SSD OPTIONAL FUNCTION SECTION

SSD Optional Function

Two SSD options explained below are provided. The respective options can be set and used independently, and naturally, when both options are used together, the performance can be improved much more.

- Customized Volume Size (CVS) option which allows to set a free volume size and reducing occurrences of waiting for host I/O issues
- Dynamic Cache Residence (DCR) option which improves the data access performance by making data at addresses in the specified area resident in the DKC cache

Contents

1. Customized Volume Size (CVS) Option

1.1 Outline ----------SSDOPT01-10
1.2 Features ----------SSDOPT01-10
1.3 Specifications ----------SSDOPT01-30
1.4 Maintenance functions ----------SSDOPT01-60

2. Dynamic Cache Residence (DCR) Option

2.1 Outline ----------SSDOPT02-10
2.2 Features ----------SSDOPT02-10
2.3 Specifications ----------SSDOPT02-50
2.4 Maintenance functions ----------SSDOPT02-60
2.5 Notes on maintenance when DCR is used ----------SSDOPT02-70
2.6 Effects of DKC failures on DCR ----------SSDOPT02-70
2.7 Automatic cancellation of DCR ----------SSDOPT02-70
2.8 Explanation of DCR cache and procedure for setting operation ----------SSDOPT02-80

1. Customized Volume Size (CVS) Option

1.1 Outline

As regards the main frame host, the multiplicity of an I/O request is restricted to one per volume because UCBs are mutually exclusive.
Therefore, when two or more files to which I/Os are applied frequently exist in the same volume, a contention for the logical volume occurs. If this occurs, the files mentioned above are stored separately in different logical volumes and an action is taken to avoid contention for access. (Or means to prevent I/Os from generation is required.)
However, the work for adjusting the file arrangement giving consideration to the accessing characteristic of the file will be a burden on users of the DKC and it is not welcomed by them.
To solve this problem, the Customized Volume Size (CVS) option is provided. (Hereinafter, it is abbreviated to CVS.)
The CVS provides a function for freely defining the logical volume size. By doing this, even in a storage system with the same capacity, the number of volumes can be increased easily. As a result, a file with a high I/O frequency can be easily allocated to an independent volume. That is to say, the trouble to consider a combination of stored files in a volume can be saved.

1.2 Features

- The contention for UCB in the main frame host can be avoided and the IOSQ time can be reduced.
- The capacity of the ECC group can be fully used.
- By combining with the Dynamic Cache Residence (DCR) option, high performance equivalent to that of a semiconductor storage device can be realized.

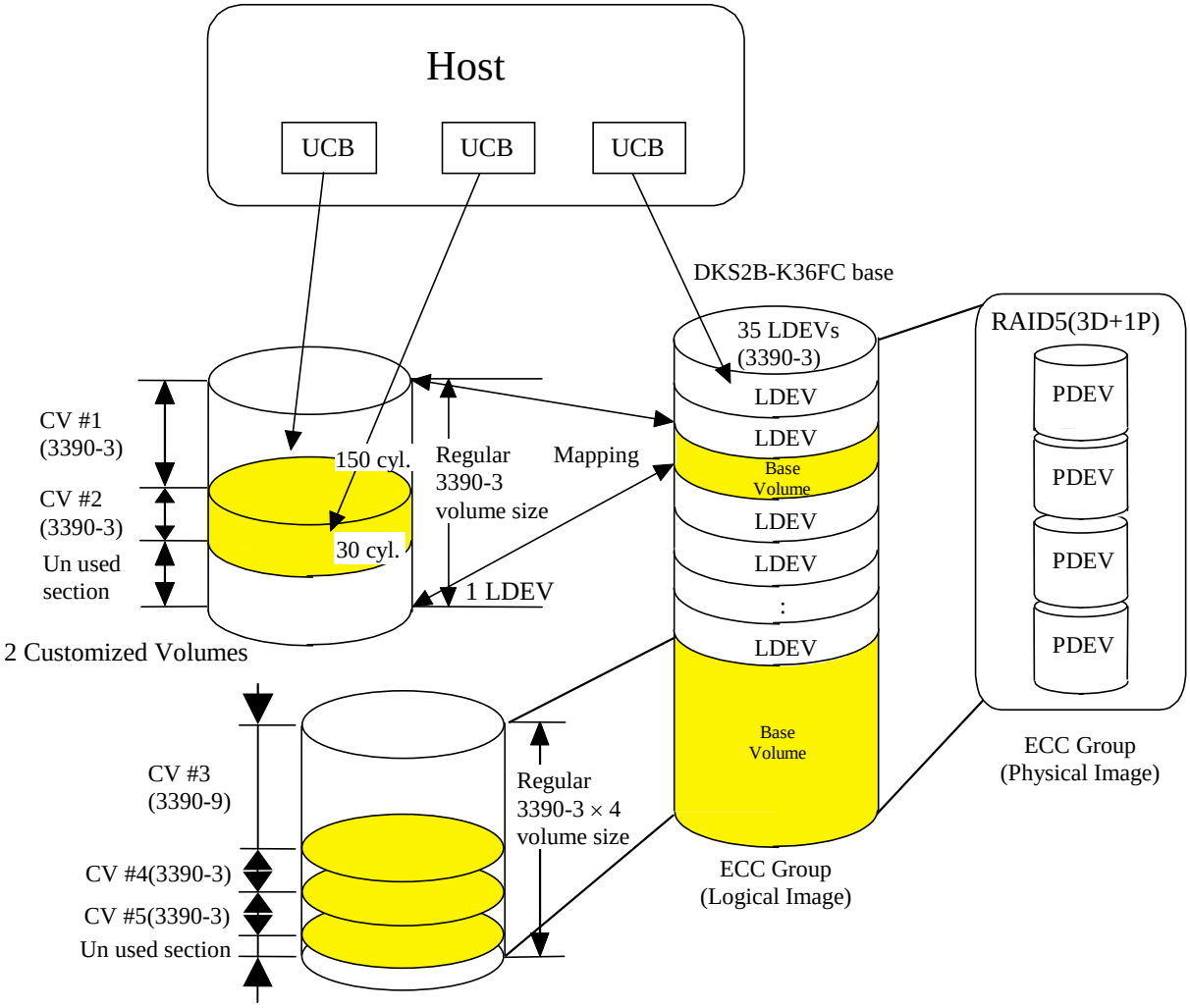


Fig. 1.1 Conceptual Diagram

1.3 Specifications

The CVS option consists of a function to provide variable capacity volumes and a function to provide free arrangement of volumes on the ECC group.

(a) Function to provide variable capacity volumes
This is a function for determining the capacity as specified by users in the same track format as that of a conventional volume.

(b) Emulation type intermix on the ECC group
As regard the DKC465I, the logical volume type in the ECC group is restricted to the same type only. However, the CVS option enables different logical volume types to coexist.

Table 1.1 Specifications of the Function for Providing Variable Capacity Volumes

Item No.	Item	Description
1	Track format	3390 track format
2	CV emulation type	3390-3/3R/9/L/3A/3B/3C, OPEN-3/8/9/E/V
3	Maximum number of CVs on one ECC group	256
4	Minimum size of CV	See note1
5	Maximum size of CV	Base volume size (different by each emulation type) In case of OPEN-V, maximum size is 60 GB.
6	Position on the disk where the CVs can be placed	Any position is allowable.
7	Emulation type intermix	Intermix of 3390-3/3A/3B/3C, 9 and L is allowable. Intermix of OPEN-3, OPEN-8, OPEN-9 and OPEN-E is allowable. Intermix of OPEN-V and the above emulation type is not allowable.

Note1:

(1) Mainframe volume

Number of physical cylinder allocated on ECC group is calculated by following formula.

Physical mapping onto ECC group(track) =

User logical cylinders(A)×15 + Additional control cylinder(B)×15 + Adjustment(C)

∴ 1 cylinder = 15 tracks

where, (A) is specified value through SVP/Web Console, (B) is defined from below chart depending on emulation type.

The value of (C) depends on (A) + (B) and (C) has the value such as ranging from 0 to 47, 0 to 55.

i) RAID5 (3D + 1P), RAID1

$$(C) = \uparrow \frac{((A) + (B)) \times 15}{48} \uparrow \times 48 - ((A) + (B)) \times 15$$

ii) RAID5 (7D + 1P), RAID1

$$(C) = \uparrow \frac{((A) + (B)) \times 15}{56} \uparrow \times 56 - ((A) + (B)) \times 15$$

∴ ↑ ↑ means round up to the next integer.

e.g. ↑3.96↑ = 4

emulation type	3390-3	3390-3R	3390-9	3390-L	3390-3A	3390-3B	3390-3C
Minimum size of CV in cylinders	1	1	1	1	1	1	1
(B)	6	6	25	23	6	6	6

emulation type	OPEN-3	OPEN-8	OPEN-9	OPEN-E	OPEN-V
Minimum size of CV in cylinders	50 (35 MByte)	50 (35 MByte)	50 (35 MByte)	50 (35 MBtye)	50 (46 MBtye)
(B)	8 (5.76 MByte)	27 (19.44 MByte)	27 (19.44 MByte)	18 (12.7 MByte)	0

Example: Physically allocated number of cylinders in case of 100 cylinders specified with 3390-3 is calculated below:

Physical mapping onto ECC group = 100 × 15(A) + 6 × 15(B) + 42(C)

So number of tracks physically allocated on ECC group is equal 1632(108.8 cylinders).

(2) Open volume

You can set the data by Mbytes. And the Program will assign the actual value as the following expression. The typical values will be described in the table 1.2.

Except for OPEN-V

X = (your setting size) × 1024 ÷ 720 (if there is remainder, add 1 to X.)

Y = (X × 96 × 15 × 512) ÷ 1024 ÷ 1024

OPEN-V

X = (your setting size) × 16 ÷ 15 (if there is remainder, add 1 to X.)

Y = (X × 128 × 15 × 512) ÷ 1024 ÷ 1024

(note) X is a value of converting capacity into number of cylinders. Y is a value of converting value into number of mega bytes.
In case of open volume, physical mapping onto ECC group(track) is also calculated by the expression of (1).

Table 1.2 CV capacity

Setting Size (Mbyte)	Actual size (Mbyte)	Note
35	35.15625	Minimum CV capacity (except for OPEN-V)
46	46.875	Minimum OPEN-V capacity
500	500.625	
1000	1000.546875	
2000	2000.390625	
2347	2347.03125	Maximum size of CV for OPEN-3
7007	7007.34375	Maximum size of CV for OPEN-8
7042	7042.5	Maximum size of CV for OPEN-9
13893	13893.04688	Maximum size of CV for OPEN-E
61439	61439.0625	Maximum size of OPEN-V

NOTICE

When you set HMDE volumes to customized volumes and reset them to the normal volume again, these volumes could not be set as HMDE volumes. Please refer to the following table.

Emulation Types for HMDE volumes	Emulation types after changing from Customized volume to normal volume
3390-3A	3390-3
3390-3B	
3390-3C	

If you want to reset these volumes as HMDE, please call technical support center to set them to HMDE volumes by SVP.

1.4 Maintenance functions

Features of the maintenance functions of the CVS option is that they allow execution of not only the conventional maintenance operations instructed by the SVP but also the maintenance operations instructed from the remote console. (See Item Nos. 2 to 5 in Table 1.3.)
Unlike the conventional LDEV addition or reduction, the operation for the ECC group is made unnecessary, so that the volumes can be operated from the remote console.

Table 1.3 Maintenance Function List

Item No.	Maintenance function	TSC	CE	User	Remarks
1	Concurrent addition or deletion of CVs at the time of addition or deinstallation of ECC group	—	✓	—	Same as the conventional addition or deinstallation of LDEVs
2	Addition of CVs only	✓	✓	✓	Addition of CVs in the free area
3	Conversion of normal volumes to CV	✓	✓	✓	
4	Conversion of CV to normal volumes	✓	✓	✓	
5	Deletion of CVs only	✓	✓	✓	Only the optional CVs are deleted and incorporated into the free area. No deinstallation of ECC group is involved.

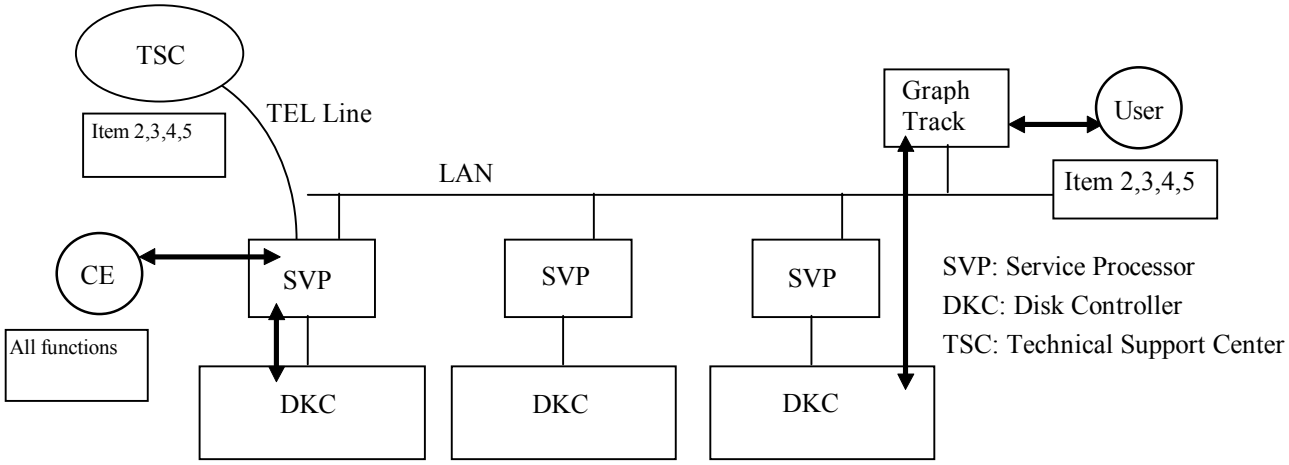


Fig. 1.2 Maintenance Execution Route when CVS Is Used

2. Dynamic Cache Residence (DCR) Option

2.1 Outline

Because cache capacity is usually smaller than the total HDD capacity in RAID subsystem, cache cannot be keep all data at the same time which is stored whole HDDs in the subsystem.
Cache management control cache extent so as to be allocate more capacity to more frequently accessed data to enbed this gap by LRU algorithm.
By this control, data with a low access frequency is hard to remain in the cache, de-staged into HDDs, therefore the more frequent accesses to physical HDDs occur, decrease the access performance and unpredictable response time appear.
The DCR option provides a function for making data being resident on the cache and realize high access performance.

2.2 Features

- Feature of the DCR consist of two modes, one is called “PRIOrity mode(hereinafter,it is abbreviated to PRIO)” and the other “BIND mode(hereinafter,it is abbreviate to BIND)”.
PRIO is a basic mode(100% Read Hit) of this feature, which fits typical user needs and BIND is supplementary for special customer(100% Read/Write Hit. Replace SSD) needs.
- To use the DCR, addition of the cache memory is required for the service as a "DCR cache".
- DCR supports PreStaging function. The PreStaging is a function which read data on Logical Volume onto cache by receiving Remote Console or SVP instructions.

2.2.1 PRIO

[Processing]

- When a read command to the data which is assigned as DCR extent is issued and first meet “miss” in cache, the data is staged into cache by usual staging mechanism.
The data remains in the DCR cache permanently for future access to guarantee read-hit performance even after data is transferred to the host, regardless usually cache LRU management.
If a write command is issued to the data remaining in the cache in the way above, the data is updated with an out-of-DCR cache segment provided for write data duplication, and the data is de-staged into the HDD. Then the cache segment for write data duplication is returned to the out-of-DCR cache segment group.
In this case, new data is left in the cache extent together with old data.

[Performance impact]

- Theoretically, because above de-staging into HDD is processed by usual asynchronous de-staging mechanism, succeeding host access has a possibility to meet the same cache slot collision by locking both the host access and asynchronous de-staging process.
However in that case, by minimizing the collision time implemented in micro-code, performance impact will be neglectably small for usual customer jobs.
- In the case the subsystem cache becomes overloaded, a responding performance degradation occurs because the non-DCR cache segment must be used for the data assurance during a period between write data reception and de-staging operation completion.

[Maximum DCR capacity]

- Addition of the "DCR cache" is required for the DCR, and a number of disk trucks equivalent to the capacity of the added "DCR cache" is the maximum number definable for the DCR. Besides, STR recommends to keep the standard cache capacity for the non-DCR portion as a minimum out-of-DCR cache, to avoid considerable performance degradation for original data because of the newly installed DCR (Standard Cache capacity is decided by the subsystem capacity).
To keep this rule, when the customer want to install DCR feature for the subsystem, he need to install additional cache capacity as a DCR area out of pre-defined standard cache capacity.
Table 2.1 indicates the additional cache capacity as a DCR area out of pre-defined standard cache capacity.
The additional cache capacity requires out-of-DCR cache capacity.
Table 2.1 is the relationship between number of using extents and additional cache capacity.

Table 2.1 Necessary addition of cache memory

	Number of Priority Mode extents	
	~ 512	~ 1024
Additional Standard cache memory capacity	512MB×2	1024MB×2

Caution
A required cache capacity in PRIO mode:
standard cache capacity + DCR cache capacity + the above cache capacity
A required cache capacity in BIND mode:
standard cache capacity + DCR cache capacity

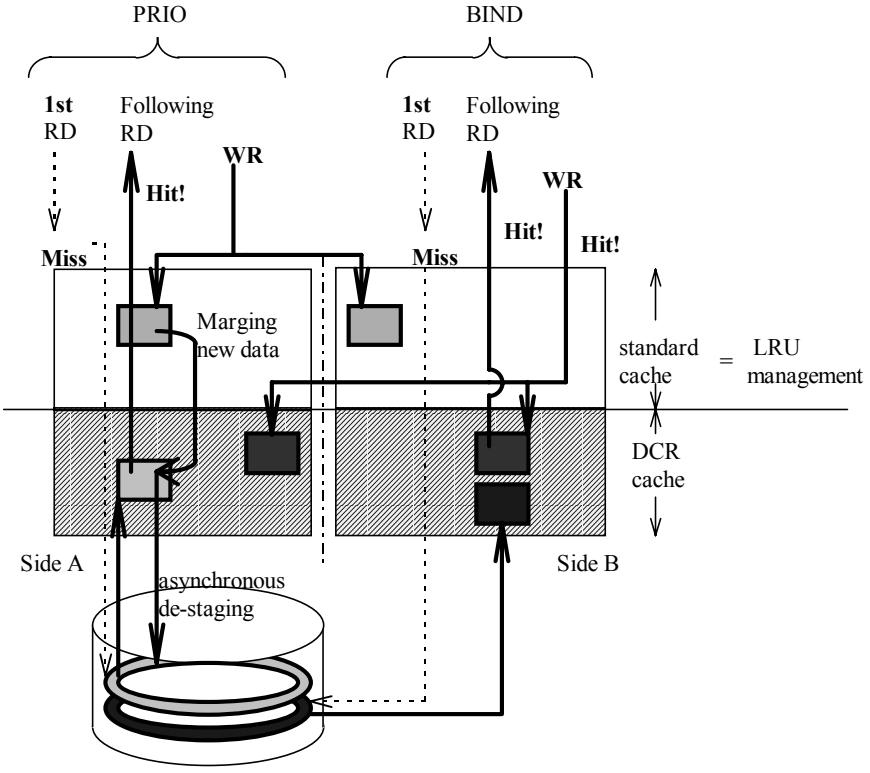


Fig. 2.1 Processing for DCR Extent

2.2.2 BIND

[Processing]

- As described above, there is a possibility that a responding performance degradation occurs in the PRIO mode because of a collide slot lock between the host access and asynchronous de-staging process or a waiting for the cache segment to be empty caused by an overload of the subsystem cache.
  This is a negligibly small impact on the performance in the typical user environment, however, in some environments, if any, in which the performance is very critical and the maximum response time must be assured, the above factors to degrade the response may be the issues.

[Performance impact]

- In BIND, the difference for PRIO, not only read data but also being all write data for the assigned DCR extent are resident in cache, no de-stage process occurred by any write command. Thus, by protecting any asynchronous de-stage process for the DCR data, read operation become a perfectly hit process.
- However, as a compensation for the perfect hit performance, additional cache element is necessary in the DCR cache area in the BIND mode three times larger than (for RAID1, as large as) that in the PRIO mode.

[Maximum DCR capacity]

- User data can be set 1/3 of cache capacity for BIND in case of RAID5, or 1/2 of that in case of RAID1.

2.2.3 Assignment of DCR extent and guard logic

- Above two modes can be assigned for each DCR extent individually.
  For example, the user want to assign 1GB cache for PRIO and 256MB cache for BIND, DKC allocate the extents in additional cache area by 1GB for PRIO and 256MB times 3 equals 0.75GB for BIND. Total 1.75GB cache should be additionally installed. The user can assign repeatedly many DCR extents with choosing each mode for each extent. Say, 1GB for PRIO + 512MB for PRIO + 256MB for BIND + 128MB for BIND comes to 2.6GB total DCR area necessary. The real capacity of cache needs adjusting to the cache unit to be added.
- SVP micro-code accept many extents allocation until the cumulative capacity for DCR does not allow the rest is over 512MB.
  In other words, DKC check that the rest cache capacity out of DCR is over 512MB by each DCR extent allocation, if the addition of DCR break the rest 512 MB boundary, SVP reject the allocation of DCR extents with an error message.
- From that point, the user theoretically can assign all cache capacity except 512MB for DCR regardless their configurations, STR strongly recommend the user should keep standard cache capacity out of DCR depend on the configuration according to the manuals by operation. Guard boundary by SVP is only 512MB for all configurations.
- For the outline and the setting operation procedure of the DCR cache, see page SSDOPT02-70.

2.2.4 DCR PreStaging

- The processor reports SIMs when the PreStaging abnormal end.
(following Table)

Error		REF CODE			SIM	Level	Host report	Remarks
		22	23	13	28			
Drive Error (Normal R/W)	PreStaging abnormal end	48	21	xx	FE	Service	No	xx: abnorml end reason code

x‘10’: No DCR PP
x‘20’: Subsystem busy
x‘40’: Staging time over
x‘50’: Cache blockade
x‘60’: LDEV blockade
x‘70’: Staging failure
x‘80’: P/S OFF
x‘90’: PreStaging canceled
x‘a0’: Cache over loaded

- In the case the subsystem cache becomes overloaded, a responding performance degradation occurs during a PreStaging execution. We strongly recommend to issue PreStaging request to stage data onto cache at the timing of normal load, or SIM REF CODE = 4821a0 may be reported and ended in failure.
- DKC rejects PreStaging requests during PreStaging execution. Please retry PreStaging requests after PreStaging termination.

2.3 Specifications

Table 2.2 Specifications of the Function

Item No.	Item	Description
1	Maximum number of areas to be made resident	For the PRIO and BIND modes together: 1024 areas/logical volume 1024 areas/subsystem
2	Unit of area specified to be resident	Mainframe : 1 track, Open : 96 logical blocks (128 logical blocks in case of OPEN-V)*2
3	Minimum/Maximum size of extent	1 track/logical volume size
4	Online change of resident area	Allowable (from the SVP and remote console)
5	Addition of cache capacity	Mandatory(Program Product: Charged with cache)
6	Maximum usable cache capacity*1 as DCR	Capacity of the cache memory added as the DCR cache. The "standard cache capacity" must be ensured by the rule.

*1: Convert as follows:
For the 3390 emulation: 1 track = 66KB
For the OPEN-3/8/9/E/L emulation: 1 track (96 logical block) = 49.5KB
For the OPEN-V emulation: 1 track (128 logical block) = 66KB

*2: In the case of open volume, the DCR program recognizes logical blocks in 96 block increments (128 block increments for OPEN-V). Table 2.3 shows the example of logical block address for 96 logical blocks.

- Recommendation about DCR for OPEN volume OPEN-3, OPEN-9, OPEN-E, OPEN-V:
  If DCR function is used for OPEN-3 or OPEN-8 or OPEN-9 or OPEN-E or OPEN-V, whole volume should be specified for DCR.
  It is because file of open system does not correspond to RAID track structure.

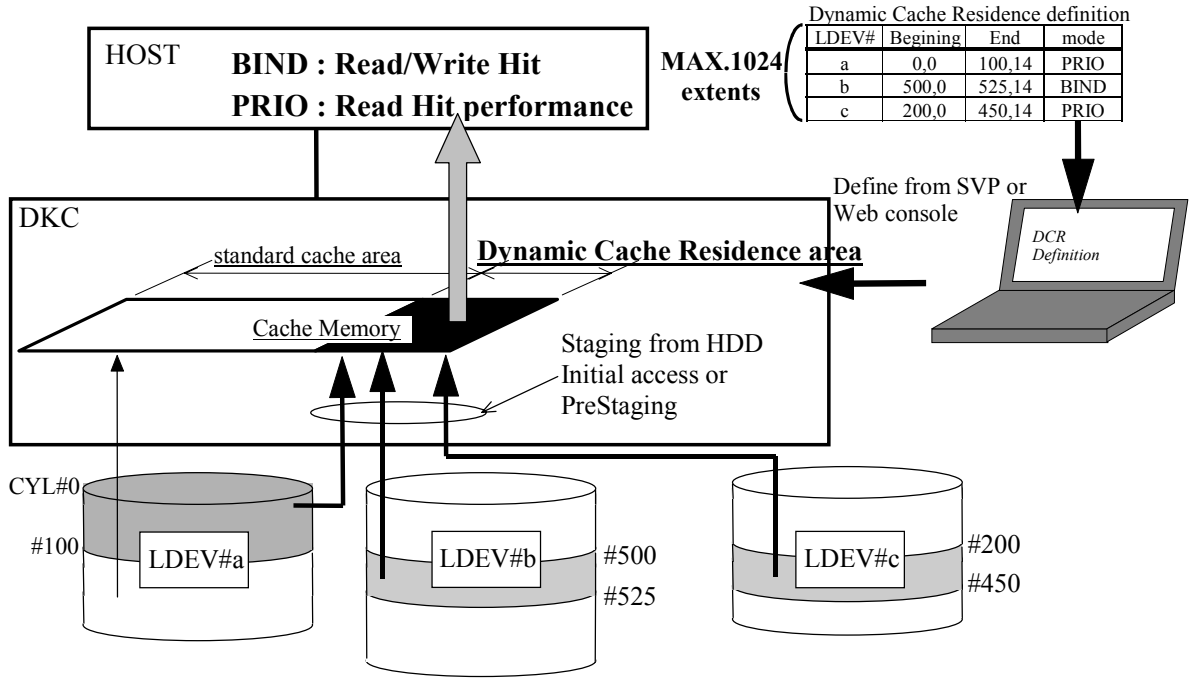


Fig. 2.2 Conceptual Diagram of DCR

Table 2.3 Starting and ending for 96 logical blocks

Starting LBA	Ending LBA	Remarks
0	95	
96	191	
192	287	
288	383	
384	479	
480	575	
576	671	
672	767	
768	863	
864	959	
860	1055	
1920	2015	
2880	2975	
3840	3935	
4800	4895	
5760	5855	
6720	6815	
7680	7775	
8640	8735	
9600	9695	
19200	19295	
28800	28895	
38400	38495	
48000	48095	
57600	57695	
67200	67295	
71904	71999	
72000	72095	
76800	76895	
86400	86495	
96000	96095	
192000	192095	
288000	288095	
384000	384095	
480000	480095	
576000	576095	
672000	672095	
768000	768095	
864000	864095	
960000	960095	
1920000	1920095	
2880000	2880095	
3840000	3840095	
4800000	4800095	
4806624	4806719	OPEN-3 Maximum LBA
4806720	4806815	
5760000	5760095	
6720000	6720095	
7680000	7680095	
8640000	8640095	
9600000	9600095	
10560000	10560095	
11520000	11520095	
12480000	12480095	
13440000	13440095	
14350944	14351039	OPEN-8 Maximum LBA
14371104	14371199	
14422944	14423039	OPEN-9 Maximum LBA
14423040	14423135	
28452864	28452959	OPEN-E Maximum LBA

2.4 Maintenance functions

Table 2.4 Maintenance Function List

Item No.	Maintenance operation	TSC	CE	User	Description
1	Addition of the DCR area	—	✓	✓	• Adds the DCR area. • Unit of the area to be specified by the SVP is a track.
2	Deletion of the DCR area	—	✓	✓	• Deletes the continuous DCR area.
3	Change of the DCR area	—	✓	✓	• Changes the DCR area size.
4	Status display of the DCR area	—	✓	✓	• Displays the specifications of the DCR area.
5	Addition and de-installation of the DCR cache	—	✓	—	• Because insertion and pulling off of the cache module into/from the DKC is required
6	Indication of DCR PreStaging	—	✓	✓	• Indicates the DCR PreStaging.

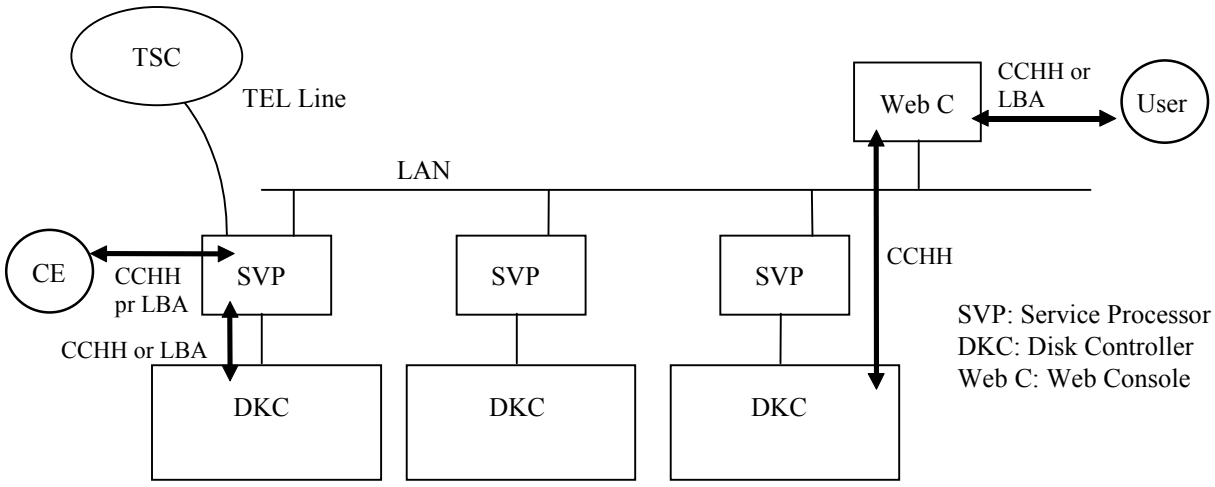


Figure 2.3 Maintenance Execution Route when DCR Is Used

2.5 Notes on maintenance when DCR is used

When performing the following maintenance, it is attended with a temporary regression of the cache memory or shared memory. Since this regression disables retention of the cache capacity required for the DCR, the DCR function is automatically suppressed for a while until the maintenance is completed.
A service person is required to obtain an approval of a user before starting the maintenance because there is a high possibility that the DCR function suppression may result in a degradation of responding performance.

(1) Cache replacement
(2) Addition and de-installation of the cache (including addition and de-installation of the DCR cache)
(3) Addition of the ECC Gr. and LDEV requiring addition of the SM because of the addition of the CU
(4) Addition and de-installation of the SM
(5) Cluster maintenance

We recommend that CE or Customer should execute the following action after DKC power supply restoration or equipment restoration, when the DKC power supply is down by power failure or mistake during "DCR area" release.
(Because it is a high possibility that the action may result in a degradation of responding performance, CE should execute the following action on a customer's authority.)

Action : (1) CE or Customer should release all DCR areas in a DCR area released Volume.
(2) CE or Customer should set up again all DCR areas with the exception of the released DCR area in the DCR area released Volume.

Reason : When DKC power is off during "DCR area" release, it is possible that DKC left a release DCR data on Cache.
DKC does not faulty operation by leaving "released DCR data" use a excessive cache memory.
We recommend that CE or Customer should execute the above-mentioned action after DKC restoration, because DKC perfectly execute the "DCR area" release process.

2.6 Effects of DKC failures on DCR

The DCR function is automatically suppressed when any of the following failures occurs. The suppression continues until the regressed operation owning to the maintenance is canceled in the cases of Items (1) to (3) or continues until an automatic recovery of the shared memory by the microprogram is terminated normally in the case of Item (4).

(1) Cache failure
(2) Shared memory failure
(3) One-side cluster down
(4) One-side shared memory blockade (SIMRC = FFEE)

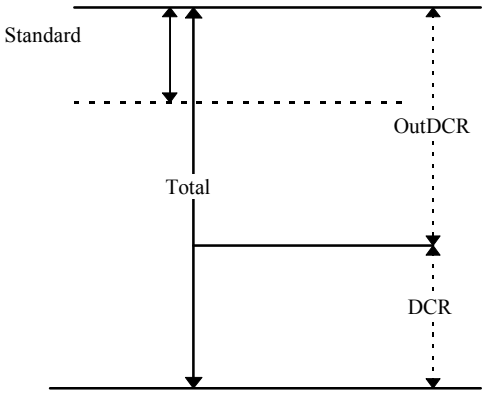
2.7 Automatic cancellation of DCR

The DCR setting of a volume to be de-installed by the functions of Deletion of CVs (LDEV), Conversion of CV to normal volume, and Conversion of normal volume to CV is automatically canceled as a part of the de-installation processing by the SVP microprogram.

2.8 Explanation of DCR cache and procedure for setting operation

2.8.1 Explanation

- A cache module must be defined and installed for the DCR before using the DCR.
- The DCR extent can be set only for the defined "DCR cache capacity."
- The "out-of-DCR cache capacity" must be retained more than "standard cache capacity" which is defined in accordance with the disk capacity in order to assure the performance in the non-DCR area.
- Therefore, DCR extent definition more than "DCR cache capacity" is rejected according to the SVP guarding logic. Also, the "DCR cache capacity" definition lower than the minimum "standard cache capacity" (256 MB × 2) is also rejected.



Total : Equipment Cache capacity

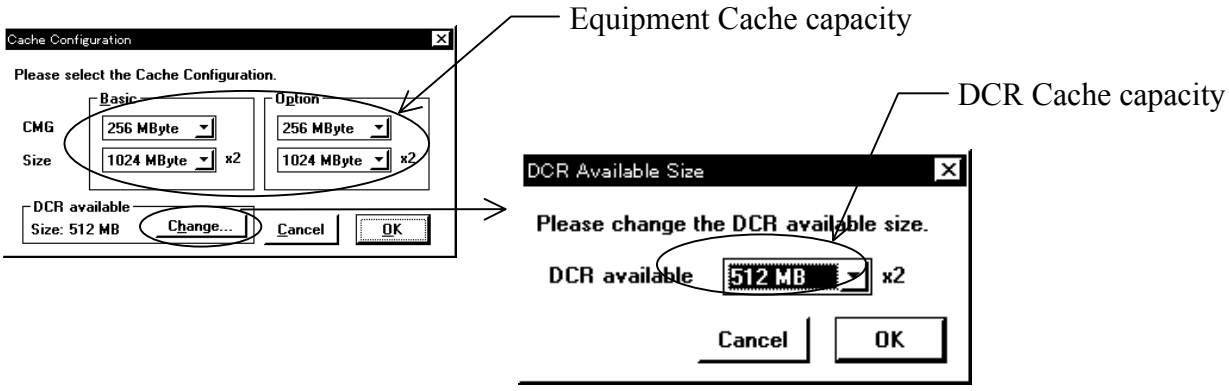
OutDCR : Out-Of-DCR Cache capacity

DCR : DCR cache capacity

Standard : Standard Cache capacity

2.8.2 Setting operation procedure

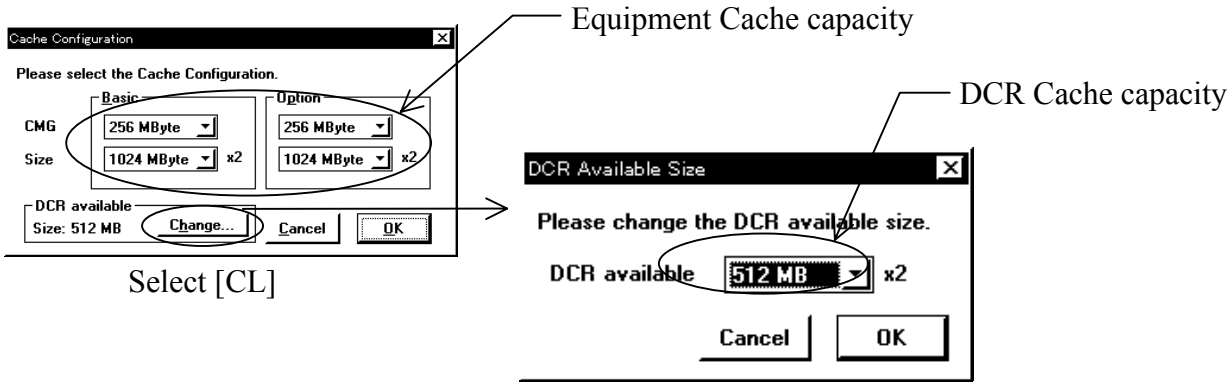
(1) Setting DCR cache capacity in Define Config & Install sequence
Set the DCR cache capacity in the equipment cache capacity.



Note :

- Set the DCR cache capacity so that it is less than the "equipment cache capacity minus standard cache capacity."

(2) Adding operation of DCR cache capacity in cache addition sequence
When adding the cache, set the DCR cache capacity in the equipment cache capacity after the addition.

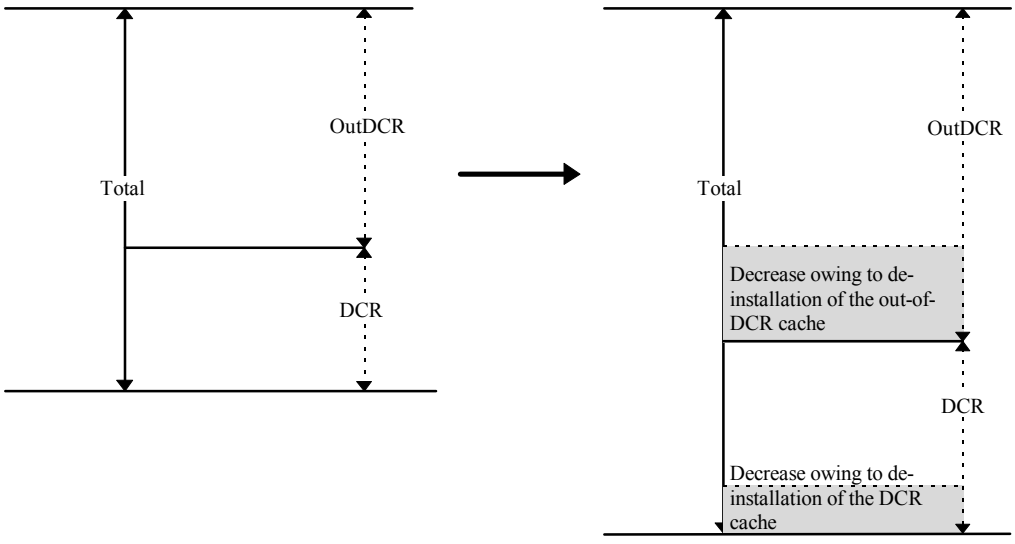


For example, to change a status with the cache of 1.0 GB × 2 installed and 256 MB × 2 of it set to the DCR cache to a status with the cache of 2.0 GB × 2 installed and 512 MB × 2 of it set to the DCR cache by adding the cache of 1.0 GB × 2,

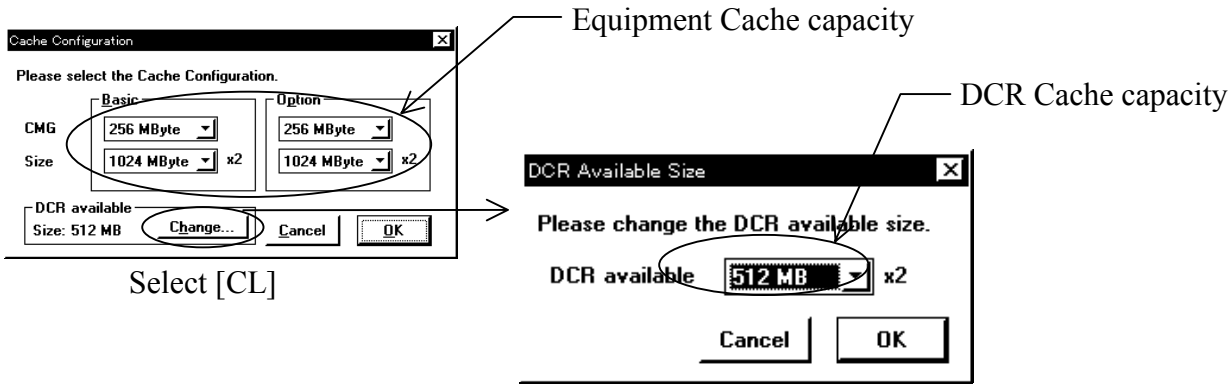
① set the equipment cache capacity to 2.0 GB × 2 in the "Cache Configuration" dialog box, and

② press the "Change..." button to open the "DCR Available Size" dialog box and set the DCR cache capacity to 512 MB × 2.

The DCR cache capacity can be set up to the cache capacity to be added. In the above example, the DCR cache capacity can be set up to 768 MB × 2 by adding 512 MB × 2.



(3) DCR cache capacity decreasing operation in cache de-installation sequence
When de-installing the cache, set a capacity to be left as the DCR cache in the equipment cache capacity.



For example, to change a status with the cache of 2.0 GB × 2 installed and 512 MB × 2 of it set to the DCR cache to a status with the cache of 1.0 GB × 2 installed and 256 MB × 2 of it set to the DCR cache by de-installing the cache of 1.0 GB × 2,

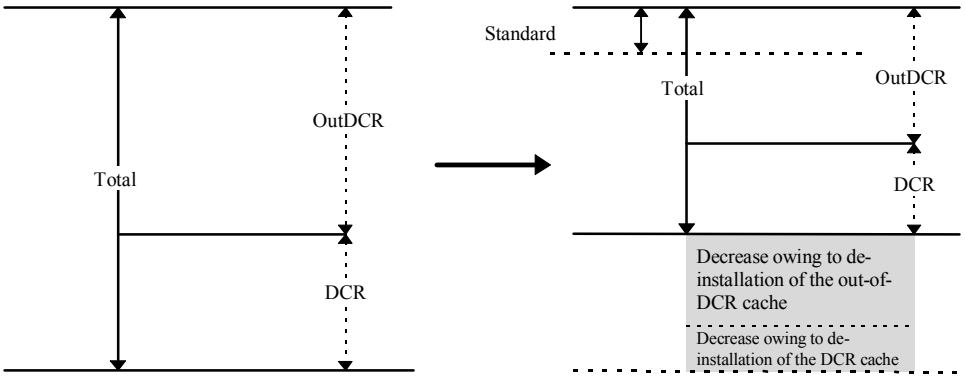
① set the equipment cache capacity to 1.0 GB × 2 in the "Cache Configuration" dialog box, and

② Press the "Change..." button to open the "DCR Available Size" dialog box and set the DCR cache capacity to 256 MB × 2.

The maximum decreasable capacity of the DCR cache is equal to the capacity of the installed cache to be de-installed. The maximum decreasable capacity of the DCR cache capacity in the above example is 1.0 GB × 2. As a result, the DCR cache capacity after the de-installation becomes 0 MB × 2.

Notes :

- In the case in which the de-installation of the DCR cache causes the capacity used by the DCR (see (i) on page SSDOPT02-100) actually defined as the DCR extent to be above the DCR cache capacity, the cache de-installing process is suspended by the SVP guarding logic. Before executing the DCR cache de-installation, cancel the DCR setting to decrease the actual capacity used by the DCR.
- It is required to avoid de-installation of the out-of-DCR cache which causes its capacity to be below the standard cache capacity.



(i) The "cache capacity used by the DCR" actually used as the DCR extent is displayed on the DCR Configuration screen in [SVP] - [Install] - [Refer Configuration] for confirmation.