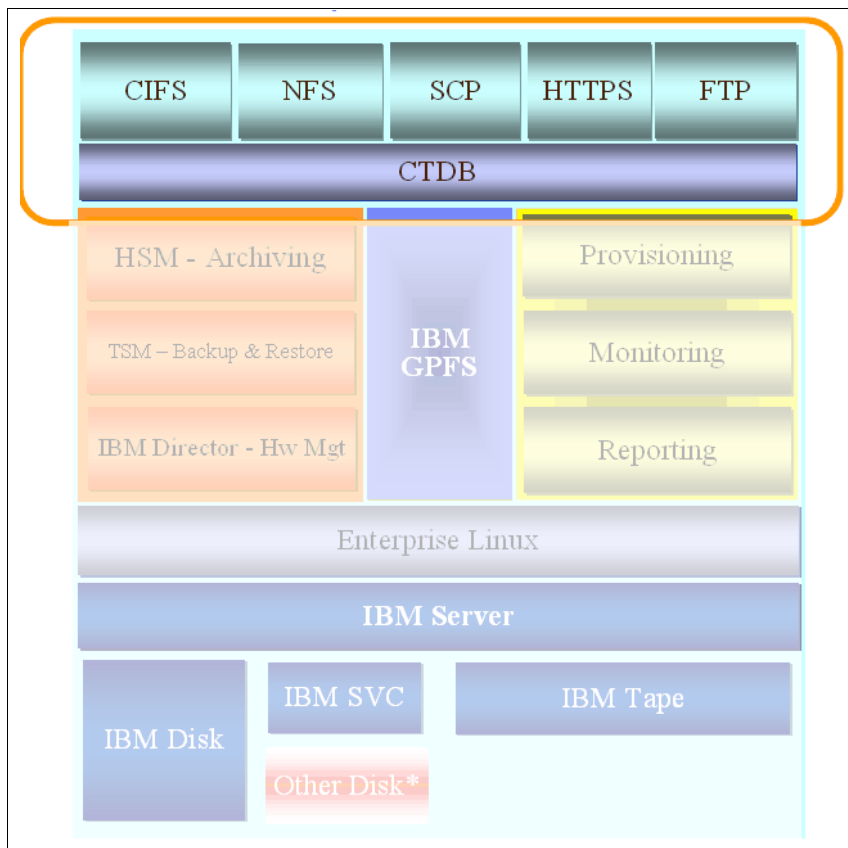


Figure 23-5 Host connectivity

23.4.2 Hardware

The hardware components of a SBSC solution consists of servers, SAN switches, and disk storage. Tape storage for archiving data is an optional part of any SBSC solution.

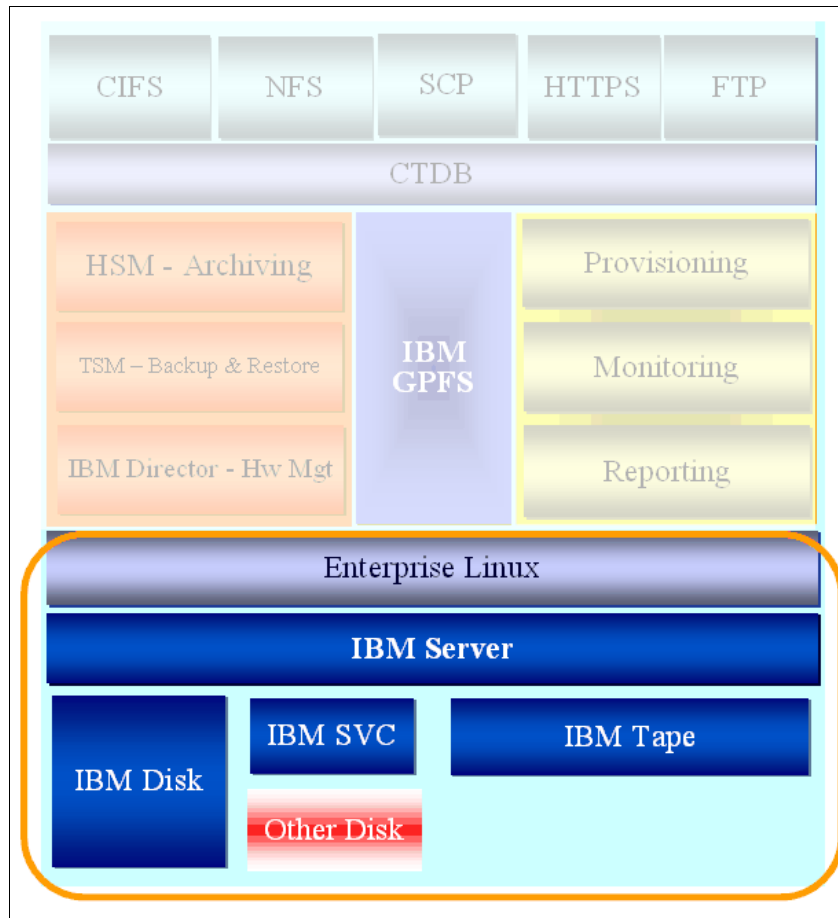


Figure 23-6 hardware components

Storage server

The servers used in a SBSC solution are clusters of BladeCenter HS21 or HS21 XM blades or rack mounted X3650M2 servers. Production servers might be connected by Gigabit or 10 Gigabit Ethernet. Connectivity to the switches and storage units is by Fibre Channel only for the blades, and Fibre Channel or SAS attachment for the rack mounted servers.

As a minimum three System x servers are required. All servers in a configuration will be the same type. A separate management console might be added or that function may be located in one of the System x servers. The BladeCenter HS21 or HS21 XM blades might be housed in a BladeCenter H with Gigabit Ethernet or a BladeCenter H with 10 Gigabit Ethernet.

SAN hardware

The SAN switch hardware currently used for SBSC systems can be seen in Table 23-1. The base configuration of these switches activates less than the maximum number of Fibre Channel ports on the switch but for SBSC configurations all ports will be activated.

Table 23-1 SBSC SAN switch hardware

Switch model	Description
2005-B16	SAN16B-2, 16-port 4Gbps Fibre Channel Switch
2498-24E	SAN24B-4, 24-port 8Gbps Fibre Channel Switch
2498-40E	SAN40B-4, 40-port 8Gbps Fibre Channel Switch

Storage units

Data is stored on one or more of the IBM storage system units described in Table 23-2.

Table 23-2 SBSC storage units

Storage unit	Description
XIV (2810-A14)	This large solution provides a capacity of 180TB and uses only 1 TB SATA disk drives. The system has up to 120 GB cache and 24 FC host ports.
DS8000	SoFS will support the use of DS8100, DS8300 and DS8700 systems, but a special request must be made. Configurations will be customized to the particular solution. This large solution provides up to 512 TB using FATA disks or 307.2 TB using FC disks, up to 32 GB cache, up to 128 FC host ports
IBM SVC	SoFS will support attachment of IBM SVC managed devices, this means any devices not on the supported hardware list can be used for SoFS storage devices if they are supported by SVC.
DCS9950	SoFS will support the use of the DCS9950, but a special request must be made. Configurations will be customized to the particular solution. This mid size solution provides up to 960TB using SATA disks or 448TB with FC, RAID-6, 8 FC host ports, 3 GB/sec read/write speed for sequential I/O
DS4800	This mid size solution provides up to 168 TB using SATA disks or 67.2 TB using FC disks, up to 16 GB cache, 8 FC host ports
DS4200	This mid size solution provides up to 84 TB using SATA disks, 2 GB cache, 4 FC host ports.
DS3400	This dual controller model of the DS3400 is the only one supported in SoFS configurations. Any number of DS3400 systems that can fit into the SAN network are supported in SoFS configurations. The preferred LUN configuration with 4 enclosures is RAID 5 in a 3+P arrangement so that each drive is in a separate enclosure. As a result, the system can survive the loss of a complete enclosure without losing a LUN.
DS3200	This entry level solution provides up 14.4 TB using SAS disks, up to 1 GB cache, and 6 SAS host ports. See the section and figure below on SAS configurations.

Note: Supported hardware device information supplied by Techline and current at 1st Oct 2009. Consult your IBM or business partner sales team for updated information.

23.4.3 Software components

IBM GPFS 3.2 product on servers running Red Hat Enterprise Linux 5 is the heart of the SBSC system, this provides the SBSC cluster nodes with a single-system view of the file system. See Figure 23-7 for an overview of the software components of SBSC system.

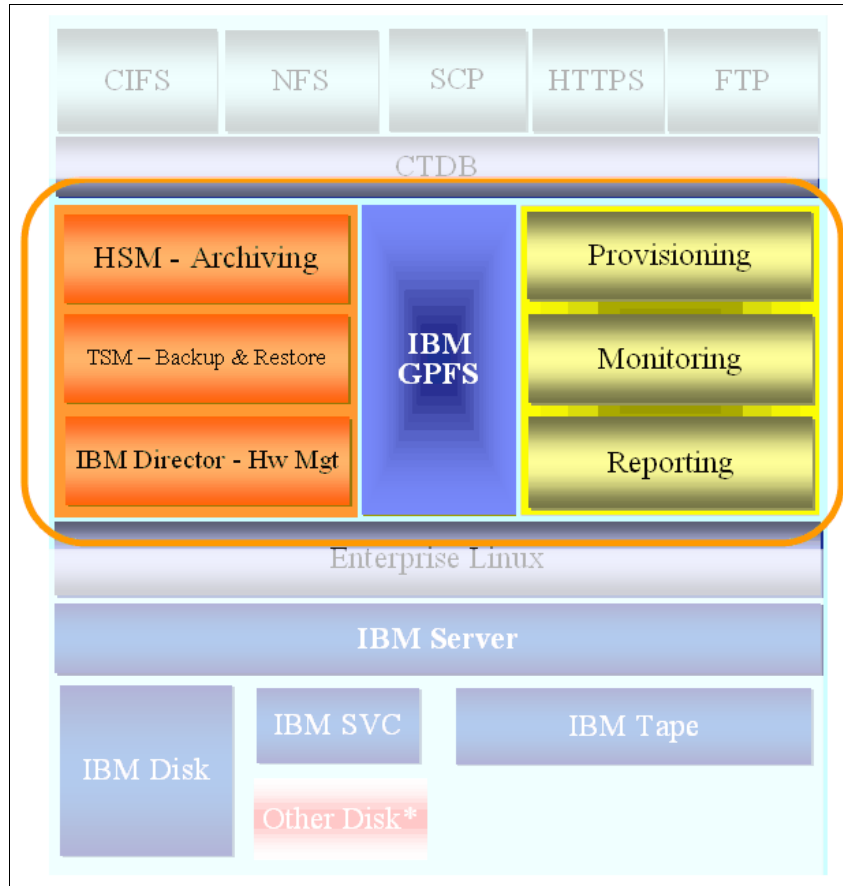


Figure 23-7 software components

Other software components included and their function are shown in Table 23-3.

Table 23-3 SBSC software components

Component	SBSC function
IBM Director	Hardware management
IBM Tivoli Storage Manager	Backup and restore
IBM Tivoli Hierarchical Storage manager	Archiving
DS Storage Manager SMclient	Storage device management
SBSC GUI	User interface
SBSC Toolset	Management of SBSC system

23.5 Data protection and disaster recovery

There are a number of choices the operator can make in the area of data protection and disaster recovery. SBSC allows for appropriate choices for each type of data, giving maximum flexibility to enable the right choice for each tier of data.

23.5.1 Data protection

The data protection choices match the choices available on the storage system used in the SBSC solution. These typically include RAID 5, RAID 6, RAID 10, and the unique XIV data protection algorithm.

23.5.2 Simple SBSC single site solution: No disaster recovery

A simple SBSC solution allows users to access data locally using a LAN, and remotely using a WAN. While the high availability is built into a single system, there is no disaster recovery built into this solution.

23.5.3 SBSC with synchronous replication on-site

The operator can set up synchronous replication of part or all of the data to other nodes within the same SBSC site as shown in Figure 23-8

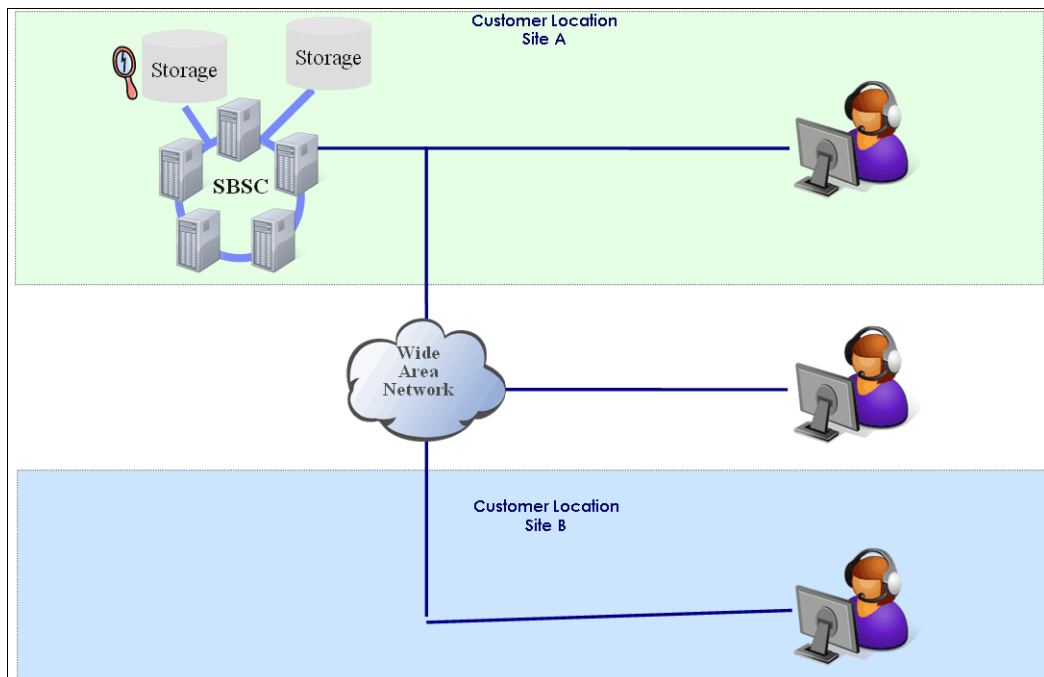


Figure 23-8 Sync replication on a single site

23.5.4 SBSC with cross site sync or async replication

The operator can set up replication of data between SBSC clusters on separate sites as shown in Figure 23-9. The replication can be synchronous if the distance can be traversed by Fibre Channel cables. It is asynchronous if the distance is longer.

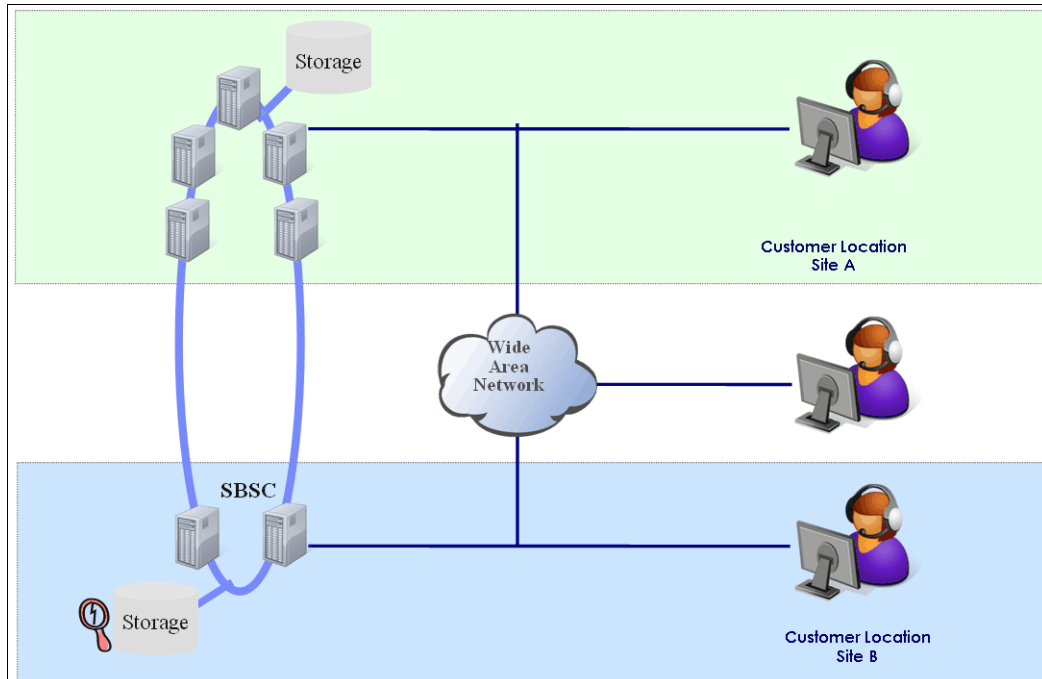


Figure 23-9 Cross site replication

23.5.5 SBSC with a cross cluster mount

Two SBSC sites can be connected and each file system mounted on the other to provide a solution that works as one large SBSC that covers two sites as shown in Figure 23-10. Within that complex, mirrors of data can be set up on separate sites within the single SBSC solution. The cross cluster mount requires a TCP/IP connection between the clusters. This IP connection can be used for data and metadata traffic. As an option there is also the ability to send data traffic over a Fibre Channel (SAN) connection between the clusters.

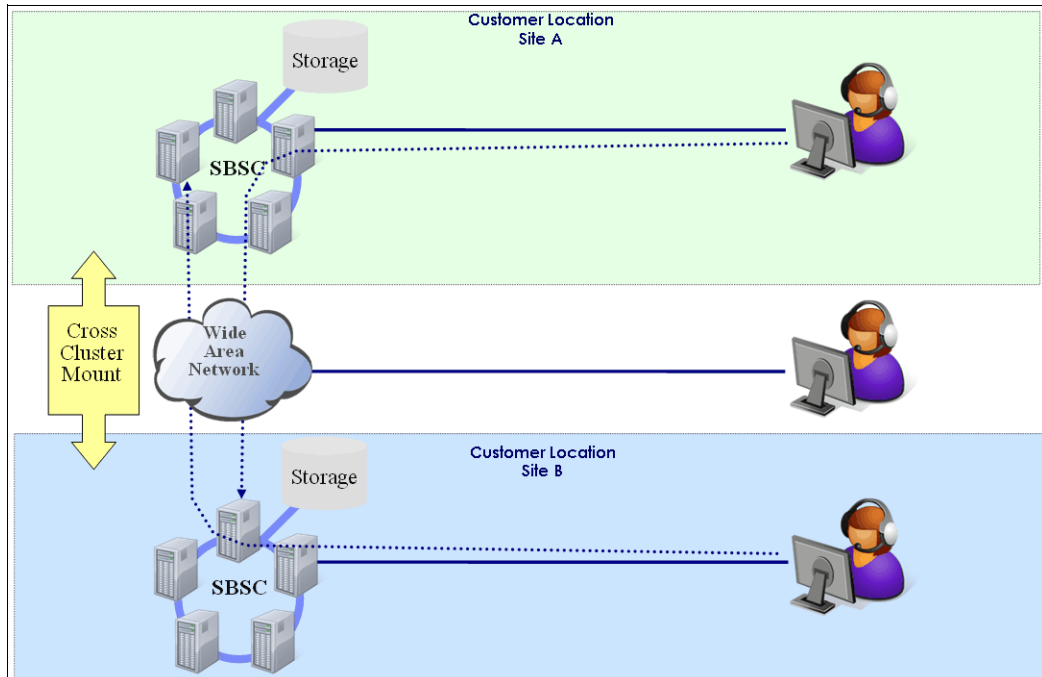


Figure 23-10 Cross cluster mounts

23.5.6 Snapshot technology

A snapshot of an entire GPFS file system (Figure 23-11) might be created to preserve the file system's contents at a single point in time. A snapshot contains a copy of only the file system data that has been changed since it was created, using a copy on write technique.

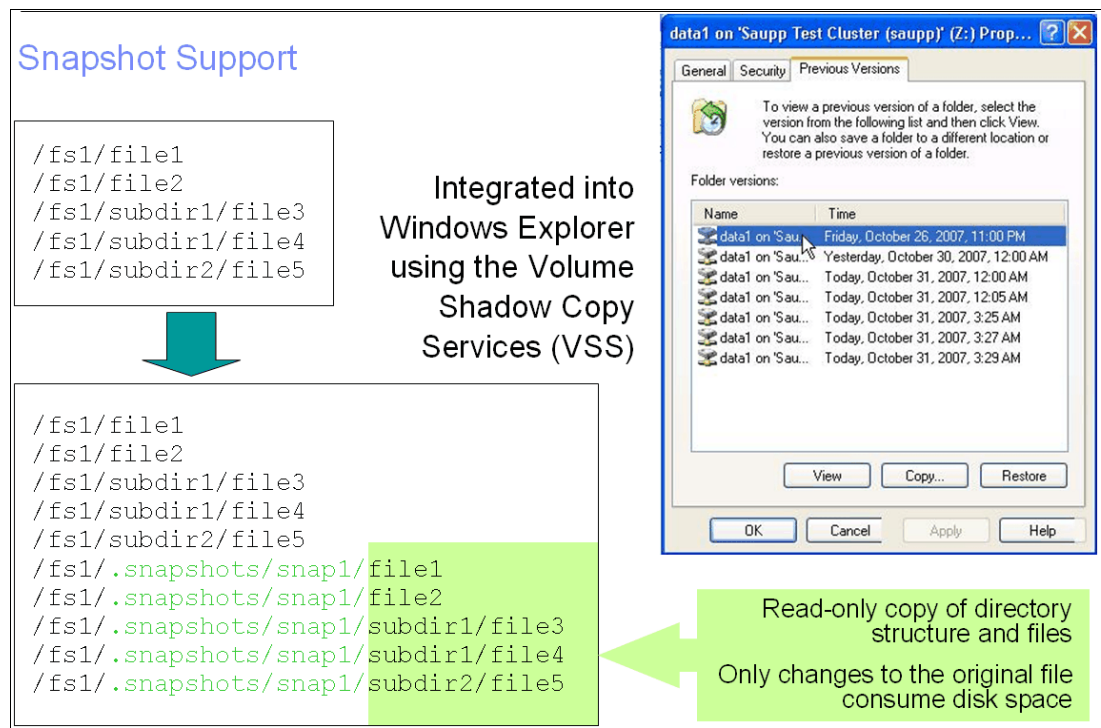


Figure 23-11 Snapshot overview

Snapshots of a file system are read-only; changes might only be made to the active (that is, normal, non-snapshot) files and directories. The snapshot function allows a backup or mirror program to run concurrently with user updates and still obtain a consistent copy of the file system as of the time that the snapshot was created. Snapshots also provide an online backup capability that allows easy recovery from common problems such as accidental deletion of a file, and comparison with older versions of a file. Because snapshots are not copies of the entire file system, they must not be used as protection against media failures.

23.6 IBM General Parallel File System (GPFS)

The IBM General Parallel File System (GPFS) is at the heart of the SBSC environment. See the Chapter 20, “IBM GPFS and SoFS” on page 543 in this Redbooks publication for a brief overview of GPFS and its functionality.

23.7 More information

See the following websites for more information:

http://www.ibm.com/ibm/cloud/smart_business/

<http://www-935.ibm.com/services/us/index.wss/offerfamily/its/a1031705>

Included within SBSC is a CIFS component which is based on the Samba open source project and a newly developed, high speed, cluster-aware version of the TDB metadata store database used by Samba. This Cluster Trivial Database (CTDB) allows Samba to scale well across multiple cluster nodes. A more detailed explanation of this is available in the paper: *IBM Scale out File Services: Reinventing network-attached storage*; see the website:

<http://www-935.ibm.com/services/us/its/pdf/Sofs-am-journal-final-070108.pdf>



A

Internet Small Computer System Interface (iSCSI)

The iSCSI protocol is the encapsulation of the industry standard SCSI protocol within TCP/IP packets. This is an alternative to Fibre Channel SAN technology which is essentially the same system of encapsulating SCSI protocol within an external 'carrier'. The significant difference being that the iSCSI san does not require a specialised or dedicated carrier network.

Introduction to iSCSI

A brief overview of the history of network attached storage, an overview of current iSCSI state and the potential benefits of implementing iSCSI follows in this chapter.

History

There was a time when storage equated with direct-attached storage (DAS) devices. There was little controversy in defining the straightforward DAS. Products in this category included devices like vanilla IDE or SCSI hard drives and RAID arrays. However problems with DAS were always the severe limitations on sharing with other machines and the physical limitations of the connection technology.

Networked storage options like NAS and storage area networks (SAN) solved the main issues with DAS devices by farming out data storage to dedicated machines. However NAS and SAN both have their own limitations.

NAS has the benefit that it can be deployed over existing network infrastructure and can go anywhere the network is deployed. However protocols such as SMB or NFS do not support the block storage of DAS. This is fine for generic file or print-sharing, but many applications such as databases require block storage.

SANs allow users to consolidate storage on block devices; however this approach requires additional and often significant investment in hardware and training. SANs are also limited to the size of the network they are deployed on. Due in part to stringent addressing limits of the FC switch standard and also based on the Fibre Channel distance limitations and the uniqueness of the cabling and infrastructure. Multiple distant sites often require multiple SANs. iSCSI is a potential solution to address both of these limitations.

Overview

The iSCSI protocol is a transport for SCSI over TCP/IP. Until recently, the standard IP protocol infrastructure (that is, Ethernet) cannot provide the necessary high bandwidth and low latency needed for reliable storage access. A special communications infrastructure, mainly Fibre Channel running FCP (SCSI encapsulated over Fibre Channel), was developed to allow Storage Area Networks (SANs) to be developed and deployed. With the recent advances in Ethernet technology, it is now practical (from a performance perspective) to access storage devices over an IP network. 1 Gigabit Ethernet is now common, and is competitive with 1 and 2 Gigabit Fibre Channel from a performance perspective.

As the newer 10 Gb Ethernet adoption rates increase, the IP based SAN performance will be equal to any of the currently available FCP networks. iSCSI is becoming an ever more popular method of implementing a storage network and is especially suited to cases where large numbers of blade servers require SAN connectivity for storage and even boot devices.

iSCSI transactions consist of an iSCSI initiator (either hardware or software) transmitting a request (i.e. read/write/etc) to an iSCSI target. This iSCSI target processes the request and responds with the appropriate information (data/sense/etc).

iSCSI initiators are typically application servers or even end users while iSCSI targets are typically either SAN access points or actual storage controllers. Because an iSCSI request is an encapsulation of a SCSI request, the SCSI concept of command descriptor blocks (CDBs)

is applicable to iSCSI. CDBs define the type of SCSI operation, the logical block address where to start, the length of data involved, and various other control parameters.

Figure A-1 gives a conceptual overview of the iSCSI Protocol layers.

As the figure illustrates, The iSCSI solution requires an initiator (host) a target (storage device generally) and a carrier network.

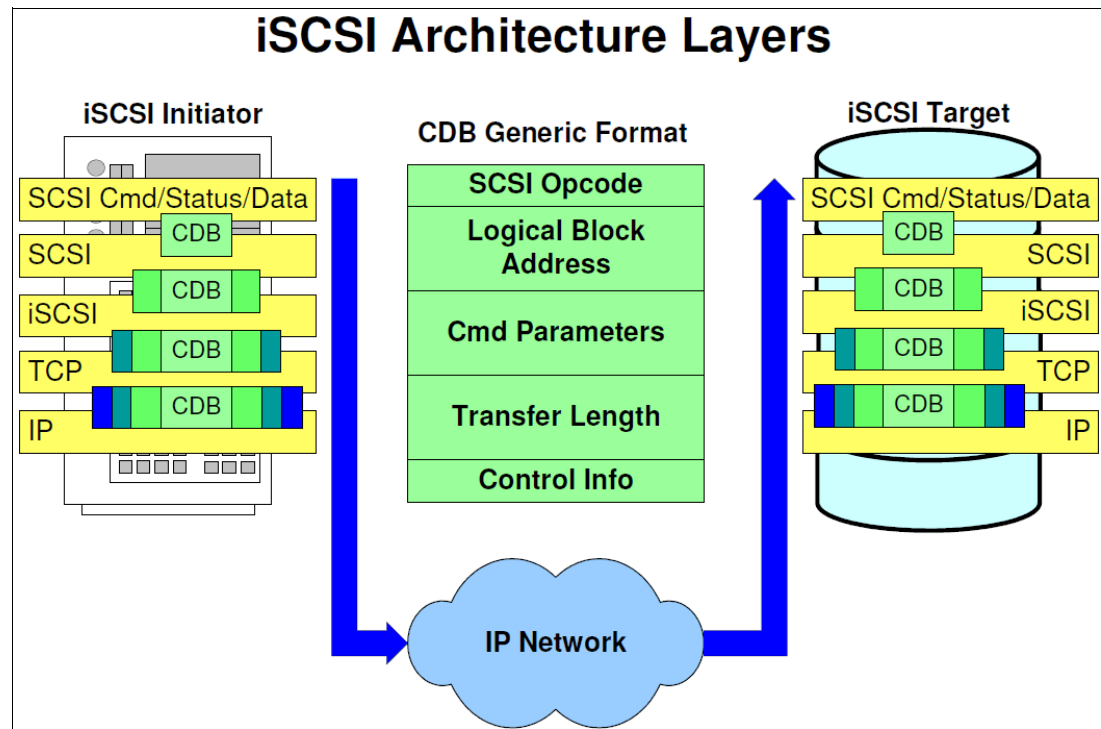


Figure A-1 iSCSI protocol layers

Benefits

The benefits of iSCSI over the more traditional Fibre Channel SAN is that an iSCSI san does not require a dedicated, specialised carrier network. iSCSI is a low-cost data transport solution that enables organizations of all sizes to create Storage Area Networks (SANs) using standard IP protocol instead of Fibre Channel protocol. Thanks to recent advances, an iSCSI network built with 10 Gbit Ethernet components, now has the necessary high bandwidth and low latency to create a network that can be competitive with any currently available Fibre Channel SAN in terms of latency and performance.

The cost benefit case for implementing iSCSI is strong. It is especially well suited to environments where large numbers of small servers need to be attached to storage. iSCSI servers do not require expensive host adaptors, or in the case of software initiators *any* additional adaptors. They do not take up expensive, and restricted, Fibre Channel SAN ports and do not require additional skills outside of the existing network support staff arena to maintain.

To Learn more about the IBM iSCSI solutions and offerings, see the following sites:

<http://www-947.ibm.com/systems/support/supportsite.wss/allproducts?brandind=5000029&taskind=1>

<http://www-03.ibm.com/systems/bladecenter/hardware/openfabric/virtualfabric.html>

<http://www-03.ibm.com/systems/i/advantages/integratedserver/iscsi/advancedtasks.html>



Redundant Arrays of Independent Disks (RAID)

RAID is a collection of techniques that treats multiple disks as combined units (arrays) and allows various levels of fault tolerance to be implemented in the disk subsystem.

RAID stands for *redundant array of independent disks*. It is a strategy created to bridge the gap between computer I/O requirements and the latency and throughput restrictions of single disk drives, while also allowing for greater degrees of fault tolerance.

RAID is the technology of grouping several physical drives in a computer into an array that can be defined as one or more logical drives. Each logical drive appears to the operating system as a single drive. This grouping technique greatly enhances logical drive capacity and performance.

SCSI has the ability to process multiple, simultaneous I/O requests. With RAID, I/O performance is improved because all drives can contribute to system I/O throughput.

In order to know RAID strengths and limitations, a clear understanding of the various RAID architectures is required.

RAID levels

The most commonly used current RAID levels are:

- ▶ 0, 1, 3, 5, 6
- ▶ Enhanced 1, Enhanced 5
- ▶ 00, 10, 50

Data striping

Data striping is the process of storing data across all the disks in an array:

- ▶ A data stripe is the collection of stripe units across the array. A stripe unit is the amount of data written on a disk before writing on the next disk.
- ▶ Choosing the correct stripe unit size will maximize the performance of the ServeRAID adapter, improving overall server performance.
- ▶ Common RAID levels that use striping are 0, 1 Enhanced, 5, and 5 Enhanced.

Parity is defined as redundant information about user data, which allows it to be regenerated in the event of a disk failure. In the following illustrations, data can mean a byte or block, not necessarily an entire file.

RAID 0: Data striping, no parity

This RAID level:

- ▶ Distributes (stripes) the data evenly across all disks in the array.
- ▶ Offers the fastest performance for multiple concurrent requests.
- ▶ Does not provide any level of fault tolerance.
- ▶ Offers most storage capacity of all the RAID levels.
- ▶ Total disk capacity of n (where n is the number of drives).

As shown in Figure B-1, in the event of a single disk failure, the data residing on the disk cannot be regenerated, and because of data striping, all of the striped data becomes unavailable.

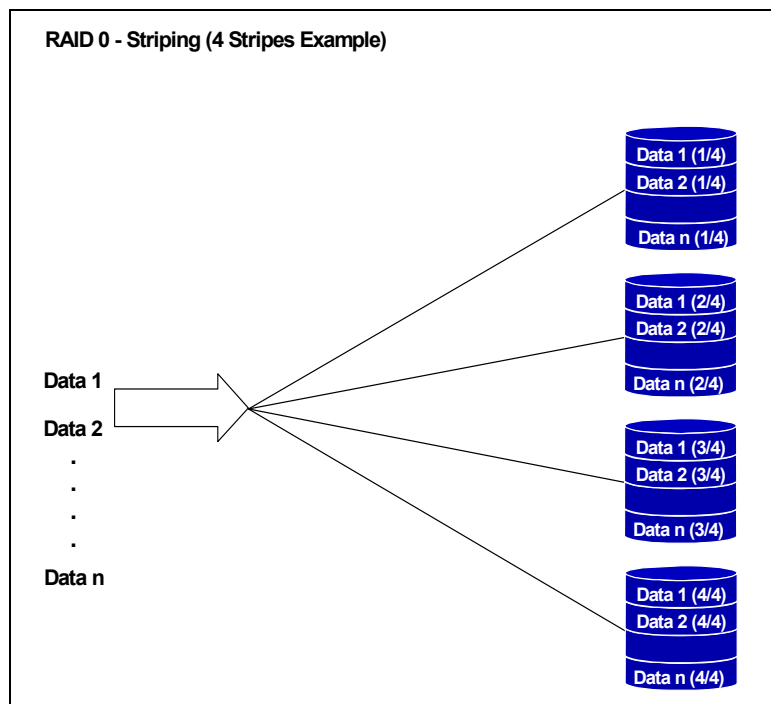


Figure B-1 RAID 0

Tip: Never use RAID level 0 for critical applications requiring high data availability. Consider it only for applications that benefit from the performance capabilities of this level.

RAID 1: Data mirroring

This RAID level:

- ▶ Data is written simultaneously to two drives:
- ▶ RAID-1 can be faster than a single drive.
- ▶ High cost of implementation.
- ▶ Allows for the loss of one disk in the array.
- ▶ Total disk capacity is $n/2$ (n is the number of disks).

The RAID 1 implementation (Figure B-2) employs data mirroring to achieve redundancy.

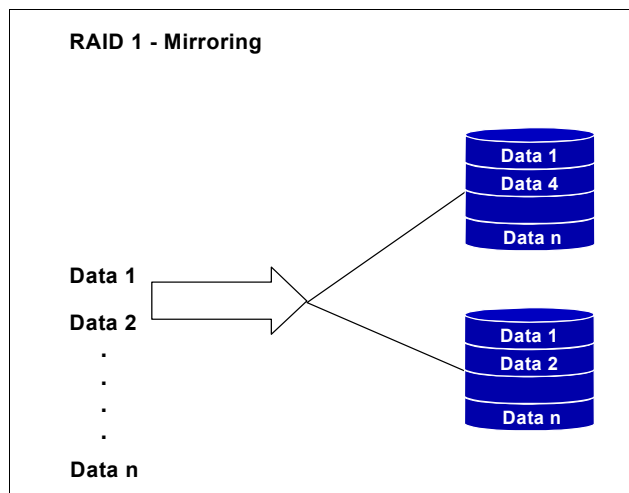


Figure B-2 RAID 1

Tip: Consider RAID 1 when cost is not a factor for applications requiring high data availability and high performance.

IBM Enhanced RAID-1

RAID-1 Enhanced (RAID-1E) is an IBM exclusive that combines data mirroring and data striping:

- ▶ Mirrors and stripes data across all disks (mirrored stripe)
- ▶ Allows disk mirroring with 3 or more disk drives
- ▶ Approximates RAID-0 performance for reads
- ▶ Allows for the loss of 1 disk in the array
- ▶ Total disk capacity of $n/2$ (n is the number of disks)

RAID 3: Sequential access to large files

RAID-3 is a parallel process array mechanism, where all drives in the array operate in unison. Similar to data striping, information to be written to disk is split into chunks (a fixed amount of data), and each chunk is written out to the same physical position on separate disks (in parallel). This architecture requires parity information to be written for each stripe of data.

Performance is very good for large amounts of data, but poor for small requests because every drive is always involved, and there can be no overlapped or independent operation. It is well suited for large data objects such as CAD/CAM or image files, or applications requiring sequential access to large data files. Select RAID-3 for applications that process large blocks of data. It provides redundancy without the high overhead incurred by mirroring in RAID-1

RAID-Double Parity based on RAID 4

Each traditional RAID4 group has a number of data disks and one parity disk, with aggregates and volumes containing one or more RAID4 groups. Whereas the parity disk in a RAID4 volume stores row parity across the disks in a RAID4 group Figure B-3.

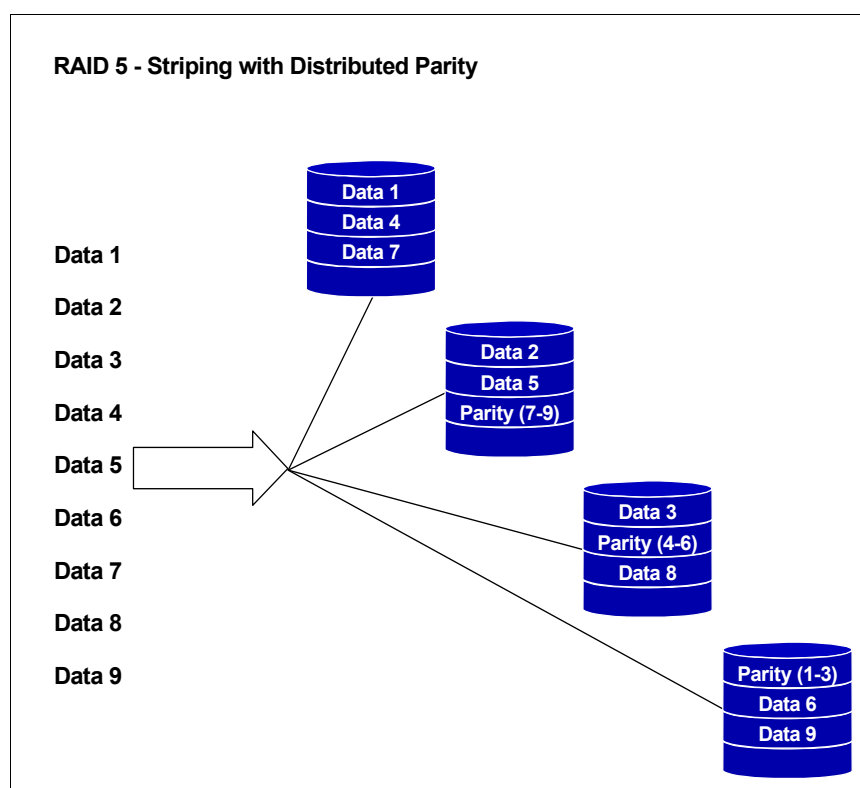


Figure B-3 RAID 4

In short, given the current landscape, with larger disk drives affecting data protection, customers and analysts demand a better story about affordably improving RAID reliability from storage vendors. To meet this demand, IBM System Storage N series uses a new type of RAID protection named RAID-DP™. RAID-DP stands for RAID Double Parity, and it significantly increases the fault tolerance from failed disk drives over traditional RAID. When all relevant numbers are plugged into the standard mean time to data loss (MTTDL) formula for RAID-DP versus single-parity RAID, RAID-DP is more reliable on the same underlying disk drives. With this reliability, RAID-DP approaches RAID1 mirroring for fault tolerance, but at RAID4 pricing. RAID-DP offers businesses the most compelling total cost of ownership storage option without putting their data at an increased risk.

At the most basic layer, RAID-DP adds a second parity disk to each RAID group in an aggregate or traditional volume. The additional RAID-DP parity disk stores diagonal parity across the disks in a RAID-DP group. With these two parity stripes in RAID-DP, one the traditional horizontal, and the other diagonal, data protection is obtained even in the event of two disk failure events occurring in the same RAID group. See Figure B-4 for a representation.

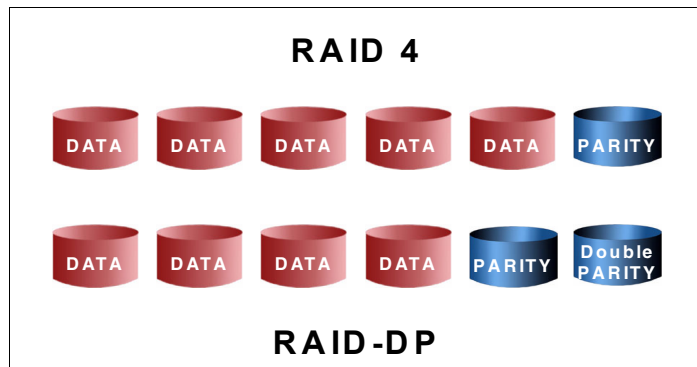


Figure B-4 RAID4 and RAID-DP comparison

So, with RAID-DP, the traditional RAID4 horizontal parity structure still:

- ▶ Offers optimal balance between price and performance
- ▶ Cost of RAID-DP is the capacity of 2 disk
- ▶ Allows for the loss of 2 disk in the array
- ▶ Protects against a single bad block failure while operating in a degraded mode
- ▶ Read performance approximates RAID-0
- ▶ Data and checksum (parity) are stored on 2 disks
- ▶ Requires a minimum of 4 disks

Total capacity of $n-2$ (n is number of disks)

RAID 5: Independent access, data striping with distributed parity

- ▶ Offers optimal balance between price and performance
- ▶ Cost of RAID-5 is the capacity of 1 disk
- ▶ Allows for the loss of 1 disk in the array
- ▶ Loss of a drive will cause system degradation
- ▶ Read performance approximates RAID-0
- ▶ Data and checksum (parity) are evenly spread across drives
- ▶ Requires a minimum of 3 disks
- ▶ Total capacity of $n-1$ (n is number of disks)

In RAID 5 the access arms can move independently of one another (Figure B-5). This enables multiple concurrent accesses to the array devices, thereby satisfying multiple concurrent I/O requests and providing higher transaction throughput. RAID 5 is best suited for random access data in small blocks.

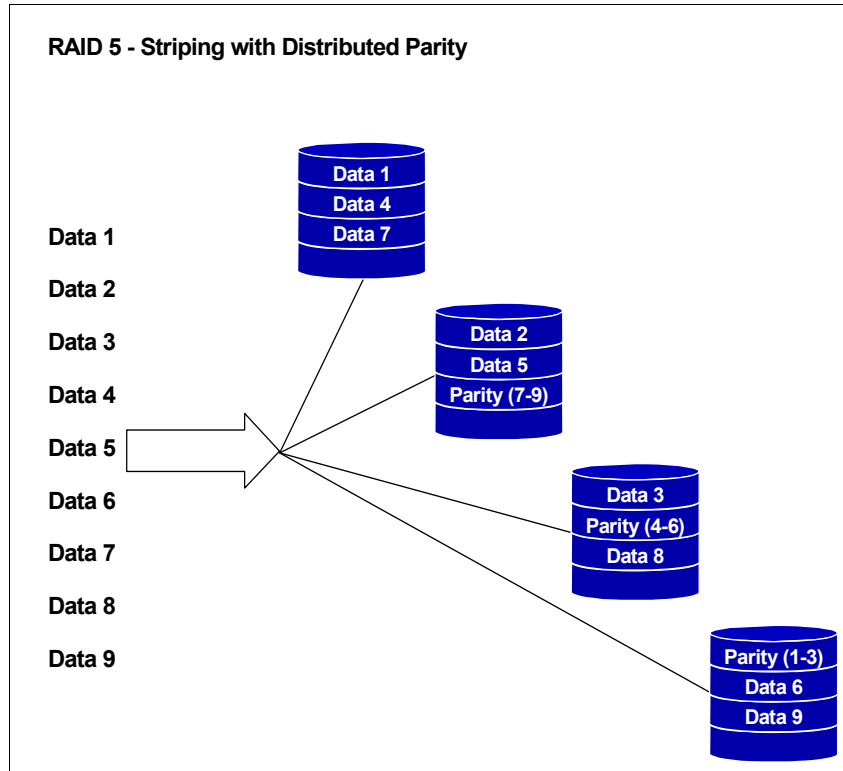


Figure B-5 RAID 5

Tip: Consider RAID 5 for environments requiring high data availability and with applications that process relatively short data records or a mixture of large sequential records and short random blocks.

RAID 5 Enhanced

- ▶ RAID-5E is standard RAID-5 with a hot spare built into the array (distributed hot spare)
 - RAID-5 uses a physical dedicated hot-spare disk
 - RAID-5E uses a logical spare
 - The once-dedicated hot spare is now distributed across the array
- ▶ RAID-5E works with a minimum of 4 disks
- ▶ Performance is good (better than RAID-5 with hot spare defined when compared with same number of total disks)
 - No idle hot-spare disk needed; provides an additional drive to stripe data across
 - up to 20% faster in small RAID-5 configurations.

RAID 6: Independent access, data striping with two distributed parities

- Offers optimal balance between price and performance, but needs more complex controller design compared to RAID5
- Cost of RAID-6 is the capacity of 2 disks
- Allows for the loss of 2 disks in the array
- Protects against a single bad block failure while operating in a degraded mode
- Read performance approximates RAID-0
- Data and checksum (parity) are evenly spread across drives
- Requires a minimum of 4 disks

Total capacity of $n-2$ (n is number of disks).

RAID 6 allows for additional fault tolerance by using a second independent distributed parity scheme (dual parity). Data is striped on a block level across a set of drives (P_n), similar to RAID 5 configurations, and a second set of parity is calculated and written across all the drives (Q_n), as illustrated in Figure B-6.

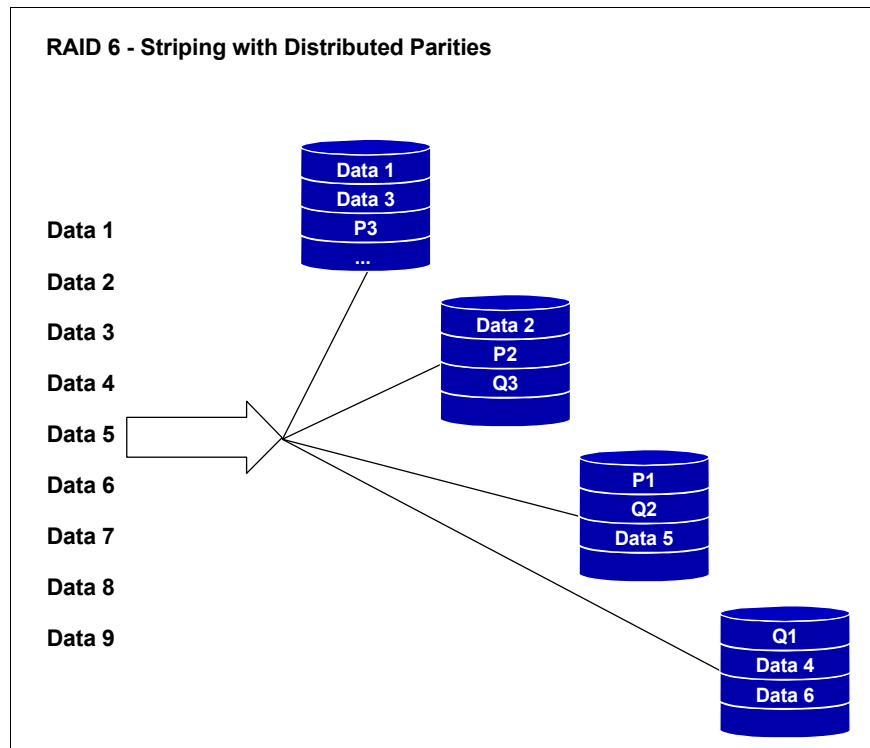


Figure B-6 RAID6

RAID 6 protection provides more fault tolerance than RAID 5 in the case of disk failures and uses less raw disk capacity than RAID 10.

Comparing RAID 6 to RAID 5 performance, gives about the same results on reads. For random writes, the throughput of a RAID 6 array is around only two thirds of a RAID 5, given the additional parity handling. Workload planning is especially important before implementing RAID 6 for write intensive applications including copy services targets and FlashCopy SE repositories. Yet, when properly sized for the I/O demand, RAID 6 is a considerable reliability enhancement.

RAID-00

RAID-00 is RAID-0 plus RAID-0

- ▶ RAID-00 works with a minimum of two disks
- ▶ Performance: excellent
- ▶ Protection: none
- ▶ Capacity: all
- ▶ Maximum configuration is 60 disks (ServeRAID-4H)

RAID-00 (Figure B-7) is a combination of multiple, local RAID-0 arrays.

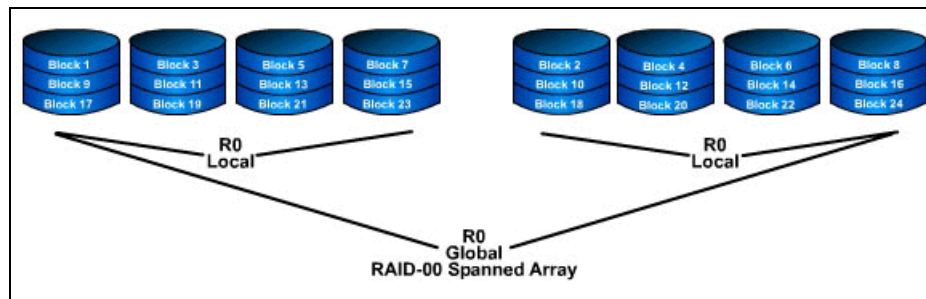


Figure B-7 RAID-00

RAID-10

RAID-10 is RAID-1 plus RAID-0:

- ▶ RAID-10 works with a minimum of four disks
- ▶ Performance: very good
- ▶ Protection: one drive in each local array might fail
- ▶ Capacity: $n/2$
- ▶ Maximum configuration is 16 disks (ServeRAID-4H)

The RAID 10 (Figure B-8) consists of a set of disks for user data plus their mirrored disks counterparts. There is no parity disk to rebuild a failed disk. In case one disk becomes unusable, then its mirror will be used to access the data and also to build the spare.

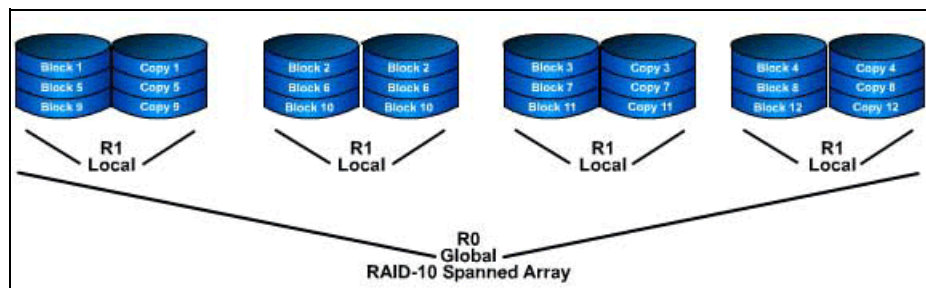


Figure B-8 RAID-10

RAID 10 is also known as RAID 0+1, because it is a combination of RAID 0 (striping) and RAID 1 (mirroring). The striping optimizes the performance by striping volumes across several disk drives (for example, in the ESS Model 800 implementation, three or four DDMs). RAID 1 is the protection against a disk failure by having a mirrored copy of each disk. By combining the two, RAID 10 provides data protection with I/O performance.

RAID-50

RAID-50 as shown in Figure B-9 is RAID-5 plus RAID-0:

- ▶ RAID-50 works with a minimum of 6 disks
- ▶ Performance: good (improves with size of local array)
- ▶ Protection: one drive in each local array might fail
- ▶ Capacity: $n \times x$ (where x = number of local arrays in RAID-0 stripe)
- ▶ Maximum configuration is 60 disks (SERVERAID-4H)

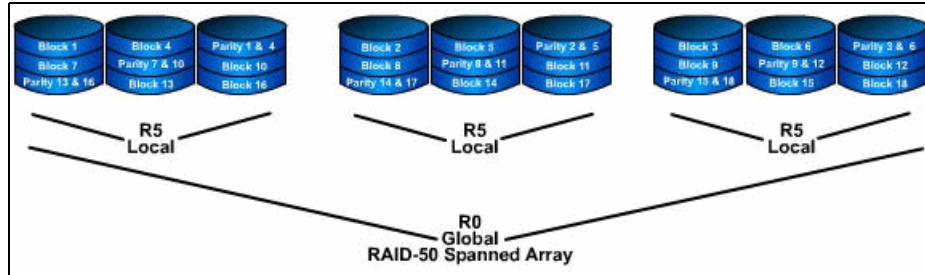


Figure B-9 RAID-50

For more information about RAID Levels see this website:

<http://www.acnc.com>



Introduction to SMI-S and CIM/WBEM

SMI-S

Efficiently managing multivendor Storage Area Networks (SANs) is a key concern for end-users and integrators alike. In mid-2002 the Storage Networking Industry Association (SNIA) launched the Storage Management Initiative (SMI) to create and foster the universal adoption of a highly functional open interface for the management of storage networks. The SMI's goal is to deliver open storage network management interface technology in the form of an *SMI Specification* (SMI-S).

The Storage Management Initiative Specification (SMI-S) is a design specification of the Storage Management Initiative (SMI) that is launched by the Storage Networking Industry Association (SNIA). The SMI-S specifies a secure and reliable interface that allows storage management systems to identify, classify, monitor, and control physical and logical resources in a storage area network (SAN). The interface integrates the various devices to be managed in a SAN and the tools used to manage them.

SMI-S is based on a number of existing technologies or industry standards that include the following:

- ▶ Common Information Model (CIM)
- ▶ Web-Based Enterprise Management (WBEM)
- ▶ Service Location Protocol (SLP)

Intended to be an industry standard, SMI-S extends the generic capabilities of the CIM, the WBEM, and the SLP to implement storage networking interoperability. For example, the WBEM provides provisions for security, resource-locking management, event notification, and service discovery.

CIM

An object model for data storage and management that is developed by the Distributed Management Task Force (DMTF). CIM makes it possible to organize devices and components of devices in an object-oriented pattern.

WBEM

Web Based Enterprise Management (WBEM) is a set of management and Internet standard architectures developed by the Distributed Management Task Force (DMTF) to unify the management of enterprise computing environments traditionally administered through traditional management stacks like SNMP or CMIP. WBEM provides the ability for the industry to deliver a well-integrated set of standard-based management tools utilizing emerging web technologies. The DMTF has developed a core set of standards that make up WBEM, which includes a data model, the Common Information Model (CIM) standard; an encoding specification, xmlCIM Encoding Specification; and a transport mechanism, CIM Operations over HTTP.

- ▶ **Common Interface Model (CIM) Standard:** CIM is a formal object oriented modeling language that is used to describe the management aspects of systems.
- ▶ **xmlCIM Encoding Specification:** This is a grammar to describe CIM declarations and messages used by the CIM protocol
- ▶ **CIM Operations over Hypertext Transfer Protocol (HTTP):** HTTP is used as a way to enable communication between a management application and a device that both use CIM.

The WBEM architecture defines the following elements:

- ▶ **CIM Client:** The CIM Client is a management application like TotalStorage Productivity Center that uses CIM to manage devices. A CIM Client can reside anywhere in the network, because it uses HTTP to talk to CIM Object Managers and Agents.
- ▶ **CIM Managed Object:** A Managed Object is a hardware or software component that can be managed by a management application by using CIM. e.g. IBM SAN Volume Controller.
- ▶ **CIM Agent:** The CIM Agent is embedded into a device and interfaces with the management application. It interprets CIM requests and responses to it.
- ▶ **CIM Object Manager (CIMOM):** A common conceptual framework for data management that receives, validates and authenticates the CIM requests from client application. It then directs the requests to appropriate component or CIM provider.

Note: The terms CIM Agent and CIM Object Manager (CIMOM) are often used interchangeably. At this time only few devices come with an integrated CIM Agent, most devices need a external CIMOM for CIM enable management applications (CIM Clients) to be able to talk to the device.

- ▶ **Device Provider:** A CIM Provider is the element that translates CIM calls to the device specific commands, it is like a device driver. A provider is always closely linked to a CIM Object Manager.
- ▶ **CIM Server:** A CIM Server is the software that runs the CIMOM and the CIM provider for a set of devices. This approach is used when the devices do not have an embedded CIM Agent. This term is often not used, instead people often use the term CIMOM but they mean the CIM Server.

SLP

Service Location Protocol (SLP) is a new Internet Engineering Task Force (IETF) proposed standard protocol that was designed to simplify the discovery and use of network resources. In a corporate intranet, users need to access services and resources on the network. Often, it is not clear to the users what useful services are available to them. These resources can include TN3270E servers, web servers, printers, fax machines, file systems, databases, and any other future services that might become available. IBM Communications Server for AIX, Version 6 supports SLP in conjunction with TN3270E servers and clients such as Host On-Demand SLP clients.

Typical component Interaction

Figure C-1 on page 612 shows a typical interactions between the components in the CIM/WBEM environment. The following description explains the steps involved in the interaction.

(1) The client application locates the CIMOM by calling a Service Location Protocol (SLP) directory service.

Note: The SMI-S specification introduces (Service Location Protocol) SLP as the method for the management applications (the CIM clients) to locate managed objects.

For more information about SLP, refer the IBM technote TIPS0523 at following link:

<http://www.redbooks.ibm.com/abstracts/tips0523.html?Open>

(2) When the CIMOM is first invoked, (3) it registers itself to the SLP and supplies its location, IP address, port number, and the type of service it provides. (4) With this information, the client application starts to directly communicate with the CIMOM.

The client application then (5) sends CIM requests to the CIMOM. As requests arrive, the CIMOM validates and authenticates each request. (6) It then directs the requests to the appropriate functional component of the CIMOM or to a device provider. (7) The provider makes calls to a device-unique programming interface on behalf of the CIMOM to satisfy (8)-(9)-(10) client application requests.

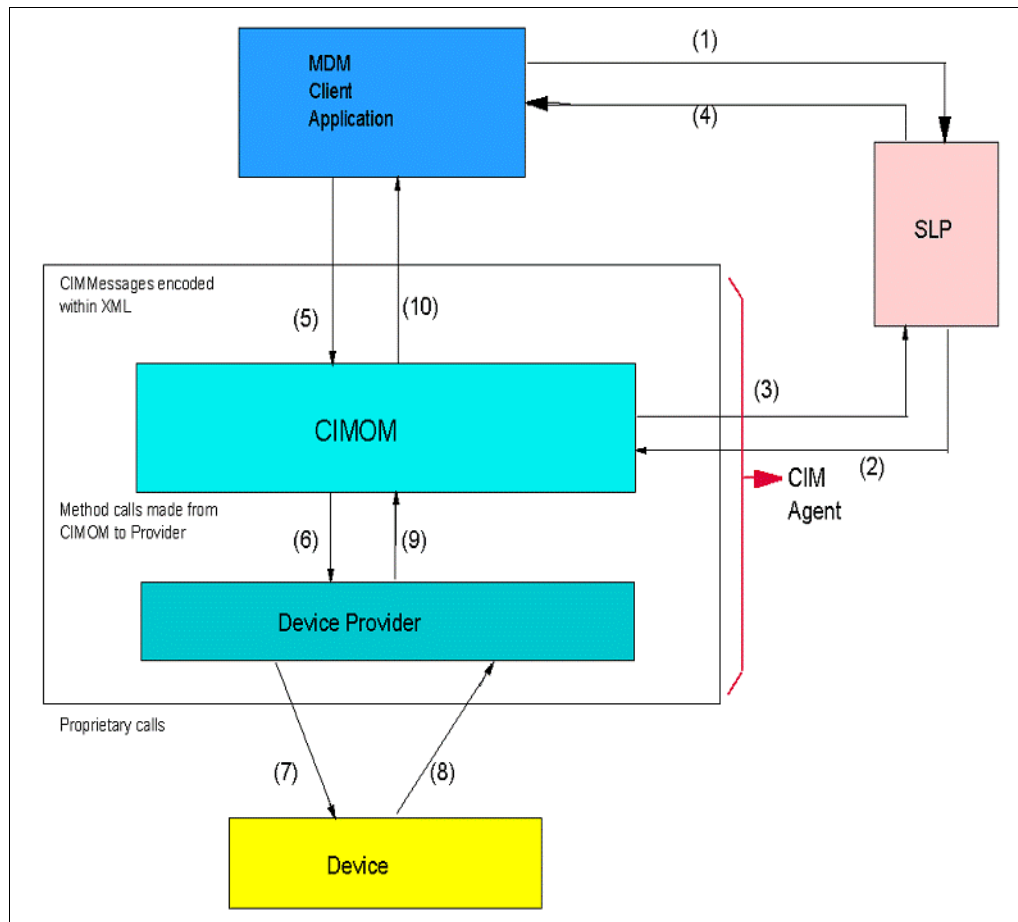


Figure C-1 Component interaction in CIM environment

More information

For more information, see following links:

<http://www.snia.org/home>

http://www.snia.org/forums/smi/tech_programs/smis_home/

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks publications

For information about ordering these publications, see “How to get Redbooks publications” on page 614. Note that certain documents referenced here might be available in softcopy only:

- ▶ *Deployment Guide Series: IBM TotalStorage Productivity Center for Data*, SG24-7140
- ▶ *Designing an IBM Storage Area Network*, SG24-5758
- ▶ *DS8000: Introducing Solid State Drives*, REDP-4522
- ▶ *DS8000 Thin Provisioning*, REDP-4554
- ▶ *IBM b-type Data Center Networking: Design and Best Practices Introduction*, SG24-7786
- ▶ *IBM b-type Data Center Networking: Product Introduction and Initial Setup*, SG24-7785
- ▶ *IBM/Cisco Multiprotocol Routing: An Introduction and Implementation*, SG24-7543
- ▶ *IBM DS8000 and z/OS Basic HyperSwap*, REDP-4441
- ▶ *IBM j-type Data Center Networking Introduction*, SG24-7820
- ▶ *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- ▶ *IBM Midrange System Storage Copy Services Guide*, SG24-7822
- ▶ *IBM Midrange System Storage Hardware Guide*, SG24-7676
- ▶ *IBM SAN Survival Guide*, SG24-6143
- ▶ *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065
- ▶ *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010
- ▶ *IBM System Storage DS8000: LDAP Authentication*, REDP-4505
- ▶ *IBM System Storage DS8000: Remote Pair FlashCopy (Preserve Mirror)*, REDP-4504
- ▶ *IBM System Storage DS8000 Series: IBM FlashCopy SE*, REDP-4368
- ▶ *IBM System Storage DS8700: Architecture and Implementation*, SG24-8786
- ▶ *IBM System Storage DS8700 Disk Encryption Implementation and Usage Guidelines*, REDP-4500
- ▶ *IBM System Storage N series*, SG24-7129
- ▶ *IBM System Storage Tape Library Guide for Open Systems*, SG24-5946
- ▶ *IBM System Storage TS7650 and TS7650G with ProtecTIER*, SG24-7652
- ▶ *IBM Tivoli Storage Manager Implementation Guide*, SG24-5416
- ▶ *IBM Tivoli Storage Productivity Center V4.1 Release Guide*, SG24-7725
- ▶ *IBM TotalStorage Virtual Tape Server: Planning, Implementing, and Monitoring*, SG24-2229
- ▶ *IBM TotalStorage 3494 Tape Library: A Practical Guide to Tape Drives and Tape Automation*, SG24-4632

- ▶ *IBM Virtualization Engine TS7500: Planning, Implementation, and Usage Guide*, SG24-7520
- ▶ *IBM Virtualization Engine TS7510: Tape Virtualization for Open Systems Servers*, SG24-7189
- ▶ *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659
- ▶ *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116
- ▶ *Implementing an IBM/Cisco SAN*, SG24-7545
- ▶ *Implementing the IBM System Storage SAN Volume Controller V5.1*, SG24-6423
- ▶ *Introducing the IBM Grid Access Manager*, SG24-7612
- ▶ *Introduction to Storage Area Networks*, SG24-5470
- ▶ *IP Storage Networking: IBM NAS and iSCSI Solutions*, SG24-6240
- ▶ *The IBM TotalStorage NAS Integration Guide*, SG24-6505
- ▶ *The IBM Virtualization Engine TS7510: Getting Started with i5/OS and Backup Recovery and Media Services*, SG24-7510

Other publications

These publications are also relevant as further information sources:

- ▶ *IBM System Storage Productivity Center V1.4: Introduction and Planning Guide*, SC23-8824
- ▶ *IBM Tivoli Storage Productivity Center and IBM Tivoli Storage Productivity Center for Replication Installation and Configuration Guide*, SC27-2337

Online resources

The following website is also relevant as a further information source:

<http://www.ibm.com/systems/storage/>

How to get Redbooks publications

You can search for, view, or download Redbooks publications. Redpapers publications, Technotes, draft publications, and Additional materials, as well as order hardcopy Redbooks publications, at this website:

ibm.com/redbooks

Help from IBM

IBM Support and downloads:

ibm.com/support

IBM Global Services:

ibm.com/services

Index

Numerics

1269-3S1 57, 59
1269-CP2 57
1726-21x 32
1726-22x 32
1726-31x 34
1726-32x 34
1726-41x 35
1726-42x 35
1750-511 99
1750-522 94
1750-EX1 100
1750-EX2 94
1814-20A 49
1814-52A 63
1818-51A 52–53
1818-53A 52–53
1818-D1A 63
2145-8A4 388
2145-CF8 391
2805-MC4 527, 532
2810-A14 137
2812-A14 137
2858-A10 183
2858-A11 183
2858-A12 183
2858-A20 183
2858-A21 183
2858-A22 183
2859-A10 179
2859-A20 179
2862-A10 179
2862-A20 179
2866-A11 185
2866-A21 185
2866-G11 191
2866-G21 191
2867-A11 185
2867-A21 185
2867-G11 191
2867-G21 191
3580 Model H4S 237
3580 Tape Drive 236, 239
3953 9
5060 45
7214 Tape and DVD Enclosure 11

A

ABARS 566, 574
Access Gateway 329
Adaptive differencing technology 440
Adaptive Multi-stream Prefetching (AMP) 118
Adaptive Multistream Prefetching (AMP) 102
Adaptive Replacement Cache (ARC) 92

Advanced Brocade Discovery 477
Advanced Encryption Standard (AES) 68
Advanced features 142
Advanced Library Management System (ALMS) 269
Advanced Performance Monitoring 327
Advanced Security 327
Advanced Security Activation 335
Advanced Single Instance Storage (ASIS) 197
Advanced Web Tools 326
Advanced Zoning 327
Agents 484
Aggregate Backup And Recovery Support 566
Aggregated views 477
American National Standards Institute (ANSI) 306
API 94, 439, 445
application programming interface (API) 223
Application-Managed Encryption (AME) 238, 256, 265
application-specific integrated circuit (ASIC) 365–366
Arbitrated Loop 53
ASIC 53
Asynchronous PPRC 133
Asynchronous PPRC-XD 132
ATM 468
Automated Cartridge System Library Software (ACSL) 444
Automated data migration 540
Automated IO path failover 31
Automatic Transfer Switch (ATS) 138
AutoSupport 195
Availability management 565

B

back-end 410
backup vm 419
Bare Machine Recovery 444
base volume 55
battery backup 32
BCDS 574
BCS 573
beginning of the tape (BOT) 232
Block aggregation 377
block aggregation 375
block level virtualization 58
Boss node 410
Business continuity 406
Business continuity and resiliency 4
Business Intelligence and Reporting Tools (BIRT) 438, 476

C

Cache 55
cache 31, 33, 36
cache block 74
cache memory 45, 50

- cache mirroring 31
- Capacity on Demand (CoD) 267
- CCE 356
- CDB (Command Descriptor Block) 316
- Change and Configuration Management database (CCMDB) 478
- CHAP authentication 400
- CICS 567
- CIFS 304, 419, 481, 540, 548, 586
- CIM 488, 609
- CIM Agent 94, 389, 488, 530
- CIM Client 488
- CIM Object Manager (CIMOM) 488, 610
- CIM Server 610
- CIMOM 389, 509
- CIMS 497
- Cisco MDS 9124 13, 328, 330
- Cisco MDS 9134 13, 333
- Cisco MDS 9222i 14, 349
- Cisco MDS 9506 340
- Cisco MDS 9509 12, 342
- Cisco MDS 9513 12, 344
- Cisco Nexus 5000 16
- Cisco Nexus 5010 16, 367
- Cisco Nexus 5020 16, 367
- CLI 128, 202
- Cloud Computing 2
- Cluster fail over (CFO) 197
- cluster trivial database (CTDB) 586
- clustered file system 581
- Clustering 441
- collocation 469
- command queue depth 54
- Command Tag Queuing 90
- command-line interface (CLI) 484
- COMmodity PArts Storage System 389
- Common Information Model (CIM) 488, 609
- common information model (CIM) 381
- Common Internet File System (CIFS) 173, 196, 202
- Compass architecture 389
- Concurrent Copy (CC) 564
- Configuration node 410
- Consistency Group 130, 154
- Constant switching capacity 147
- Continuous Data Protection 5
- controller card 95
- Controller service mode 76
- Converged Enhanced Ethernet (CEE) 14
- Converged Network Adapters (CNA) 321
- Copy Services 77, 129, 387
- cross site replication 555
- CUP Activation 335
- CWDM 468

D

- DAT72 292
- Data 148, 232
- Data agents 487
- Data Center Fabric architecture 338
- Data Center Fabric Manager 356

- data center infrastructure efficiency (DCiE) 6
- Data Deduplication 4, 282, 415
- Data Facility Sort (DFSORT) 561
- Data Facility Storage Management Subsystem (DFSMS) 288, 561
- Data Manager 486, 491
- Data Manager for Chargeback 497
- Data Manager for Databases 496
- Data Manager policy 492
- Data migration 407
- Data Mirroring 601
- Data Modules 140
- Data ONTAP 174, 176, 443
- Data server 483
- Data Set FlashCopy 130
- Data Striping 600
- Database 484
- DCFM 356
- DCOLLECT 567
- DCS9900 43, 45–46, 56
- DCS9900 3S1 SAS/SATA Storage Expansion Unit 57
- DCS9900 Controller Model CP2 57
- DDMs 99–100
- DDS Gen 5 292
- DDS-3 293
- DDS-4 293
- Deduplication 9, 279
- Dense Wavelength Division Multiplexing (DWDM) 468
- device adapter (DA) 112
- Device class 431
- Device provider 489
- Device server 483
- DFDSS 572
- DFHSM 572
- DFSMS 21, 562, 571
- DFSMS Optimizer 21, 571, 576
- DFSMS Transactional VSAM Services 567
- DFSMSdfp 562
- DFSMSdss 562, 564–565
- DFSMSshm 562, 565, 574
- DFSMSrmm 288, 562, 566–567, 576
- DFSMSstvs 562, 567
- DFSORT 561, 568
- DFSORT/VSE 569
- direct-attached storage (DAS) 596
- Directory agent (DA) 490
- Disaster Recovery Manager (DRM) 441
- disaster recovery plan (DRP) 441
- Discovery Library Adapter 478
- disk diagnostics 168
- disk drive module (DDM) 114
- Disk Magic 509
- Disk Manager 486, 497
- Disk sanitization 195
- Disk scrubbing 167
- Disk storage expansion unit 203
- disk storage expansion unit 208
- disk to disk (D2D) 277
- disk to disk to tape (D2D2T) 277
- Distributed Management Task Force (DMTF) 488, 610

- DLT 236
- Domain Name System (DNS) 195
- Drive interface 55
- Drive migration 74
- drive queue depth 54
- DRM 441
- DRP 442
- DS Storage Manager 70, 126
- DS Storage Manager 38
- DS3000 29–30
- DS3200 8, 29–30
- DS3300 8, 29–30, 34
- DS3400 8, 29, 31, 35
- DS4000 43
- DS4100 45
- DS4300 45
- DS4500 45
- DS4700 43, 45–46
- DS4700 Express 8
- DS4700 Express Model 46
- DS4800 43, 45
- DS5000 7, 43, 46, 482
- DS5020 7, 43, 45–46, 49
- DS5100 43, 45–46, 52
- DS5300 45–46, 52
- DS6000 7, 86
- DS6000 expansion enclosure 97
- DS6800 90
- DS8000 7, 101
- DS8100 87
- DS8100 Turbo Model 931 121
- DS8300 87
- DS8300 Turbo Model 9B2 101
- DS8300 Turbo Models 122
- DS8700 87, 104
- DSCIMCLI 528, 530
- DSMSERV INSERTDB 418
- Dual active RAID controllers 31
- Dual controller 33
- dual controller 32, 34–36
- dual RAID controller 43, 47
- DVD storage 11
- DWDM 336, 468
- Dynamic Capacity Expansion 54, 77
- dynamic infrastructure 2–3
- Dynamic multithreaded transfer 440
- Dynamic RAID Level Migration 54
- Dynamic Segment Size 54
- Dynamic Volume Expansion 54, 77
- Dynamic Workspace Linking (DWL) 572

E

- edge-switch 12
- EEFMT3 246
- EFCM 327
- EFMT3 246
- electronic tape vaulting 466, 468
- Embedded CIM agent 480
- Embedded processing power 147
- EMC PowerPath 480

- Encryption and Security Management 4
- end of the tape (EOT) 232
- Endpoint device alerts 505
- Energy efficiency 4–5
- Engenio provider 530
- Enhanced cache 75
- Enhanced monitoring 168
- Enhanced performance 74
- Enhanced Remote Mirror 77
- Enterprise administration 441
- Enterprise Fabric Connectivity Manager 327
- enterprise resource planning (ERP) 413
- Enterprise SAN directors 11
- Enterprise Storage 4
- Enterprise Tape Library (ETL) 563
- Environmental Services Modules (ESM) 37
- ERM 77, 80
- ESCON 468
- Ethernet Direct Connection 53
- Ethernet Switched Network 53
- Ethernet switches 140
- EX5000 45
- EXN1000 192
- EXN4000 192
- EXP100 45
- EXP3000 8, 29, 32–33, 35, 37
- EXP5000 52, 63
- EXP5060 64
- EXP520 45–46, 50, 63
- EXP710 45, 47
- EXP810 46–47, 50, 61
- EXPIRE INVENTORY 416
- Extended Address Volume 103
- Extended Fabric 327
- Extended Fabric Activation 335
- Extended Remote Copy (XRC) 129, 563
- Extended Remote Copy XRC 134
- Extent 410

F

- F_Port 314
- Fabric agents 487
- Fabric alerts 504
- Fabric level virtualization 375
- Fabric Manager 327, 486, 501
- Fabric OS 329, 331, 335
- Fabric Shortest Path First (FSPF) 336, 340
- Fabric Watch 327
- FC 11, 46
- FC DDM 50
- FC encryption-capable DDM 51
- FC HIC 53
- FC-AL 96, 106
- FC-FC Routing 346, 348
- FCIP 317, 337, 347
- FCIP Tunneling 348
- FCoCEE 11, 337–338
- FCoE 321, 356, 368
- FCoE adapter 314
- FDE 67

- Fibre Channel (FC) 43
- Fibre Channel over IP (FCIP) 304, 317
- Fibre Channel Protocol (FCP) 304
- Fibre Channel Switched Fabric 53
- Fibre extenders 468
- Fibre tape failover 464
- FICON 12, 304, 315, 332, 338, 468
- FICON Accelerator 335
- file aggregation 376–377
- file constraints 495
- file system extension 495
- File Transfer Protocol (FTP) 202
- FilerView 195
- Filesets 547
- FL_Port 315
- FlashCopy 29, 31, 48, 55, 74, 77, 387, 397, 404, 520, 565
- FlashCopy Expansion License 40
- FlexCache for NFS 196
- FlexClone 196
- Flexible snapshots 149
- FlexPort scalability 327
- FlexScale 196
- FlexShare 195
- FlexVol 195
- Forwarding Information Base (FIB) 361
- Front-end 410
- Full Disk Encryption (FDE) 67, 103
- Full Disk Encryption drive 106
- Full Disk Encryption for FC Disks 45
- full function removable media management system 567
- Full Height Tape Drive 10
- FX_Port 315

G

- G_Port 315
- GAM 538
- GbE 16, 344
- GbE SFPs 336
- GBICs (Gigabit Interface Converters) 312
- Gbps 11
- GHS 74
- Gigabit Ethernet (GbE) 14
- Gigabit Ethernet IP (GbE) 342
- Global Copy 55, 129
- Global copy 82
- Global Hot Spare drives 53
- Global Hot Spares 74
- Global Hot-Spares 31
- Global Mirror 55, 129, 387, 395, 405, 521
- Global mirror 81
- global namespace 581
- global spare capacity 153
- Goal distribution 163
- GPFS 21, 220, 380, 420, 479, 543–544, 581
- GPFS cluster 544
- graceful Routing Engine switchover (GRES) 365
- graceful shutdown 162
- Grain 410
- graphical user interface (GUI) 484

- Grid Access Manager 537
- Grid Access Manager Software 21
- grid architecture 146
- Groups 503
- GUI 38

H

- HACMP 459, 461
- half-height Linear Tape-Open Tape Drive 10
- Half-High (HH) 260
- Half-High SCSI Tape Enclosure 298
- Hardware Management Console (HMC) 106
- Hardware parallelism 146
- headless system 174
- HH Tape Drive 257
- Hierarchical Storage Management 441
- Hierarchical Storage Management (HSM) 418, 429, 448
- Hierarchical Storage Manager 582
- High availability clustering 459
- high density (HD) 255
- high Linear Tape-Open Tape Drive 11
- High Performance Computing (HPC) 56, 64
- High-Availability Cluster Multi-Processing 481
- High-Performance Computer (HPC) 14, 357
- HMC 103
- Host adapters 96, 106
- Host Bus Adapter (HBA) 312
- Host interface 55
- Host Interface Cards (HIC) 50, 53
- host logins 54
- host server 38
- host storage partitions 54
- Host volume 521
- hot-swappable 31, 47
- HPC 358
- HSM 220, 418, 420, 461–462
- HTTP 488, 540
- HTTPS 488
- HyperFactor 280
- Hypertext Transfer Protocol (HTTP) 196, 202, 488
- Hypertext Transfer Protocol over Secure Socket Layer (HTTPS) 488

I

- I/O enclosures 112
- I/O group 390, 392, 410
- IBM 20, 82, 360
- IBM 1U rack-mount external tape enclosure 299
- IBM 7206 Model 336 External DAT72 292
- IBM b-type Ethernet switches and routers 14
- IBM Converged Switch B32 14, 356
- IBM c-series Ethernet switches 15, 359
- IBM DB2 Cloning Tool 21, 571
- IBM DB2 Cloning Tool for z/OS 572
- IBM DDS Generation 5 USB tape drive 293
- IBM DDS Generation 6 Internal USB Tape Drive 294
- IBM Enhanced RAID-1 601
- IBM Enterprise Disk Systems 85
- IBM Ethernet Router J02M/J06M/J011M 16

- IBM Ethernet Routers 366
- IBM Ethernet Switch B24X 14, 357
- IBM Ethernet Switch J08E 363
- IBM Ethernet Switch J16E 363
- IBM Ethernet Switch J48E 15, 365
- IBM General Parallel File System 543
- IBM General Parallel File System (GPFS) 21, 380, 479
- IBM g-series Ethernet switches 15
- IBM High Availability Cluster Multi-Processing (HACMP) 441, 459
- IBM i Host Kit 53
- IBM Information Archive 215, 217
- IBM Information infrastructure 4
- IBM j-type Ethernet switches 15
- IBM LTO Ultrium family 236
- IBM m-series Ethernet/IP routers 15
- IBM Multipath Subsystem Device Driver (SDD) 90
- IBM Network attached storage (NAS) 16
- IBM r-series Ethernet switches 14, 358
- IBM SAN Volume Controller (SVC) 19, 378
- IBM Scale Out Network Attached Storage (SONAS) 201
- IBM Service Management 6
- IBM Smart Business Storage Cloud 580
- IBM Smarter Planet 2
- IBM s-series Ethernet Switch 15, 362
- IBM storage architecture 89
- IBM Subsystem Device Driver (SDD) 394
- IBM Support Assistant (ISA) 478
- IBM System Storage 3952 Tape Frame Model F05 245
- IBM System Storage 7212 Storage Device Enclosure 296
- IBM System Storage 7214 Storage Device Enclosure 297
- IBM System Storage DS Storage Manager 29, 70
- IBM System Storage DS8000 86
- IBM System Storage DS8000 Storage Manager 530
- IBM System Storage DS8000 Turbo 378
- IBM System Storage Global Copy 132
- IBM System Storage Global Mirror 133
- IBM System Storage Metro Mirror 131
- IBM System Storage N series 173
- IBM System Storage N series Gateway 16, 187
- IBM System Storage Productivity Center (SSPC) 20, 22, 527
- IBM System Storage ProtecTIER 275
- IBM System Storage SAN Volume Controller (SVC) 383, 530
- IBM System Storage SAN24B-4 328
- IBM System Storage software 369
- IBM System Storage solutions 2
- IBM System Storage TS1120 Tape Drive 245
- IBM System Storage TS3100 Tape Library 254, 257
- IBM System Storage TS3200 Tape Library 254, 259
- IBM System Storage TS3310 Tape Library 254
- IBM System Storage TS3400 Tape Library 254
- IBM System Storage TS3500 Tape Library 254, 267
- IBM System Storage TS7650 ProtecTIER Deduplication Appliance 279, 379
- IBM System Storage TS7650G ProtecTIER Deduplication Gateway 9, 379

- IBM System Storage z/OS Global Mirror 134
- IBM Tivoli Advanced Allocation Management 575
- IBM Tivoli Advanced Audit for DFSMSHsm 21, 571, 574
- IBM Tivoli Advanced Backup and Recovery 21, 571
- IBM Tivoli Advanced Catalog Management for z/OS 21, 571, 573
- IBM Tivoli Automated Tape Allocation Manager 21, 571
- IBM Tivoli Automated Tape Allocation Manager for z/OS 575
- IBM Tivoli Continuous Data Protection for Files 450
- IBM Tivoli Integrated Portal 479
- IBM Tivoli OMEGAMON XE 572
- IBM Tivoli Omegamon XE 21, 571
- IBM Tivoli Provisioning Manager 510
- IBM Tivoli Storage FlashCopy Manager 420
- IBM Tivoli Storage Manager 413, 440
- IBM Tivoli Storage Manager 6 for Microsoft SharePoint 445
- IBM Tivoli Storage Manager architecture 423
- IBM Tivoli Storage Manager Backup/Archive client 426
- IBM Tivoli Storage Manager Extended Edition 441
- IBM Tivoli Storage Manager FastBack 449
- IBM Tivoli Storage Manager FastBack for Bare Machine Recovery 450
- IBM Tivoli Storage Manager FastBack for Microsoft Exchange 449
- IBM Tivoli Storage Manager for Advanced Copy Services 421, 446
- IBM Tivoli Storage Manager for Copy Services 421, 447
- IBM Tivoli Storage Manager for Databases 445
- IBM Tivoli Storage Manager for Hardware 446
- IBM Tivoli Storage Manager for HSM for Windows 446
- IBM Tivoli Storage Manager for Mail 445
- IBM Tivoli Storage Manager for SAP V6.1 447
- IBM Tivoli Storage Manager for Space Management 448
- IBM Tivoli Storage Manager for Storage Area Networks 426, 448
- IBM Tivoli Storage Manager for System Backup and Recovery 449
- IBM Tivoli Storage Manager reporting and monitoring 435
- IBM Tivoli Storage Manager server 425
- IBM Tivoli Storage Optimizer for z/OS 21, 571, 576
- IBM Tivoli Storage Productivity Center 379, 482, 512
- IBM Tivoli Storage Productivity Center (TPC) 20, 473
- IBM Tivoli Storage Productivity Center Basic Edition 20, 513, 530
- IBM Tivoli Storage Productivity Center for Data 20, 515
- IBM Tivoli Storage Productivity Center for Disk 20, 515
- IBM Tivoli Storage Productivity Center For Replication 516
- IBM Tivoli Storage Productivity Center for Replication 20, 479
- IBM Tivoli Storage Productivity Center Standard Edition 20, 524
- IBM Tivoli Storage Productivity Center version 4.1.1 474
- IBM Tivoli Storage Productivity Center with Advanced Provisioning 20, 525
- IBM Tivoli Tape Optimizer for z/OS 21, 571, 576
- IBM TotalStorage 3494 Tape Library 255

- IBM TotalStorage 7207 External Tape Drive 295
- IBM TotalStorage ETL Expert 563
- IBM TotalStorage Master Console (TSMC) 245
- IBM TotalStorage Productivity Center 129, 474
- IBM TotalStorage SAN256B 335
- IBM Virtualization Engine 275, 378
- IBM Virtualization Engine TS7700 9, 285–286
- IBM Virtualization Engine TS7740 9
- IBM XIV Storage System 7, 137
- ICEGENER 568
- ICETOOL 568
- ICMP 493
- IETF standard 489
- ILM 204
- image backup 455
- In-band 38, 502
- in-band 143
- Incremental FlashCopy 130
- Infiniband host connectivity 58
- InfiniBand switches 203
- Infiniband switches 212
- Information Archiving 4
- Information infrastructure 4
- Information Lifecycle Management (ILM) 581
- Infrastructure simplification 407
- Integrated Routing 331
- Integrated Solutions Console 414
- Integrated Solutions Console (ISC) 439, 441
- Intelligent information life-cycle management 540
- Intelligent Write Caching (IWC) 103, 118
- Interface Modules 140
- Interface node 208
- Interface nodes 203
- Interleaved 229
- Internet Engineering Task Force (IETF) 611
- Internet Fibre Channel Protocol (iFCP) 304, 318
- IPv4 15
- IPv6 15, 102, 420
- ISA 478
- iSCSI 7, 30, 45–46, 50, 53, 304, 316, 388, 400, 595
- iSCSI Gateway 348
- iSCSI HBA 313
- iSCSI HIC 53
- ISL 335, 339, 344
- ISL Trunking 328–329, 332
- ISPF 567
- ISV 240

J

- J02M 366
- J06M 366
- J11M 366
- Journal volume 521
- JUNOS 15, 364, 366

K

- KSDS 573
- KVM 221

L

- LACP 359
- LAN 413
- LAN-free 452
- LAN-free data transfer 441
- LANFREECOMMMETHOD 428
- LC connectors 311
- LDAP 102, 202, 476
- Library and device support 441
- Library-Managed Encryption (LME) 256, 265
- Licensed Machine Code (LIC) 135
- Lightweight Directory Access Protocol (LDAP) 143
- Lightweight Third Party Authentication (LTPA) 476
- Linear cache growth 147
- Linear density 234
- Linear Tape-Open (LTO) 228
- lithium-ion battery 53
- logical block address (LBA) 480
- logical drives 73
- logical unit (LU) 410
- logical unit number (LUN) 410
- Logical Volume Managers (LVM) 374
- logical volumes 54, 151
- longitudinal position (LPOS) 231
- Loop (arbitrated) topology 307
- Low Voltage Differential (LVD) 238, 257, 260
- LPARs 108
- LPOS 231
- LTO 10, 248, 255, 259–260
- LTO CM 228
- LTO tape 10
- LTO2 drives 10
- LTPA 476
- LUN mapping 191
- LUN masking 93
- LUN Provisioning 495

M

- MAN 13
- managed disk 482
- Managed disk (MDisk) 410
- managed disk group (MDG) 393, 410
- managed disks (MDisks) 480
- Management Information Base (MIB) 502
- Management node 201, 203, 208
- management station 38
- Massive parallelism 141
- Master console 410
- Maximum drives supported 55
- Maximum LUN size 191
- maximum number of partitions 73
- MCDS 574
- MDG 393
- MDisks 393
- Media reuse 434
- MEDIA10 246
- MEDIA9 246
- MERGECAAT 573
- Metro Mirror 55, 81, 129, 387, 395, 405, 520

- Metro Ring Protocol (MRP) 359, 362
- MetroCluster 197
- Metropolitan Area Network (MAN) 347
- MIB 502
- Microsoft Cluster Server (MSCS) 441, 459, 462
- Microsoft Virtual Disk Service (VDS) 94
- Microsoft Volume Shadow (VSS) 94
- Microsoft Volume Shadow Copy Services (VSS) 421
- Microsoft's Volume Snapshot Services (VSS) 447
- Minimum LUN size 191
- mirrored cache 391
- Mirrored VDisk 398
- mirrors 74
- movement policies 566
- MPLS 15, 366
- MSCS 462, 464–465
- multipath architecture 234
- multipath device driver 394, 481
- multipathing device driver 75
- Multiple Relationship FlashCopy 130
- multiprotocol router 13, 346
- Multistore 196
- multi-tiered storage 91

N

- N_Port ID Virtualization 315
- N_Ports 314
- N3000 16, 179
- N3300 176
- N3600 176
- N6000 183
- N6040 176
- N6060 176
- N6070 176
- N7000 17, 185
- N7700 176
- N7900 176
- NAS 423, 583
- NDCL 169
- NDMP 189, 443, 454
- Near-line storage 175
- NearStore option 197
- Network Attached Storage (NAS) 304
- Network Data Management Protocol (NDMP) 441
- Network File System (NFS) 173, 202
- Network Information Service (NIS) 195
- Network shared disks (NSDs) 544
- Network-attached storage (NAS) 454
- Network-free rapid recovery 440
- NFS 196, 222, 304, 419, 540, 548, 586
- NL_Ports 315
- Node 410
- Non-Disruptive Code Load (NDCL) 142
- nonrewritable (NENR) 217
- no-query restore (NQR) 417
- Novell Cluster Services (NCS) 441
- NPIV 315
- N-Port identifier virtualization 334
- NPV 334
- NQR 417

- NSD 544
- NTP 103
- NTP synchronization 389
- NVRAM 186
- NVram destage backup battery 391

O

- Object store with global name space 539
- OCDS 574
- ODBC 417
- off-site vaulting 469
- OLTP 49
- OM3 fiber optic cables 45
- open database connectivity (ODBC) 417
- Open Systems Management Server (OSMS) 502
- Operations Manager 19
- Operations Manager Core 197
- Operations Manager SRM 197
- OSMS 502
- OUTFIL 568
- Out-of-band 38, 502

P

- PAM 198
- parallelism 146
- Partition Expansion License 39
- Partitioning 29
- PB 17
- PCI-DSS 104
- Performance Accelerator Module (PAM) 198
- performance management 477, 499
- performance thresholds 500
- petabyte (PB) 9, 17
- PiT 79
- POD license 331
- point-to-point topology 307
- Ports on Demand (POD) 329
- Power over Ethernet (PoE) 360, 365
- POWER5 87, 105
- POWER6 87, 104
- PowerHA 459
- PPRC 468
- premium features 39
- primordial storage pool 500
- Private Cloud 580
- processor complexes 111
- processor memory 53, 111
- processors 95
- Progressive backup methodology 440
- progressive incremental backup 435
- Proportional interface growth 147
- ProtectTIER 9, 279, 282
- ProtectTIER Enterprise Edition software 281
- ProtectTIER Native Replication 282, 284
- Protection Manager 18
- Provisioning profiles 478
- Public Cloud 580

Q

QoS 15
query archive 419
query backup 419
querysummary 419
queue depth 73

R

RAID 33, 35, 221, 599
RAID 0 600
RAID 1 601
RAID 3 602
RAID 5 603
RAID 5 Enhanced 604
RAID 6 605
RAID controller 50, 55
RAID level migration 31
RAID-00 606
RAID-10 606
RAID4 602
RAID-50 607
RDAC 76
RDBMS 496
RDBMSs 482
reclamation 434
redundant array of independent disks 599
remote disk mirroring 468
Remote I/O 105
Remote Mirroring 48, 143
Remote Support Manager 54
Remote Switch 328
Remote Switch Activation 335
Remote Volume Mirroring 55
Report Generator 567
Reporting views 477
Reverse FlashCopy 388, 397
RIO-G 105, 112, 116
RLM 186
RMAN 445
Role pair 521
Rollup views 477
Routing Engines (REs) 16, 366
RSCT 420
RSM 82, 219, 221

S

SAN 413
SAN discovery 418
SAN Planner 478
SAN Volume Controller cluster 403
SAN Volume Controller software 403
SAN04B-R 13
SAN04B-R multiprotocol router 347
SAN06B-R 14
SAN24B-4 13
SAN256B 11
SAN384B 11, 337
SAN40B-4 12, 331
SAN768B 11, 338

SAN80B-4 12, 332
SANDISCOVERYTIMEOUT 418
SAP 421
SAP BRTools 17
SARC 92
SAS 8, 30, 32, 37, 179, 238, 255, 257
SAS host connectivity 29
SATA 7–8, 30, 32, 37, 106, 137
SATA DDM 51
SBOD controller card 97
SBSC 579
SBSC host attachment 586
SC connectors 311
scale out file services 548
SCSI 441
SCSI tape failover 464
SDDDSM 75
SDLT 236
SDM (System Data Mover) 562
SE VDisk 399
Secure Copy Protocol (SCP) 202
SecureAdmin 195
Security 4, 6
Self-healing 142
Sequential Prefetching in Adaptive Replacement Cache (SARC) 92
Serial Attached SCSI (SAS) 33, 35
serial connection 55
ServeRAID 37
Server-to-server virtual volumes 458
Service agent (SA) 490
service level agreements (SLAs) 52
Service Location Protocol (SLP) 489, 498, 514, 609, 611
Service management 4
Service Oriented Architecture 2
service processor (SP) 112
servo bands 229, 242
Servo tracks 229
SFPs (Small Form Factor Pluggable Media) 312
shortwave 330
silent installation 479
Single Character Display (SCD) 238
Single command-line interface (CLI) 447
single controller 33, 35–36
Single Mailbox Recovery for Exchange (SMBR) 197
Single sign-on 476
single virtual cluster node 583
SLC 241
SLP 489
smart business storage cloud (SBSC) 579
smart planet 6
SMB 11
SMI Specification (SMI-S) 609
SMI-S 94, 263, 482, 489, 498, 501, 506, 514, 609
SMI-S (Storage Management Initiative Specification) 87
SMI-S agent 477, 480
Snapdiff 419
SnapDrive 197
SnapLock 18, 196
SnapManager 197

- SnapManager for Hyper-V 17
- SnapManager for Microsoft Exchange 18
- SnapManager for Microsoft Office SharePoint Server 2007 18
- SnapManager for Microsoft SQL server 18
- SnapManager for Oracle 18
- SnapManager for SAP 17
- SnapManager for Virtual Infrastructure 17
- SnapMirror 19, 179, 196, 417
- SnapMover 195
- SnapRestore 196
- snapshot 143, 151, 160, 195, 421, 556, 592
- Snapshot reserve capacity 155
- SnapValidator 197
- SnapVault 179, 196
- SNIA 94, 489, 609
- SNMP 219, 493
- SOFS 543, 548
- SoFS Hardware Components 551
- SoFS single site solution 553
- SoFS Software Components 552
- Software 147
- Solid State Disk 45, 65
- solid state drives (SSDs) 54, 103, 391
- Solid-state Storage Architectures 4
- SONAS 201
- Space Efficient (SE) 397
- Space Efficient VDisk 399
- Space Management 565
- Split-cluster 388
- split-mirror/point-in-time backup 453
- srvoptsetencryptiondisabled 419
- SSD 65, 104, 106, 388
- SSPC 22, 107, 389, 527–528
- Storage Archive Manager (SSAM) 222
- Storage Area Network (SAN) 304
- Storage Class Memory (SCM) 391
- Storage controller 201, 203, 208
- Storage entity views 477
- Storage Hierarchy 565
- Storage Management 4
- Storage Management Initiative - Specification (SMI-S) 489, 609
- Storage Manager 45, 70
- Storage Networking Industry Association (SNIA) 381, 489
- Storage node 203, 208
- Storage Optimizer 477, 506, 508
- Storage partitions 31, 55
- Storage Pool Striping 102
- Storage pools 151, 154, 158, 432, 546
- Storage pools and volumes 431
- Storage resource groups 478
- Storage Resource Management (SRM) 514
- Storage usage for Network Appliance quotas 495
- storage usage quotas 495
- Storage virtualization 4, 372
- Streaming Lossless Data Compression (SLDC) 241
- Structured Query Language (SQL) 476
- subsystem device drivers (SDD) 481

- Support bundle 73
- SVC 19, 378, 383, 388
- SVC entry level edition software (EE) 388
- Switch alerts 504
- Switch Control Board (SCB) 16, 366
- Switch Performance Monitors 504
- switched FC-AL 91
- Switched Fibre Channel Arbitrated Loop (FC-AL) 91
- Switched topology 308
- Synchronous PPRC 131
- synchronous replication 554
- SyncMirror 196
- SysBack 449
- System quiesce 162
- System-Managed Encryption (SME) 256, 265
- System-Managed Storage (SMS) 576
- Systems Manager 196

T

- TagmaStore CIM agent 500
- Tape Autoloader 10
- Tape Manager 486, 505
- Tape resource sharing 440
- Tape vaulting 466
- Tape Virtualization 276
- Tape-Open Cartridge Memory (LTO - CM) 249
- Target volume 521
- Task assistants 76
- TCIM 219
- TCO 12
- TCP/IP port 425
- Telco 33, 35–36
- Thin provisioning 103, 142–143, 156
- Thin server 174
- Threshold migration 420
- Tiered storage 387
- Tivoli Disaster Recovery Manager 458
- Tivoli Enterprise Console (TEC) 493
- Tivoli Enterprise Portal (TEP) 572, 575
- Tivoli Integrated Portal 475, 481, 484
- Tivoli Key Lifecycle Manager (TKLM SW) 69
- Tivoli Monitoring 436
- Tivoli Storage Manager 20, 413, 582
- Tivoli Storage Manager Administration Center 438
- Tivoli Storage Manager application program interface (API) 439
- Tivoli Storage Manager DRM 466
- Tivoli Storage Manager enterprise configuration 457
- Tivoli Storage Manager family 440
- Tivoli Storage Manager monitoring agent 436
- Tivoli Storage Manager security 439
- Tivoli Storage Manager Version 6.1.2 414
- Tivoli Storage Productivity Center 475
- Tivoli Storage Productivity Center for Replication 475
- TKLM 108, 127
- top-of-rack (TOR) 16
- Topology Viewer 503
- TPC 20, 144, 483
- Transceivers 312
- True virtualization 142

- TS1120 240, 245
- TS1130 9, 240, 246
- TS2230 10, 236
- TS2240 11, 236–237
- TS2340 236, 238
- TS2900 10, 254–255
- TS3100 10
- TS3200 10
- TS3310 10, 262
- TS3400 9, 265–266
- TS3500 9
- TS7650 10, 279
- TS7650 ProtecTIER Deduplication Appliance 10
- TS7650G 281
- TS7650G ProtecTIER Deduplication Gateway 281
- TS7700 9, 285
- TS7700 Virtualization Engine 9
- TS7720 285
- TS7740 285
- TSM Administration Center - TsmAC 414
- tsmdlst utility 417

U

- Ultrium 10–11, 228, 234
- Ultrium 1 235
- Ultrium 3 235
- Ultrium 4 235
- unification engine 179
- UPS 391
- User agent (UA) 490

V

- Vaulting 466
- VCS 464
- VDisk 390
- VDisk Mirroring 388
- VERITAS 459
- VERITAS Cluster Services (VCS) 441
- Virtual backhitch 240
- virtual cartridges 279
- Virtual Chassis 365
- Virtual disk (VDisk) 386, 410, 480, 482
- Virtual Disk Service (VDS) 76
- virtual drives 279
- Virtual File Manager (VFM) 197
- Virtual I/O 45, 270
- Virtual Interface (VIF) 177
- virtual libraries 279
- virtual local area network (VLAN) 177
- Virtual SAN (VSAN) 330, 334, 341–342, 344, 350
- Virtual Switch Redundancy Protocol (VSRP) 359, 362
- Virtual Tape Library (VTL) 278, 282
- Virtualization 4–5, 371
- VMware Consolidated Backup 456
- VMware ESX 17
- Voice over IP (VoIP) 15
- VoIP 15, 359–360
- VOLSER 572, 574
- Volume layout 152

- Volume locking 160
- volume retention 566
- Volume Shadow Copy Service (VSS) 76
- VolumeCopy 29, 31, 48, 55, 77, 79
- VolumeCopy License 40
- VSAM 567
- VTOC 572
- VTOCIX 572
- VVDS 572–573

W

- WAN 13, 413
- WAN-optimized replication 539
- Wavelength Division Multiplexing (WDM) 323
- WBEM 610
- WDM 323
- Web-Based Enterprise Management (WBEM) 488, 609
- Wide Area Network (WAN) 13
- Windows RDAC 75
- Workload profiles 478
- WORM 238, 255
- Write Once Read Many (WORM) 217, 235
- write order consistency (WOC) 81
- Write path redundancy 161

X

- XCLI 144–145
- XFP 359
- XFPs 336
- XIV 87
- XIV Storage Manager GUI 144
- XIV Storage System 378, 480
- XIV Storage System architecture 141
- xmlCIM 488
- XRC 347
- XRC Performance Monitor 563

Z

- z/OS 562
- z/OS Global Mirror 563
- z/OS NFS 563
- z/OS storage management tools 21, 571
- Zone control 501
- Zoning 501



Redbooks

IBM System Storage Solutions Handbook

(1.0" spine)

0.875" <-> 1.498"

460 <-> 788 pages



IBM System Storage Solutions Handbook



Redbooks®

Learn about IBM storage products and solutions

Get flexible, scalable storage solutions for the on demand world

Find the products that answer your business needs

This IBM Redbooks publication provides overviews and pointers for information about the most current IBM System Storage products, showing how IBM delivers the right mix of products for nearly every aspect of business continuance and business efficiency. IBM System Storage products can help you store, safeguard, retrieve, and share your data.

We cover the following topics:

- ▶ Part 1 describes disk products that include IBM System Storage DS Series (entry-level, midrange, and enterprise offerings), IBM XIV Storage System, IBM System Storage N Series, and the IBM Information Archive.
- ▶ Part 2 presents an overview of the IBM TotalStorage and System Storage Tape Drives, IBM Tape Automation Products, and IBM Tape Virtualization Solutions Products.
- ▶ Part 3 describes storage networking infrastructure and presents the switches and directors to form SAN solutions. Also provided is an overview of the recently introduced IBM Ethernet switches and routers.
- ▶ Part 4 discusses the IBM System Storage software portfolio.
- ▶ Part 5 describes the z/OS storage management software and tools.

This book is intended as a reference for basic and comprehensive information about the IBM System Storage products portfolio. The book provides a starting point when establishing your own enterprise storage environment.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks