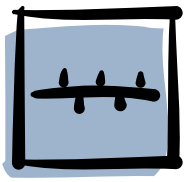
Figure Legend

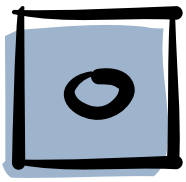
LAN Switch



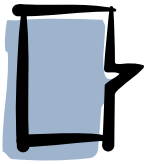
Ethernet

Router

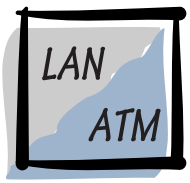
ATM Switch



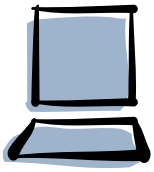
Token Ring



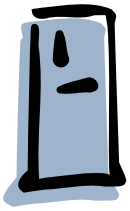
PBX



Hybrid LAN / ATM Switch

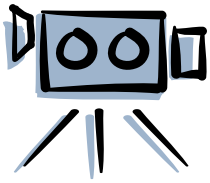


Workstation



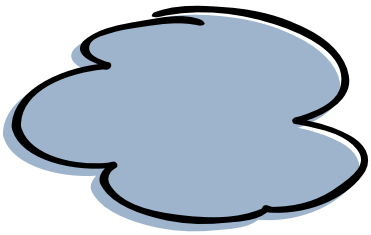
File Server

WAN Comm Link



Video

FDDI



Network

Table of Contents

I. Introduction 5
Why is switching important? 5
Notes to the reader 7

II. A Brief History of Networking and Switching 11
Era I: Mainframe networks 11
Era II: Minicomputer networks 12
Era III: Early PC networks 13
Era IV: Enterprise PC networks 19
Era V: Intelligent fabric networks 24
What's an intelligent fabric? 24

III. Switching Today 31

IV. LAN Switching 33
Where are LAN switches useful? 33
Basics of LAN switching 37
Any-to-any (translational) switching 40
Token ring switching 41
Inverse multiplexing 45
The future of LAN switching 46

V. Virtual LANs 49
Where are virtual LANs useful? 49
General issues surrounding VLANs 50
How VLANs are defined 52
VLANs standards 55
The future of VLANs 57

VI. Layer-Three Switching 59
Where is layer-three switching useful? 59
Protocol stacks 60
Basics of routing 62
Hardware-based routing 63
Cut-through switching 65
Label switching 67
The future of layer-three switching 69

VII. ATM Switching 71
A (very) brief history of ATM 71
ISDN and BISDN 71
Where is ATM useful? 73
Cells and frames 75
Virtual circuits 76

Table of Contents

Quality of Service 78
Physical interfaces 79
Network interface protocols 81
Interswitch protocols 82
Inside ATM switch 84
Traffic management 87
Making connections 87
How does MPOA work? 91
The future of ATM switching 93

VIII. Gigabit Ethernet 95

Where is Gigabit Ethernet useful? 95
Physical layer 97
MAC layer 98
Gigabit and layer-three switching 98
Gigabit and QoS 98
Gigabit and ATM 99
Cost 99
Familiarity 100
The future of Gigabit switching 101

IX. Advanced Services in Switched Networks 103

Intelligent fabrics: why switching and services are linked 103
Authentication services 103
Firewall services 104
Mobility services 104
Address services 105
Address translation 105
Directory services 106
Priorization ad QoS service 107

X. Switching and Wide Area Networking 109

XI. Management 117

Background issues 117
Standard management software platforms 118
Management of switched networks 119

XII. About Xylan 123
Glossary 141
Index 183
Acknowledgements 193
Reply 195

Why is switching important?

Very few areas of technology change as rapidly as networking. Almost every new capability in computing has an impact on networks. And networking technology itself grows and changes at an extraordinary rate.

In the last decade, networking has exploded, from a tool used by experts within organizations to an ordinary part of many people's lives, both at work and at home. The most obvious example today is the Internet, and its graphical component, the World Wide Web. Schoolchildren, marketeers, and political leaders all use the Web on a daily basis. But other vast, high-speed networks are equally critical to the day-to-day functioning of any advanced society. Most payments for goods and services flow across networks. Doctors schedule office visits and track our health history on networks. Networks send naval ships to distant seas, and taxicabs to front doors.

Underlying the explosion in network use is a powerful set of infrastructure technologies. It would be impossible for the Internet – and the tens of thousands of other networks in the world – to operate today with the infrastructure of ten years ago.

One important set of infrastructure technologies is generally referred to as *switching*. The term has been applied to a very wide range of standards and techniques, but there are some common elements.

- Switching delivers much higher data rates than earlier devices, such as hubs, bridges, and routers.
- Every machine connected to a switch has its own dedicated connection.
- With a few exceptions, switches can generally connect to devices operating at different rates.
- All switching technologies share a focus on the use of hardware to move information. Although hardware in switching is generally supplemented by software, it is central to the basic goal of moving information as quickly as possible.

If you watch old movies, or possibly if you have lived in a very remote rural area, you may be familiar with party lines. A party line is a telephone circuit that's shared by a number of subscribers. Just as with Ethernet and token ring, only one person can use it at a time; the others have to wait until the first subscriber is finished in order to take their turns. And just as with Ethernet and token ring, it is very easy to listen to someone else's conversation.

Party lines were once the normal method of connecting subscribers, but have now been almost completely replaced by dedicated connections. As private (switched) connections to the public telephone network became less expensive, this was inevitable. It's equally inevitable, as private (switched) connections to campus networks become less expensive, that they will replace today's shared-bandwidth networks; bandwidth, security, and quality of service guarantees will all increase automatically. What will be left for a long time will be the residue of these networks: Ethernet and token ring drivers and interfaces. But the bandwidth-sharing capabilities they were initially designed to provide will lie unused.

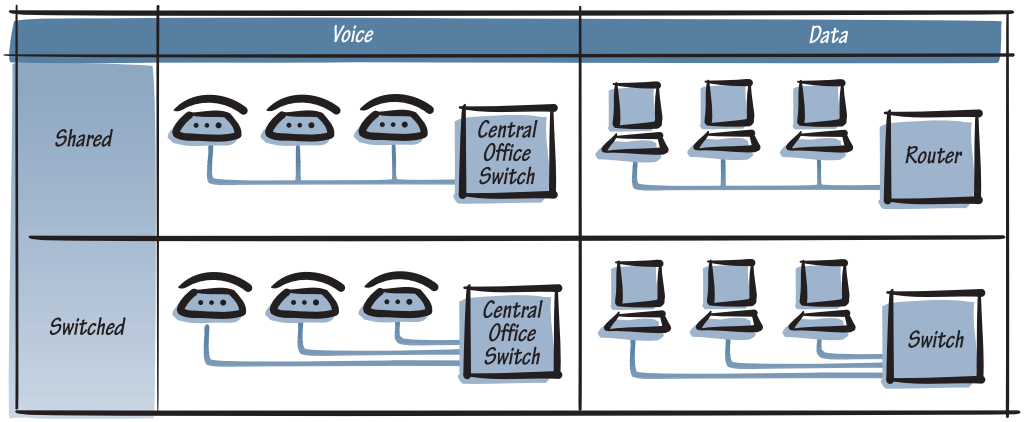


Figure 1.1. Shared-bandwidth and fully-switched network models

Here's another metaphor. Fifty years ago when automotive traffic moved from one city to another it did so on relatively narrow highways. Every time a car passed through a city, a town, or even a village, it would encounter traffic lights, stop signs, railroad crossings, local congestion, and complex intra-city streets. As traffic loads increased and inter-city travel became more important, this system became a major problem for efficient traffic flow. Now, of course, inter-urban superhighways create a mesh network that spans the developed countries of the world.

Today's complex, relatively slow campus router backbones are straining under the load that's being placed on them by Web servers and other graphical applications. The replacement of these complex, node-by-node routed systems with high-speed switching backbones is already under way.

In short, the transition from hubs and routers to switched networks is driven by historical inevitability. The rate of the change will be based on the cost of the new infrastructure. This cost is essentially governed by chip sets; these costs are dropping at a prodigious rate.

Notes to the reader

How to use this book

The Switching Book is intended for use in two ways.

- If you are new to networking, or to switching technologies, then you might want to read the textual chapters from beginning to end. The sequence is meant to provide basic background for people who may be new to our industry, or who need a refresher on some of the basics.
- If you are already knowledgeable about some of what is covered here, just dip in wherever you like. We've included a detailed table of contents to make that as easy as possible.

Scope

This book focuses on switching within a customer's premises: a suite of offices, a building, a campus. Both end-user wide area communications and carrier networks use switching in important ways. We refer to these areas a number of times, but do not cover them in detail.

This is not an intensive treatise on any of the topics covered. To explore in detail ATM, LAN switching, layer-three switching, Gigabit Ethernet, and related topics would require thousands of pages, supplemented on an almost weekly basis. Instead, this is meant to be a gentle, concise introduction to the topic as a whole. Obviously, there are much more complete references available on many of the topics covered here; readers who wish more detailed explanations are referred to the Suggestions for Further Reading at the back of the book.

In some cases we have simplified in order to improve clarity. For example, we express the data rates for standard interfaces without a long string of decimals (for example, we say that E1 operates at 2 Mbps, when the precise number is 2.048 Mbps).

Also, except for a brief "About Xylan" section, this is not a sales guide to Xylan products. For more information on Xylan or its products, please call us or visit our Web site.

Terminology

- We have included a glossary at the back, so we only define some terms in the text.
- We use the term "frame" to refer to a layer-two protocol entity, and "packet" to refer to a layer-three protocol entity. This seems to us to be more precise. However, it's quite common to talk about Ethernet "packets", and the reader should expect to see this in some other publications.
- We sometimes use the term "campus" to mean a single user premise, which can be a few offices, a building, or a large facility with many buildings. This is a standard usage in networking, although not in everyday English.
- We refer to Internet protocols by their IETF (Internet Engineering Task Force) RFC (Request for Comment) number (e.g., RFC 1483). Technically speaking, these protocols are authorized by the IAB (Internet Activities Board), rather than the IETF itself.
- To simplify terminology, we sometimes use North American nomenclature. So, for example, we use "OC-3" rather than "STM-1". Although the two are similar, they are not identical, and it is important for those outside North America to verify that prospective vendors support the international standards.

Timeliness and accuracy

This book is being completed in early 1998. We've tried very hard to make sure that all of our statements are accurate. If you disagree with anything we say here, we would love to hear from you. There's a reader response card at the back of the book. Or call or e-mail us.

Also, networking technologies, and the networking industry, change very rapidly, and some sections of this book will need to be supplemented as time goes on.

In order to understand switching, one needs to know a little about the history of networking.

Era I: mainframe networks

The first computers were very large, very expensive, and required a great deal of human support. Consequently, these mainframe computers were used primarily by large organizations, and in almost all cases they were shared by a number of people. Because mainframes are uniquely well suited to certain massive computing tasks, many are still in use today, in organizations that also use large numbers of smaller computers.

Continuing issue. In a traditional mainframe network the intelligence was centered in the mainframe computer and in closely related computers called front-end processors. The network infrastructure had very little intelligence, and the desktop machines had almost none. As we will see, the balance of power changed a great deal as other types of computing emerged.

Continuing issue. Early mainframe networks made use of what we will call the shared-bandwidth model. In this model a number of machines take turns with a given piece of bandwidth. This can be thought of as a form of time-division multiplexing. In mainframe networking the workstations, and / or the controllers to which the workstations are connected, often shared local cables and wide-area telephone circuits.

The use of shared-bandwidth has changed significantly in the different types of computing, largely due to the impact of new technologies on cost and performance.

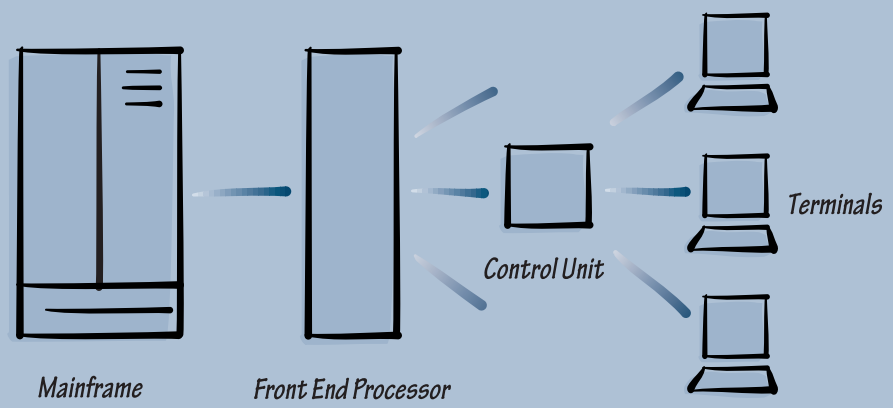


Figure 2.1. Mainframe network model

Era II: minicomputer networks

Minicomputers are similar to mainframes, but are smaller, less expensive, and use different computing technologies. Minicomputer networks were similar to mainframe networks in some important ways: one or several large computers sat at the center of the network, with terminals feeding into them. Most minicomputers are now phased out, because servers based on microprocessors have become powerful enough to take on their tasks.

Minicomputer networks had a continuing, and powerful, effect on networking.

Continuing issue. The network infrastructure acquired substantially more intelligence. As minicomputers emerged, so did the first microprocessors. Although these were too weak to be of much general-purpose use in computing, they were very adept at focused communications tasks. Statistical multiplexers, which connected minicomputers to remote terminals, were based on microprocessors. So were data PBXs, which connected terminals to multiple minicomputers, and which allowed a large number of terminals to contend for a limited number of minicomputer ports. However, the desktops were still "dumb terminals".

Continuing issue. Switching was used extensively for the first time, primarily in data PBXs, but also in some statistical multiplexers. Switching made it easier to manage these networks, and provided much greater aggregate bandwidth.

Continuing issue. In the later part of the minicomputer period, an important new technology was used to link minicomputers together, and to provide terminal access to them. Local area networks allowed terminal servers to connect to hosts at very high speeds; Ethernet at 10 Mbps was typical. The terminals still connected to these servers in traditional, slower ways. But the infrastructure which was to drive PC networks began to develop.

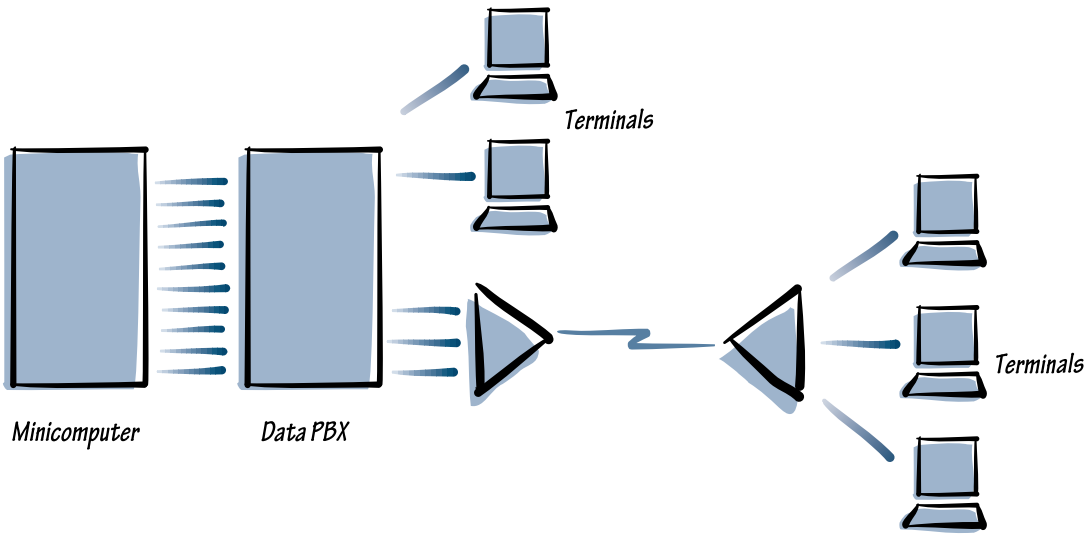


Figure 2.2. Minicomputer network model

Era III: early PC networks

It can reasonably be said that there have been three great inventions in the history of computing: the first stored-program computers; the personal computer; and the Web. PCs (we will use this term to mean all desktop computers) changed computing completely, in ways that are too obvious to describe here. What interests us specifically is the way in which the PC changed networking.

Early PC networks were generally divorced from the mainframe and minicomputer environments that operated in parallel in the same organizations. Typically, these networks were used in small components of an organization: the chemistry department at a university; the marketing group at a manufacturer; a research team at a government department. So the network technologies that developed tended to be small in scope.

Continuing issue. Now the desktops were intelligent, too. Communications were no longer mediated by intelligent devices supporting dumb terminals. Instead, each PC ran a set of communications protocols called a protocol stack; this provided a complete set of services for moving information between machines. Oddly, the network infrastructure tended to become "dumb" again. The intelligence in the PCs was sufficient for workgroup communications, and in many cases the entire network infrastructure was nothing more than a piece of cable.

Continuing issue. Virtually all of the early PC networks relied on the shared-bandwidth model. This was driven by the mission of the PC networks: to move files quickly between machines, and to allow access to shared resources such as printers. File transfer needs significant bandwidth, but only occurs intermittently, so taking turns with a high-speed pipe works quite well. To moderate access to the shared medium, a protocol layer called the Media Access Control (MAC) layer is used.

Ethernet

The most important LAN type to develop in this period was Ethernet, which was invented at Xerox's Palo Alto Research Center. Ethernet is a very simple mechanism. Every computer is connected to a cable. If a computer wants to transmit, it checks to see if the wire is in use. If not, it transmits; if it is, it waits and checks again a little later. Occasionally, two computers will try to talk at almost exactly the same moment. When the transmissions bump into each other on the wire, there's a collision. Both computers sense this, abort their transmissions, and wait for a partially random time before trying again.

Ethernet works best with a small number of machines that need to send bursts of information at long intervals. As the number of machines grows, and as the computers need to send more frequently, the collision process slows down effective throughput.

The original form of Ethernet – and still the most common – works at 10 Mbps. Fast Ethernet and Gigabit Ethernet, which we'll discuss later, work at 100 Mbps and 1000 Mbps.

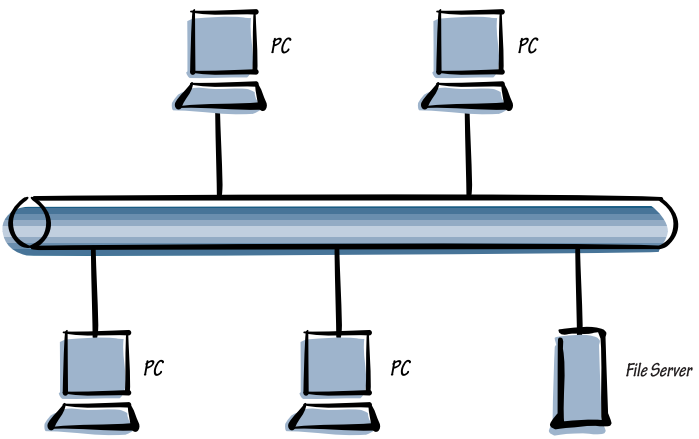


Figure 2.3. Ethernet LAN model

Token ring

Token ring has long been the principal competitor to Ethernet. It uses a more elaborate MAC than Ethernet, in order to provide more efficient use of the bandwidth, and fairer access. Token ring is commonly associated with IBM, which made extensive use of it.

In token ring all computers are connected in a ring. When a computer wants to transmit it waits for a special data pattern, called a token, to come around the ring in an available state. It then attaches its data to the token and sends it onto the ring. When the receiving station receives the information it sets a flag to note that it has copied it. The originating station removes the data from the token, sending the token back onto the ring in an available state.

There are no collisions in token ring; a ring can carry virtually 100% of its bandwidth in actual data. And, since access to the token passes around the ring, each computer has equal ability over time to transmit. The drawback of the mechanism is cost. Since the token ring MAC is powerful and complex, it's expensive to implement in chips; as a result, token ring prices have typically been at least 100% greater than Ethernet prices.

The first token ring LANs operated at 4 Mbps; a later version at 16 Mbps is now more common.

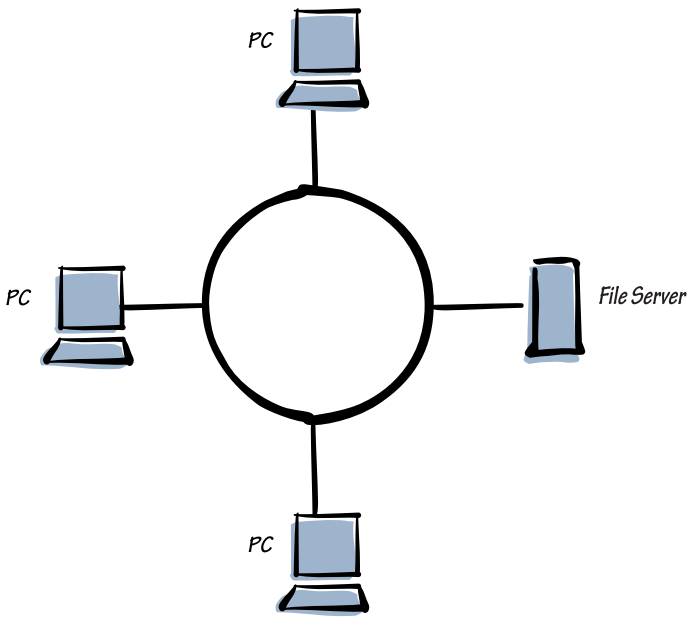


Figure 2.4. Token ring LAN model

Other early LANs

A number of other alternatives received substantial use in the early days of PC-based LANs. Apple Computer provided a capability called LocalTalk with its Macintosh computers. This was similar to Ethernet, but was much slower. However, since it was built into the PC, it was virtually free, and it was easy to use. Virtually all LocalTalk has now been phased out in favor of Ethernet or token ring.

Another protocol, developed by Datapoint, was called ARCnet. This was an early form of token ring, and is now almost completely gone.

It's interesting to note that acceptance and widespread use of computer network protocols is not always technically rational. As with other technology competitions (for example, Betamax vs. VHS) there is a feedback effect: as more users select a technology, more vendors implement it, resulting in more users selecting it. An early agreement among important vendors often serves as a "seed" to begin a process of acceptance, even when a technically superior alternative exists. There is wisdom in this process. For a user, it's not only important for a technology to move information quickly, or fairly, or inexpensively. It also needs to be supported by a wide range of vendors, so that network elements will work together, and so that users can force price competition among manufacturers.

Cabling options

The first LANs often used unique cabling types. Ethernet operated over thick coaxial cable (10Base5, because it could extend 500 meters) and thin coaxial cable (10Base2, because it could extend 185 meters). The cable was configured as a physical bus; the cable snaked around the building, and each computer connected at its nearest drop. Token ring operated over a special kind of shielded twisted pair cable, as well as over a more typical unshielded twisted pair.

Bridges

When there is too much traffic for a LAN to support, users start to experience slow response times. In an extreme case, the network will periodically lock up. In order to avoid this, the network must be segmented.

An important early device for segmenting a network was (and still is) the *MAC-layer bridge*. A bridge connects to two or more LANs. It observes the traffic on each, and automatically learns the MAC-layer addresses of the devices attached to the LAN. When it sees a frame on LAN A with a destination address that is on that LAN, it does nothing. But when it sees a frame on LAN A with a destination address that is not on that LAN, it copies and forwards it. If it knows the location of the destination address (for example, LAN C) it sends the frame only to that LAN. But if it does not, then it sends it to all attached LANs; this is called flooding.

There are special types of frames called *broadcast* and *multicast* frames. These are sent to all stations (broadcast) or to a group of stations (multicast). There's no way for a bridge to know which stations should receive a given broadcast or multicast, so it copies and forwards them to all LANs to which it is connected. This is also a flooding process. As we will see, this is one of the significant differences between bridges and routers.

Bridges have topological limitations as well. Imagine two bridges, with two Ethernet connections between them. When one bridge observes a broadcast on any of its ports, it will forward it out on all other ports – including the ports that go to the other bridge. When the other bridge sees those broadcast frames, it will then copy and broadcast them on all links – including the other link to the first bridge. This will continue endlessly. Bridged networks cannot have loops. To avoid this, an IEEE protocol known as *802.1d Spanning Tree* is used. It runs in the software of the bridges, and detects loops. The bridges reconfigure themselves into a topology with no loops: a tree. If a link fails, the bridges will, if possible, reconfigure themselves to use the links that had been held idle.

Spanning Tree is effective, but it does not allow spare bandwidth in the network to be used very efficiently. And creating a new tree topology in response to a problem in the network can take a relatively long time.

Bridges have another useful function. They filter bad frames: those which are too long, too short, don't meet the rules of the protocol, or in which a transmission error is detected. A bridge prevents such frames from being transmitted to, and possibly causing problems for, other devices on the LAN.

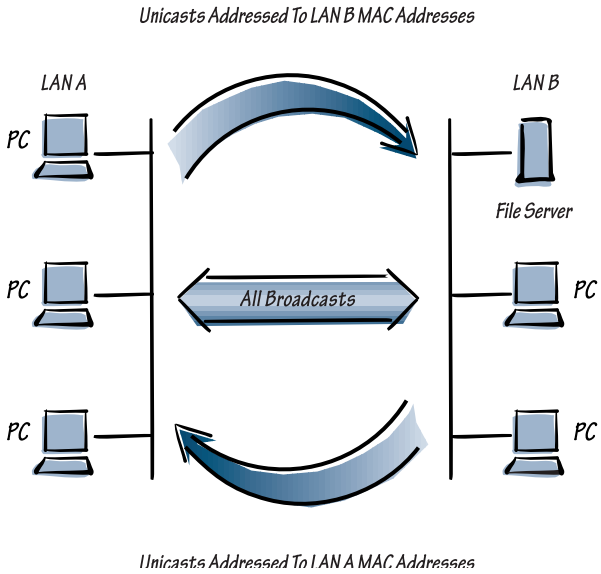


Figure 2.5. How a bridge forwards traffic

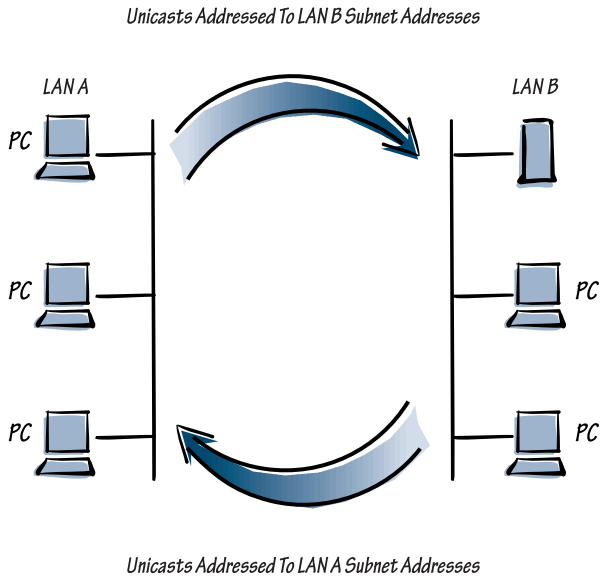


Figure 2.6. How a router forwards traffic

Era IV: Enterprise PC networks

As PCs became more powerful, they began to take a larger role in the total computing structure of an organization. MIS departments began to take control of PCs and the networks that served them and undertook the difficult task of standardizing and interconnecting them. PC networks now moved into their next major phase.

Most traffic still flowed within a workgroup. The applications had not fundamentally changed; people in a department, working on the same files and applications, would share a file server. Some functions – such as e-mail and mainframe access – required broader connectivity. But in general the workgroup model was still in place.

"Intelligent" hubs

As networks extended beyond the workgroup, and as more users were supported, early LAN wiring systems became increasingly ineffective.

Comprehensive building cabling systems have been in place for a long time to support the telephone network. They use unshielded twisted pair cable, arranged in a physical tree topology. Wire pairs extend from each office or cubicle to a wiring closet (this is often called "horizontal" or "station" cabling). Wiring then runs from each wiring closet to a central point in the building (this is often called "riser" cabling). Structured cabling is a convenient way to install and maintain cable plant, and it was natural that data networks would use it.

To make this work with Ethernet, a new wiring alternative known as 10BaseT was developed. This operates over unshielded twisted pair, at distances of up to 100 meters. Rather than connecting the devices in a physical bus, 10BaseT uses a tree topology. Each computer is connected, over two dedicated pairs of wire, to a hub, usually located in a wiring closet. The hubs are then interconnected. Early hubs were purely repeaters; they regenerated the signal from one set of wires to another, but did not process or manage it in any way.

But hubs quickly took on three other roles:

- They were a convenient place to install bridges.
- Vendors added microprocessor-based management modules, which could enable and disable ports, and observe and report on traffic flows.

- The distances between hubs often exceeded the 10BaseT 100-meter limit, and fiber optic cable was a superior alternative to coax for use in building risers. Hubs began to convert between cable types, typically supporting unshielded twisted pair to the computers, and fiber optic cable (10BaseFL) to other hubs.

Routers

Bridges are simple, easy to use, and inexpensive. But, as noted above, they must flood a certain number of frames throughout the network. In a small, lightly-loaded LAN this is rarely a problem. But there are circumstances in which flooding becomes a problem:

- The total number of flooded frames is more or less linear to the number of machines connected to a LAN. As the LAN becomes very large, flooding becomes a problem, because each machine receives every flooded frame.
- Wide area connections tend to operate at much slower rates than LANs. In the United States and Canada, 56 Kbps is common; elsewhere, 64 Kbps is common. Flooding can much more quickly affect such a link. For example, flooded frames which occupy an insignificant 1% of an Ethernet would more than fill a 56 / 64 Kbps circuit.
- Sometimes a device, due to misconfiguration, hardware failure, or a software bug, will generate a large number of erroneous frames. If these are valid frames with invalid addresses, they may be flooded, causing what is sometimes called a broadcast storm. Until this situation is corrected it can bring down the entire network.

Routers provided an alternative to bridges. While a bridge operates at the MAC layer, a router operates at the network layer, the next level up in a protocol stack. In fact, network-layer protocols are based around routers. Unlike a bridge, a router forwards data based on layer-three addresses, rather than MAC-layer (layer-two) addresses. Rather than simply forwarding broadcast and multicast frames, a router responds to them. (Note that we are talking here about layer-two broadcasts and multicasts. There are similar packets at layer three, which a router forwards.)

Routers solve the problem of too much flooding in a large or heavily loaded network, and over low-speed wide area connections. And they provide firewalls within a network, ensuring that a broadcast storm will only affect one area of the network. As a result, routers rapidly became the standard tool for interconnecting hubs, and for connecting sites together across wide area networks.

Routers are very intelligent devices; a large router contains a number of high-speed processors. Routers have traditionally needed a lot of processing power, for two reasons. It used to be common for networks to support several protocol stacks, each of which a router needed to execute; and using software for packet forwarding required substantial power.

However, the ratio of routers to workstations is generally very low; one router supports many hubs, and one hub supports many workstations. Although the advent of routers increased the intelligence in the network infrastructure, it did so in a limited way, and only in the center of the network. And most of that intelligence was used up in packet forwarding.

FDDI

Ethernet and token ring were initially used to interconnect hubs, as well as for workstation connections. But in some networks the backbone needed more throughput, and a new, higher-speed alternative was needed. FDDI (Fiber Distributed Data Interface) was developed to serve this need. As the name implies, it typically runs on fiber, although a twisted pair version is also available. FDDI is essentially token ring running at 100 Mbps, with some additional enhancements to support throughput and reliability in large rings. Notably, it is a dual counter-rotating ring. When an FDDI ring breaks, the remaining ring segments wrap, keeping traffic flowing.

Fast Ethernet

Another high-speed alternative to Ethernet and token ring is Fast Ethernet. This is basically Ethernet running at 100 Mbps instead of 10 Mbps. There are copper and fiber versions, with distance specifications similar to those for Ethernet. Although Fast Ethernet was originally used to interconnect hubs and switches, it is now increasingly common as a workstation interface. Many network interface cards support both 10 Mbps and 100 Mbps Ethernet, automatically operating at the appropriate rate.

Early LAN switches

We've made it to switches!

Routers are very powerful, but they are also very expensive to build. In the early 1990s several small companies started to make alternatives to routers which they called LAN switches. These were essentially multiport bridges, which used the standard Spanning Tree protocol. They supported a number of Ethernet ports, and some also supported a limited number of FDDI uplink ports.

One exception to this description is that some early switches used a technique called cut-through to reduce latency. Normally, a bridge (or a switch acting as a bridge) can begin to transmit a frame out the destination port only after the entire frame has been received at the input port. This permits it to check for errored frames. In a cut-through switch, the switch starts to transmit the frame out the destination port as soon as the destination address portion has been received at the input port, or as soon as the header is received completely. This technique reduces latency (delay).

However, cut-through has two serious drawbacks. It prevents the switch from completely checking for invalid frames. And, more importantly, it forces all ports to operate at the same rate. This means that an Ethernet switch can't have Fast Ethernet or FDDI uplinks; there can be no gaps in a frame's transmission, which is inevitable when going from a low-speed port to a high-speed port, unless the entire frame has been buffered.

Note that later in this book we're going to use the term "cut-through" again, in a completely different sense (related to moving traffic across an ATM network).

Early LAN switches were very expensive. As a result, they were largely confined to very specific applications, and to a minor role as router alternatives. As we will see, this has changed.

Standards

Public standards are now crucial to networking. Understanding the activities of the IETF, IEEE, ITU, ATM Forum, Frame Relay Forum, and other standards bodies is important for anyone who wishes to work in the networking field.

Much of this emphasis on standards emerged in the two eras of PC and LAN networking. In the mainframe and minicomputer periods it was common to obtain most or all of a network's components from the manufacturer of the central computer. But the new LAN technologies came primarily from new suppliers, who were generally too small to offer every element of the network. So users needed to integrate products – hubs, routers, network interface cards, wide area transmission, servers, network operating systems – from multiple suppliers. This required standards, and both vendors and users moved rapidly to create standards. Although there were significant exceptions, such as Cisco's IGRP routing protocol, the time when a vendor could successfully create an entire protocol structure, as IBM had done with SNA, were over.

Standards Body	Networking Standards Coverage	Examples
ATM Forum	Anything Related To ATM	ATM LAN Emulation PNNI
BellCore	Safety And Other Carrier Concerns	NEBS, SONET
EC (European Commission)	Electric Emissions	CE-Mark
Frame Relay Forum	Anything Related To Frame Relay	FRF.5, FRF.8, FRF.9
IAB / IETF	Anything Related To TCP/IP	TCP, SNMP
IEEE	LAN Physical And MAC-Layer Specifications	802.3, 802.5
ITU (formerly CCITT)	Wide Area Networking	OC-3, V.35, ISDN, SDH
UL	Electrical Safety	UL 1950

Table 2.1. Some sources of networking standards

Era V: Intelligent Fabric Networks

We are now at the beginning of a major shift in campus networking, which will continue for some years, and which will result in the complete replacement of all current networking infrastructure: network interface cards, routers, hubs, driver software, lower-layer protocols – even most of the early switches that have been installed so far.

By the time this networking revolution is complete, campus networks will be built from a single "intelligent fabric". It will:

- provide automatic lower-layer translation, making it easy to mix MAC-layer technologies and to migrate from one to another
- heal itself when failures occur, automatically shifting to backup components, links, and nodes, to keep traffic flowing without interruption while repairs are made
- combine today's separate data, voice, and video networks, automatically providing each type of information with the delay characteristics and priority that it needs
- simplify and unify the management of user access to resources
- scale to almost any size and almost any data rate with ease
- support globalized computing, with any user in an organization able to access appropriate resources, regardless of the user's location
- deliver a wide range of services, including: security firewalls; user authentication and resource access control; network directories; automatic address management; automatic compression across wide area circuits; and others

What's an intelligent fabric?

Smart workstations, smart network

In mainframe and minicomputer networks the terminals were dumb and the network, beyond the front-end processors, was dumb. All the intelligence resided in the host computers. With the appearance of personal computers this changed – intelligent desktops and servers shared the task of managing the protocols. But with the exception of a relatively small number of routers and a few management modules in the hubs, the network itself was dumb. It was just too expensive to broadly implement a high degree of intelligence in the network infrastructure.

Workstations are continuing to become more powerful. Processors, memory, hard drives, and other basic components are following Moore's Law and increasing in power steadily. And dedicated coprocessors are driving graphics, sound, video, and other functions.

But the cost of adding significant power to the network infrastructure itself has changed, and will change even faster in the next few years. *ASICs* (application-specific integrated circuits – chips designed by a manufacturer for its own products), commercially available switching chips, and inexpensive *RISC* processors are combining to make the network itself, not just the workstations and servers, highly intelligent. To take one example, a single nine-slot Xylan OmniSwitch can contain 37 RISC processors and 92 ASICs. This provides the power to build an entirely new type of network.

Availability and quality of service

When you pick up a telephone you expect it to automatically and rapidly provide a connection with the characteristics required for a voice conversation. It's very unusual to get a trunk (fast busy) blockage, and almost unknown to get no dial tone. The network does what it needs to do when it's needed. Of course, this is a much more difficult task in a multimedia network that needs to support bursty graphical data, constant bit rate voice, and multicast, high-variability video. An intelligent fabric is needed to negotiate the requirements of each traffic stream, rapidly set up the bandwidth in the network, and reliably move the data.

Scalability

What's the largest network in the world? Most data people would say "the Internet". Of course, the correct answer is the public switched telephone network. Hierarchical switching allows networks of almost any size to be built.

Integration and migration

Today's workstations use Ethernet and token ring. Tomorrow's workstations will use Fast Ethernet and ATM.

Today's servers and backbones use FDDI, Fast Ethernet, and OC-3 ATM. Tomorrow's servers and backbones will use OC-12, OC-48, and Gigabit Ethernet.

All of these technologies make sense in the right application; almost every network can optimize performance and cost by combining them. And most network managers will continue to transition to new technologies as standards mature, prices drop, and traffic increases. An intelligent fabric makes it easy to mix any combination of technologies – Ethernet, token ring, FDDI, CDDI, Fast Ethernet, Gigabit Ethernet, 25M ATM, OC-3 ATM, and OC-12 ATM – in a single network mesh, with data moving rapidly and automatically between the different layer-two standards.

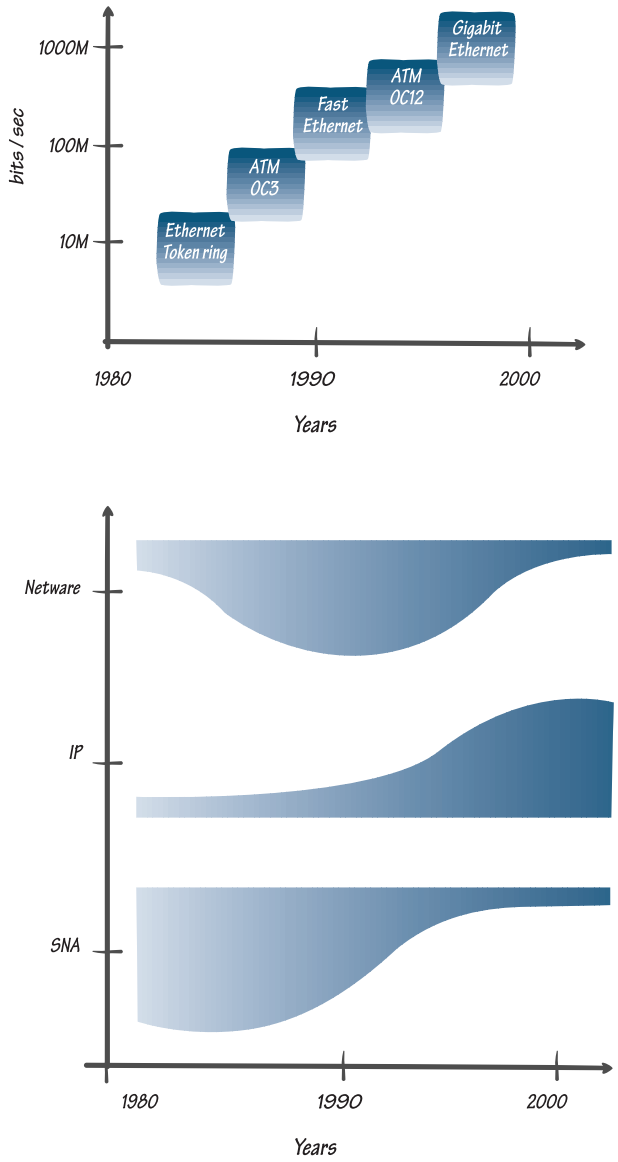


Figure 2.7. Networking technology is shifting rapidly

Self-healing

In an Ethernet network an errant workstation that generates a giant frame can bring down every machine on a segment, over and over again. A token ring workstation with a loose wire can do the same thing by beaconing. And the cost of routers has led many managers to build star, rather than mesh, topologies. As a result, we have become accustomed to periodic failure of portions of the data network. Ironically, average network availability is lower now than in the days of mainframe or minicomputer networks – at the same time that organizations have become increasingly dependent on their networks.

With an intelligent fabric network one, two, or many alternate paths can back up every inter-switch path. Combining this topological redundancy with redundant components in each switch – logic, power, and cooling – results in a network that can recover automatically from failure, with little or no impact on users.

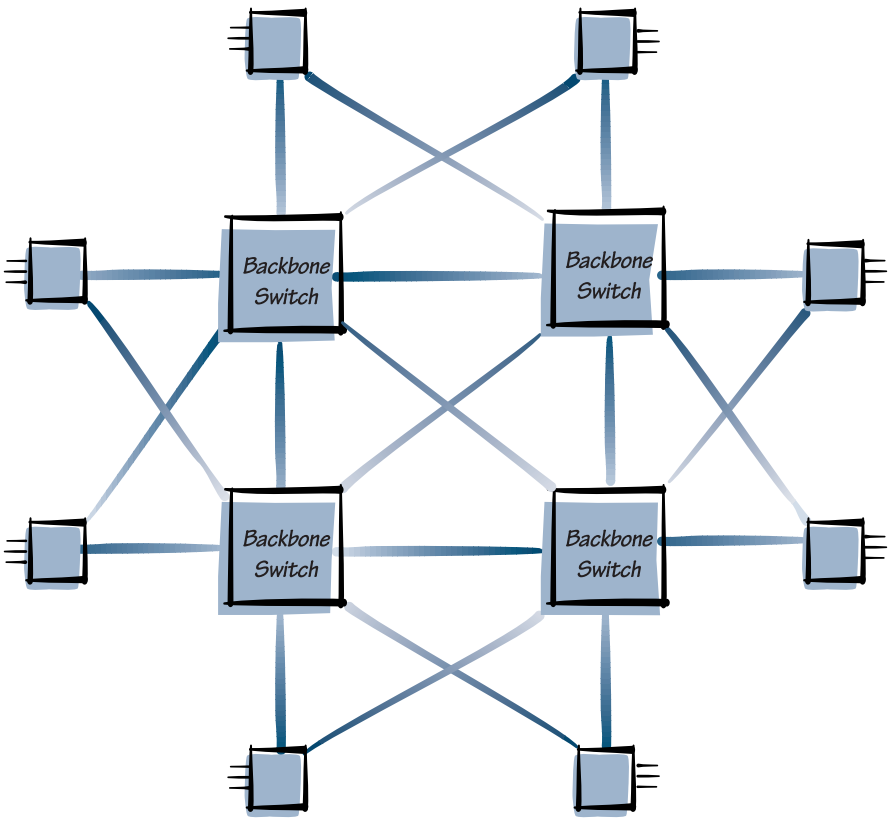


Figure 2.8. Intelligent fabric networks enable powerful mesh topologies

Security

It used to be true that network security was like the weather; everyone talked about it but no one did anything about it. The Internet has changed that. Every workstation needs access to the Internet, and every Internet connection is an invitation to attack by mischievous or malicious intruders.

The problem is compounded by the need for internal users to access many resources, and by the way modern users move around a campus. All of this access must be controlled.

At a minimal level, as campus network switches take their place as general internetworking platforms, they will need to incorporate powerful IP firewalls. But beyond that, an intelligent fabric can monitor and control access to resources flexibly, and on a network-wide basis. With network-wide user authentication, the switching network can provide network managers with a valuable new tool for controlling resource access across the entire network, based on human, rather than machine, identities.

Switching Today

III.

Current campus networks are in a state of transition. The majority of workstations are still connected to hubs, and the majority of those hubs are still interconnected with routers. But:

- A large minority of workstations is now connected to LAN switches, and many of those are interconnected with ATM, or with backbone switches using Fast Ethernet. LAN switches alone are now a multi-billion dollar business, and they are one of the fastest-growing segments of the networking industry.
- Gigabit Ethernet is about to become a widely accepted tool for campus backbones.
- The number of campuses based around ATM switches is growing rapidly.
- Layer-three switching – inside Gigabit Ethernet and ATM switches – is on the verge of replacing software-based routers in many networks.
- Intelligent fabric networks are beginning to be implemented, delivering advanced services to corporations, government institutions, universities, and others.

The remainder of this book will look in some detail at each of the new technologies.

LAN switches have evolved a great deal since the first LAN switches, which we described earlier. They have become a fundamental building block of modern networks.

Where are LAN switches useful?

LAN switches interconnecting hubs

The first use of switches, as we noted above, was as an alternative to routers, for interconnecting hubs. This is still a useful role for LAN switches, although switches are now often replacing hubs, rather than interconnecting them.

Routers tend to be very complex to configure and manage. This is natural, given the power that they provide, and the number of protocols that they need to support. A LAN switch, on the other hand, is very simple; in some cases all that the network manager needs to do is to plug it in. So in a smaller network it is often quite reasonable for a LAN switch alone to handle the movement of data between hubs.

As a network becomes larger it needs the broadcast handling capabilities that are provided through routing. However, it's important to differentiate between *routers* as a product and *routing* as a function. Routing can be provided in a number of platforms. In fact, one of the trends in current networking products is to combine routing and LAN switching in a single product.

We will discuss this below in the layer-three switching and Gigabit Ethernet sections.

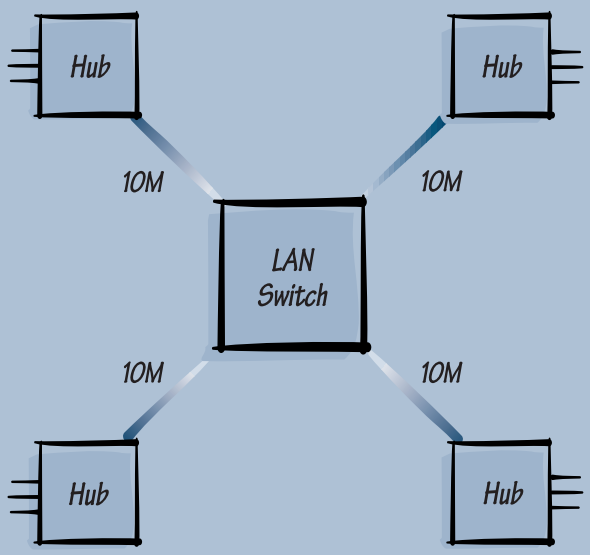


Figure 4.1. LAN switches interconnecting hubs

LAN switches replacing hubs – pricing and bandwidth

LAN switches have become dramatically less expensive since their introduction. Early LAN switches cost about $2,000 per Ethernet port; low-end LAN switches under $100 per Ethernet port are now available. As a result, it has become possible to connect each device – workstation, printer, server – to its own switch port, completely eliminating hubs. In fact, even Fast Ethernet is now sufficiently inexpensive that fully switched 10/100 networks are being installed.

The most obvious advantage of a fully switched network is that each device has its own dedicated bandwidth. Imagine an Ethernet LAN with 50 users. The total theoretical bandwidth in the network is ten Mbps. But a shared Ethernet can't sustain more than a 50% - 70% load, depending on number of machines and cable distances. So the real capacity of the network is five-seven Mbps. Now imagine this as a switched network, with Fast Ethernet to every desktop. The total network capacity is now 500 Mbps – an increase of two orders of magnitude. This compares to the difference between an early IBM PC and a Pentium Pro.

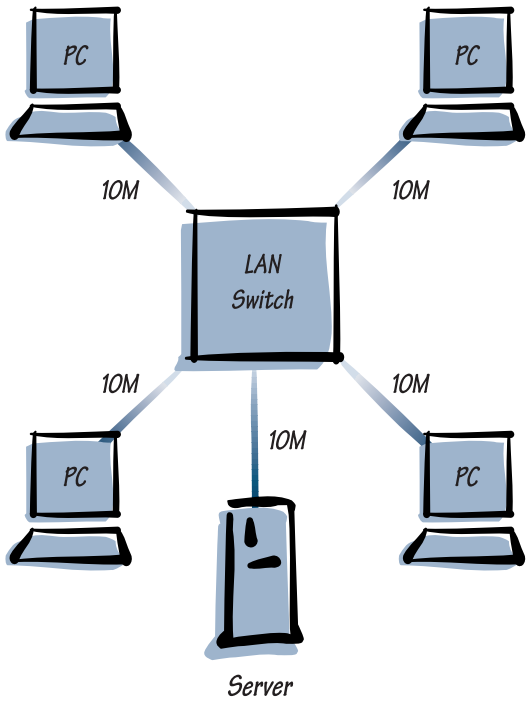


Figure 4.2. LAN switches replacing hubs

LAN switches replacing hubs – rate adaptation

But simply giving dedicated bandwidth to each user is not enough. When a number of users access the same server, the server needs to operate substantially faster than each individual workstation. Think about it this way. Imagine a 16-port Ethernet hub, with 15 workstations and a server connected to it. Obviously, the throughput is limited to the 10 Mbps in the single Ethernet segment. Now imagine that the hub is replaced with a switch with 16 Ethernet ports. Every machine now has its own dedicated ten Mbps of bandwidth. The result? Virtually no change in throughput; the bottleneck has simply shifted from the hub bus to the ten Mbps port to the server.

It's like a freeway system; the freeways need to be wider and faster than the surface streets that feed into them. So rate adaptation is an important part of what LAN switches do. In some cases this is very simple, as when a switch supports both 10 Mbps and 100 Mbps Ethernet, or both 4 Mbps and 16 Mbps token ring. At other times it requires more elaborate conversion, as described below.

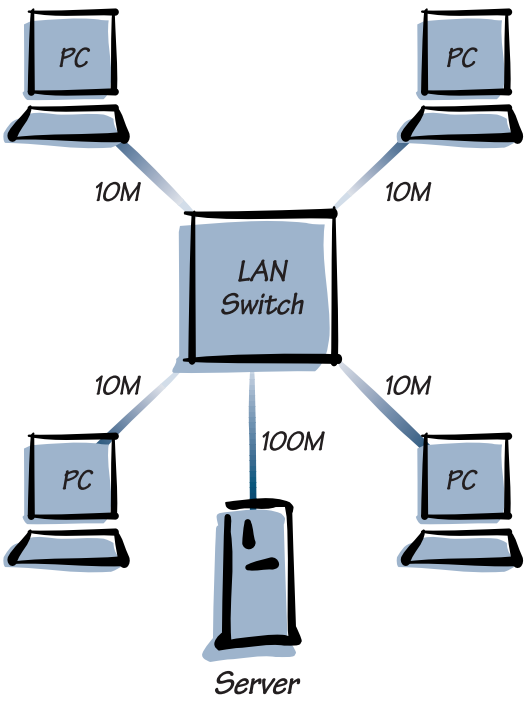


Figure 4.3. LAN switches enabling high-speed servers

LAN switches replacing hubs – access to high-speed backbones

It is not only servers that need to operate at high data rates. In a mid-sized or large network with multiple switches, a backbone is needed to interconnect resources. Rate and format conversion in the switches means that workstations can use inexpensive network interface cards and cabling, while taking advantage of high-speed backbones like ATM and Gigabit Ethernet.

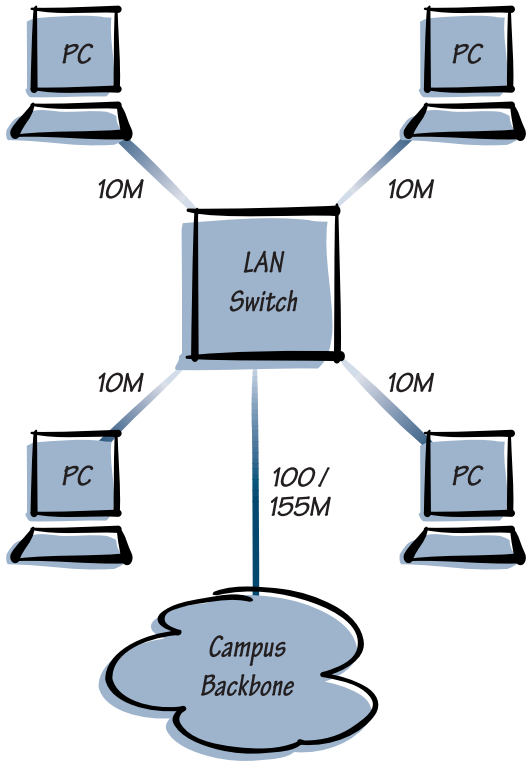


Figure 4.4. LAN switches enabling high-speed backbones

LAN switches replacing hubs – enhanced management

A fully switched network provides another important advantage. In a hub-based network, a variety of problem conditions can be propagated from one machine to another. A token ring station can start to beacon, perhaps because of a loose cable, bringing down the entire ring. An Ethernet station can experience a software failure and send giant frames or runts onto

the network, sometimes crashing other machines. In contrast, a switched network resolves these problems automatically; each switch port examines every frame, and drops all errored frames.

And unlike a hub, a switch inherently looks at every frame, and so is able to provide more information. Switches report to the network manager on traffic flows, errored frames, alarm thresholds exceeded, and other conditions.

LAN switches replacing hubs – security

In a hub-based LAN, all data is available to all stations connected to the segment or ring. Anyone with access to a data jack in an office can monitor and copy all transmissions, learning passwords and layer-three addresses, and observing potentially sensitive information.

This is not possible with a LAN switch. A unicast is sent directly from the originating device to the switch, which then sends it directly to the destination device. Tapping into an unused network outlet would reveal nothing more than an occasional routing protocol broadcast, of no real value to an intruder.

Basics of LAN switching

MAC addresses

Every device on a LAN – a workstation, a server, a router, a printer, etc. – has a MAC-layer address. When a device transmits a frame, it includes a header that contains a source address (its own address), a destination address, and control information. The LAN switch learns the addresses that are present on the LAN by observing source addresses. It transmits the frame out the correct port, based on the destination address.

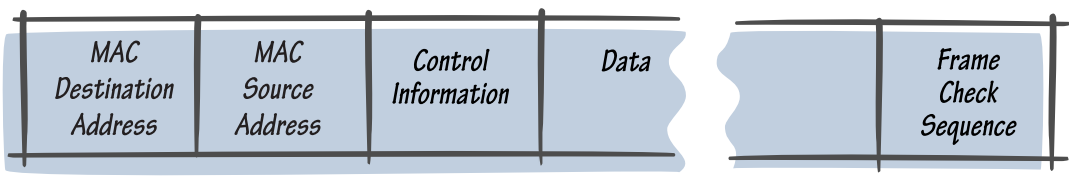


Figure 4.5. Basic MAC frame structure

Broadcast, multicast, and unknown destination forwarding

As noted above in the discussion of bridges (remember that LAN switches are basically bridges), a LAN switch forwards all broadcast frames, all multicast frames, and all unicast frames with unknown destination addresses to all ports, a process known as flooding. This process occurs on all ports in a switch that does not implement virtual LANs, and on selected ports in a switch that does implement them. Virtual LANs are discussed in detail below.

In addition to virtual LANs, various techniques have been developed to control excessive broadcasts:

- Some switches, without taking on full routing functionality, intercede in certain protocols, such as NetWare's SAP and RIP protocols; this is sometimes called "spoofing". It is especially valuable over low-speed lines, which can easily become loaded with broadcasts advertising route updates or server availability.
- Some switches allow a manager to define a maximum broadcast level, and discard all broadcasts that exceed that threshold. If the level is chosen carefully, it will never be exceeded except during a broadcast storm.

Address cache

LAN switches must store learned MAC-layer addresses, and information associated with them; they typically use a special kind of high-speed memory called CAM (Content-Addressable Memory). This address store is referred to as a *cache* or *forwarding table*. If a LAN switch is to interconnect hubs, it may need to learn a large number of addresses, and therefore needs a large cache. But if all the devices in the network are directly switched,
then the cache need only be slightly larger than the number of devices directly connected to the switch.

Chips

Some early LAN switches were based around a high-speed RISC processor. These are almost completely gone now. LAN switches still use processors for other purposes, but the actual switching is performed largely by hardware.

A manufacturer has two basic choices when designing any kind of high-speed switch. They can use commercially available chip sets, some of which are effective and flexible. Or they can design their own ASICs.

Generally, as a technology matures, commercial chip sets are almost as cost-effective as ASICs. The advantage that an ASIC offers is that it allows the vendor to include advanced capabilities that are not available commercially. In a LAN switch, these include VLAN identification and trunking, security access flags, and codes to assist in protocol translation.

An important advance in recent ASICs is the inclusion of one or more RISC processors in the ASIC. This provides an optimal combination of the high speed and low cost of an ASIC, and the ability of a processor to load new code, allowing it to support new functions and protocol standards.

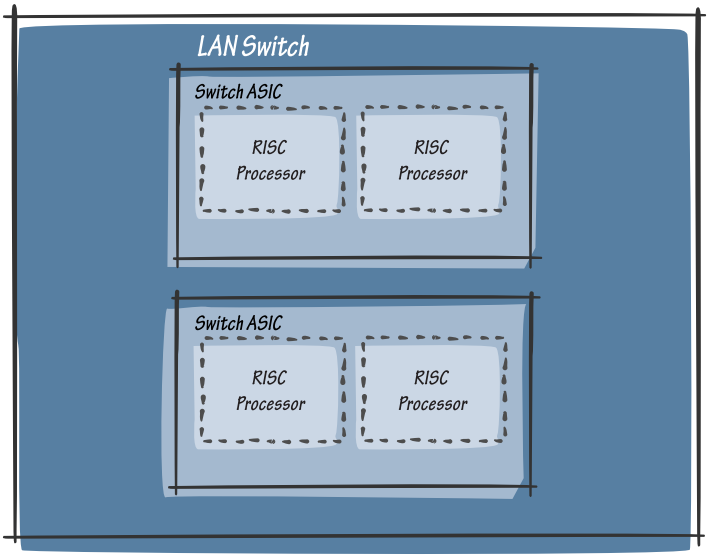


Figure 4.6. New switch architecture: RISC processors embedded in ASICs

Any-to-any (translational) switching

Advantages

As noted above, a switch that can convert from one LAN protocol to another is useful in a number of ways:

- Workstations can operate at one rate, and servers at a higher rate.
- The backbone does not need to use the same protocol as the workstations. For example, an ATM backbone offers substantial advantages of reliability and extensibility. But ATM is expensive as a workstation technology. LAN switches that offer translation can allow the desktops to use Ethernet, token ring, or Fast Ethernet, while the backbone and servers use ATM.
- Workstations and servers that use different LAN protocols (for example, Ethernet and token ring) can communicate with each other. In many organizations a diverse set of LAN technologies has evolved over time; LAN switches can unify the network.
- The network can migrate from one protocol to another. For example, a network with token ring workstations and an FDDI backbone can gradually shift to Ethernet workstations and a Fast Ethernet backbone. Later, the same network can migrate to Fast Ethernet workstations and a Gigabit Ethernet backbone. While these shifts are taking place, all of the workstations can communicate with all of the servers, regardless of which LAN technology they use.

MAC header differences

Each MAC-layer protocol has a different header structure. Switches that translate from one protocol to another must transform the header structure.

A major example of this is the order of bits in the addresses. In Ethernet and Fast Ethernet the bits are ordered using canonical addressing; the first bit transmitted is the least-significant bit. In token ring and FDDI the bits are ordered using non-canonical addressing; the first bit transmitted is the most-significant bit. The bit sequence needs to be reversed when converting from a canonical to a non-canonical protocol.

Frame size differences

Frames must have a maximum size, so devices can allocate appropriate buffer space, and so that multiple devices can get reasonable access to the LAN. Each LAN protocol specifies a maximum frame size. For Ethernet and Fast Ethernet it's 1,518 bytes. For token ring it's 17,800 bytes, although 8,192 bytes is a more typically implemented maximum. For FDDI it's 4,500 bytes.

When converting from one LAN protocol to another, a switch must deal with varying frame sizes. Some protocols can be "fragmented", permitting large frames to be subdivided. In other cases the network manager must configure each device with a "worst-case" maximum, so that over-sized frames are not transmitted.

Token ring switching

The basic functioning of token ring was described above, in the section on the history of networking and switching.

Switching token ring devices

Token ring switches, like Ethernet switches, can be used to replace hubs. In this case each token ring device connects directly to a switch port.

Token ring differs from Ethernet in that devices (such as workstations) and *lobe* ports (on hubs and switches) operate differently from each other. In Ethernet, hub / switch port operation is essentially identical to that of a device. But in token ring, when a device is active, it must send a DC voltage to the lobe port to which it is connected; this tells the hub or switch to open a relay and add the device to the ring.

Just as cost reduction has been the fundamental shift that has permitted Ethernet switches to replace Ethernet hubs, cost has been the barrier to token ring switches replacing token ring hubs. Token ring is very powerful when many token ring devices are sharing a ring. Bandwidth is used very efficiently; every workstation has equal access to the ring; the ring recovers rapidly from certain types of failures; and the protocol tracks valuable management information. But this power also makes token ring very complex, and most of these advantages are meaningless when every token ring device is connected to its own switch port. However, the cost of building powerful, expensive token ring chips must still be carried

forward. Although there have been various proposals aimed at simplifying the functions of token ring in a switched environment, no standard has been approved.

In fact, many token ring users have decided to convert to other LAN options, such as Ethernet, Fast Ethernet, or ATM. The most important reason is switching. Ethernet and Fast Ethernet are inexpensive to switch, and although ATM is even more expensive than token ring, it offers substantial functional advantages in a switched environment.

Switching token ring hubs

For users who want to leverage their token ring investment, but would like to gain some of the advantages of switching, another option is available. Token ring switches can be used to connect token ring hubs to each other.

One significant difference between token ring and Ethernet is in the upper-layer protocol stacks that tend to use them. Token ring, unlike Ethernet, is often used in conjunction with NetBIOS and SNA, protocols which are not routable, and which must be bridged / switched. Many token ring networks have used routers as large bridges. So it's easier for these applications to substitute token ring switches for routers.

There are two ways to connect a switch to a token ring hub (often called a MAU):

- The switch can pretend to be a device, and connect to each hub on a lobe port. In this case, the switch must provide the DC current to keep the hub port open. One advantage of this kind of connection is that the hub automatically recognizes a cable break between the switch and the hub, and the ring within the hub remains viable.
- Or the switch can connect to the hub on the Ring In / Ring Out ports, pretending to be the next hub in the ring. This makes sense in many applications, because the existing network topology places a token ring hub in each wiring closet, with a Ring In / Ring Out connection to another hub in a central location. The switch simply replaces the central hub.

It's interesting to note that there is no public standard for Ring In / Ring Out connections over fiber optic cable. Each manufacturer has come up with its own specification, or has used someone else's as a de facto standard.

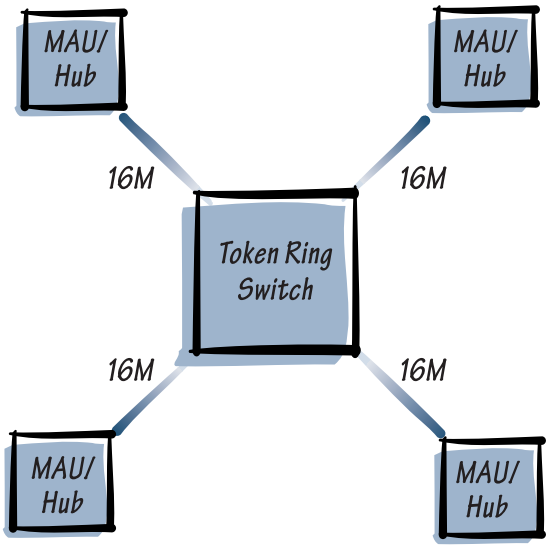


Figure 4.7. Token ring switches interconnecting hubs / MAUs

Token ring source routing

Many, although not all, token ring applications use source routing. This is a MAC-layer protocol that finds an optimal path between two devices, and then uses that path to move information between them. Source routing operates in the end devices and in the bridges / switches. It allows loops to exist in the network, and will find an alternate route when a failure occurs at some point in the network.

A token ring switch implements some combination of four options with regard to source routing:

- ***None.*** Many token ring applications don't use source routing, and can be switched / bridged in the same way as Ethernet. This is referred to as transparent bridging. Routing, such as IP or IPX routing, can also be used.
- ***Source routing.*** Some switches support source routing, but do not support transparent bridging. This is effective in many applications that use SNA or NetBIOS. However, since most organizations now use at least some TCP/IP traffic, pure source routing bridges tend to be limiting.

- ***Source route / transparent without conversion (SRT).*** When some stations on a token ring LAN use source routing, and others do not, and they do not need to communicate with each other, an SRT switch is useful. It examines each frame; if the frame contains a routing information field (RIF), then the switch forwards it using source routing; if it doesn't, then it uses transparent bridging, based on the destination MAC address.

- ***Source route / transparent with conversion (SRTB).*** When source routed and non-source routed stations need to communicate with each other, an SRTB switch is useful. It emulates the source routing protocol and converts from one type to the other.

High-speed token ring switching

As noted above, the token ring protocol is complex and difficult to implement in chips. This is especially true at rates substantially above 16 Mbps. But for networks that wish to continue to use token ring, data rates of 100 Mbps and greater would be very useful for network trunking.

Some manufacturers are working on token ring switching at high data rates; their efforts break down into three alternatives. All share support for large token ring frame sizes and for source routing:

- ***A full token ring MAC at 100 Mbps.*** This would be end-to-end compatible with current token ring technology at lower rates; a workstation running 16 Mbps token ring could connect to a server running 100 Mbps token ring. This is an alternative to using layer-two translation to FDDI, or more complex translation to Fast Ethernet, as a 100 Mbps server interface.

- ***The ability to pass token ring frames across a 100 Mbps trunk.*** A token ring workstation at one end of the trunk could connect to a token ring server at the other end. Or, using layer-two translation, the remote end could be FDDI or Fast Ethernet.

- ***The ability to pass token ring frames across a gigabit trunk.*** This is the same as the second option, except that the trunk would run at 1,000 Mbps. This has the advantage of a high workstation-to-trunk bandwidth ratio, with much less chance of link blockage during heavy loading. Some vendors may simply make it possible for native token ring frames to operate across a standard Gigabit Ethernet trunk. This is beneficial in a network which will eventually migrate to Ethernet or Gigabit Ethernet at the desk, as the same high-speed pipes can be used for both types of frames.

A group of vendors (the High Speed Token Ring Alliance – HSTRA), is, as of this writing, working on specifications for high-speed token ring, which could be added to the IEEE 802.5 standard. It seems likely that products will start to appear in 1998, although a public standard is not yet certain.

Inverse multiplexing

Periodically in the history of networking, inverse multiplexing has been used as a trunking mechanism when the rate of the available trunking technologies is not sufficient to support the applications. Inverse multiplexing bundles multiple connections together into a single virtual link. Traffic is distributed across them, with some degree of load balancing; when one link fails, the others take up its load, without relying on Spanning Tree's slow convergence process. The drawback to inverse multiplexing is that it requires multiple physical ports, and distributing traffic across multiple queues make less efficient use of a given amount of bandwidth.

Prior to the widespread use of Fast Ethernet, inverse multiplexing of Ethernet was used to provide high-speed pipes into servers. This is no longer widely used.

A similar situation exists now with Fast Ethernet. Gigabit Ethernet is not yet widely available, and inverse multiplexing of Fast Ethernet links provides much of its bandwidth benefits. However, as Gigabit Ethernet costs decline, and as it becomes more readily available, it seems likely that it will replace the multiple Fast Ethernet option.

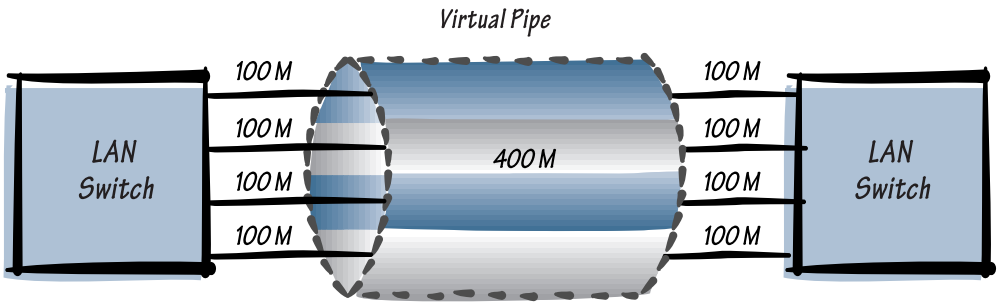


Figure 4.8. Inverse multiplexing between switches

The future of LAN switching

A few predictions:

- The cost of LAN switching will continue to decline. Ethernet prices declined very rapidly in 1996 and 1997, and will come down somewhat more slowly in 1998 and 1999. Fast Ethernet prices will come down more rapidly, and 10/100 switches with modular ATM / Gigabit Ethernet uplinks will be the fastest-growing category of switches.

- LAN switches will increasingly replace hubs. Just as dedicated voice lines have almost completely replaced telephone party lines, so dedicated LAN connections will inevitably replace hubs. The continuing decline in the price of switch ports will fuel this change.

- LAN switching sales will continue to grow for a long time. In 1997, for the first time, worldwide revenues from switch sales exceeded those from hubs. However, the total number of hub ports shipped was still much greater. There is a large, and still growing, base of hub ports that are potential targets for LAN switches.

- Multi-level switches will increasingly integrate LAN switching with other technologies, including ATM, Gigabit Ethernet, and layer-three switching. Inexpensive, powerful switches for the wiring closet allow network managers to design highly failure-resistant mesh networks.

- Both modular and non-modular LAN switches will continue to be used. In some applications, the best switch for the wiring closet is the least expensive. In others, it's worthwhile to invest more for additional power, port count, media and protocol flexibility, maintainability, and failure resistance.

- Multi-level LAN / ATM / layer-three switches will eventually merge with voice switches to provide a single cohesive voice / data / video network. This will happen as applications become widely available that integrate all three technologies, and as users find synergy in doing so.

Where are virtual LANs useful?

A virtual LAN is a broadcast domain defined in software. Switches forward broadcasts (and multicasts, and unicasts with unknown destination addresses) within a VLAN, but not between them.

Virtual LANs are fundamentally a tool for establishing a hierarchical network. In a hub and router network, the hierarchy is automatic: an Ethernet segment or a token ring equates to a broadcast domain. And a broadcast domain equates to a router port. In a TCP/IP network, it also equals an IP subnet.

But in a fully switched network every device has its own private Ethernet or token ring.

- Should each device have its own IP subnet, and every frame be routed? Since routing is more complex than LAN switching, it's inherently slower. Forcing all traffic through a routing process – even a high-speed layer-three switching process – tends to degrade network performance.
- Or should there be one large flat network, with everything switched, and nothing routed? In that case, every device receives every broadcast and every multicast from every other device. In a large network this can place a heavy load on bandwidth, and interrupt each workstation's processor excessively.

Virtual LANs provide a balance. Using virtual LANs, a switching network switches within broadcast domains, and routes between them.

VLANs are not needed in every LAN switching network. Many networks are small enough that they can be managed as a single broadcast domain. Others can simply route between LAN switches, with each switch a broadcast domain.

General issues surrounding VLANs

VLAN implementations vary radically among switching manufacturers. There are two reasons for this:

- There are VLAN standards, but they are limited (see "VLAN standards", below). So vendors have had to develop their own definition of what a VLAN should do, resulting in widely varying capabilities.
- A VLAN mechanism must rapidly examine and make complex decisions about high-speed streams of frames, slowing the frame down as little as possible. So VLANs must be implemented in hardware, with minimal real-time software support. Once a manufacturer has built a chip, designed a product around it, and shipped the product to thousands of customers, they are reasonably reluctant to make fundamental design changes that will outmode the product.

VLANs across multiple switches

Some VLAN mechanisms exist only within one switch. This was sometimes useful when switches primarily interconnected hubs. But in a fully switched network, it's less useful, since all of the devices connected to a single switch could reasonably exist within a single broadcast domain in almost any network configuration.

In many organizations, all or many of the servers have been centralized, but there is often still a connection between the people in one area of the building or campus, and a particular server in the central computer room. It's likely that the users and the server are connected to different switches. Data moving between those users and that server must either be routed, or switched in a single VLAN across multiple switches.

A related issue is mobility. Some manufacturers originally touted VLANs as primarily a feature which provided mobility: with VLANs, you could add or move user workstations at will, without having to reconfigure layer-three addresses in the machines. Although this is true, the more fundamental benefit of VLANs is broadcast isolation. And in many IP networks, DHCP takes care of IP address management automatically. However, DHCP applies only in an IP network, and for other protocol stacks this is still a convenient function of VLANs. It requires, of course, that the same VLAN be available in the user's new switch as in his / her old switch.

VLANs could theoretically traverse more than one switch by assigning a dedicated trunk to each VLAN. This would be expensive and clumsy. The alternative is a *tagging protocol*, which provides some identifier for each frame that flows within the network, so that the switch at the other end can place it into the correct VLAN. Some tagging protocols are vendor-specific and others are based on public standards.

VLAN throughput and latency

As with any communications technology, VLAN implementations exhibit important differences in performance. One good way to examine the differences in switch performance is to read analysis published in leading networking magazines. Look carefully at the quantitative data, and at the construction of the tests, and remember that some tests are designed, not to simulate real network environments, but to "break" the products being tested. Vendors should be asked to explain any questionable test results.

VLANs supporting multiple technologies and rate conversion

A multi-switch VLAN implementation must support multiple MAC types in a single VLAN, because the trunks between switches are almost always high-speed standards like Fast Ethernet, ATM, Gigabit Ethernet, or FDDI, while workstations are generally Ethernet, Fast Ethernet, or token ring. The same thing is true of server connections.

Some VLAN implementations also translate from token ring to Ethernet, Fast Ethernet, or Gigabit Ethernet. This means that a token ring workstation in one part of a building can connect to a Fast Ethernet server in another part of the building, across a Gigabit Ethernet backbone. For token ring users who have decided to migrate to Ethernet, but who can’t change an entire large network overnight, this helps to simplify the transition.

Multiple VLANs per device

Many workstations run multiple protocol stacks at the same time. One common combination is TCP/IP for access to the Internet, in-house Web servers, and corporate file server applications, and NetWare for access to departmental applications and e-mail. At a minimum, it would be useful for each workstation to belong to an IP VLAN, and also to an IPX VLAN.

VLANs can also be used very flexibly to control access to resources; a workstation, or even a person, can be assigned access to some VLANs and denied access to others.

How VLANs are defined

Direct configuration vs. policy-based management

There are two basic ways to configure VLANs:

Each port can be directly configured for VLAN membership, typically in a single VLAN. This can be burdensome in a large network, and the configuration needs to be redone every time a device is moved or added.

Or some form of policy-based management can be used. Most of the more flexible VLAN definitions require policy-based management.

VLAN as a group of switch ports

A VLAN can be simply a list of physical switch ports; some implementations offer only port-based configuration. As noted above, this tends to be more difficult to manage.

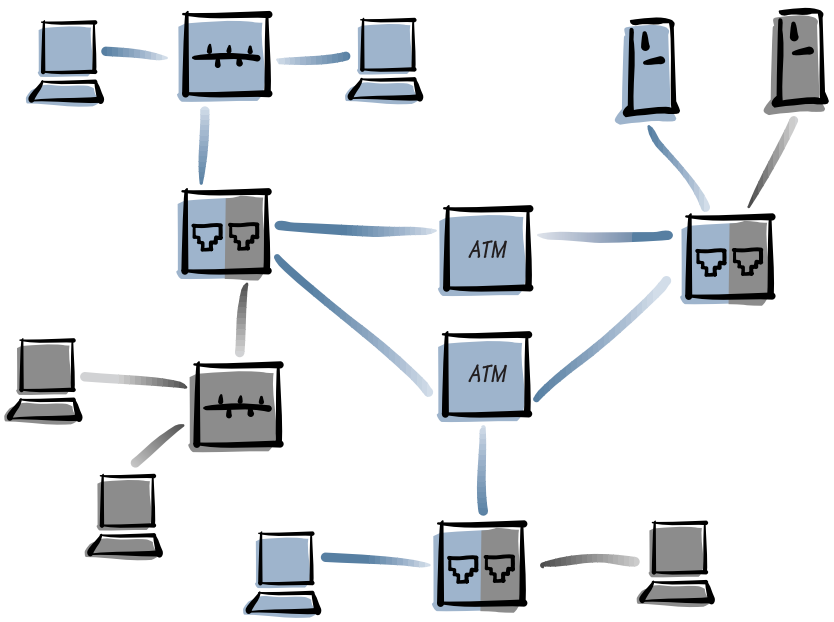


Figure 5.1. VLAN defined as a group of switch ports

VLAN as a list of MAC addresses

A VLAN can sometimes be a list of MAC addresses. In some cases the list operates within a particular switch. In others, the person managing the network does not need to know where in the network the devices are located; the list is sent to all switches, which automatically assign the appropriate ports to the VLAN.

MAC addresses are cryptic and uninteresting, and it's easy to make a mistake. When a network interface card is replaced for repair or upgrade, the MAC address changes, and must be added to the VLAN list.

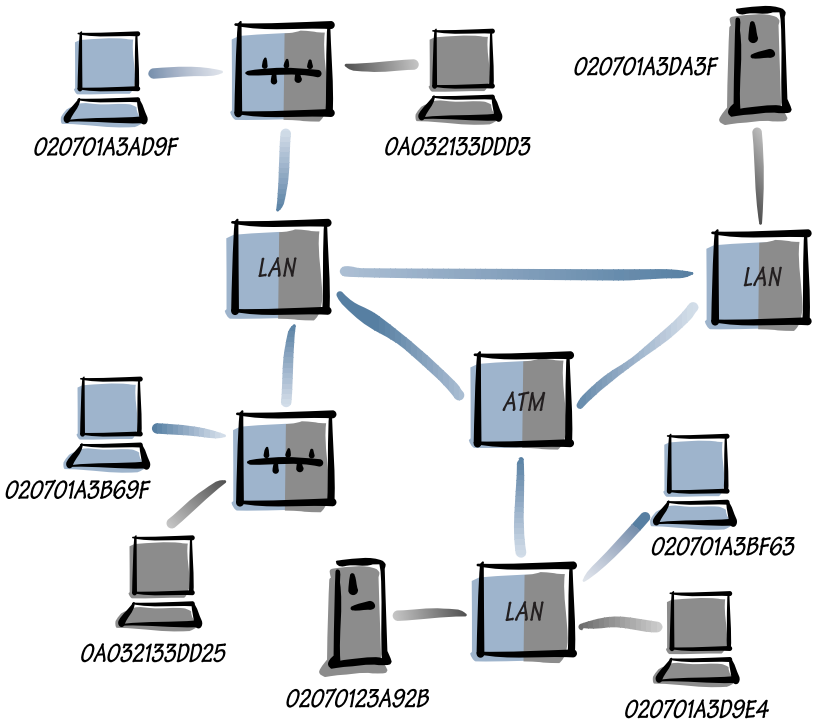


Figure 5.2. VLAN defined as a list of MAC addresses

VLAN as a protocol type

It can be very useful to group all workstations that use a particular protocol into a single VLAN. For example, some organizations have a relatively small percentage of their workstations running DECNet, or AppleTalk. These machines can be assigned to a common VLAN, keeping their broadcasts from the workstations only running IP, IPX, or other protocols.

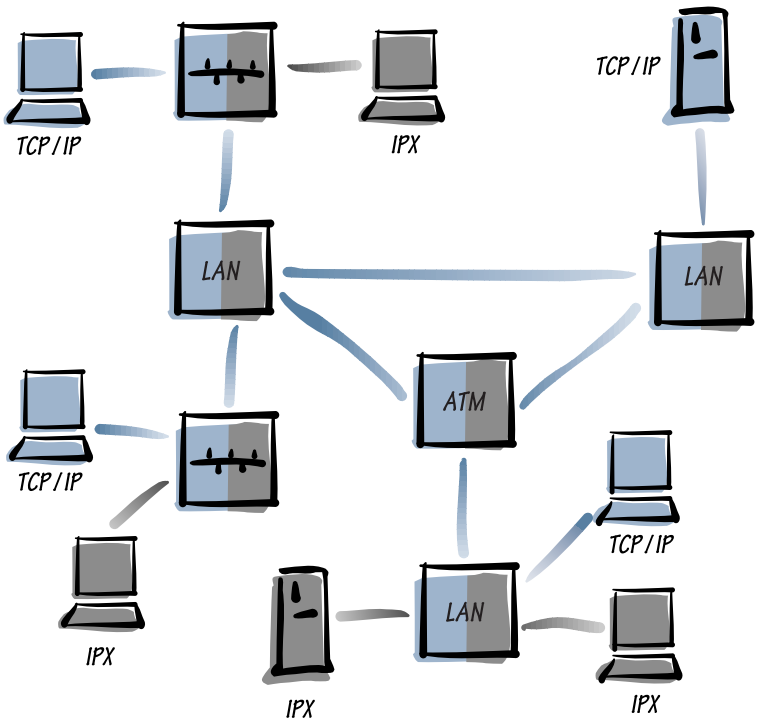


Figure 5.3. VLAN defined as a protocol type

VLAN as a subnet

In a routed TCP/IP network, a basic unit of management is the IP subnet, which is a broadcast domain with a common subnet portion in the IP address of its members. It makes obvious sense to define VLANs in an IP network by the IP subnet address of each device. This is a particularly easy way to manage a large number of IP-based workstations.

Note that this type of definition is generally in conflict with IP address distribution using the DHCP protocol. It's obviously a problem to assign IP addresses automatically and broadly, and to also use assigned IP addresses as the basis for VLAN membership. However, it is possible to combine the two processes, having DHCP servers automatically assign a range of addresses within a predefined set of devices.

Some manufacturers also support VLANs defined by IPX network, and sometimes by AppleTalk zone.

VLAN as a multicast group

In some implementations, a VLAN can be defined as a group of workstations, all of which wish to receive the same multicast distribution.

VLAN standards

As noted above, one of the most important changes of the last few years was the increasing emergence of public standards. It would be ideal if a broad VLAN standard were available, to which all vendors could develop. Unfortunately, because VLANs are so tightly coupled with hardware, it is very difficult to arrive at a standard which mandates an advanced set of features; a number of manufacturers would be unable to implement the new standard without changing hardware, and would oppose it.

However, there are some important partial VLAN standards.

ATM LAN Emulation

LAN Emulation is discussed in detail below. As we'll see, it provides services to LAN-based equipment connected to an ATM network, services which are needed because LANs operate differently from ATM. One way in which they are different is broadcasts. Every MAC-layer protocol uses broadcasts as a basic tool. ATM is inherently point-to-point, so broadcasts have to be emulated.

LAN Emulation permits this to happen across an ATM network. An ATM emulated LAN is a broadcast domain defined in software. In other words, it's an ATM VLAN. Just as with more general VLANs, broadcasts, multicasts, and unicasts with unknown destinations are switched within an emulated LAN, and data must be routed between them. And a trunking capability allows frames belonging to multiple emulated LANs to be carried across the same physical link.

And just as with more general VLANs, there are substantial differences in vendor implementations. For example, some vendors require port-by-port assignment of ports to emulated LANs. Others allow policies to be set in network management, which then automatically assigns users to emulated LANs.

Modified 802.10

An attempt was made at one time to use an existing protocol – the IEEE 802.10 protocol for metropolitan area security – as a VLAN trunking protocol. The idea was to employ two octets that were left unused in the standard's header structure. However, the IEEE subcommittee failed to agree on this, and although some vendors have implemented it anyway, that use has remained proprietary.

802.1Q and 802.1p

A more important standards-making effort started at the same time in the IEEE 802 committee, and has been successful. This is the 802.1Q standard. It is now complete, and will be widely implemented within the next year. 802.1Q is tightly coupled with the 802.1p standard.

802.1Q defines:

- a standardized form of port-based virtual LANs (note that this is only a limited form of VLANs, without the advanced policy-based capabilities that are needed in some applications)
- a modified Ethernet frame, with three bits that specify up to eight priority levels; 12 bits that specify up to 4,096 different VLANs; and one bit which is reserved for non-Ethernet frames types which are being switched across Ethernet
- a VLAN trunking mechanism
- a protocol for distributing VLAN information to switches

802.1p is basically used to prioritize frames. Up to eight levels of priority are available.
It defines:

- a generalized protocol (GARP – Generic Attributes Registration Protocol) for signaling between workstations and the network
- a version of GARP (GARP Multicast Registration Protocol) which allows devices to request membership in a specific multicast group
- a version of GARP (GARP VLAN Registration Protocol) which allows devices to request membership in a specific virtual LAN

The future of virtual LANs

A few predictions:

- Port-based VLANs are too difficult to manage to be widely used; they will be generally replaced with policy-based VLANs.
- Many applications in which VLANs would otherwise be useful will soon be implemented instead with layer-three switching. However, the two techniques can be complementary, and will sometimes be used together.
- VLANs will continue to provide a useful structure for many types of services, including authentication and prioritization.

Where is layer-three switching useful?

We saw above that LAN switching is basically very high-speed, hardware-based bridging. In the same way, layer-three switching is basically very high-speed, hardware-based routing. Some forms of layer-three switching use exactly the same protocols as traditional routing, while others use new protocols to make high-speed processing easier.

What problems does layer-three switching solve?

Throughput

Consider the following:

- A single high-speed backbone, like Gigabit Ethernet or ATM OC-12, can carry more than one million pps. Most networks that use these technologies would have a number of these links.
- LAN switching can generate huge amounts of packets at the network edge. 1,000 workstations on dedicated Fast Ethernet connections could theoretically send and receive 280 million pps.
- Graphical applications are increasingly making use of all this bandwidth.

Traditional software-based routers can't keep up with this load. Hardware-based mechanisms are needed. And they can also be supplemented by new protocols, designed to operate at very high rates.

Latency

Another advantage of layer-three switching is low latency. By shifting the process of packet forwarding from software to hardware, it's possible to reduce packet processing time at each router in a path. It is not unusual for a software-based router to take a millisecond to process a packet; layer-three switches measure this time in microseconds.

Network integration

Software-based routers are massive, complex code-processing engines. Layer-three switches consist of a few chips. It's possible to integrate them into the same switches that support the workstations and the backbone, combining LAN switching, ATM switching, Gigabit Ethernet switching, and routing in a single platform. This makes the job of managing the network easier and adds substantial flexibility.

Cost

Because software-based routers use general-purpose CPUs, they're expensive to build. Layer-three switching reduces the cost of routing a packet by 80% or more.

Network reliability

One of the drawbacks to a campus network based entirely on layer-two switching is that reconfiguration, typically handled with the Spanning Tree protocol, tends to be slow. A network that uses layer-three switching in the backbone is able to more rapidly route around failures in cable, AC power, switch ports, or entire switches.

Protocol stacks

Even a preliminary discussion of protocols would require a large volume of its own; a complete analysis would take a number of volumes. The following section is meant to be a small introduction to the subject, as a basis for understanding layer-three switching.

Early communications software was monolithic; a single piece of code provided all of the needed functions. It became clear that this approach was a problem, for several reasons:

- It's difficult to create such a massive body of code, and it's difficult for third-party software developers to supply code components to manufacturers.
- It's difficult to maintain; any change may require substantial revision.
- It's hard to adjust to varying circumstances, such as different physical media, new wide area technologies, and so on.
- Standards are important in the development of a global industry, and it's hard to standardize a huge piece of code.

As a result, layered protocols, on which all modern data communications is based, were developed. Functions are arranged into logical layers; the collection of layers is called a protocol stack. Examples include TCP/IP, SNA, DECNet, NetWare, and AppleTalk. In a protocol stack, each layer provides a well-defined set of services to the layer above it, and relies on services from the layer below it. Information goes down the stack, from a user application at the top, to a piece of wire at the bottom, across a network, and then back up a stack in the machine at the other end.

One common model for layered protocol stacks was developed some years ago by ISO (the International Standards Organization); it is referred to as the OSI (Open Systems Interconnect) model. This model maps closely to most protocol stacks up to layer four; above this layer they diverge widely:

- Layer one, the lowest layer in a stack, is the physical layer. It describes physical connectors, electrical and light levels, and similar capabilities. An example of a layer-one protocol is 10BaseFL (Ethernet running over multimode fiber optic cable).
- Layer two is divided into the MAC (Media Access Control) sub-layer and the LLC (Link Layer Control) sub-layer. Layer two is where LAN protocols such as Ethernet, token ring, and FDDI are defined. Layer two is responsible for getting information across a LAN or a wide area circuit.
- Layer three is the packet, or network, layer. Generally, it does the best it can to move packets across a network, which may span a building or the globe. Several types of protocols operate together at layer three. A packet forwarding protocol, such as the widely used IP (Internet Protocol) moves packets from one network node to another. A routing protocol, such as RIP (Routing Information Protocol) is used to determine the best path toward another point in the network. Other layer-three protocols provide network signaling and other functions. Traditionally, layer-three protocols have been associated with routers; they are now also associated with layer-three switches.

- Layer four is the transport layer. It provides end-to-end connections across the network. Below the transport layer, each protocol operates in both end stations and intermediate network nodes. Transport-layer (and higher) protocols operate only in end stations. Typical layer-four services include guaranteed delivery of data across a network, and segmentation of large messages into packets small enough to be handled by the lower-layer protocols. An example of a layer-four protocol is TCP (Transmission Control Protocol).

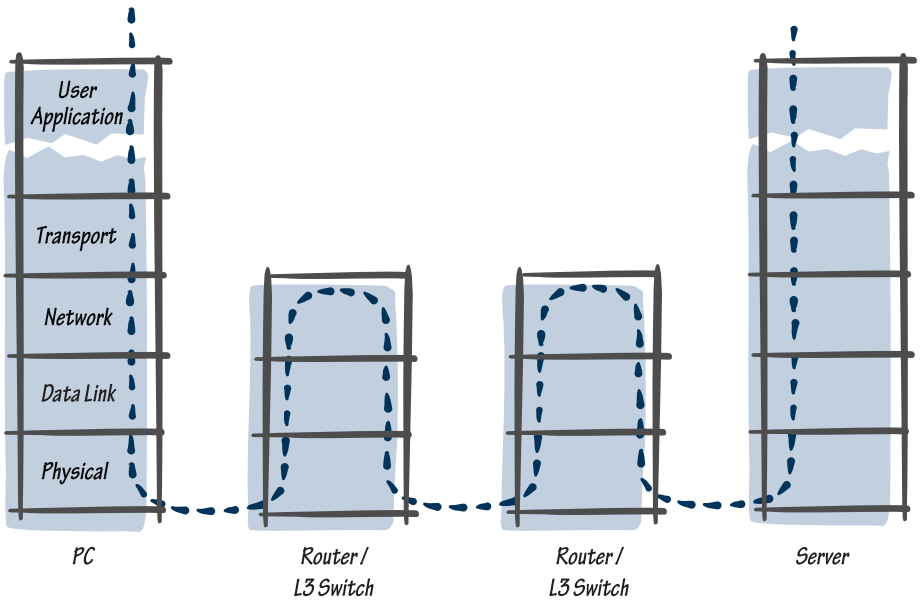


Figure 6.1. Information flows in layered protocols

Basics of routing

As we discussed earlier, routing is used to interconnect layer-two broadcast domains, which may be a physical LAN (as with a hub) or a virtual LAN (as with a switch). The router acts as a proxy for the data that's being sent from one LAN to another. For example:

- A workstation running TCP/IP over token ring sends a token ring frame to the local router's MAC address.

- The router strips off the MAC-layer information and examines the IP packet contained within it.
- If the packet is destined for another LAN connected directly to the router, the router adds the appropriate MAC header and forwards the packet onto that LAN.
- If the packet is destined for a remote LAN, the router adds the appropriate MAC header and forwards the packet to the router that supports that LAN – or to an intermediate router that will pass it on.

It's relatively easy for a router to forward a packet to a LAN to which it is connected. It simply needs to know the subnet addresses of all the LANs to which it is attached. But how does a router know where to forward a packet that needs to go to a remote LAN?

Routers do this with one or more *routing protocols*, such as RIP, RIP II, OSPF, BGP4, and IGRP. Using these protocols, the routers tell each other which subnets are connected to them, and the state of the inter-router links. With this information a router is able to construct a *routing table*, which it consults when forwarding packets remotely.

Hardware-based routing

Description

One option for layer-three switching is hardware-based routing, also called packet-by-packet layer-three switching.

All routing – software-based or hardware-based – has two basic elements:

- One is the operation of the routing protocols, and from these the calculation of the routing table. Routing protocols are continually evolving; the task of operating them is complex; and the amount of information to be moved is relatively small (routing table updates might only be sent every 30 seconds, for example). So, even on high-speed links, it's not difficult to handle this function in software.
- The other is the real-time forwarding of packets on appropriate links. This is a much simpler process, but one which needs to occur at very high rates. If a router handles 10,000 pps, each packet can be processed for one tenth of a millisecond. If it handles 1,000,000 pps, each packet must be processed in a microsecond. No general-purpose processor can keep up with this.

As with LAN switching, the forwarding process can be implemented in custom chips, called ASICs. By coupling these routing ASICs with general-purpose processors, it becomes possible to inexpensively build massive amounts of routing power into a network, and still have plenty of flexibility.

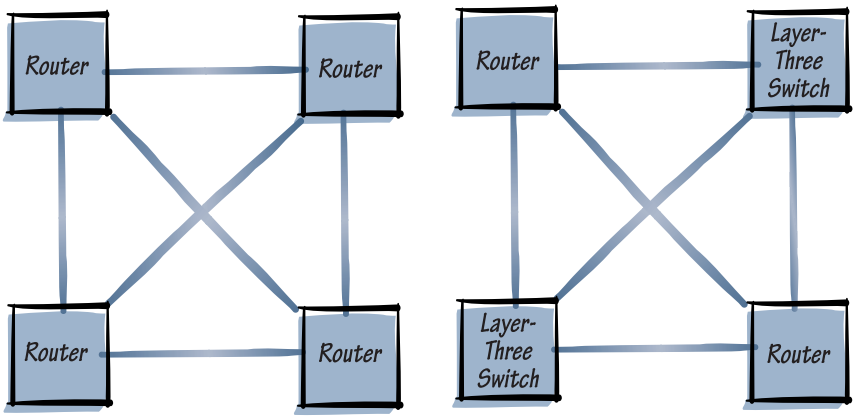


Figure 6.2. Some layer-three switches are interchangeable with routers

Advantages

Packet-by-packet hardware-based routing offers most of the advantages described above. It can deliver very high throughput. It can operate with very low latency. It can provide rapid rerouting around failed links and nodes; in fact, this is one of the primary reasons that layer-three switching is so important to Gigabit Ethernet, which would otherwise have to depend on Spanning Tree.

And it is inexpensive. Hardware-based layer-three switches are already 80% less expensive, on a per-packet basis, than software-based routers, and over the next year these prices will drop by another order of magnitude. It's obvious that the days of the traditional, software-based, multiprotocol router as a campus backbone standard are quickly drawing to a close.

Issues

Not all layer-three switches will have the same capabilities:

- Many layer-three switches support only IP. Some support IPX as well. The argument that the IP-only vendors make is that only TCP/IP applications are growing rapidly, and the existing routers can easily handle other protocols, including IPX. While this is often true, it's clumsy to route one protocol stack on one set of equipment (integrated LAN / layer-three switches) and another protocol stack on another set of equipment (software-based routers). It's easier, if possible, to route all protocols in the same equipment.
- Some layer-three switches only support the RIP routing protocol. If another protocol – like OSPF – is being used already in the network, it's optimal for the new switches to integrate with it.
- Some layer-three switching implementations are limited to Ethernet, Fast Ethernet, and Gigabit Ethernet. For many applications this is not a limitation. But for networks that have a large base of token ring workstations, or which want to transition gradually from an FDDI backbone, layer-three switching support for other MAC protocols is important.

Cut-through switching

ATM offers an important advantage: a virtual circuit structure. This makes it possible to set up directly switched connections between two end devices, with no traditional routing between them. Each device's transmitted packets are mapped into the same virtual circuit and delivered by the switching network. This results in very high throughput and very low latency.

If the two devices need to be constantly connected, the process is very simple. But for transient connections sophisticated protocols are needed:

- to match LAN addresses to ATM addresses
- to distribute broadcasts and multicasts
- to store and relay packets that are transmitted before a session is established
- to establish the connection
- to tear the connection down

It's possible to move information across an ATM network using standard layer-three protocols, such as IP and OSPF. However, this requires repeated conversions between packets and cells, and many of ATM's benefits are lost.

Another option is to make all devices members of the same layer-two broadcast domain, with unicasts switched directly between devices, and broadcasts repeated to all devices. However, this doesn't scale well; just as with a frame-based network, at some point the network must be divided into subnets.

An important alternative method is to use protocols to set up and tear down the connection, and to switch user data directly between the end devices. This is called cut-through switching. Two competing alternatives have been proposed.

IP Switching

A number of vendors have licensed code from a company called Ipsilon (now a part of Nokia), which also offers the protocol implemented in switches and router servers. Basically, Ipsilon's IP Switching routes packets in the ordinary way as a default. However, whenever the switches detect that a long-term flow is in progress, they set up directly switched connections between the two end devices. For example, an SNMP packet would just be routed, while a dedicated ATM circuit would be set up to support an FTP transmission.

Very few users have chosen to install IP Switching, and it appears that the increasing availability of MPOA from major vendors will prevent further widespread acceptance.

NHRP and MPOA

These form an alternative to IP Switching, and have the advantage of being sanctioned by the IAB and the ATM Forum. Together they operate in a similar way to IP Switching, although MPOA will eventually support protocols other than IP. They are discussed in more detail in the section on ATM.

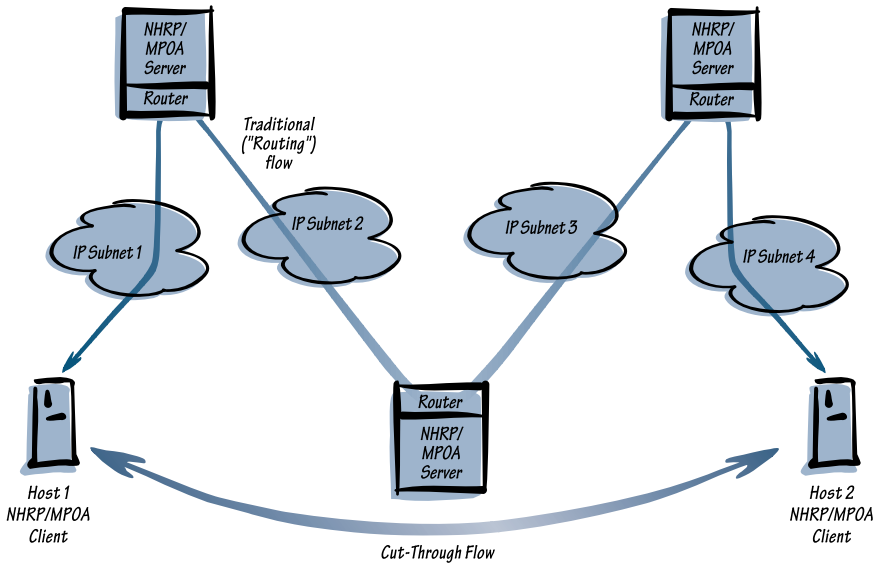


Figure 6.3. Cut-through switching model

Label switching

The basic definition of switching is that it is a simplified function performed in hardware. Some switching alternatives do this by splitting the complex but less time-critical component – protocol processing – from the simpler forwarding process. Protocol processing is kept in software, and forwarding is put into hardware. An example is traditional routing implemented in hardware: as noted above, the RIP / OSPF / BGP4 tables are often constructed with software, while the packet forwarding process is done with hardware.

Label switching uses the same concept. Each packet is prepended with a label. A table is constructed in each switch; for each label it shows the best port to use to move a packet toward the destination entity which corresponds to that label. This destination would often be a subnet, but could also be as large as a network, or as small as an application. When a packet is received, the switch reads only the label and forwards the packet out the proper port.

The determination of routes is done in a manner similar to other advanced link state routing protocols, such as OSPF. In fact, it is possible to use one of these protocols as the route-determination mechanism.

Label switching offers several advantages:

- Its simplicity should deliver high throughput and reduced cost.
- The ability for labels to apply to network entities at various levels makes it extensible to very large networks. A label could be as large as an autonomous network, or as small as an individual application flow.
- The technique does not mandate an underlying virtual circuit structure, although it can make use of one if it exists, so it can integrate cell switching and frame switching.
- Since the label is independent of the network-layer address, and the route determination process can use various routing protocols, it can support various traffic types, and can evolve flexibly.

One proposed form of label switching is Cisco's tag switching. This may be merged into the MPLS (Multi-Protocol Label Switching) working group in the IETF; Cisco is one of the vendors who have contributed to the MPLS. As of this writing it is expected that MPLS will be ratified in early 1999.

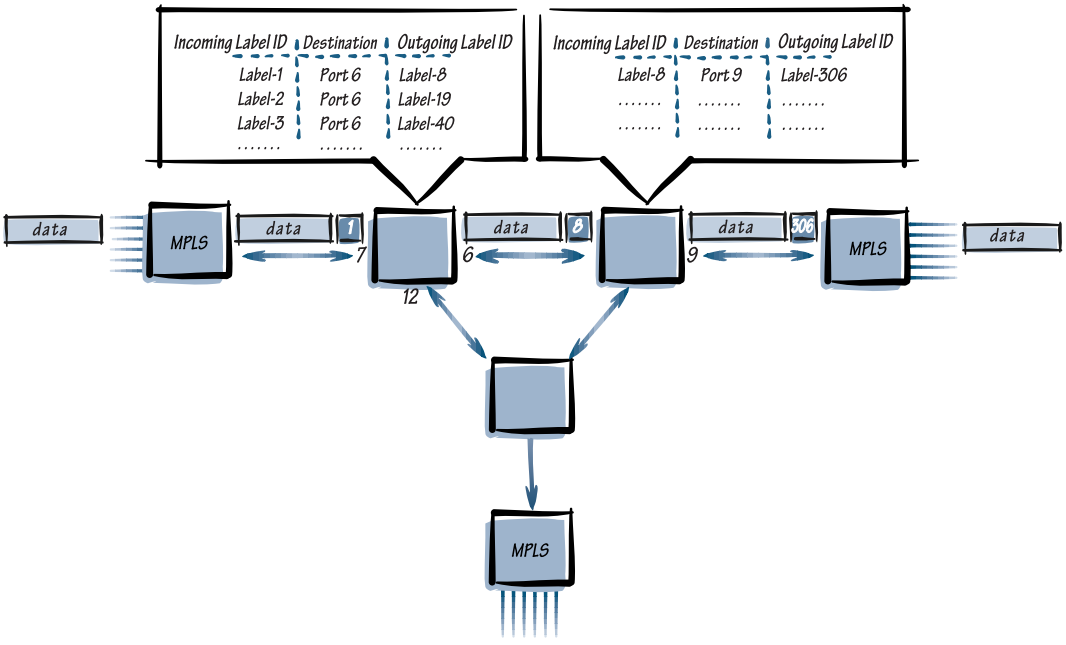


Figure 6.4. Label switching model

The future of layer-three switching

A few predictions:

- Over the next few years, these techniques will be widely used. There will be a rapid transition to layer-three switches from software-based routers.
- For most mid-sized end users, a combination of Ethernet, Fast Ethernet, Gigabit Ethernet, and packet-by-packet layer-three switching will be an optimal solution. Layer-three switching will be a standard feature in high-end LAN switches and in Gigabit Ethernet switches.
- Some end users will benefit from the advanced capabilities of an ATM backbone, while still using Ethernet, Fast Ethernet, and token ring at the desktop. These users will gain high performance levels from MPOA.
- Layer-three switches will continue to get faster. However, for most users, throughput greater than 10 Mpps will not be meaningful for some time.
- Gigabit Ethernet prices will decline to some extent, but not to the extent that Fast Ethernet costs have. Much of the cost of Gigabit Ethernet is in the fiber optic transceivers, which are not as amenable to rapid cost reduction as copper transceivers.
- Label switching will be of value in the operation of the Internet, but not for most end users. Since approval of MPLS is unlikely to occur soon, its future is less certain than that of MPOA or hardware-based routing.

ATM is another subject to which a number of books have been dedicated. We will provide a brief overview of ATM, in order to put ATM switching in perspective.

In the last several years ATM has gone through a surprising change of public opinion: from telephone company long-term plan; to the inevitable future of all communications; to a hopelessly complex technology destined to be replaced by Gigabit Ethernet; to, finally, a more balanced perspective as an important part of both campus and carrier networks.

A (very) brief history of ATM

The telephone companies around the world have long had a goal of providing an integrated network that can carry multiple types of information across a single infrastructure. This would have a number of advantages:

- It would reduce costs by eliminating expensive duplicate overlay networks for various types of traffic.
- It would reduce costs by dispensing bandwidth statistically rather than in fixed allocations.
- It would provide a conduit for enhanced services; carriers make much greater profits on enhanced services than they do by simply selling bandwidth.
- A single multimedia network would be easier to maintain.
- A multimedia network would make possible applications that cannot operate on separate networks, but which are possible when multiple media types are integrated in user stations.

ISDN and BISDN

The first attempt at an integrated multimedia network was ISDN (Integrated Services Digital Network), which provided user data rates of 64 Kbps and multiples of 64 Kbps. Unfortunately, the consultative process by which carriers must develop standards took a number of years, and by the time ISDN was finally standardized it suffered from two problems. The world had started to use LANs at 10 Mbps and above, and a 64 Kbps channel was perceived as slow. And the standard was full of alternatives, so that implementations that met the standard often could not interwork.

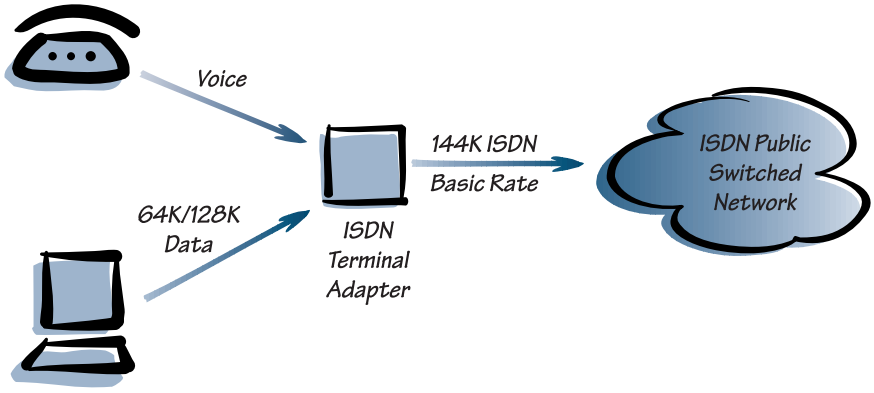


Figure 7.1. Multimedia over ISDN

The carrier community resolved to attempt again, and began to design what it called BISDN (Broadband Integrated Services Digital Network), which would provide enormously greater throughput than ISDN. Once again, the creation of the standard took a long time, and it is possible that BISDN would have become a very slow, completely carrier-based effort.

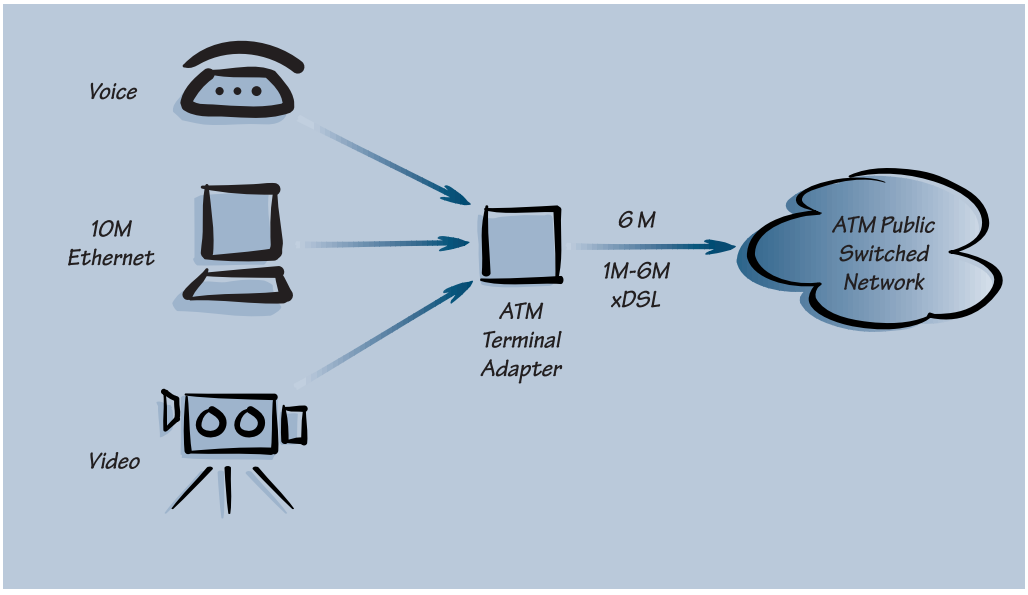


Figure 7.2. Multimedia over ATM

ATM Forum

Then something changed. The networking vendor community needed a high-speed mechanism for communications on the campus, for backbone networks and for high-speed desktops. They looked at the emerging specifications for BISDN and decided that this could fill the role.

The Frame Relay Forum, a symbiotic effort between vendors and carriers, had been very successful in making frame relay an important standard for WAN communications. In that case, vendors had taken standards which the carriers were developing slowly, had rapidly brought them to a reasonable state of completion, had handed the revised and enlarged standards back to the carriers for ratification, and had built products to deliver the standard. They believed that the same process could work for BISDN, which by now was being referred to as ATM (Asynchronous Transfer Mode).

It worked. The ATM Forum moved rapidly to complete a wide range of standards, including physical interfaces, user-to-network software interfaces, network-to-network software interfaces, traffic management, and others.

Today, the ATM Forum is to some extent a victim of its own success. It has over 900 members, and producing new standards has become more difficult. ATM is now an important business, and vendors have a great deal at stake as standards change. There is an argument that very large vendors sometimes manipulate the standards-making process, slowing it or shifting it to suit their needs. In general, though, the degree of cooperation among major vendors in the ATM Forum has been, and continues to be, an important and positive element in the networking industry.

Where is ATM useful?

End-to-end ATM

The original vision of ATM on the campus was a complete one. Every workstation would eventually have its own ATM network interface card. All of the switches in the wiring closets and the backbone would be ATM-only switches, and the interface to the wide area would be ATM. New workstation and server software would make optimal use of the new network.

That didn't happen; or, at least, it hasn't happened yet, and for all but a few networks it won't happen for a long time. Instead, users have preferred to keep the Ethernet, token ring, and Fast Ethernet network interface cards in their workstations, and to operate with the existing networking software.

ATM's strengths on the campus

ATM offers some strong advantages:

- It can be scaled to very high rates. As networks need to support higher rates at the workstation, and as new applications demand greater bandwidth, ATM can be scaled up almost indefinitely.
- It provides a virtual circuit structure, which can be used for device-to-device connections with very high throughput.
- ATM has advanced mechanisms for failure recovery. One of these operates at the physical level, for multiple links between two switches. Another operates at the network level, load-balancing and reconfiguring the network in response to circuit or switch failures.
- Integral to ATM is the ability to deliver the quality of service characteristics needed by various traffic types. In particular, real-time voice and video require very low delay and very low variation of delay; ATM handles this quite well.
- ATM switched connections have very low latency.
- As noted below, ATM is becoming an important component of wide area networking. As a result, integration of the campus with the wide area is eased when both use this technology.

Real ATM today

These strengths are greatest when the entire campus, including desktops, uses ATM. But this is unrealistic for most users, due to:

- the cost of ATM network interface cards
- the cost of ATM switch ports
- the substantial changes in workstation software that would be needed
- the limited availability of workstation software that can effectively synergize voice, data, and video

- the need to learn an entirely new set of protocols and technologies
- the perception that ATM is overly complex

So ATM's campus role is largely focused on backbones and servers, where many of its strengths are most important. As a result, in most ATM-based campus networks, switches in the wiring closets convert from ATM to LAN protocols such as Ethernet, Fast Ethernet, and token ring.

Some users are following this modified ATM path, with ATM in the backbone and frame-based communications to the desktop. Others will use Gigabit Ethernet as an alternative for the backbone. It seems likely that both the ATM and Gigabit Ethernet approaches will be widely adopted over the next few years, and that the goal of a single universal communications technology will be incomplete.

This split is geographically correlated. In the United States, users tend to favor Gigabit Ethernet, although ATM is also quite strong. Outside of the United States, users tend to favor ATM, although Gigabit Ethernet is also quite strong. Even within the U.S., some areas are more strongly disposed to ATM, and others to Gigabit Ethernet.

Cells and frames

A basic characteristic of ATM is that it uses fixed-length cells, rather than variable-length frames. Every ATM cell is 53 bytes long: a 48-byte payload, and a 5-byte header. A train of multiple cells is used to send longer messages. When a cell's payload is less than 48 bytes long, it's padded to that length.

Fixed-length cells are used because they are better suited to hardware-based switching than are variable-length frames.

An amusing side note on cell size is that an early ATM debate in the carrier industry pitted 32-byte cells against 64-byte cells. The final 48-byte cell was a compromise, brokered in part by the U.S. Department of State.

The drawback to cells, compared to frames, is that the associated header information occupies greater overhead on the transmission medium; this is sometimes called the "cell tax". It is relatively less important on high-speed connections, but more important on low-bandwidth circuits, such as 56 / 64 Kbps, or even DS-1 and E1.

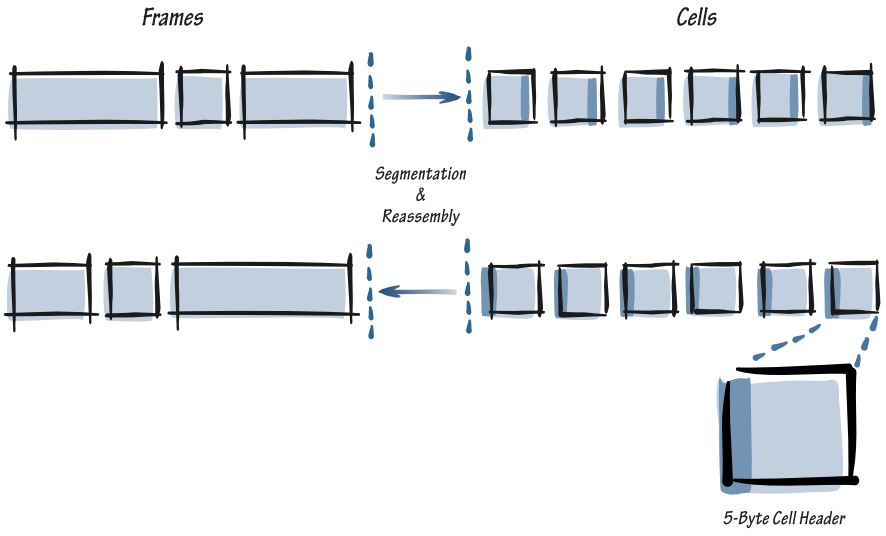


Figure 7.3. Frame-to-cell conversion (segmentation and reassembly)

Virtual circuits

In electronics, a circuit is a continuous, discrete, defined path along which electricity flows. A virtual circuit in networking is a continuous, discrete, defined path which does not have electrical continuity, but along which data flows. When data flows across a virtual circuit, a connection is in progress. Virtual circuits and connections (in this sense) are not essential; most local area networks, including Ethernet and token ring, do not use them. But they offer certain advantages:

- Since a virtual circuit is a defined path, its characteristics (such as maximum delay and delay variation) can be determined ahead of time.
- A virtual circuit can be assigned a fixed amount of bandwidth, or at least a minimal amount.
- It's easier to send and receive flow control information across a virtual circuit, making buffers more effective.
- Virtual circuits simplify the process of building fast switches. When a virtual circuit passes between two switches it is assigned a number, by which frames or cells belonging to it are identified. It's relatively easy to switch on the basis of these virtual circuit numbers.

Virtual circuits are central to ATM, as they are to frame relay. In fact, virtual circuits give ATM its connection orientation, which is one of the most important differences between ATM and LANs, such as Ethernet. Virtual circuits can be switched virtual circuits (SVCs), which are set up dynamically, as needed; or permanent virtual circuits (PVCs), which are configured by a network administrator to be connected all the time.

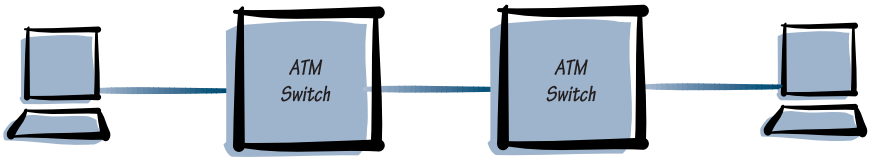


Figure 7.4. Permanent virtual circuits in ATM

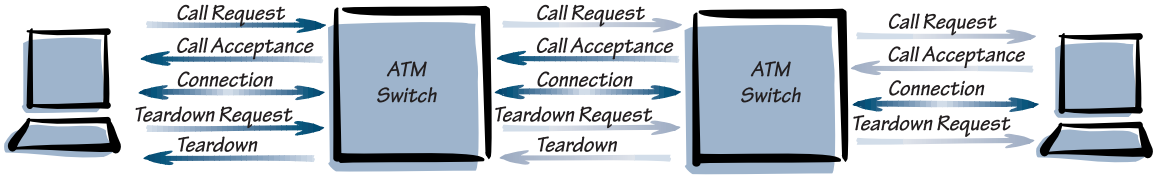


Figure 7.5. Switched virtual circuits in ATM

Quality of service

As noted above, fixed-length cells were chosen in large part because of their ability to effectively support real-time voice and video.

Transit delay

Short cells are used because applications like real-time video and voice are optimized when latency is minimal. Anyone who has ever experienced the annoyance of an inter-continental telephone call routed over a satellite knows the problem with excessive latency. There are two components of latency: serialization delay, and processing delay. Short cells mean that each cell can be rapidly moved into a switch, minimizing serialization delay. Hardware-based switching minimizes processing delay.

Delay variation

Another characteristic that needs to be controlled for some types of traffic is delay variation. For example, real-time voice needs little bandwidth, but must be delivered with very low variation of delay; data file transfers should go as quickly as possible, but can arrive in bursts. ATM can guarantee each what it needs.

Even within video, various kinds need different delay variation guarantees. Stored video, such as instructional programming, can be delayed slightly, since it's not real-time. This buffering can smooth out delay variation.

Classes of service

The ATM standards define several classes of service, based on the parameters discussed above, and on others.

- CBR (Constant Bit Rate). Emulates a fixed bit rate, time division multiplexed circuit, with clock frequency and phase maintained end-to-end. Typical uses: emulating a leased-line circuit, and carrying traditional 64 Kbps PCM voice.
- rt-VBR (Real Time Variable Bit Rate). Clock frequency can vary, but maximum delay and maximum variation of delay between cells is guaranteed. Typical use: real-time videoconferencing.

- nrt-VBR (Non-Real Time Variable Bit Rate). Only mean delay is specified. Typical use: stored video.
- ABR (Available Bit Rate). The network attempts to maximize throughput; congestion control is provided through explicit rate flow control. Typical use: user workstations equipped with ATM network interface cards.
- UBR (Unspecified Bit Rate). "Send and pray"; each device can send whenever it wants, there is no congestion control, and when the network is congested it can either buffer or discard cells. Typical use: a LAN switch with ATM uplinks.

Application prioritization

A related issue, which is often confused with QoS, is prioritization. Take the example of two workers in the same company, one browsing the Web for the day's news, the other preparing a quote for a customer. The former will need more bandwidth, but it's likely that the latter communication is much more important to the organization. It would be optimal if the network could, when necessary, provide the more critical application with prioritized access to bandwidth.

Obviously, if bandwidth is effectively unlimited, this prioritization is unnecessary. Such a condition is possible, but not typical, especially as applications grow into a network's bandwidth. QoS does not address this issue, but various vendors have done so, and it seems likely that this will be an important issue over the next few years.

Physical interfaces

SONET / SDH

The ITU and BellCore have standardized a series of high-speed rates that are becoming important in the ATM world, both for wide area and local area communications. This is the first time in the history of networking that both areas have shared common rates.

- OC-1 (STS-1), at 51 Mbps, is essentially unused.
- OC-3 (STS-3) operates at 155 Mbps. It can operate over multimode fiber optic cable, single mode fiber optic cable, and unshielded twisted pair cable. Maximum cable distance in public specifications is 2,000 meters over multimode cable, 15,000 meters over single mode cable (intermediate reach optics), and 40,000 meters over single mode

cable (long-reach optics). However, in many cases these can be exceeded, depending on fiber, connector, and equipment characteristics. This is the most common ATM rate.

- OC-12 (STS-12) operates at 622 Mbps. OC-12 operates over multimode fiber optic cable and single mode fiber optic cable. Maximum cable distance in public specifications is 500 meters over multimode cable, and 15,000 meters over single mode cable (intermediate reach optics). As with OC-3, these specifications can be exceeded. As workstations move to Fast Ethernet at 100 Mbps, OC-12 will become increasingly common as a backbone rate.

- OC-48 (at 2.488 Gbps) and higher rates are rare at present, except within carrier network backbones; cable distances over 200,000 meters are possible over single mode fiber optic cable. It is likely that these very high data rates will become more important as increased throughput at the edge demands greater bandwidth in the core.

Traditional WAN interfaces

An existing carrier infrastructure makes it necessary to operate ATM over earlier types of circuits.

- DS-1 operates at 1.5 Mbps over unshielded twisted pair copper cable, and E1 at 2 Mbps over unshielded twisted pair or coaxial pair copper cable. These are relatively low rates for ATM, and some networks would use frame relay and traditional voice interfaces instead. Both operate over copper cable, although they can be multiplexed into higher-speed fiber optic circuits.

- DS-3 operates at 45 Mbps, and E3 at 34 Mbps, both over coaxial pair copper cable. Both are copper interfaces. These will rapidly be replaced by OC-3, which is often equally cost-effective for the service provider to deliver.

Workstation interfaces

There are some networks that use OC-3 over fiber optic cable to the desktop. But this is generally too expensive, especially since most organizations have fiber in risers, but not in station cabling. For networks that want to run ATM to the desk, there are three alternatives.

- TAXI was an early standard meant to be less expensive than OC-3 for campus connections. Maximum cable distance is 3,000 meters. It's not commonly used now.

- 25M (actually 25.6 Mbps) uses two pairs of unshielded twisted pair copper cable, including most of the wiring installed for data today. Maximum cable distance is 100

meters. Only a few customers and vendors have chosen to implement 25M ATM, in part because of the availability of inexpensive Fast Ethernet network interface cards and switch ports.

- OC-3 over UTP. Maximum cable distance is 100 meters, and it requires Category 5 unshielded twisted pair cable. This is a reasonable compromise of high throughput and support for much existing LAN cabling.

Network interface protocols

UNI (User-to-Network Interface)

UNI defines a set of protocols which operate between an ATM-equipped device (such as a workstation, server, or LAN switch) and an ATM switch. UNI is based on the signaling model in the ITU Q.2931 standard. UNI relies heavily on ATM's virtual circuit structure, implemented with VCIs (virtual circuit identifiers) and VPIs (virtual path identifiers); virtual paths are bundles of virtual circuits.

Signaling occurs between the two devices using a virtual circuit that is dedicated to this function. Typically, ITU-standard E.164 addressing (20 bytes total per address) is used, although two other addressing standards are also possible, and used, within the 20-byte address field.

UNI uses various types of AAL (ATM Adaptation Layer) parameters. For example, AAL1 is typically used for voice, and AAL5 for data. The end device and the network use these to validate the network's ability to support the QoS qualities required by the application.

UNI 3.0 and 3.1 are commonly used today; UNI 4.0 will add a number of capabilities, including support for ABR (Available Bit Rate), which provides very advanced congestion control, and much more flexible multicasting.

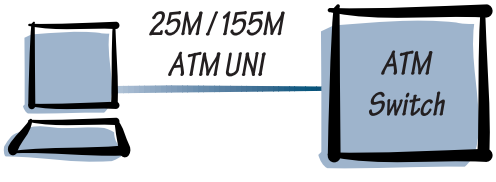


Figure 7.6. Original concept of UNI application

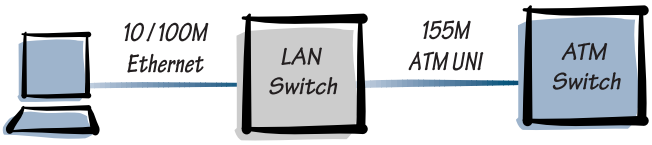


Figure 7.7. UNI as typically applied today

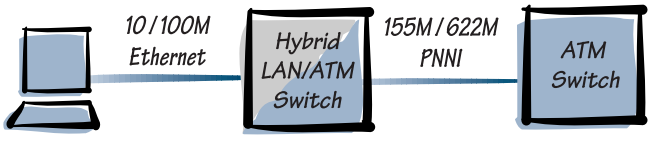


Figure 7.8. Powerful PNNI alternative

Inter-switch protocols

IISP (Interim Interswitch Signaling Protocol) is a limited protocol that operates between ATM switches, using statically defined connections. It has been largely replaced by PNNI (see below). IISP is sometimes called PNNI phase 0.

PNNI (Private Network to Node Interface) is an advanced, dynamic routing protocol that operates between ATM switches. It is based on link-state protocols, such as OSPF, with extensions that enable switches to advertise their own capabilities, such as capacity and delay. It supports a hierarchical network structure; topology information from one group of switches is aggregated and presented as a single node to the next higher-level peer group.

PNNI routes SVC (switched virtual circuit) requests through an ATM network, including support for QoS. It is able to automatically recover to an alternate link in the event of link failure, and to balance loads across multiple links.

PNNI operates automatically; no manual input is required to establish routes.

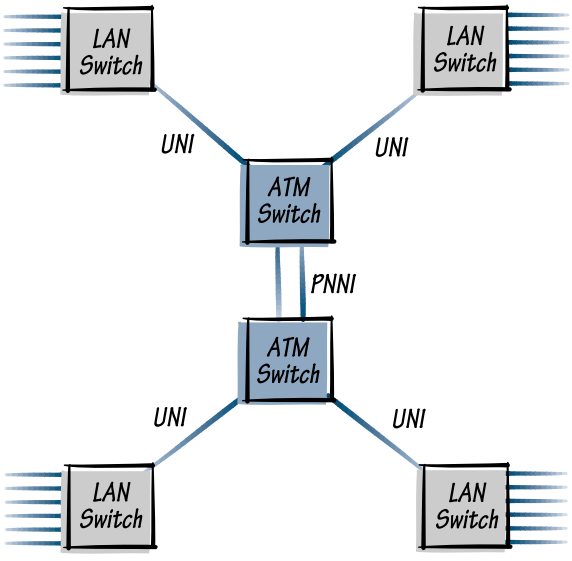


Figure 7.9. Wiring closet to backbone: star topology using UNI

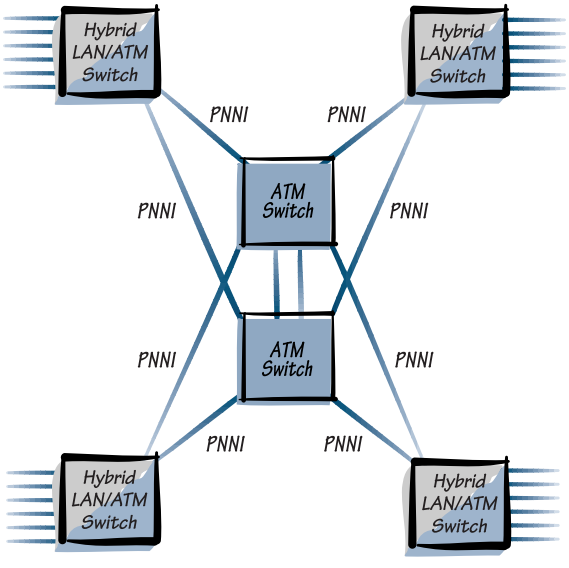


Figure 7.10. Wiring closet to backbone: mesh topology using PNNI

Inside an ATM switch

There are substantial differences in the way ATM switches have been architected. These can have a strong effect on performance.

Buffers

Early ATM switches were designed with several important assumptions:

- Most workstations would be equipped with ATM network interface cards. This implies that congestion control between a switch and a workstation is possible; ATM-equipped workstations can respond to congestion control requests, whereas Ethernet, Fast Ethernet, and token ring workstations cannot. This would minimize the need for buffers. But, as we know, the great majority of workstations connected to ATM backbones today are frame-based, and cannot respond to congestion control. So buffer requirements can be substantial.

- Video and voice traffic would be constant bit rate. This implies that video and voice connections would be given a guaranteed, fixed amount of bandwidth, so they would need almost no buffering; space would always be reserved on trunks. Instead, it appears that most video on an ATM network will be compressed, such as MPEG II. This implies variable bit rate and therefore burstiness: more buffering.

- A large percentage of network traffic would be voice and video. Instead, the explosion that's occurred in networking has been in graphical data applications, led by the Web. Naturally, all of this traffic tends to be bursty – and requires substantial buffering.

As a result of the assumptions incorporated in their architectures, many ATM switches tend to have less buffer capacity (measured in number of cell buffers) than would be optimal. This is likely to change as new ATM switches are developed.

Queues

Fabric blocking occurs when a switch's fabric capacity is less than the sum of its inputs. If so, the switch, even when lightly loaded, can drop cells. It's more difficult to build a really fast fabric, so most ATM switches are limited to a relatively small number of ports.

Many people assume that if the fabric rate of a switch equals or exceeds the sum of the input ports, no blocking is possible. Unfortunately, this is far from true. If a switch has a simplistic queuing structure, it's possible for cells destined for a congested port to block other cells,

which are sitting behind them at the same input port, but which are destined for other output ports. This condition is called *head-of-line blocking*.

To prevent this, switches need sophisticated input queues, able to sort out cells by destination, and optimally by QoS type as well. An ideal switch queuing architecture would provide, at each input port, a separate queue for each type of QoS supported at each output port.

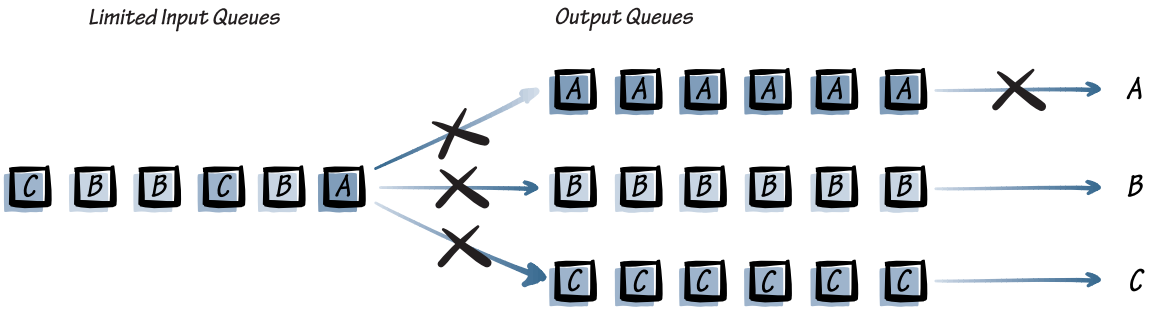


Figure 7.11. Head-of-line blocking with single queue

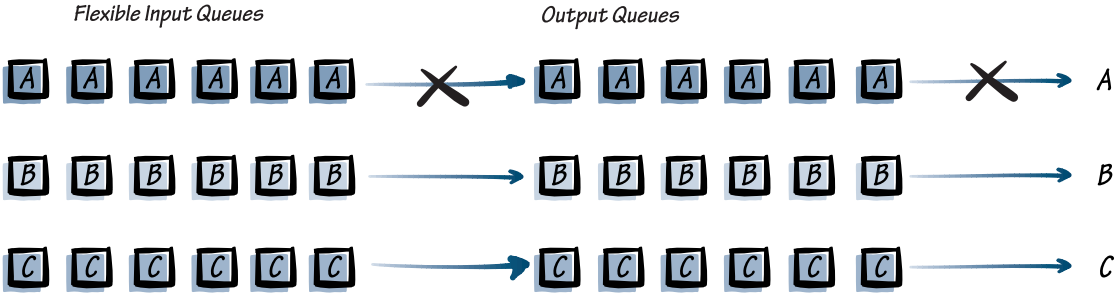


Figure 7.12. Avoiding head-of-line blocking with multiple queues

Cell discard

Even with extensive queuing and congestion control, a switch may sometimes receive more data than it can buffer. Additional cells will have to be discarded. With voice or video, just randomly discarding cells is a reasonable approach.

With data, however, there is a problem. The discarded cell was originally part of a frame, which probably was split into a number of cells. If even one of those cells is discarded, the source device will retransmit all of them. In fact, with TCP/IP, and most other protocols, every subsequent cell in every subsequent packet sent by that same source to that same destination will be retransmitted. Obviously, just randomly discarding one cell can result in a much larger load being placed on the switch; this is like the problem in the film *Fantasia* of the sorcerer's apprentice and the brooms.

The solution is to intelligently discard cells. Three mechanisms are available.

- EPD and PPD (Early Packet Discard and Partial Packet Discard). When an ATM switch is congested, these mechanisms use header information to locate cell "trains" (a cell train is a sequence of cells that originated as a packet). A cell train that has already entered the buffer is allowed to complete. New trains of cells are prevented from filling an already-congested buffer. EPD and PPD can dramatically improve frame-level throughput.
- RED (Random Early Detection). RED allows a switch to invoke PPD/EPD in a round-robin fashion, across all connections destined for a congested output port. This desynchronizes retransmission timers in layer-four protocols, thus dampening waves of retransmissions. RED can significantly improve overall network throughput.

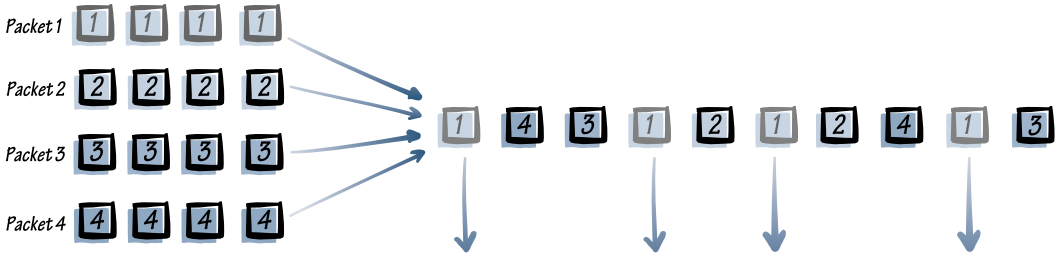


Figure 7.13. Intelligent cell discard

SAR

SAR (Segmentation and Reassembly) is the conversion of frames into cells, and cells into frames. In many implementations this is a function of a LAN switch with ATM uplinks; in others it is imbedded into the ATM switch itself. Early SAR chips could convert fast enough to fill, or almost fill, an OC-3 connection; newer OC-12 chips are now available.

Traffic management

One of the significant advantages of a connection-based network, like ATM and frame relay, is that traffic flows can be effectively managed, on a link level, and across the network. In ATM this function is provided in several interlinked protocols, including UNI, PNNI, and Traffic Management. There are two basic classes of traffic management: policing and congestion control.

Call admission control is the allocation of bandwidth to a connection as a part of the setup of the connection. It applies to QoS types, such as CBR and VBR, which must be guaranteed an exact or a minimal level of bandwidth.

Congestion control is more dynamic. An ATM device determines that it is in danger of exceeding its buffer capacity, and requests a connected ATM device to reduce or halt transmission. A major debate in the ATM world a few years ago centered around rate-based vs. credit-based congestion control; the rate-based scheme was adopted as being more implementable.

Effective congestion control, and sufficient buffer capacity, are essential in maintaining high levels of throughput in an ATM network. This is one of the significant advantages that ATM offers in comparison to Gigabit Ethernet, which has no mechanism for policing, and only an extremely limited link-level flow control capability.

Making connections

LAN Encapsulation Over ATM

IETF RFC 1483 describes a way to transport connectionless protocol units, such as packets and frames, across ATM's AAL5 service. It serves as a basis for various connections, including IP Over ATM, LAN Emulation, NHRP, and MPOA.

Classical IP

IETF RFC 1577 defines how to implement Classical IP on top of ATM Adaptation Layer 5.

RFC 1755, a companion specification, defines the ATM signaling involved in IP over ATM operation.

ATM LAN Emulation

There are some fundamental differences between LANs and ATM:

- LANs are connectionless; ATM is connection-oriented. A device on a LAN expects to be able to send to another device at any time. ATM requires that a connection be set up first.
- LANs use broadcasts for housekeeping functions; ATM relies on point-to-point and point-to-multipoint connections.
- LANs use six-byte MAC addresses. ATM uses 20-byte addresses.

As noted earlier, in a data protocol stack each layer relies on the services provided by the layer below it. So, for example, IP relies on the services provided to it by the LLC and MAC layers below it. ATM LAN Emulation was designed to provide these services in an ATM network that would not natively provide them. It's central to LAN Emulation that, whenever possible, data is switched directly between end stations using switched virtual circuit (SVC) connections.

There are three basic components in LAN Emulation:

- Emulated LAN (ELAN). A related collection of virtual circuits, clients, and servers running over an ATM network. An emulated LAN may be either Ethernet or token ring and appears to a protocol stack in a workstation or server to be a "real" LAN. There can be one or many emulated LANs running at the same time on an ATM network.
- LAN Emulation client (LEC). If a workstation uses an ATM network interface card, then the LAN Emulation client runs in the workstation. If a workstation uses an Ethernet, Fast Ethernet, or token ring network interface card, then the client runs in a proxy device, such as a LAN switch with an ATM uplink. There are typically many clients in a network. Note that there must be a separate LANE client for each emulated LAN supported by a workstation or switch.
- LAN Emulation server. This runs in a dedicated server, or in an ATM switch, or in a LAN switch with ATM uplinks. There may be only one LAN Emulation server in a network, or there may be several. It's possible for one server to back up another, for failure-resistance.

Within a LAN Emulation server, there are three separate processes. Each of these is referred to as being a server, and they could run on separate machines, but rarely do.

- BUS (Broadcast and Unknown Server). This takes a broadcast or multicast frame, or a unicast frame with an unknown destination address, and sends it to all members of an emulated LAN. It also stores the first frames sent in a unicast transmission, while the LES establishes a direct virtual circuit connection between the two end devices.
- LES (LAN Emulation Server). The LES translates MAC addresses to ATM addresses. Note that a LES is a software process that operates, along with other software processes, within a LAN Emulation server, which is a physical device.
- LECS (LAN Emulation Configuration Server). The LAN Emulation Configuration Server tells each LAN Emulation client the address it should use to access the LES for a given emulated LAN. In some vendors' implementations it may perform a number of other functions as well.

The current version of LAN Emulation is LANE 1.0. A new version – LANE 2.0 – was ratified in 1997 by the ATM Forum. It supports multiple LANE servers, including a protocol to operate between them so that the backup server can mirror the active one. It also supports efficient methods for handling broadcast-intensive traffic, such as multicast video, and adds hooks for QoS.

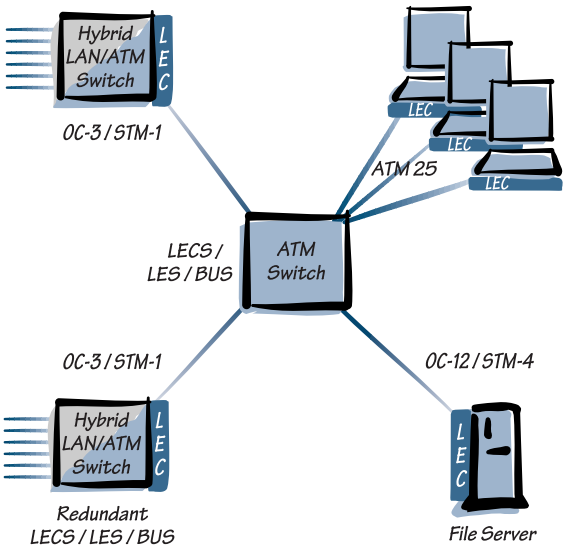


Figure 7.14. LAN Emulation model

NHRP

NHRP (Next Hop Resolution Protocol) is a proposed IETF standard designed to exploit the advantages of an ATM network for IP traffic, by minimizing the number of router hops in an ATM / IP network. Reducing router hops lowers latency and increases throughput.

In an IP network, the path between a source and a destination often passes through multiple routers; these are referred to as hops. If ATM is being used as the network backbone, each of these hops adds substantial latency, since traffic must be converted from cells to frames, routed, and then converted back to cells. This is obviously inefficient.

NHRP uses a client / server architecture. When an NHRP client, such as a switch in a wiring closet, wants to connect to another device, it asks an NHRP server to translate the destination IP address – which it knows – to the destination ATM address. To do this, a local NHRP server may talk to other NHRP servers. Once both end devices know each other's ATM addresses, they can communicate directly across switched connections. This is also called *zero hop routing*.

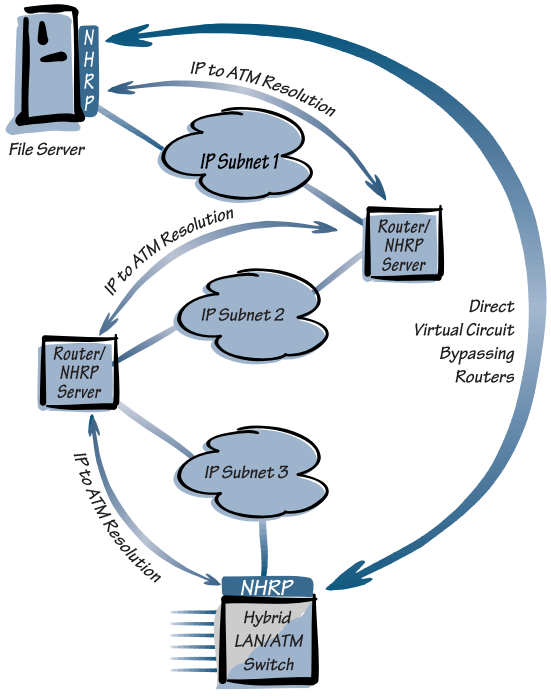


Figure 7.15. NHRP model

MPOA

Just as ATM LAN Emulation is a form of bridging across an ATM network, MPOA (Multiprotocol over ATM) is an ATM Forum standard that will eventually route multiple protocols in "native" mode over an ATM network.

MPOA is designed to bypass the limitations of two major alternative technologies: LAN Emulation and routing. LAN Emulation, like any bridging technology, is effective within subnet boundaries, but not between them. Emulated LANs must be interconnected with routers. Just as with LAN switching, some networks are too large to service with LANE alone. But routers, including layer-three switches, must operate on each packet in a stream. Even with layer-three switching this means some amount of latency, and some limitation to throughput. And because LANE "spoofs" upper-layer protocols into thinking they are operating over a LAN, it adds to protocol overhead and delay in the network. MPOA is meant to bypass this.

MPOA uses a route server to establish directly switched connections between end devices attached to the same ATM network. The route server can be a dedicated device, or it can be built into an ATM switch. It is able to communicate with traditional routers using standard protocols like RIP and OSPF.

Upper-layer protocols connect directly to the MPOA software, so that large packet sizes and QoS signaling can be supported.

How does MPOA work?

- When a workstation or server with an MPOA-compatible NIC – or a LAN switch providing an MPOA proxy function – needs to send a packet, it looks to see if the packet is contained in its local Internet address summarization group (basically a protocol subnet). If it is, the packet is bridged, using LAN Emulation.
- The MPOA client detects whether an IP flow is using an IP destination address which doesn't belong to its local subnet.
- A threshold counter is associated with each IP flow going to a different subnet. If the threshold (expressed in packets per second) is exceeded, a shortcut VC must be set up.

- The MPOA client checks the destination layer-three address, and looks to see if it has the corresponding ATM address stored in its cache. If it does, then it forwards the packet, using that shortcut VC.
- If the address is not in its cache, then the device contacts a route server. The route server maintains its own cache. If the destination is listed there, it tells the originating device immediately. If not, then the route server determines the address and notifies the originating device.

Note that the route server handles only addresses – not data. This is to maximize throughput, and to keep latency to a minimum.

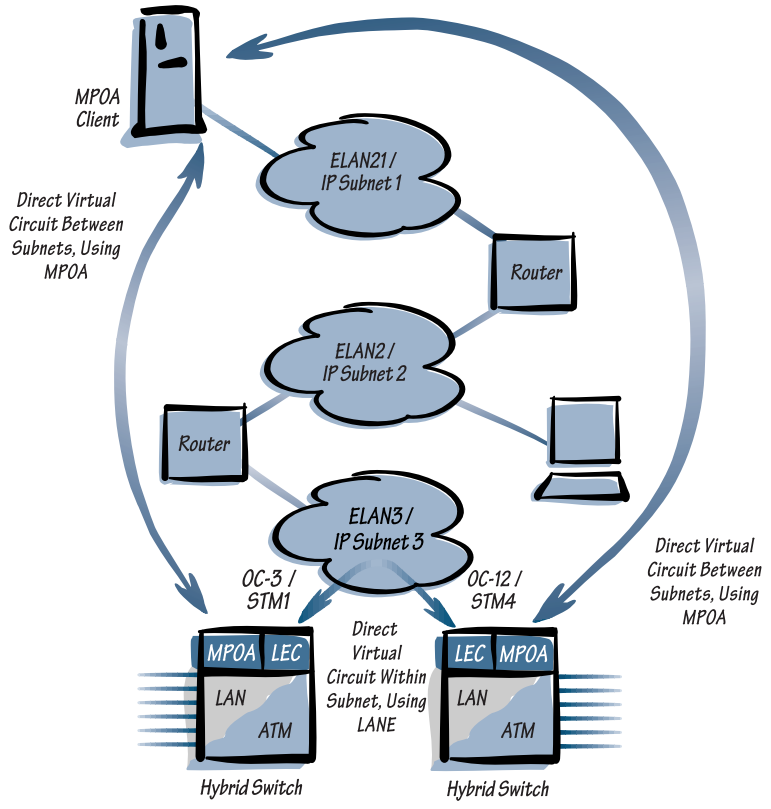


Figure 7.16. MPOA model

The future of ATM switching

A few predictions:

- ATM will continue to become less expensive, including desktop interfaces, OC-3, and OC-12.
- ATM use will continue to expand within carrier networks, eventually replacing older legacy data services, and ultimately unifying all transmission.
- Switching platforms will more tightly integrate ATM switching and LAN switching. This will permit protocols like PNNI to be used more extensively to build mesh networks.
- MPOA will gradually replace LAN Emulation for support of LAN protocols. Eventually, new layer-three and layer-four protocols that are optimized for a high-speed switched network will start to take over.
- Policy-based management will increasingly be used to set network configurations, such as LAN Emulation membership.
- Gigabit Ethernet and ATM will both be widely used as campus backbones. Any given campus will choose one or the other. However, a Gigabit Ethernet backbone may have uplinks into an ATM wide area network.
- ATM switching, LAN switching, and voice switching will be integrated into a comprehensive multimedia campus network that will replace current voice and data campus networks. This will be based around new switches, but may make use of existing switches.

Where is Gigabit Ethernet useful?

Until recently it appeared that ATM would provide the next level of throughput, eventually dominating networking, spanning both local and wide area needs.

ATM will certainly play an important role in the future of networking. Carriers will use it to create integrated multimedia wide area networks. And on the campus ATM offers a broad, powerful tool kit, especially for organizations that want to implement extensive interactive video.

But ATM is an entirely new technology. A user who adopts ATM must retrain network personnel, purchase new test tools, and perform complex equipment evaluations. And ATM is more expensive than some alternatives. Unless an organization has significant resources in both network planning and operations, this is a daunting task.

An alternative has arisen, using a traditional technique: Ethernet. Recent advances in semiconductor technology allow the Ethernet protocol to provide much higher throughput.

Gigabit Ethernet is a fairly simple concept. Take Ethernet, which runs at 10 Mbps, and multiply it by ten, and you get Fast Ethernet. Take Fast Ethernet, which runs at 100 Mbps, and multiply it by ten, and you get Gigabit Ethernet. The information that goes at the front of each frame (the MAC header) stays the same. One or two capabilities are added, but by and large it's just Ethernet. There are two big advantages to this:

- Network managers, network technicians, and vendor support engineers already know how Ethernet works. This reduces training costs and makes the organization's operational cost of a major network mistake less likely.
- It's easy for a switch to translate among Ethernet ports running at various rates. Since it's easy to do, it can be done at high rates, and at relatively low cost.

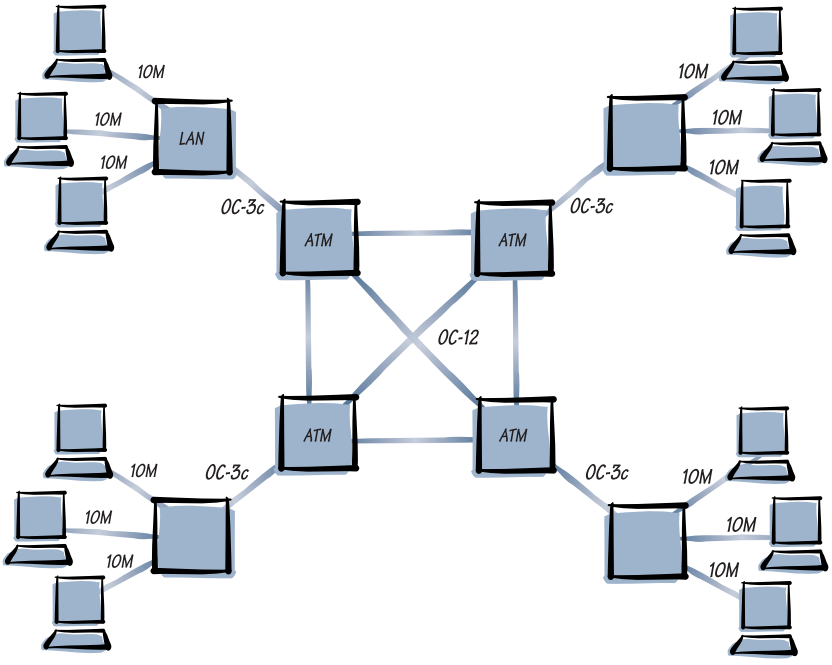


Figure 8.1. ATM backbone

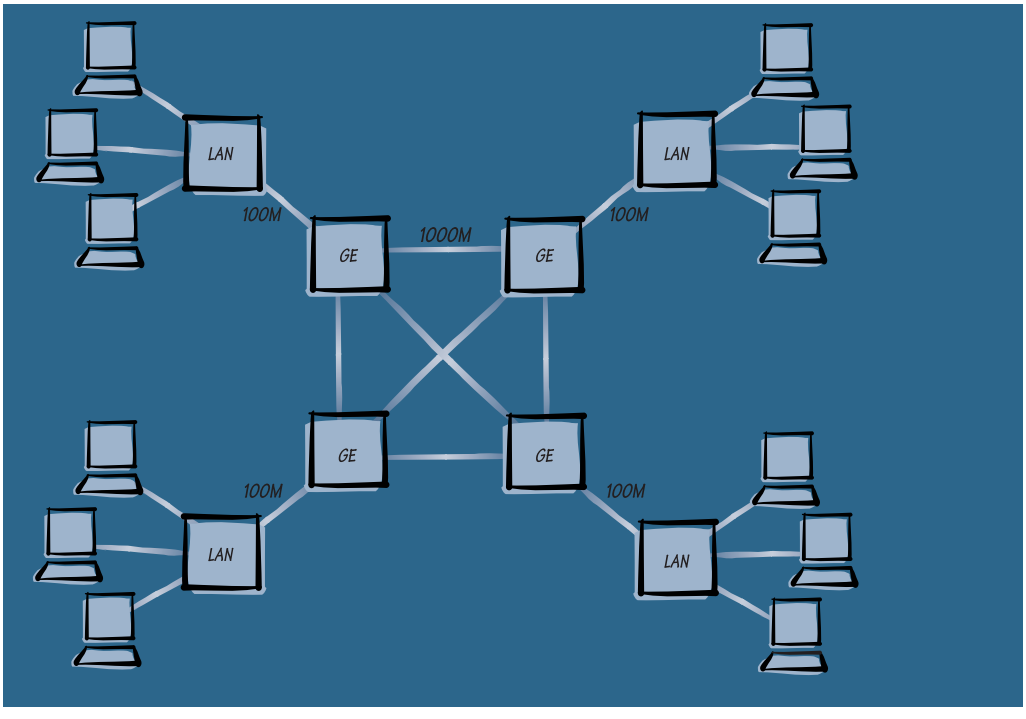


Figure 8.2. Gigabit backbone

Physical layer

Fibre Channel

Gigabit Ethernet is defined in the IEEE 802.3z specifications.

Gigabit Ethernet's fiber optic and twinax versions use the physical-layer specification of another high-speed standard called Fibre Channel. This made it easier to achieve a workable standard in a short time frame.

Cabling

There are several options for Gigabit cabling.

- ***Multimode fiber optic (1000Base-SX and 1000Base-LX)***. SX uses an 850 nm frequency, and LX uses 1330 nm. On cable with a 62.5 micron core, maximum distance for Gigabit Ethernet is 300 meters (the SX specification) or 500 meters (the LX specification). On cable with a 50 micron core, maximum distance for Gigabit Ethernet is 550 meters (the SX and LX specifications).

  Fiber has the inherent advantage of immunity to electrical interference, making it ideal in high-noise environments. Since many organizations have multimode fiber installed in building risers today, and since the backbone is the most important initial application for Gigabit Ethernet, multimode cable will probably predominate.

- ***Single mode fiber optic (1000Base-LX)***. Maximum distance for Gigabit Ethernet is 3,000 meters. The increase in distance over multimode is primarily due to lower dispersion in the cable.

  Given the distance limitations of multimode fiber, it's clear that inter-building cabling, and even some building riser cabling, will generally use single mode fiber. Since most users don't have this cable installed today, this will require cable installation. Any organization that is pulling fiber optic riser cable should consider pulling both multimode and single mode in the same bundle.

- ***Twinax (1000Base-CX)***. This uses a Fibre Channel standard for transmission over twinaxial copper cable at distances up to 25 meters.

- ***Unshielded twisted pair (UTP) copper (1000Base-T)***. This is the 802.3ab specification, which is far from complete. Assuming Category 5 cable (the only type supported in the specification), maximum distance for Gigabit Ethernet will be 100 meters. This is based on four-pair cable, with each wire carrying a 125 Mhz signal in each direction.

Most workstations will not need Gigabit Ethernet connections for some time, and not many workstation locations have four pairs of data cabling available. The most likely use for the UTP standard will be server connections, since many of these will be located in data centers where new cable can be easily pulled. On the other hand, fiber can also be easily pulled in a computer room. The choice will probably come down to cost and availability. Copper will probably be somewhat less expensive, but that remains to be proven. And fiber will be the first option actually available.

MAC layer

At the MAC layer, Gigabit Ethernet is identical to Ethernet and Fast Ethernet, with the exception that it uses a simple form of flow control, specified in the 802.3x standard. A device can send a command that tells the device at the other end to stop transmitting, and later send a command to tell the other device that it can start sending again. This somewhat archaic form of flow control may work well in simple networks. It is uncertain how well it will work in more complex nets.

Gigabit and layer-three switching

Gigabit Ethernet is essentially a high-speed version of Ethernet, so a Gigabit Ethernet network without routing would be one large broadcast domain. For a mid-sized network this might be quite acceptable; the total number of broadcasts would not burden the stations. To the extent that workstations are operating on Fast Ethernet, the broadcast burden would be even lower.

For a larger network, Gigabit Ethernet, like other LAN technologies, operates better with some form of broadcast control through subnetting. Since gigabit rates are so high, traditional software-based routers could not generally handle the packet loads. So layer-three switching is a natural component of a gigabit network. In general, Gigabit Ethernet switches have layer-three switching built in, generally at a rate greater than one million packets per second.

Gigabit and QoS

There is a very active debate in the networking industry over QoS and Gigabit Ethernet. The current Gigabit Ethernet draft standard contains no explicit support for QoS. Some argue that, given enough bandwidth, explicit support for QoS is unnecessary; when there is no contention for link bandwidth, or for bandwidth in a switch, then every packet will be delivered at the network's maximum speed. In this simple form the argument has little merit;

variable packet sizes, bursty networks, and the need to control delay variation as well as delay mean that, at least for certain types of traffic, explicit QoS support is needed.

However, the high bandwidth in a Gigabit Ethernet network certainly helps in delivering QoS. Several vendors have moved ahead of the standard in implementing prioritized queuing and other mechanisms. Early implementations in large real Gigabit Ethernet networks may help clarify which techniques are useful in which applications.

Gigabit and ATM

Let's set up a scorecard for Gigabit Ethernet and ATM in campus backbone applications.

Report Card	Gigabit Ethernet	ATM
Cost	B	C
Familiarity	A	C
Standards Marurity	B-	B
QoS Support	C	A
Support of Ethernet Workstations	A	B
Support of Token Ring Workstations	D	B
WAN Integration	C	A

Figure 8.3. Gigabit vs. ATM: a report card

Cost

Although ATM costs have declined, it is still much more expensive than Gigabit Ethernet. OC-12 switch ports often cost about three times more than Gigabit Ethernet ports, and OC-3 switch ports often cost about three times more than Fast Ethernet ports.

Familiarity

Gigabit Ethernet is essentially very fast Ethernet, and it's typically coupled with layer-three switching, which is essentially very fast routing. So there's not much new to learn for most users. ATM is complex; the complexity is driven by the powerful capabilities built into the standard, but some users don't need those additional capabilities.

Standards maturity

The great bulk of what most users will need from ATM is already completed in the ATM Forum. Gigabit Ethernet is somewhat further behind, but the simplicity of that standard effort argues that it should be relatively smooth.

QoS support

ATM has a demonstrated capability to carry a wide range of information types, including those needing guarantees of low delay and low variation of delay. It is possible that prioritization in Gigabit Ethernet switches will deliver, at least for most users, practically the same capabilities. But this will not be certain until a number of large Gigabit Ethernet networks, carrying substantial amounts of real-time traffic, are actually running.

Support of Ethernet workstations

Gigabit Ethernet naturally has an advantage, since the similarity of its MAC to that of Ethernet makes switching between the two protocols extremely simple. ATM, although different from Ethernet, has explicit and effective support for Ethernet built into both LAN Emulation and MPOA.

Support of token ring workstations

Here Gigabit Ethernet is at a substantial disadvantage. Only a few vendors have announced plans to support token ring across gigabit-rate links, and if a standard for this emerges, it will take some time. The reason more vendors are not proceeding on this capability is not so much technical as market-based. Since most token ring users appear to be migrating away from the technology, it is less profitable for vendors to develop very high-speed token ring.

ATM supports token ring in both LAN Emulation and MPOA.

WAN integration

As noted above, ATM is increasingly used by carriers as the basis for high-speed networks, especially in a metropolitan area. LAN / WAN integration is relatively straightforward, if both the campus backbone and the high-speed access to the service provider's network are ATM. Gigabit Ethernet's distance limitations make it infeasible as a wide area technology. However, some Gigabit Ethernet switches will provide an ATM gateway with a high-speed SAR process to connect the two.

The future of gigabit switching

A few predictions:

- Most early Gigabit Ethernet switches will offer little more than high speed and layer-three switching for IP. This will be too limited for many applications. Services and protocols will differentiate gigabit switches for some time, especially because almost all of these switches will have enough switching capacity for almost all applications.
- Both Gigabit Ethernet and ATM will be widely used, with campuses standardizing on one or the other.
- Gigabit switches will incorporate ATM uplinks to the wide area.
- Layer-three switching will be a standard feature of Gigabit Ethernet switches. This will be primarily hardware-based routing.
- There will be some vendor support for Gigabit token ring, although a standard is unclear.

Intelligent fabrics: why switching and services are linked

When the network infrastructure was relatively weak, with only a few processors for a large base of workstations, it was reasonable that the network would focus simply on moving information effectively from one place to another. However, we are now seeing the emergence of intelligent fabrics, which have huge amounts of aggregate processing power, distributed across LAN switches, Gigabit Ethernet switches, layer-three switches, ATM switches, and dedicated servers. This fabric is capable of providing a wide range of services. These services will be the single most important trend in networking over the next several years.

Authentication services

Somehow, users need to be granted or denied access to certain resources, such as servers that handle financial and human resource information, based on their identity.

It's reasonable to base that access on a machine identity, if that identity is constant. But with DHCP, IP addresses are handed out dynamically, so these are ineffective for authentication. A MAC address can be spoofed, and a MAC address changes every time a network interface card is replaced. And in many networks a user may show up in various places, at various times – sometimes even using different workstations. So a physical location in the network may be inadequate for authentication.

The alternative is to authenticate the user dynamically, by querying her / him, obtaining a user ID and password, and matching it against a database. The database would contain a list of resources to which that user is allowed access, and the network would then permit those connections.

One of the advantages of this process is that it occurs once at the beginning of a session, and can effectively control access to high-speed services. This is impossible for a firewall, which is inherently rate-limited, since it uses packet-by-packet filtering.

Firewall services

For access to the Internet, for dial-in routers, and for certain other applications, a more intensive validation process is useful. This is provided by an IP firewall, which is essentially a packet filter that compares each packet against a set of filters or policies. Some firewalls keep track of connections in a state table, and can determine whether a packet is a valid part of a proper flow.

Firewalls are often implemented today in standalone computers and in software-based routers. As these traditional routers are phased out in favor of layer-three switches, and as switch hardware becomes increasingly powerful, firewalls will tend to move in the switching fabric. This will reduce costs and make management easier, especially if firewalls are used, not only for Internet access, but also for protection of internal resources.

Mobility services

In a traditional workplace a user came to work, sat in an assigned office to work, sat in conference rooms just to talk, and went home. All connections into the network were from that office.

In many organizations this has changed; users now carry portable computers, and they may plug into the network in another cubicle, in a conference room, at another location of the organization in a distant city, dialing in from a hotel room, or telecommuting from home. Some of these connections are even wireless, with very little physical identity to the transmission. Wherever they are located, these users need to have access to the same resources as they would in their offices.

A number of tools must interoperate in order to make this possible. IP firewalls, dial access security, user authentication, layer-three address distribution, and directory services all make mobility safe and convenient.

Address services

Address assignment

Layer-three addresses, such as IP addresses, were originally tied closely to a user's physical location in the network. The subnet portion of the address would equal an Ethernet segment or a token ring. This was inconvenient when users moved around a network, since the subnet address was tied to a specific physical location, and had to be reassigned every time the workstation was moved. And, as noted above, a correspondence between subnet number and router port becomes meaningless in a fully switched network.

For TCP/IP, the problem is addressed with DHCP (Dynamic Host Configuration Protocol), which automatically assigns an IP address when a workstation appears on the network. In order for this to happen flexibly, the switched network needs to support DHCP Relay, which is based on the earlier BootP Relay; this allows a DHCP server to provide addresses to machines across subnet boundaries.

Over time, this function will be integrated into the switching fabric. By doing so, the network is able to respond more quickly to address requests, and with greater failure resistance.

Address translation

Another limitation in the current IP subnet structure is its rigidity. Since almost every network that runs TCP/IP connects to the Internet, the same subnet addresses are used for both internal and external communication. This is a problem, because the number of subnet addresses is limited and dwindling, and intra-organization communications can be made more difficult because the assigned subnet addresses are in non-contiguous blocks.

A solution to this problem is NAT (network addresses translation), which converts from one set of addresses to another. This allows one large, contiguous address space to be used within the organization, and another InterNIC-assigned address space to be used for Internet connections. Today this function is present in some IP firewalls and in some standalone devices. Over the next few years it will become a common capability in switching fabrics.

Directory services

When a computer that runs TCP/IP, NetWare, or another routable protocol needs to communicate with a process on another machine, it must first know the destination's network address. Making these addresses readily available, and coupling them with a variety of other information, is a rapidly advancing area of networking.

DNS

DNS (Domain Name System) is an IETF standard (RFC 1035 and later updates) that provides a globally-accessible table of domain names and their corresponding IP addresses. When a workstation or server knows a domain name of a destination, and needs the IP address, it uses DNS to look it up.

LDAP

DNS is fine if the process which needs an IP address already knows the domain name. But what if the domain name is not known? For example, what if an email user needs to look up another user's name and email address?

LDAP (Lightweight Directory Access Protocol) is a recent IETF standard (RFC 1777 is the latest at this writing) for storing information in a central directory that is shared by many different services. It is a subset of DAP (Directory Access Protocol), which is part of the ITU X.500 protocol. An LDAP directory, which is hierarchical, can be distributed among many servers. Netscape, Microsoft, Oracle, Sun, Novell, and a number of other vendors all support LDAP.

With LDAP, a workstation user can access email, group calendars, discussion groups, and intranet access with a single ID and password; all these applications would draw from the same LDAP database for authentication, user preferences, and quality of service. The directory can also be useful for mundane information like employee phone numbers, email addresses, and organization structures.

LDAP is to a computer network as a phone book (white and yellow pages) is to the telephone network.

By integrating an LDAP client into a switch, and refreshing it from a central LDAP server, it's possible to dramatically reduce the time needed to take advantage of directory services. It also allows the switch to use information stored in the directory for switching decisions.

Prioritization and QoS services

As noted above, a number of networks (notably ATM) devote considerable attention to supporting the quality of service characteristics needed by some applications, especially interactive video and voice. This process is somewhat more straightforward in a connection-oriented network like ATM or frame relay, since the flows are more deterministic.

An essential element in QoS provisioning is the signaling between the application (running in a workstation or server) and the network. This is the weakest link in the QoS chain today; applications and operating systems have not generally been enabled to request QoS characteristics.

RSVP (Internet Resource ReSerVation Protocol) is an IETF protocol that allows a device to ask the network to provide specific QoS characteristics for an application which the device wants to run. RSVP carries the request through the network; at each node through which the application's data will pass, it requests a resource reservation. It is designed to support both unicast and multicast connections. Although it was originally designed for use across the Internet, RSVP, like many IETF protocols, is useful across other networks as well.

RSVP is included in Microsoft's Winsock version 2.0, and should become much more widely used when it's integrated into Windows NT. This is scheduled for release 5.0 of that operating system.

Most of this book focuses on the use of switching in local networks. But almost no campus networks exist in isolation, and remote connections are needed to other sites (large and small) and to individual users at remote locations. Sometimes these connections are relatively transparent as when a campus network with an ATM backbone uses high-speed ATM connections across the wide area to other campuses. At other times, substantial conversion is needed to accommodate differences in bandwidth and protocols.

- Bandwidth in a LAN is limited only by the capital cost of electronics and cable. Bandwidth in the WAN is paid for by the month, and is very expensive.
- Building a LAN involves an end user organization and its hardware and software suppliers, and often a network integrator. A WAN needs all of these, but adds one or more carriers. This makes standards even more important in the WAN than in the LAN, since carriers must base their services on standards.

A specialized product, such as a router or an ATM edge switch can provide the interface between a switched campus network and a wide area network. Or WAN interfaces can be integrated into the campus backbone switches, providing a more cohesive set of services.

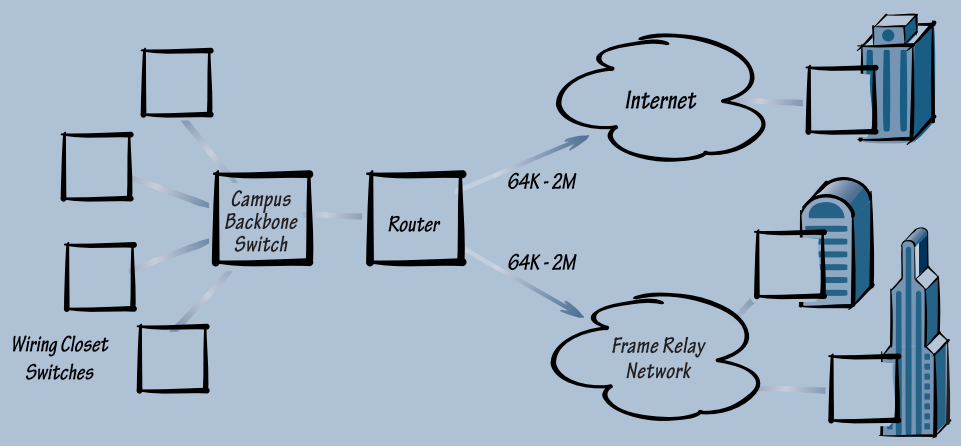


Figure 10.1. Wide area connections using a standalone router

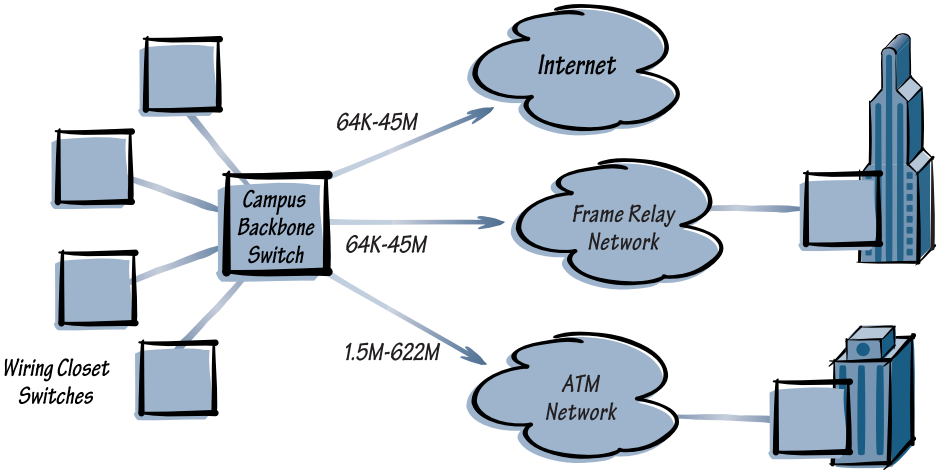


Figure 10.2. Wide area connections integrated into the campus switching system

Frame relay

Frame relay is a packet-switching service. Individual packets are moved from one frame relay switch to another, eventually reaching their destination. The standard specifies what happens at the edge of the network, where the subscriber's equipment connects to the carrier. Frame relay typically operates at rates from 56 / 64 Kbps to T1 / E1.

Frame relay is gradually replacing two earlier technologies, X.25 and leased lines.

- X.25 is a packet-switching standard that forms the basis for many public networks, especially outside of the United States. X.25 checks for errors at each switch, while frame relay relies on the end stations to do this, assuming that the underlying circuits will have relatively low error rates. Therefore, frame relay has higher throughput and lower latency than X.25, and is generally offered at much higher data rates.

- Leased lines give the end user complete control over the line facility; the carrier simply delivers bits at the destination at exactly the same rate as they are transmitted at the originating end. This has traditionally been the most common form of WAN in the United States, and is widely used elsewhere as well. Its limitation is low efficiency: circuits are not shared, and on average run at a very low rate of utilization. Conversely, since many users' data can share a frame relay circuit, it tends to be much more efficient – and therefore less expensive.

Because frame relay uses packet switching, it also has another advantage over leased lines: each location can connect with the data rate that it requires. For example, an insurance firm's central site might use one or more T1 / E1 connections, regional offices 384 Kbps, and branch offices 56 / 64 Kbps.

The principal difficulty when connecting a campus switching network to frame relay is the difference in data rates. A single gigabit circuit, for example, is 100 times as fast as five E1 lines. There are several techniques for addressing the difference in rates:

- Routing helps over low-speed connections by minimizing broadcast traffic. The standard for routing across a frame relay network is IETF RFC 1490.
- Virtual LANs can also limit the amount of broadcast traffic that traverses the wide area.
- Spoofing a protocol (like SAP or RIP) means sending updates across the wide area only when a change occurs; until it does, the remote end just keeps issuing the same broadcast. This can dramatically reduce broadcast traffic, especially in a network with many servers.
- Compression can increase effective bandwidth when the information to be sent can be coded in a more efficient way. This is almost always true with text and with software code; it is less true with previously compressed graphics, such as JPEG files. Compression based in software is generally effective only over links operating at 128 Kbps and below; above this, hardware-based compression is needed, and can be effective at very high rates.

A fundamental advantage of compression is that it operates on all data, not just broadcasts. Generally, a layer-two switching process with compression will be much more effective across a low-speed circuit than will a layer-three switching process without compression.

Wide area ATM

As we discussed above, ATM was originally designed by carriers as a complete network for voice, video, and data. It still provides carriers all of the advantages for which it was designed.

ATM is offered in the wide area in several service forms.

- The simplest is ATM bearer services, in which the carrier provides switched ATM connections among various points. At present, most such services are rate-insensitive; the customer pays a fee based on the rate of the access line, rather than on the number of cells they actually send.

ATM bearer services tend to provide relatively low profits to the carriers, especially in early implementations in which the cost of ATM carrier infrastructure is spread across a few customers. And they require the customers to become sufficiently familiar with ATM technology to be able to source and configure access equipment.

- A number of service providers are delivering TLS (Transparent LAN Services) across ATM. In the simplest form of TLS, the carrier installs ATM circuits to each of a customer's sites, and installs a LAN / ATM switch at each site. The customer connects each LAN into a switch port, and the carrier then manages the network on a turnkey basis. This allows the carrier to charge more, since it is providing a service, rather than simple bandwidth. And the customer need not be expert in ATM; indeed, they need not even be aware that the underlying bearer service is ATM.

- One variant of TLS extends to providing access to various services. The most common is Internet access. Since TLS access rates are very high (typically 34, 45, 155, or 622 Mbps) Internet access at LAN rates becomes possible.

- Another variant is to add voice and/or video support, typically using Circuit Emulation, an ATM Forum specification for sending fixed-rate voice and video across an ATM network. This helps justify the cost of the ATM circuit. For example, ten PBX DS-1 trunk lines, carrying 240 voice channels, is 15 Mbps of bandwidth – only 10% of the bandwidth in an ATM OC-3 circuit.

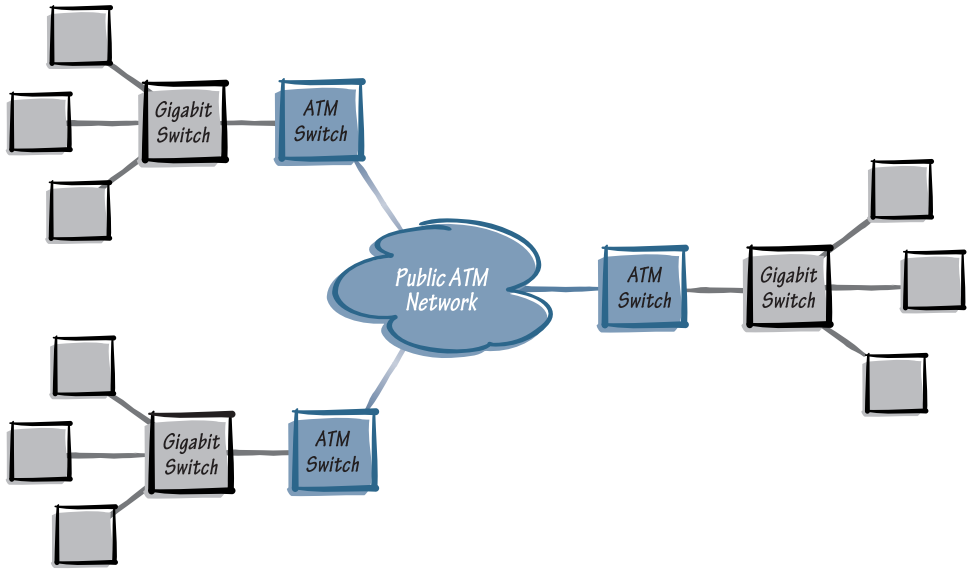


Figure 10.3. ATM bearer service

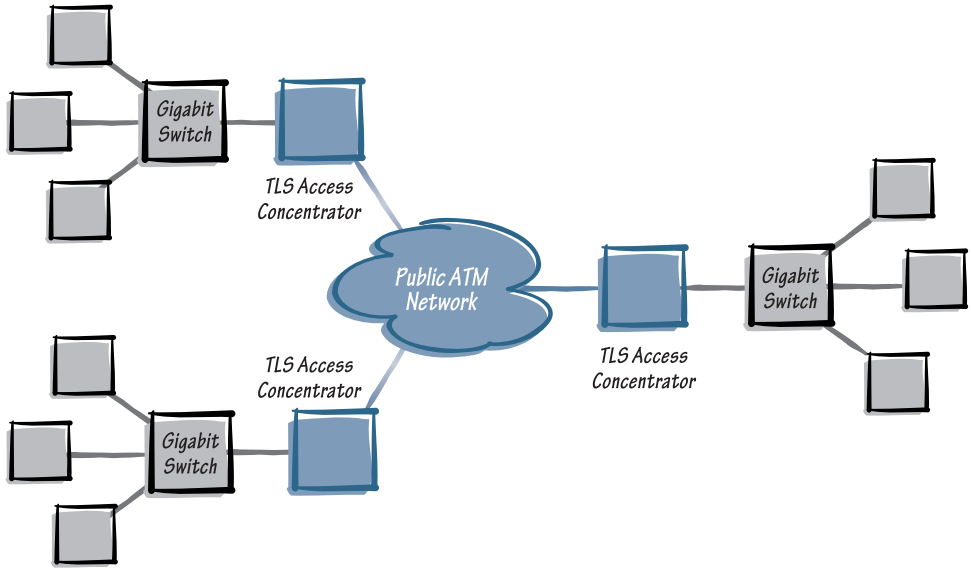


Figure 10.4. Transparent LAN service

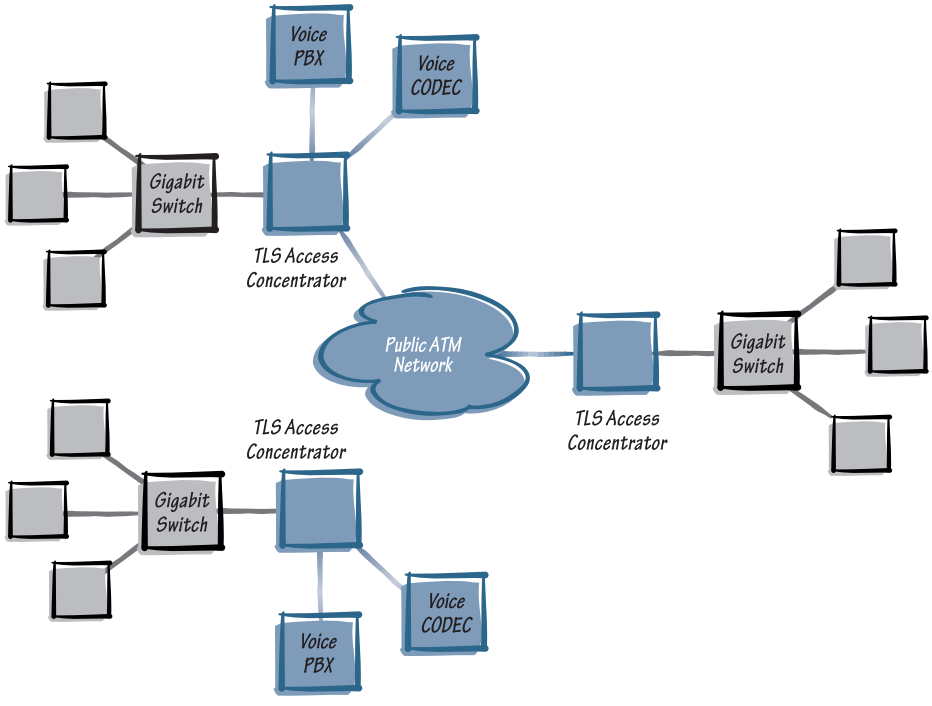


Figure 10.5. Transparent LAN service + circuit emulation

ISDN

ISDN was the precursor to ATM; it supports voice and data on digital lines operating at 64 Kbps to 2.048 Mbps. Because of its complexity, and because bulk data transmission quickly moved to higher rates, ISDN has never been broadly successful. However, remote subscriber access to the Internet has grown rapidly in the last few years, providing an important new application for ISDN.

ISDN is a circuit-switched (dial) technology; it is most cost-effective when used for short periods. Its principal application to switched campus networks is as a backup to frame relay, ATM, or leased lines.

The Internet

Obviously, almost every campus switched network must be connected to the Internet. Several functions are needed to accomplish this safely.

- IP routing capability.
- An IP security firewall. As noted above, a firewall is essential for a safe connection to the Internet.

- A wide area protocol, typically PPP (Point-to-Point Protocol), although frame relay is an alternative.
- Serial interfaces operating at appropriate access rates, typically from 56 Kbps up to DS-1 in the United States, and from 64 Kbps up to E1 elsewhere, although higher rates are possible.

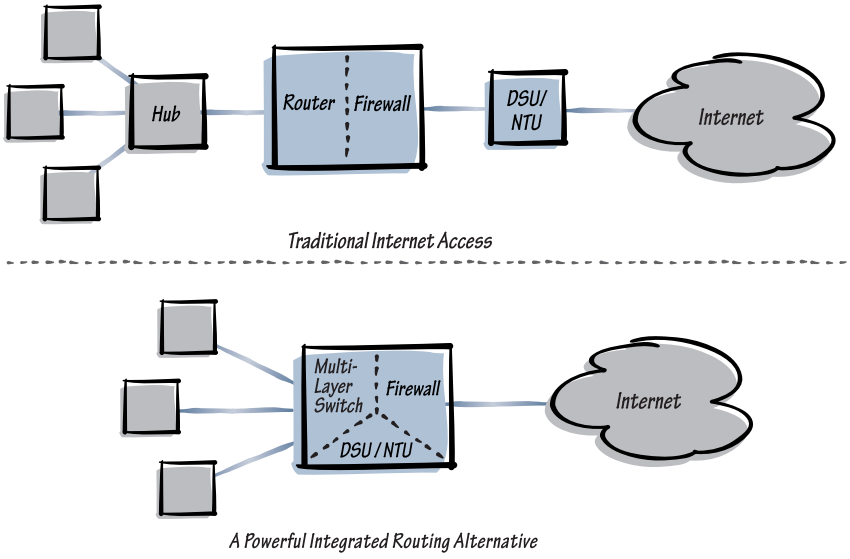


Figure 10.6. Internet access basics

There are basically two ways to install these services. A traditional router with firewall capability can sit between the switched network and the Internet. Or the capabilities can be built into the campus switches themselves; this reduces latency and simplifies management.

The basic problem faced by managers of switched campus networks is the difference of three to four orders of magnitude between switched campus fabrics (with many gigabits per second of bandwidth), and serial access rates (typically at one or two megabits per second). Alternatives are emerging that address this differential. These sometimes offer the end user the ability to dynamically select from one of several Internet service providers. For example, as noted above, wide area ATM services, including Transparent LAN Services, offer DS-3, E3, and OC-3 connections to the Internet.

Background issues

SNMP and standard MIBs

In the early days of networking, all management was based on proprietary systems developed by equipment vendors, and a management system would rarely support another vendor's equipment. This changed almost overnight with the introduction of SNMP by the IAB, which created a basic standard adhered to by the entire networking industry. It defines a number of elements of management.

- Management information is stored in a MIB (management information base); the MIB is stored in an agent in the device being managed, and in the central management station.
- A management station can query devices for status and statistics. It sends a "Get" when it wants to know the state of something in the managed device, and it sends a "GetNext" when it wants to know the state of the next item (in a series) in the managed device.
- A management station can configure devices. It sends a "Set" when it wants to set the state of something in the managed device.
- A managed device sends a "Response" to respond to one of these requests.
- A managed device can send unsolicited information about a change of state, by sending a "Trap" to the management station.

An important reason for SNMP's success is that the data items it defines are highly extensible, through a series of standard MIBs, and through the ability of a vendor to add its own proprietary MIBs to cover items not yet standardized. Virtually every manageable device in the world of switching supports SNMP.

SNMP operates over TCP/IP's *UDP* protocol; each message is sent as a separate datagram, without guarantee of delivery. It uses a very simple structure, as its name implies.

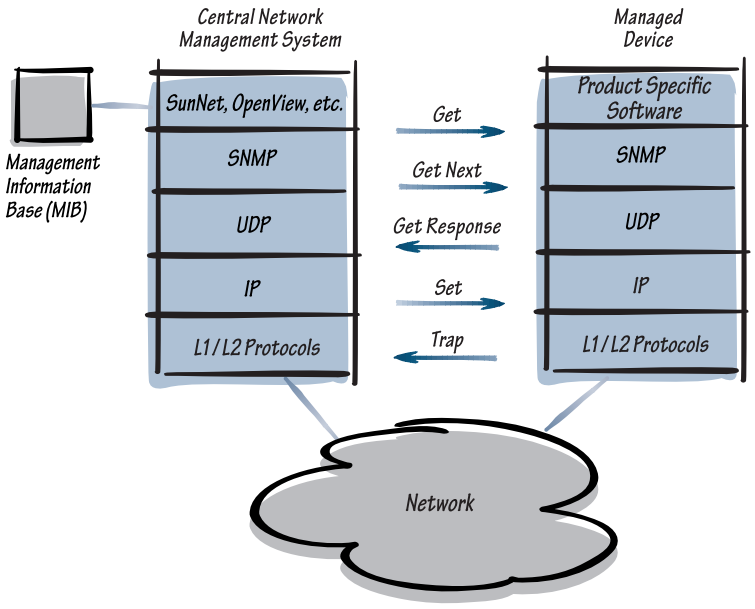


Figure 11.1. SNMP model

Standard management software platforms

SNMP does not define the user interface to the management station. This is handled by several standard management platforms, which provide an overall software platform, and which define standards to which vendors can write in creating management application software. The most successful of these platforms are H-P's OpenView, Sun's SunNet Manager, IBM's NetView for AIX, and Cabletron's Spectrum. Many equipment vendors write their management applications so that they can operate on more than one of these platforms. This allows users to mix applications from multiple vendors on a single platform.

RMON

There are various devices that connect to a LAN and measure traffic characteristics; these are often referred to as monitors. When they are also able to capture and decode traffic, they are called network analyzers. Until several years ago, a network manager had to physically attach one of these devices to a LAN segment.

The creation of a standard SNMP MIB called *RMON*, in IAB RFC 1757, changed this. RMON is a standard, structured method for retrieving and sending data from a remote monitoring agent.

Although RMON could be used in a standalone monitor, it was also incorporated into hubs, and later into switches. Virtually all LAN switches now include an RMON capability, which is able to observe every port in the switch and report statistics and alarms to one or more network management stations. SNMP agents that support RMON can support one, several, or all of the optional nine groups for Ethernet and ten groups for token ring. Since supporting all groups of RMON takes a considerable amount of memory, most switch vendors support just four groups of RMON.

Management of switched networks

Management of networks, especially large networks, has always been difficult. On one hand, insufficient information makes it difficult for managers to analyze performance and isolate problems. On the other, excessive information overwhelms the manager with too much data – the "sipping from a fire hose" problem.

Switches can make networks more difficult to manage.

- Switches can increase the amount of management data enormously. For example, in a hub, configuration is primarily a matter of turning ports on or off. But in a LAN switch, each port has a number of items that can be configured. And it is typical for every port in a LAN switch to be measured with RMON, producing huge amounts of statistical information. ATM switches also present a great deal of management data.
- Virtualization of broadcast groups (through VLANs or ATM emulated LANs) can be difficult to conceptualize. Traffic flows do not map as readily to physical paths as in a hub / router network; some traffic across a given link might be switched at layer two, while other traffic is being routed at layer three.

On the other hand, switches can also ease the manager's tasks.

- Connection-oriented networks, in which data flows between two devices on a deterministic path, are inherently easier to troubleshoot when a problem arises, compared to datagram networks in which packets could flow over various paths.

- Switches automatically solve some problems that would otherwise require trouble-shooting. An example noted earlier is beaconing in a token ring network, which can disable an entire token ring because of a poor cable connection at one workstation. Token ring hubs run complex diagnostic patterns to try to isolate beaconing, bringing down all workstations and then bringing sections back up. A token ring switch simply observes the beacon condition from the device that is sending it, blocks the beacons from damaging any other devices, and notifies the network manager.
- Switches incorporate more intelligence than hubs or bridges, and can provide more information to the network manager.

Tracking virtual groups

A common misunderstanding of virtual LANs is the idea that using them will result in unpredictable flows of traffic around a network. Actually, whether data traffic moves within a subnet (intra-VLAN) or between subnets (routed), the great bulk of the data is unicasts, rather than broadcasts. So data loading on the network is fundamentally unaffected.

However, virtual LANs do make it more difficult to determine the location of members of a subnet, since the subnet is no longer defined physically, as it is in a hub and router network. Management tools can solve this. What's needed is a database which can automatically determine the members of a virtual LAN, along with their MAC addresses, layer-three addresses, and physical switch ports, and which can display this information in a hierarchical tree structure.

Policy-based management

Information technology is in large part a progressive process of abstraction and simplification. Only a few years ago a file was explicitly downloaded using FTP; now we just click on an icon or filename, and our Web browser does the rest of the work. Policy-based management is a powerful form of simplification. The network manager makes general decisions, and the network infrastructure then intelligently implements those decisions.

For example, a manager might decide that all AppleTalk machines should be in a single virtual LAN; setting this policy could take less than a minute. The switched network would then examine every workstation and server in the network, determine those which use AppleTalk, place them into a single broadcast domain, and notify the network management system of what it had done. The entire process can take less than 5 minutes – for the entire network.

Policies are useful in many areas of network configuration, including firewalls, maintaining software on switches, assigning QoS and priority levels, and gathering network inventory information.

RMON

The RMON (Remote Monitor) management specification was designed before switching became widely used, and is oriented largely at the interchange of information within an Ethernet segment or a token ring. But in a fully switched network there is only one device on each segment or ring; what's needed is a management function that looks at the movement of information within a subnet, and between subnets.

Eleven things you may not know about Xylan

1. Xylan's OmniSwitch is the most complete networking switch in the world, integrating LANs, ATM, Gigabit Ethernet, frame relay, layer-two switching, layer-three switching, ATM switching, firewalls, and authentication.

2. Xylan has installed more than 1,000,000 switch ports worldwide.

3. Xylan was the first company to successfully integrate Ethernet, token ring, FDDI, and ATM in a single switch.

4. Xylan was the first company to integrate an IP firewall into a multi-layer switch.

5. Xylan was the first company to integrate frame relay into a multi-layer switch.

6. Xylan was the first company to integrate user authentication services into a multi-layer switch.

7. Xylan is the fastest-growing hardware company in networking; as of this writing, after 40 months of product shipments, Xylan is shipping at an annualized rate of over $300 million.

8. Xylan was the fastest-growing company of any kind in California in 1996, according to the Los Angeles Times.

9. Xylan has its own sales and support offices in over 80 cities around the world.

10. Xylan has partnerships with companies like Alcatel, Anixter, Check Point, IBM, LSI Logic, Samsung, Sun Microsystems, Tech Data, and Unisys.

11. Thousands of networks, including major corporations, large universities, and governments have standardized on Xylan's OmniSwitch as the basis of their switched network. Some of these are among the largest switched networks in the world. Xylan customers include 3M, Aetna, Bell Atlantic, BellSouth, General Motors, GE Financial, Getty Center, the government of France, Hanjung, Kuala Lumpur International Airport, Lockheed-Martin, Mitsubishi, Rogers Network Services, Sollac, Telecom Italia, Teleport Communications, Telecomm Malaysia, Texas A&M, Thomson, Time Warner, UCLA, the U.S. Navy, the U.S. Social Security Administration, and Unisys.

Switching + standards + software + chips = switched services

Xylan has now integrated a broader set of capabilities into its campus switches than any other vendor. And these services continue to be expanded and enhanced, often in unique ways. The result is a switching network that does more than just move packets around at high speed – it makes information technology easier to use.

Over the last few years, switching technology has made networks radically faster. A number of Xylan networks now aggregate more than 1,000 gigabits of switching capacity. OC-12 and Gigabit Ethernet are delivering more bits per second, and layer-three switching is delivering more packets per second, than most users will need for years to come. And switches are radically cheaper, too. Xylan is building 160 Gbps switches and driving costs down with increasingly dense chip designs. But it's focusing most of its efforts on the real problems facing network managers today.

The Next Big Thing in networking is Switched Network Services. Switched network infrastructures will continue to evolve into intelligent fabrics capable of managing information flows, securing resources, authenticating users, establishing QoS, matching users to resources, and supporting multimedia traffic – all based on policies established through network management and through directory servers using LDAP and RADIUS.

Xylan is delivering Switched Network Services across all of its product families, integrated into the XOS (Xylan Operating System) software that powers all its switches.

Address Management Services

- DHCP relay
- DHCP server
- Network address translation

Directory Services

- LDAP
- DDNS

Multicast Services

- Layer-two multicast groups
- Layer-three multicast routing
- IP multicast switching

QoS and Prioritization Services

- 802.1p
- 802.1Q
- RSVP
- Prioritized ELANs

Security Services

- User authentication
- IP firewall
- RADIUS integration

WAN Services

- FRF.9 compression
- FRF.5 / FRF.8 interworking
- RIP / SAP spoofing
- Broadcast suppression

Advanced Management Services

- Policy-based configuration
- Multi-layer path trace

Choices

Network planners are rapidly shifting from hubs and routers to LAN switches, ATM switches, and layer-three switches. The Xylan OmniSwitch combines the most complete set of switching solutions available. Why is this important?

- ***Every organization has unique needs.*** Xylan doesn't try to convert users to one particular technology. Instead, we take you from wherever you are – today – to wherever you want to be – tomorrow.
- ***You can evolve your network gracefully.*** For example, you can use an existing FDDI backbone while you gradually shift to ATM. Or you can switch among Ethernet and Token Ring workstations and servers, while gradually shifting to Fast Ethernet workstations and Gigabit Ethernet servers.
- ***Many networks need a combination of solutions.*** Ethernet and token ring. Fiber and copper. Switching and routing. Layer-two switching and layer-three switching. Small switches and large switches. Local and wide area. We provide a single, easily managed, integrated solution set that solves all your switching needs.

If a switching system doesn't give you the choices you need, today and tomorrow, is it really a strategic solution? And can you afford a system that isn't strategic?

One-stop shopping

Xylan is one of the fastest-growing companies in the networking industry. Why? Because we've focused on one thing – building the most powerful, comprehensive, fully featured switches in the world. If you want a single vendor to supply you with hubs, software-based routers, remote access devices, dial modems, and DS-1 multiplexers, you should look elsewhere. But if you want…

- LAN switching
- Layer-three switching
- Gigabit switching
- ATM switching
- WAN access switching
- Integrated security
- Virtual LANs
- Network management
- Advanced network services

… from a single vendor, then Xylan is the obvious choice. And we don't just build all these technologies. We integrate them, in a complete range of platforms, from small stackables to large modular units.

Let's be blunt. Most of Xylan's competitors bought switching technology by buying small start-up companies. So, they ended up with a set of products that have very little in common. That's not a very good way to create an integrated network.

Xylan offers a complete, integrated set of switches, including workgroup switches, gigabit backbone switches, ATM switches, layer-three switches. All designed by the same team. All managed in exactly the same way.

Low cost

Usually, powerful products are expensive. Not Xylan's. We have one of the most extensive in-house custom chip development programs in the networking industry. As a result, we are able to deliver high-speed, flexible, sophisticated solutions, at a low price.

And Xylan lowers cost of ownership in important, measurable ways.

- How much do you save if the same switches can support the interfaces and protocols you've been using, and the ones you want to start using now, and the ones you'll start using in a few years? If you can avoid an entire generation of network upgrade, and if the change can occur gradually, without interrupting users?
- How much do you save if you only have to train your people on a single set of switch technology?
- How much do you save if you only have to stock spares for a single set of switch technology?
- How much do you save your organization if your switched network is so reliable that it keeps running, even when a cable or a component fails?

Interfaces

A switching system should offer powerful options for workstations, servers, and high-speed backbones. No one does this better than Xylan.

Workstations. Low-cost, high-performance switching makes your workstations more effective without the cost of new interface cards and wiring, and without the hidden costs of interrupting your users. Give each workstation its own dedicated segment, or switch hubs and MAUs.

Servers. You can dramatically increase the effective bandwidth of your network by using switches to reduce the number of workstations on each ring or segment. But why do this if they must still contend for a server operating at the same rate? You'll achieve maximum performance gains when servers and other common resources (such as mainframes and routers) use high-speed connections.

Backbones. Some networks can be built with a single switch. But larger applications call for a number of switches linked together over a backbone. You might want to make optimal use of an FDDI network you already have. You might want to move as quickly as possible to an ATM backbone, seeing ATM as the trend of the future. Or you might prefer 100BaseT because of its simplicity and its similarity to Ethernet.

Choices. It's very simple – the OmniSwitch provides more interface options than any other switch in the world.

- Ethernet over unshielded twisted pair cable (10BaseT)
- Ethernet over fiber optic cable (10BaseFL)
- Ethernet over thin coaxial cable (10Base2)
- Ethernet over thick coaxial cable (10Base5)
- Token Ring over unshielded twisted pair cable
- Token Ring over shielded twisted pair cable
- Token Ring over fiber optic cable
- Fast Ethernet over unshielded twisted pair cable (100BaseTX)
- Fast Ethernet over multimode fiber optic cable (100BaseFX)
- FDDI (DAS and SAS) over multimode fiber optic cable
- FDDI (DAS and SAS) over single mode fiber optic cable
- FDDI SAS over unshielded twisted pair cable (TP/PMD)
- ATM DS-1 over public leased line
- ATM E1 over public leased line
- ATM 25.6 Mbps over unshielded twisted pair cable
- ATM E3 over public leased line
- ATM DS-3 over public leased line
- ATM OC-3c / STM-1 over multimode fiber optic cable
- ATM OC-3c / STM-1 over single mode fiber optic cable
- ATM OC-3c / STM-1 over unshielded twisted pair cable
- ATM OC-12 / STM-4 over multimode fiber optic cable
- ATM OC-12 / STM-4 over single mode fiber optic cable

- DS-1 AAL1 circuit emulation inputs
- E1 AAL1 circuit emulation inputs
- V.35; RS-232 / V.24 / V.28; RS-422 / RS-449; X.21; RS-530 circuit emulation inputs
- ISDN Basic Rate U and ST Interface (BRI)
- ISDN dial backup
- Integral DSU for WAN connections
- Frame relay over V.35; RS-232 / V.24 / V.28; RS-422 / RS-449; X.21; RS-530

LAN switching

Sometimes simple LAN switches can be very effective. But many networks need a more powerful approach to networking. That's what Xylan focuses on. The OmniSwitch provides a complete range of LAN switching tools.

- Any-to-any MAC-layer translation among all LAN types
- Spanning Tree Bridging (802.1D)
- 802.1Q VLAN standard trunking
- Optimized Device Switching
- Source Route Bridging
- Source Route / Transparent Bridging
- Broadcast throttling, configurable per port
- LAN encapsulation in ATM (RFC 1483)
- ATM LAN Emulation for Ethernet
- ATM LAN Emulation for token ring
- Link aggregation (inverse multiplexing) for Fast Ethernet
- Dynamic LAN Emulation (automatically assigns users to correct Emulated LANs)

Layer-three switching

For many years, routing has been a central technology for network designers. Routing provides a standards-based method for moving traffic between subnets. Routing keeps broadcasts from flooding networks. And routing makes it possible to construct hierarchical networks.

Xylan has coupled advanced routing protocols with hardware-based layer-three switching technology. The result is very high throughput, very low latency, and layer-three networking – using proven, well-understood standards.

- HRE (Hardware Routing Engine), providing 226,000 packets per second per switch
- HRE-X, providing 12,000,000 packets per second per switch
- IP
- IPX
- RIP
- RIP II
- OSPF
- IP multicast (DVMRP, IGMP)
- DHCP / BootP relay
- Classical IP Over ATM (RFC 1577)

ATM switching

For many networks, ATM provides an ideal combination of high throughput, extensible rates, advanced network services, public standards, and quality of service guarantees. Xylan's X-Cell ATM cell switching architecture delivers the most advanced technology for campus ATM networks in the world.

And X-Cell is fully integrated with all of the other capabilities of the OmniSwitch. Any combination of LAN switching, ATM switching, and wide area access modules can go into any set of slots. The OmniSwitch's frame bus links to its cell matrix through a high-speed frame-to-cell conversion process. And a single management agent handles all modules in the switch.

- Full support for CBR, rt-VBR, nrt-VBR, ABR, and UBR
- 13.2 Gbps single-stage non-blocking cell matrix
- Up to 131,000 cell buffers per port
- Up to 2,000,000 cell buffers per switch
- Up to 16,000 point-to-multipoint virtual circuits per port
- Up to 65,000 point-to-point virtual circuits per port
- Switched and permanent point-to-point circuits
- Switched and permanent point-to-multipoint circuits
- Soft PVCs
- Virtual path tunneling
- High-performance, redundant LAN Emulation server
- Dynamic policy-based activation of Emulated LANs
- ATM Forum Private Network-to-Network Interface (PNNI) 1.0
- ATM Forum Interswitch Signaling Protocol (IISP)
- ATM Forum User-to-Network Interface (UNI) 3.0 and 3.1
- ATM Forum Circuit Emulation Services (CES) 2.0
- Next Hop Routing Protocol (NHRP)
- Dynamic Input Buffering with Output Control (DIBOC)
- ATM Forum Traffic Management 4.0 with Dual Generic Cell Rate Algorithms to police user traffic
- Explicit Rate and Relative Rate flow control for ABR stations
- Explicit Forward Congestion Indication bit for non-ABR stations

Gigabit Ethernet

Xylan is building powerful Gigabit Ethernet. But more importantly, it is integrating it into a complete switching fabric. Gigabit Ethernet is a powerful network technology, but it is not a network. It must integrate with the Ethernet, Fast Ethernet, token ring, and FDDI LANs that are already installed. It must integrate with frame relay, PPP, and ATM in the wide area. It must incorporate both layer-two and layer-three switching. And Gigabit Ethernet, even more than Fast Ethernet, must be supported by high levels of redundancy.

- 22 Gbps switching fabric
- 12 Mpps of layer-three switching for IP and IPX
- Extensive routing protocols
- Up to 37 RISC processors
- Up to 92 Xylan-designed switching chips
- Up to 32 Gigabit Ethernet ports

Virtual LANs

As networks become larger, they need more than high speeds – they need to be organized into manageable domains. That's what virtual LANs do. Xylan's AutoTracker builds policy-based virtual LANs; it's easy to define the rules, and they apply automatically, no matter where in the network a device is located. Xylan's virtual LANs track devices automatically as they move around the network. They can span multiple switches, across Fast Ethernet, FDDI, frame relay, and ATM backbones. And any workstation or server can belong to as many as 31 virtual LANs.

A Xylan virtual LAN can be defined as:

- A collection of switch ports
- A list of MAC (layer-two) addresses
- A protocol type (AppleTalk, IP, VINES, etc.)
- An IP subnet address
- An IPX network number

- A multicast group
- A custom-defined field value

Security

Five years ago, most network users accessed only a single server. Now it's common to access many servers, and other resources – one of which is usually the Internet. The number of security tables that need to be managed has grown, and the opportunity for unauthorized or harmful access to the network has increased. Security is no longer a server issue; now it is a network issue.

Xylan has recognized this and responded. Today, Xylan is the first switch vendor to integrate a powerful security firewall into its switching fabric. And Xylan has developed comprehensive network-wide authentication and access control – integrated into the switching fabric.

- IP security firewall
- Anti-spoofing to block altered IP addresses
- Stateful Inspection analyzes sessions, not just individual packets
- Policy-based security definitions
- SNMP, e-mail, and pager alerts on security violations
- Network-wide authentication database
- Port-level security based on network address, MAC address, and protocol type

Wide area access

In the first half of the 1990's, multi-protocol routers were flexible, comprehensive tools for building networks. In the second half of the 1990's, that role is being taken over by high-end switches, like Xylan's OmniSwitch. So it's natural that, like routers, these switches should integrate local and wide area networking.

- Routed frame relay (RFC 1490, including Q.922 Annex A framing, routed IP, IPX, ARP/InARP, BPDU, Ethernet, token ring and FDDI)

- RFC 1293 (inverse ARP) dynamic mapping of IP network layer addresses to DLCIs
- Switched frame relay integrated with VLANs
- Multiple VLANs in a single DLCI
- Frame Relay Forum data compression (FRF.9)
- Individual compression code tables for up to 200 virtual circuits
- Configurable DLCMI support (ANSI DS-1.617 Annex D; ITU Q.933 Annex A; LMI revision 1.0)
- Point-to-Point Protocol
- ISDN automatic dial backup for leased lines
- ISDN bandwidth on demand
- IPX and SPX spoofing

Models

Some applications need large switches; some need small switches; some need a combination of the two. Xylan provides four platforms, each designed for a specific need. All use the same custom chips, all run the same software, all are managed with the same SNMP agent. Building a unified network with Xylan's products is automatic.

- Omni-9wx: modular; nine slots, any combination of technologies
- Omni-5wx: modular; five slots, any combination of technologies
- Omni-3wx: modular; three slots, any combination of technologies
- OmniStack 1000: semi-modular, with 32 Ethernet ports and two Fast Ethernet uplinks
- OmniStack 2000: semi-modular, with 32 Ethernet ports and various uplink options, which can be ATM, Fast Ethernet, or Gigabit Ethernet
- OmniStack 3000: semi-modular and stackable, with up to 96 Ethernet ports and various uplink options, which can be ATM, Fast Ethernet, or Gigabit Ethernet
- OmniStack 4000: semi-modular, with 16 Fast Ethernet ports
- OmniStack 5000: semi-modular, with 12 or 24 Fast Ethernet ports and various uplink

options, which can be ATM, Fast Ethernet, or Gigabit Ethernet

- OmniStack 1900: semi-modular, with twelve Ethernet ports and one or two uplinks, which can be ATM, Fast Ethernet, Gigabit Ethernet, FDDI, frame relay, circuit emulation, or ISDN

Network management

Xylan's X-Vision network management suite ensures that a network of OmniSwitches can be planned, monitored, and controlled.

- Graphical control of switches, modules, and ports
- Tracking and display of network statistics
- Management of ATM connections
- Policy-based IP firewall management
- Policy-based virtual LAN management
- Port mirroring
- RMON on every port
- X-Vision integrated with OpenView for Windows
- X-Vision integrated with OpenView for UNIX
- X-Vision integrated with SunNet Manager
- X-Vision integrated with NetView for AIX
- Web-based management tools

Redundancy and reliability

Your organization depends on the network being available. Network failures create enormous direct and indirect costs. Xylan believes that the network infrastructure should be highly resistant to failure – and at a reasonable cost.

- Passive backplane
- Distributed processing architecture for both frame and cell switching
- Redundant, hot-swappable Management Modules

- Redundant, load-sharing, hot-swappable power supplies, each with its own power input, switch and fuse
- 115 VAC, 230 VAC, and -48 VDC power supplies
- Hot-swappable Switching Modules
- Redundant, hot-swappable cooling fans
- Flash memory for configuration and software storage
- Compliance with Bellcore NEBS Level 3 standards
- FDDI optical bypass support
- Redundant connections to an ATM switch fabric
- Redundant connections to an FDDI backbone ring
- Redundant server connections
- Redundant connections to hubs

Directions

Xylan is not satisfied with having developed world-class switching technology. We're constantly pushing ahead. In the near future we expect to complete delivery on a number of on-going projects, including:

- Multilink PPP
- Network address translation for IP
- BGP4 routing for IP
- Resource Reservation Protocol (RSVP)
- OSPF over dial-up ISDN
- Multi-Protocol Over ATM (MPOA)

- ATM Forum User-to-Network Interface (UNI) 4.0
- ATM Forum LAN Emulation (LANE) 2.0
- ISDN Primary Rate Interface (PRI)
- High-density token ring
- High-density stackable 10/100 switching platforms
- Full nine-level RMON for Ethernet and token ring
- OmniVision integrated with Cabletron's Spectrum
- Multi-Protocol Label Switching (MPLS)
- Advanced policy-based QoS

Service and support

Whether you have one office in Opelika, Alabama, or a worldwide network that reaches from Kuala Lumpur to Zurich, Xylan has local resources to support you. We have support and sales offices of our own in more than 75 cities worldwide. And partners like Alcatel, Anixter, IBM, Samsung, and Unisys cover thousands of other locations – ready to deliver network consulting, installation, on-site service, spare parts, and software maintenance. We offer a complete range of support programs, including extended warranties, overnight parts replacement, training, regular software updates, 24-hour phone support; whatever you need to keep your network alive and well.

Offices

Figure 12.1. Xylan offices in North America

Xylan Information Team:
100 Main Street
Suite 400
Dover, NH 03820
603-740-6000
603-740-6002 (fax)
800-995-2612 (toll-free U.S. phone)

Xylan Corporation Corporate Headquarters:
Calabasas, CA
818-880-3500
Additional offices throughout the United States

Figure 12.2. Xylan offices around the world

Xylan Canadian Headquarters
Woodbridge, Ontario
905.264.2787
Additional office in Montreal

Xylan European Headquarters
Hoofddorp, The Netherlands
31.23.556.0155
Additional offices in Berlin, Frankfurt, Helsinki, Madrid, Milan, Munich, Nacka (Sweden), Paris, Prague, Rome, Sandton (South Africa), Stockholm, High Wycombe (UK), Zürich

Xylan Korean Headquarters
Seoul
82.2.565.4433

Xylan Latin American Headquarters
Calabasas, CA, USA
(818) 880-3500
Additional offices in Bogotá, Buenos Aires, Mexico City, Santiago, São Paulo

Xylan Asia / Pacific Headquarters
Singapore
65.336.2972
Additional offices in Bangkok, Beijing, Brisbane, Canberra, Guangzhou, Hong Kong, Jakarta, Kuala Lumpur, Manila, Melbourne, Sydney, Taipei, Tokyo, Wellington

Glossary

10Base2 – A variant of Ethernet, connecting stations via thin coaxial cable; maximum cable distance in one non-repeated segment is 185 meters.

10Base5 – A variant of Ethernet, connecting stations via thick coaxial cable; maximum cable distance in one non-repeated segment is 500 meters.

10BaseFL – A variant of Ethernet, connecting stations via fiber optic cabling.

10BaseT – A variant of Ethernet, connecting stations via twisted pair cabling.

100BaseFX – A variant of Ethernet which runs on multimode or single mode fiber optic cabling at 100 Mbps. This is one version of Fast Ethernet.

100BaseTX – A variant of Ethernet which runs on Category 5 unshielded twisted pair wiring at 100 Mbps. This is one version of Fast Ethernet.

1000Base-CX – A variant of Gigabit Ethernet which runs on twinaxial cable.

1000Base-LX – A variant of Gigabit Ethernet which runs on multimode and single mode fiber optic cable at a 1330 nm frequency.

1000Base-SX – A variant of Gigabit Ethernet which runs on multimode fiber optic cable at an 850 nm frequency.

1000Base-T – A variant of Gigabit Ethernet which runs on unshielded twisted pair cable.

802.x – The set of IEEE standards defining LAN protocols.

A

AAL – See "ATM Adaptation Layer".

ABR – See "Available Bit Rate".

Access Control Method – This is the main distinguishing feature between different LAN technologies. It regulates each workstation's physical access to the cable (transmission medium), and determines the order in which nodes gain access so that each user gets efficient service. Access methods include token passing, which is used in token ring and FDDI, and Carrier Sense Multiple Access with Collision Detection (CSMA/CD), which is employed by Ethernet and Fast Ethernet.

Active Monitor – A node on a token ring network which purges the ring and generates a new token (when necessary), initiates and monitors neighbor notification, and maintains the master clock.

Address Mask – Used to select bits from an Internet address for subnet addressing. The mask is 32 bits long and selects the network portion of the Internet address and one or more bits of the local portion. Sometimes called Subnet Mask.

Address Resolution Protocol (ARP) – Because host addresses and protocols vary in length and value, they are often incompatible with the corresponding 48-bit Ethernet address. The Address Resolution Protocol (ARP) allows for the dynamic distribution of the information needed to build tables which facilitate the translation of an incompatible address into a 48-bit Ethernet address. This protocol has been defined by the IETF.

Adjusted Ring Length (ARL) – Calculated to ensure that if there is a ring failure, the longest ring path is still within specifications. Generally associated with token ring, Adjusted Ring Length ensures that the secondary ring can still function properly in the event of a failure on the shortest trunk cable.

Agent – The portion of the system in the client-server model that performs information preparation and exchange on behalf of a client or server application.

American National Standards Institute (ANSI) – A U.S. standards body. ANSI is a member of the International Organization for Standardization (ISO).

Applications Program Interface (API) – Software designed to make a computer's facilities accessible to an application program. All operating systems and network operating systems have APIs. In a networking environment it is essential that various machines' APIs are compatible, otherwise programs would be exclusive to the machines in which they reside. As networking has developed, some APIs have become de facto standards, including NetBIOS and DOS 3.1.

ARQ – See "Automatic Repeat Request".

ASIC (Application-Specific Integrated Circuit) – A chip designed for a specific application, generally by the manufacturer of the product in which the chip is used.

Asynchronous – A method of transmitting data whereby each byte is clocked separately. One start bit is added to the beginning, and one or more stop bits to the end, of each character. Asynchronous transmission is the most rudimentary form of data communication, as the originating and recipient machines do not have to be in sync. It is commonly used for low-speed transmission, as with a PC's serial port. This meaning of the term "asynchronous" is completely different from that in the next definition.

Asynchronous Transfer Mode (ATM) – A high-speed, connection-oriented switching and multiplexing technology for transmitting information across a wide area or local area network. ATM divides information into fixed-length cells capable of transmitting different types of traffic simultaneously, including voice, video, and data.

ATM – See "Asynchronous Transfer Mode".

ATM Adaptation Layer (AAL) – Provides a conversion function to and from ATM for various types of information, including voice, video, and data. There are several versions of AAL, each applicable to a given information type. All of them convert elements of an information stream (such as voice frame and data packets) into cells, giving ATM the versatility to carry many different types of data, from constant-rate voice data to highly bursty messages generated by LANs, all within the same cell format.

ATM-ARP – Resolves MAC to ATM address translation.

ATM Forum – An international consortium of hundreds of companies and users chartered to accelerate the use of ATM products and services by developing specifications and promoting the technology. The ATM Forum is not a de jure standards body, but on a de facto basis it has been responsible for development of a wide range of ATM standards. It works in cooperation with standards bodies such as ANSI and ITU, submitting to them proposed standards.

ATM LAN Emulation (LANE) – See "LAN Emulation".

Attachment Unit Interface (AUI) – Defined in the IEEE 802.1 specification as the interface between an Ethernet MAU and DTE. Basically, the way an Ethernet station connects to a transceiver sitting on a thick Ethernet cable.

Attenuation – The progressive weakening of a signal as it travels away from its point of origin.

AUI – See "Attachment Unit Interface".

Authentication – A means to establish or prove identity; verifying eligibility of users, machines, or objects.

Authorization – Privileges granted and resources available.

Automatic Repeat Request (ARQ) – A type of error correction ensuring that a transmitting device automatically resends any data containing errors.

Autonomous System – Internet (TCP/IP) terminology for gateways (routers) that fall under one administrative entity and cooperate using a common Interior Gateway Protocol (IGP). See "Subnet".

Available Bit Rate (ABR) – A form of ATM transmission in which an information stream is allowed to access the bit rate left after the predictive and guaranteed service traffic (CBR / VBR) are served. ABR provides a dynamically negotiated rate and includes a congestion control capability. A typical use is for support of workstations equipped directly with ATM network interface cards.

B

Backbone – LAN or WAN connectivity between subnets across a high-speed network. Often applied to a high-speed campus network, such as ATM OC-12 or Gigabit Ethernet, that interconnects lower-speed networks, such as ATM OC-3 or Fast Ethernet. Fiber optic cable is often used.

Backplane – Describes the bus or matrix that traditionally resides at the back of a modular networking product, and into which the modules are plugged.

Bandwidth – (1) The range of signal frequencies that can be carried on a communications channel. The capacity of a channel is measured in cycles per second, or hertz (Hz), between the highest and lowest frequencies. (2) Commonly, the carrying capacity of a digital translation facility, measured in bits per second (bps).

Baseband – A technique whereby digital input is directly applied to transmission media without the intervention of a modulating device. Baseband is generally applied in an environment with high bandwidth over a short distance. It is generally considered easier and more cost-effective than broadband. Ethernet, token ring, FDDI, and ATM generally use baseband.

Basic Rate Interface (BRI) – An ISDN subscriber interface which operates over a single copper cable connection, providing one control (D) channel at 16 Kbps, and one or two bearer (B) channels, at 64 Kbps each. The two B channels are sometimes combined to provide a single 128 Kbps service. BRI is the interface commonly provided to residential ISDN subscribers.

Bellcore (Bell Communications Research) – Telecommunications research and development organization currently owned by the seven U.S. regional Bell operating companies.

BootP (Bootstrap Protocol) – A UDP/IP-based protocol that allows a booting host to configure itself dynamically, and more significantly, without user supervision. It provides a means to assign a host its IP address, a file from which to download a boot program from a server, that server's address, and (if present) the address of an Internet gateway.

Border Gateway Protocol (BGP4) – Interdomain policy routing protocol for communications between a router in one autonomous system (AS) and routers in other AS's.

Bridge – See "MAC-Layer Bridge".

Broadband – Characteristic of any network that multiplexes multiple, independent carrier signals onto a single cable. This is usually accomplished through frequency division multiplexing. Broadband technology allows several signals to coexist on a single cable; traffic from one signal does not interfere with traffic from another, since data is transmitted on a different frequency. Cable television uses broadband.

Broadband ISDN (B-ISDN) – The new generation of Integrated Services Digital Network (ISDN) which carries digital data, voice, and video over SONET networks. B-ISDN allows Asynchronous Transfer Mode (ATM) and Synchronous Transfer Mode (STM) services to operate on the same network.

Broadband LAN – A LAN which uses frequency division multiplexing (FDM) to divide a single physical channel into a number of smaller, independent frequency channels. The different channels created by FDM can be used to transfer different forms of information, such as, voice, data, and video.

Broadcast – A packet delivered to all workstations on a network. Broadcasts exist at layer two and at layer three.

Broadcast Domain – The set of end stations which receive the same broadcast packets.

Broadcast Storm – An overload condition in a network created by an incorrect packet broadcast onto the network that causes multiple hosts to respond all at once. Typically the response contains equally incorrect packets, which causes the storm to grow exponentially in severity.

Broadcast and Unknown Server (BUS) – An ATM LANE process which relays broadcast and multicast packets, and packets with unknown destination addresses, to all Emulated LAN clients. It can be implemented on any ATM device, including a file server, a switch, an access device, or a router.

Bus – (1) A conductor, or set of conductors (e.g. wires), that serves as the interconnection between a related set of devices. (2) A specific type of backplane in which all slots are connected to a common set of wires or traces on which they send to and receive from all other slots. (3) A network topology in which the signals sent by one device are received by all other devices. Each device then selects those transmissions addressed to it based on address information contained in the transmission.

BUS – See "Broadcast and Unknown Server".

C

CAC – See "Connection Admission Control."

Cache – A special area of high-speed memory used to store addresses in a switch. Also called a forwarding table.

Campus Network – A network which covers a single customer location, such as a building, a floor of a building, or all of the buildings on a large commercial, educational, or other campus.

Carrier Sense Multiple Access with Collision Detection (CSMA/CD) – A contention-based network access method in which any computer may attempt to communicate at any time. Since there is no centralized force controlling the medium, a device must first sense whether or not the medium is in use. If the medium is unused the device then transmits. If two computers sense that a channel is open and transmit at the same time, the result is a collision, after which there is a random pause determined individually by each transmitting machine. Each machine then senses the line again and, if it is available, retransmits.

CBR – See "Constant Bit Rate".

CE – See "Circuit Emulation".

Cell – A fixed-length transmission unit which forms the basis of ATM. Each cell is 53 bytes in length, divided into a 48-byte payload and a 5-byte header.

Cell Discard – The process within an ATM switch of discarding cells when the switch's buffer capacity is exceeded.

Cell Loss Priority (CLP) – A one-bit field in the ATM cell header that determines whether or not a given cell should be dropped by network equipment during periods of congestion. This explicit loss priority can be set by the source node or by the network. A CLP which equals zero receives high network priority while a CLP which equals one is dropped during periods of congestion.

Cell Loss Ratio (CLR) – The ratio of discarded cells to cells that are successfully transmitted. Specifically, CLR equals the number of discarded cells divided by the number of transmitted cells.

Checksum – A computed value which is the outcome of a mathematical function applied to the contents of a packet. This value is sent along with the packet when it is transmitted. The receiving system computes a new checksum based on the received data and compares this value with the one sent with the packet. If the two values are the same, the data was received correctly.

Circuit Emulation (CE) – A service provided across a public or private ATM network which emulates the characteristics of a leased-line service.

Circuit Switching – A communications method whereby a circuit is held open and maintained only while the sender and recipient are communicating. This is different from a dedicated circuit which is held open regardless of whether data is being sent or not, and different from a datagram / connectionless network, in which data flows without the establishment of a connection.

Classless Inter-domain Routing (CIDR) - An IAB protocol which uses variable-length subnetting techniques to distribute the allocation of Internet address space. CIDR is needed to address the exhaustion of class B network address space, the growth of Internet routing tables, and the eventual exhaustion of the 32-bit IP address space.

CLP – See "Cell Loss Priority".

CLR – See "Cell Loss Ratio".

Coaxial Cable (Coax) – Formerly common in Ethernet networks, coax comes in various types with different transmission characteristics. It is copper-based, with an inner conductor surrounded by an outer conductor, with insulation between the two, insulation around the outer conductor, and a jacket. Coax is less flexible than twisted pair cable, but more resistant to EMI and physical breakage.

Collapsed Backbone – A network architecture in which a router or switch provides a building or campus backbone using a star topology.

Collision – Concurrent Ethernet transmissions from two or more devices on the same segment. A collision is sensed by the transmitting stations as an over-voltage condition, and they retransmit after waiting a random amount of time.

Common Open Policy Service (COPS) - An IAB client/server model for supporting policy control over QoS signaling protocols with similar properties as ReSerVation Protocol (RSVP). In RSVP, the router, or network device, must respond to bandwidth reservation requests; with COPS, the router forwards the bandwidth request to the nearest COPS policy server. The server makes the end-to-end bandwidth decision; the router implements it. The result is less overhead on the router and overall lower network latency.

Congestion Control – Mechanisms that control traffic flow so that intermediate network devices and end stations are not overwhelmed. Used in connection-oriented networks such as frame relay and ATM. More sophisticated mechanisms are needed to deal with congestion in large networks carrying different types of traffic. Sometimes referred to as flow control.

Connection Admission Control –The set of actions taken by the network during the call setup phase (or during call re-negotiation phase) in order to determine whether a connection request can be accepted or should be rejected (or whether a request for re-allocation can be accommodated).

Connection-Oriented – The model of interconnection in which communication proceeds through three well-defined phases: connection establishment, data transfer, connection release. Examples of connection-oriented networks include ATM, frame relay, X.25, and Internet TCP.

Connectionless – The model of interconnection in which communication takes place without formal connection establishment. Examples include Ethernet, Internet IP, and UDP.

Constant Bit Rate (CBR) – A form of ATM transmission in which a fixed bit rate is provided, with clock frequency and phase maintained end-to-end. Typical uses include emulation of a leased-line circuit, and carrying traditional 64 Kbps PCM voice.

CRC – See "Cyclical Redundancy Check".

CSMA/CD – See "Carrier Sense Multiple Access with Collision Detection".

Customer Service Unit (CSU) – A device used at the customer premise to connect a device, such as a PBX, to a public digital network facility, such as a T1 line. Provides repeater and control functions.

Cut-through – (1) A form of switching, typically LAN switching, in which the switch begins to forward the initial portion of a packet to its destination while the remainder of the packet is still being received. This was useful when the throughput of LAN protocols was highly degraded by latency in the data path. It is uncommon today. (2) A form of switching, typically in an ATM network, in which a routing process is used to set up a connection between two devices, but the data subsequently flows directly between the two devices, without passing through the routing process. MPOA is one important example.

Cyclical Redundancy Check (CRC) – An error-checking mechanism for layer-two data transmissions. Polynomial calculations are performed using only the number of bits in the message. The bits are then sent along with the data to its recipient. The recipient checks the data it receives and repeats the calculation. If there are any discrepancies between the results of the two calculations, the recipient requests the originator to resend the data.

D

DAS – See "Dual Attached Station".

Data Communications Equipment (DCE) – Traditional data communications terminology for the equipment that enables a DTE to communicate over a telephone line or data circuit. The DCE establishes, maintains, and terminates a connection as well as performing the conversions necessary for communications.

Data Link Connection Identifier (DLCI) – A unique number assigned to a PVC end point in a frame relay network. Identifies a particular PVC end point within a user's access channel in a frame relay network, and has local significance only to that port.

Data Terminal Equipment (DTE) – Traditional data communications terminology for a device receiving and/or originating data on a network. Typically a computer or dumb terminal.

Datagram – A self-contained, independent entity of data carrying sufficient information to be routed from its source to the destination computer without reliance on earlier exchanges between the source, the destination computer, and the transporting network.

DCE – See "Data Communications Equipment".

Decryption – The inverse of "encryption."

DHCP – See "Dynamic Host Configuration Protocol".

Digital Service Unit (DSU) – A device used at the customer premise to connect a data device, such as a computer, to a public digital network facility, such as a T1 line. Provides electrical translation and line coding. Technically, this is generally a DSU / CSU, combining both functions.

Digital Signature – Electronic means to ensure message integrity, typically based on a public key cryptosystem.

Distance Vector Multicast Routing Protocol (DVMRP) – A protocol designed to support the forwarding of multicast datagrams through an internetwork. DVMRP constructs source-rooted multicast delivery trees using variants of the Reverse Path Broadcasting (RVP) algorithm. Some version of DVMRP is currently deployed in the majority of MBONE routers.

DNS – See "Domain Name System".

Domain – In networking, a subdivision of the hosts on a network. The division can be physical, as in separate building LANs, or logical, as in giving the hosts in a particular administrative area their own group name even though they are on the same network.

Domain Name System (DNS) – An IAB standard that provides a globally-accessible table of domain names (e.g., xylan.com) and their corresponding IP addresses.

DS-0 – A 64 Kbps digital channel carried within a DS-1 or E1 signal.

DS-1 – The digital signal carried on a North American high-speed facility operating at 1.544 Mbps.

DS-3 – The digital signal carried on a North American high-speed facility operating at approximately 45 Mbps.

DS-3 – The 45 Mbps transmission rate carried on a US T3 facility.

DSU – See "Digital Service Unit".

DTE – See "Data Terminal Equipment".

Dual Attached Station (DAS) – A form of FDDI connection in which a dual counter-rotating ring is supported. Typically used for connecting concentrators and servers to a main ring.

Duplex – A technique allowing bi-directional, simultaneous transmission along a channel. Generally referred to as full duplex.

DVMRP – See "Distance Vector Multicast Routing Protocol".

Dynamic Host Configuration Protocol (DHCP) – A protocol within the TCP/IP family which allows a server process to assign a layer-three (IP) address to a device, when the device requests it. DHCP replaces static configuration of IP addresses by network operators, and in some cases can substantially simplify network management.

Dynamic Routing – A procedure for sending messages across a network by which line failure or overload results in message rerouting.

E

E-1 – The digital signal carried on a high-speed facility operating outside of North America at 2.048 Mbps.

E-3 – The digital signal carried on a high-speed facility operating outside of North America at approximately 34 Mbps.

Early Packet Discard (EPD) – An intelligent cell discard process that occurs within an ATM switch when its buffer capacity is exceeded. EPD discards all cells that originated as members of a single data frame, since the entire frame would have to be retransmitted if even one cell were discarded.

EFCI – See "Explicit Forward Congestion Indication".

EGP – See "Exterior Gateway Protocol".

ELAN (Emulated LAN) – See "LAN Emulation".

Electro-Magnetic Interference (EMI) – Electrical interference with the operation of an electrical device, or with a communications transmission, caused by magnetic radiation. Typically originates from another electrical device, or from a communications transmission in a nearby cable.

Encapsulation – The technique used by layered protocols in which a layer adds header information to the protocol data unit (PDU) from the layer above.

Encryption – The process of converting information from an easily understandable format (plain text) into apparent random gibberish (ciphertext) by the use of well-defined rules and calculations known as algorithms or cipher. A process used to ensure the privacy and confidentiality of information. The reverse process is decryption.

Ethernet – The most common layer-two protocol used in LANs. Ethernet is a 10 Mbps CSMA/CD standard originally developed by Xerox to run on thick coaxial cabling. It has evolved and now runs primarily on twisted pair cabling.

Explicit Forward Congestion Indication (EFCI) – EFCI is an indication in the ATM cell header. A network element in an impending-congested state or a congested state may set EFCI so that this indication may be examined by the destination end-system. For example, the end-system may use this indication to implement a protocol that adaptively lowers the cell rate of the connection during congestion or impending congestion. A network element that is not in a congestion state or an impending congestion state will not modify the value of this indication. Impending congestion is the state when network equipment is operating around its engineered capacity level.

Explicit Rate – Explicit Rate is a type of flow control mechanism defined by the ATM Forum Traffic Management 4.0 standard. With Explicit Rate flow control ATM-attached sources are periodically issued resource management (RM) cells which stipulate a maximum cell rate (measured in cells per second) at which the device can transmit and be guaranteed that traffic will not be discarded by the network.

Exterior Gateway Protocol (EGP) – A routing protocol used by gateways in two-level Internets. EGP is used in the Internet core system.

F

Fabric Blocking – The state that can exist within a switch when its internal switching fabric is not capable of handling simultaneous maximum-rate transmissions by all inputs.

Fast Ethernet – A version of Ethernet which operates at 100 Mbps. See 100BaseTx and 100BaseFX.

Fault-Tolerance – The ability of a device to prevent or recover from network and internal failures. Key elements of fault tolerance include hot-swappable modules, redundant load-sharing power supplies, passive backplanes, and redundant cooling systems.

FDDI – A local area network based on a backbone of dual counter-rotating 100 Mbps fiber optic rings. One of the rings is normally designated as the primary ring; the other is the secondary ring. The dual ring is connected to single-attached "slave" rings through concentrators.

FDM – See "Frequency Division Multiplexing".

Fibre Channel – A form of high-speed fiber optic transmission designed primarily for communications between mainframe computers, and between mainframe computers and high-speed peripherals such as disk drives. Sometimes used for general-purpose networking.

Field Programmable Gate Array (FPGA) – A general-purpose semiconductor component which can be customized to operate physically as though it were a chip dedicated to a specific task.

File Transfer Protocol (FTP) – The protocol within the TCP/IP protocol suite which is used to transfer files between computers.

Firewall – A security mechanism which protects a server, a subnet, or an entire end user location from unauthorized access. Firewalls can be standalone devices, or they can be incorporated into routers and switches.

Flooding – Transmission of a frame to all devices in a segment or ring (in routed networks) or a virtual LAN (in a virtual LAN-based network). Flooding is performed on broadcasts, multicasts, and frames whose destination address is unknown.

Flow Control – See "Congestion Control".

Forwarding Table – A special area of high-speed memory used to store addresses in a switch. Also called a cache.

FPGA – See "Field Programmable Gate Array".

Fragmentation – The process in which a protocol data unit is broken into smaller pieces to fit the requirements of a network. The reverse process is reassembly.

Frame – A unit of information in a layer-two protocol. In LANs, a frame is a MAC-layer unit containing both control information and an entire layer-three packet. Although the term "packet" is sometimes used to mean a frame, the term "frame" is never used to describe a layer-three packet.

Frame Relay – An ITU standard for the interface to a public frame-switching network designed to provide high-speed frame transmission with minimum delay across the wide area. It operates at layer two, and is used in public and private networks, gradually replacing X.25 and leased-line networks.

Frame Tagging – A process of adding a header to the front of a layer-two frame, so that additional information needed to manage the frame through the network is provided. This information can include membership in one or more virtual LANs, priority information, and / or quality of service information.

Frequency Division Multiplexing (FDM) – Method by which the available transmission frequency range is divided into narrower bands; each of these bands is used for a separate channel. This allows several signals to be sent over the same transmission medium.

FRF.5 – A Frame Relay Forum specification for internetworking ATM and frame relay networks. FRF.5 allows ATM networks to transparently pass frame relay data link connection indentifiers over ATM virtual channel indentifiers. This allows ATM networks to act as high-speed backbones for frame relay networks.

FRF.8 – A Frame Relay Forum specification for service internetworking ATM and frame relay. This allows frame relay data link connection indentifiers to be directly mapped into ATM virtual channel indentifiers. FRF.8 allows frame relay devices to directly communicate with ATM attached devices.

FRF.9 – A Frame Relay Forum specification for data compression within frame relay.

FTP – See "File Transfer Protocol".

Full-Duplex – A communications method in which a transmission path is provided in each direction, so that each end can simultaneously transmit and receive.

G

Gateway – A combination of hardware and software that interconnects otherwise incompatible networks or networking devices. The term is sometimes used to indicate a device (uncommon now) which translates between disparate protocol stacks.

Gbps – Billions of bits per second.

Gigabit Ethernet – A variant of Ethernet which operates over multimode fiber optic cable, single mode fiber optic cable, or unshielded twisted pair, at 1,000 Mbps.

H

Half-Duplex – A communications method in which one end transmits while the other receives, then the process is reversed. This was common in wide area point-to-multipoint circuits, such as those used in many SNA networks.

Head End – A central point in a broadband network that receives signals on one set of frequency bands and retransmits them on another set of frequencies. The head end is viewed as a central hub. Every transmission on a broadband network must go through the head end.

Header – A portion added to the beginning of a message containing essential information such as the source address, destination address, and control information.

Head-of-Line Blocking – The state that exists when frames or cells within a single input queue are destined for multiple outputs, and one output is congested, thus delaying all cells.

HEC (Header Error Control) – An 8-bit Cyclic Redundancy Code (CRC) computed on all fields in an ATM header; capable of detecting single-bit and certain multiple-bit errors. HEC is used by the physical layer for cell delineation.

Horizontal Cabling – That portion of a building's cabling system which extends from the wiring closets to the individual workstations, servers, telephones, and other devices. This is generally copper twisted pair cable.

Hot Standby Router Protocol (HSRP) – A Cisco-driven IAB protocol that allows hosts to appear to use a single router and to maintain connectivity even if the actual first hop router fails. Multiple routers participate in this protocol by creating the illusion of a single virtual router. The protocol ensures that one and only one of the routers is forwarding packets on behalf of the virtual router. End hosts forward their packets to the virtual router. See also "Virtual Router Redundancy Protocol."

HTML – A form of page description language used in the World Wide Web.

HTTP (Hypertext Transfer Protocol) – An IAB protocol used in the World Wide Web which defines how requests for HTML and graphics files which make up a WWW page are handled between the web server and the client browser.

Hub – The center of a star topology network or cabling system. Typically used in older Ethernet and token ring networks. A device connected to a hub receives all the transmissions of all other devices connected to that hub. Hubs are now being replaced in many cases by LAN switches.

Hybrid Network – A LAN consisting of a number of topologies and access methods. For example, a network that includes both token ring and Ethernet.

I

IAB – See "Internet Activities Board".

ICMP – See "Internet Control Message Protocol".

ICMP Router Discovery – An extension of the IAB Internet Control Message Protocol (ICMP) that enables hosts attached to multicast or broadcast networks to discover the IP addresses of their neighboring routers without requiring static default route configurations.

IDF – See "Intermediate Distribution Frame".

IEEE – See "Institute of Electrical and Electronic Engineers".

IEEE 802.1D – See "Spanning Tree".

IEEE 802.1p – An IEEE standard for prioritizing time-critical flows and filtering multicast traffic to contain traffic in layer-two networks. The 802.1p header includes three bits for prioritization, allowing for eight priorities to be established.

IEEE 802.1Q – An IEEE standard for providing a virtual LAN capability within a campus network, used in conjunction with IEEE LAN protocols such as Ethernet and token ring.

IEEE 802.2 – A data link standard outlining how basic data connectivity over cable should be set up. Used with the IEEE 802.3, 802.4 and 802.5 standards.

IEEE 802.3 – The IEEE's specification for Ethernet, including both physical cabling and layer-two protocol.

IEEE 802.5 – The IEEE's specification for token ring, including both physical cabling and layer-two protocol.

IEEE 802.10 – The IEEE's protocol for providing security in a metropolitan area network. A variant of 802.10 has sometimes been used to provide a virtual LAN service within a campus network, although this is now generally replaced with 802.1Q.

IETF – See "Internet Engineering Task Force".

IGMP (Internet Group Management Protocol) – A protocol that runs between hosts and their immediately neighboring multicast routers; the mechanisms of the protocol allow a host to inform its local router that it wishes to receive transmissions addressed to a specific multicast group.

IGP – See "Interior Gateway Protocol".

IISP (Interim Interswitch Signaling Protocol) – An ATM Forum specification for signaling between ATM switches, using statically defined connections. Largely replaced by PNNI.

ILMI (Interim Local Management Interface) – An interim requirements definition in ATM Forum UNI 3.1. It supports bidirectional exchange of management information between UNI management entities related to the ATM layer and physical layer parameters.

Information Superhighway – A sadly meaningless phrase, generally associated with politicians, which implies something or other having to do with the Internet.

Institute of Electrical and Electronic Engineers (IEEE) – A standards-making body responsible for implementing many standards used in LANs, including the 802.x series.

Integrated Services Digital Network (ISDN) – A CCITT standard developed to cover a range of voice, data, and image services. It is intended to provide end-to-end, simultaneous handling of voice and data on a single link. Access channels include Basic Rate Interface (BRI) and Primary Rate Interface (PRI).

Intelligent Hub – A hub that adds network management capabilities, such as maintaining port statistics, determining port status, and automatically segmenting faulty ports. Also known as a second-generation hub.

Interior Gateway Protocol (IGP) – A type of protocol used to exchange routing information between collaborating routers on the Internet. RIP and OSPF are examples of IGPs.

Intermediate Distribution Frame (IDF) – In a structured building wiring system, the gathering point for cabling from a section of a building, such as a floor or a portion of a floor. Typically, multiple IDFs located in wiring closets connect to a central MDF.

International Organization for Standardization (ISO) – An international organization that develops standards. ISO is best known in networking for its seven-layer Open Systems Interconnection (OSI) reference model that conceptually organizes communications protocols into seven layers.

Internet Activities Board (IAB) – The technical body that oversees the development of the Internet suite of protocols.

Internet Control Message Protocol (ICMP) – The protocol used to handle errors and control messages at the IP layer. ICMP is actually part of the IP protocol.

Internet Engineering Task Force (IETF) – A body within the IAB which supports the development of new protocols for the Internet.

Internet Packet Exchange Protocol (IPX) – The layer-three protocol used in Novell's NetWare protocol suite. IPX provides a connectionless datagram delivery service for transport-layer protocols such as SPX and NCP. Has nothing to do with the "Internet" as that term is commonly used today.

Internet Protocol (IP) – The layer-three protocol used in the TCP/IP set of protocols which support the Internet and many private networks. IP provides a connectionless datagram delivery service for transport-layer protocols such as TCP and UDP.

Internet ReSerVation Protocol (RSVP) – An IAB standard which allows an end device and a network to negotiate specific QoS characteristics.

Internetwork – Two or more networks connected by bridges or routers.

Intranet – The use of various Internet tools and protocols, especially HTTP and HTML, within an organization.

Inverse Multiplexing – The use of multiple circuits between two devices in which the circuits are treated as a single virtual channel. Traffic is spread across the circuits, and the loss of one circuit results in reduced bandwidth rather than loss of the connection.

IP – See "Internet Protocol".

IP Datagram – The fundamental unit of information passed across the Internet at layer three. It contains source and destination addresses along with the data, and a number of fields that define such things as the length of the datagram and the header checksum.

IP Switching – A form of layer-three cut-through switching pioneered by Ipsilon Corporation, which is now a division of Nokia. In IP Switching, the first packet, or packets, of each information flow are routed as in a traditional router-based network. However, if the routers detect that the flow is likely to be long-lived (as, for example, an FTP connection), then a cut-through path is set up between the end stations.

IPX – See "Internet Packet Exchange Protocol".

ISDN – See "Integrated Services Digital Network".

ISO – See "International Organization for Standardization".

ISO Reference Model for Open Systems Interconnection (OSI) – Developed by the International Standards Organization, the seven-layer model describing the process of network communication. It is intended to facilitate communications among computers from different manufacturers and to provide a common basis for coordinating international standards. Most modern protocols map to the OSI model to some extent, especially at the lower layers.

Isochronous – Signals which are dependent on some uniform timing or carry their own timing information embedded as part of the signal.

ITU (International Telecommunications Union) – An international body of member countries whose task is to define recommendations and standards relating to the international telecommunications industry. The fundamental standards for ATM have been defined and published by the ITU (previously CCITT).

J

Jitter – A short-term timing deviation.

K

Kbps – Thousands of bits per second.

L

Label Swapping – Also known as label switching. A general term for a layer-three switching mechanism which attaches a label, or tag, to each packet. The label provides intermediate switches with the information needed to forward the packet toward its destination.

LAN – See "Local Area Network".

LAN Emulation (LANE) – A set of protocols developed by the ATM Forum which allows legacy LAN protocols, such as Ethernet and token ring, and higher-layer protocols and applications which depend on LAN protocols, to work transparently across an ATM network. LANE translates address formats, emulates the LAN broadcast function, and automatically sets up ATM connections. LAN Emulation retains all Ethernet and token ring drivers and adapters; no modifications need to be made to Ethernet or token ring end stations. Multiple emulated LANs (ELANs) within the same ATM network are common. Also, single stations can belong to multiple ELANs.

LAN Emulation Client – An end device in a LANE application. Can be a workstation or server with an ATM NIC; more commonly, a LAN switch with an ATM uplink.

LAN Emulation Configuration Server (LECS) – A process within ATM LAN Emulation which assigns individual LAN Emulation Clients to emulated LANs.

LAN Emulation Server (LES) – A process within ATM LAN Emulation which translates between MAC addresses and ATM addresses.

Latency – Delay in a transmission path or in a device within a transmission path. Also referred to as propagation delay.

LDAP – See "Lightweight Directory Access Protocol".

Leased line – A transmission facility which is leased by an end user from a public carrier, and which is dedicated to that user's traffic. Typically, frequency synchronization is maintained from one end of the circuit to the other. Leased line circuits are generally used less in recent times, while public data networks are more common.

LEC – See "LAN Emulation Client".

LECS – See "LAN Emulation Configuration Server".

LES – See "LAN Emulation Server".

Lightweight Directory Access Protocol (LDAP) – An IAB standard, based on the ITU X.500 standard, which provides a mechanism for communicating with a central directory that is shared by many different services. LDAP is likely to play a central role in managing dynamic networks.

LLC – See "Logical Link Control".

Lobe Port – In token ring, a port on a MAU or hub to which the cable from a device attaches. Lobe ports must receive a specific voltage from the attached device in order to allow the device into the ring.

Local Area Network (LAN) – (1) The network which interconnects all computing devices located within a single end user location; e.g., an integrated token ring / ATM network covering an entire campus. (2) A single layer-two network, which may be connected to other such networks within an end user location; e.g., a single Ethernet segment. To avoid confusing the two definitions, Xylan commonly refers to the former as a "campus" network.

Logical Link Control (LLC) – A sublayer of layer two which provides a connection between the layer-three protocol, such as IP or IPX, and the MAC layer protocol. LLC2, one form of LLC, provides a connection-oriented service.

Loopback – A testing method in which the transmitted data is looped back to the receiver.

M

MAC – see "Media Access Control".

MAC Address – The layer-two address of a LAN device.

MAC (Media Access Control) Layer – A sublayer of layer two that deals with the issues specific to a particular type of LAN; e.g., Ethernet or token ring.

MAC-Layer Bridge – A device used to forward data between LANs at layer two, by automatically filtering out traffic which is local to each LAN, while forwarding on traffic which is not local to each LAN. All broadcasts and multicasts, as well as all traffic with a destination address which has not been learned by the bridge, is forwarded.

MAC-Layer Protocol – See "Media Access Control".

MAC-Layer Switching – LAN data transferred through a network based on the source and destination addresses contained in the MAC header of the frame. Essentially the same as bridging, but almost always employing dedicated hardware to perform the switching.

Main Distribution Frame (MDF) – In a structured building wiring system, the central point for cabling throughout the building. Typically, multiple IDFs located in wiring closets connect to a central MDF.

MAN – see "Metropolitan Area Network".

Management Information Base (MIB) – A database of objects that can be accessed via a network management protocol. See "SNMP."

MAU – See "Media Access Unit".

Maximum Lobe Length (MLL) – The maximum allowable distance between a node and a MAU or hub on a token ring network.

Maximum Transfer Unit (MTU) – An IAB discovery protocol that polls the network for the highest MTU possible between a source and a destination. The result is an optimized frame size that prevents fragmentation and yields better end-to-end throughput.

Mbps – Millions of bits per second.

MDF – See "Main Distribution Frame".

Media Access Control (MAC) – The way in which LAN workstations share access to a transmission medium. MAC-layer protocols include Ethernet, token ring, and FDDI. Has absolutely nothing to do with the Apple Macintosh computer.

Media Access Unit – In token ring, a hub which interconnects the devices connected to the ring, and in turn connects to other MAUs through Ring In / Ring Out connections. Generally a MAU is not managed via software.

Metropolitan Area Network – A network spanning a geographical area greater than a LAN, but less than a WAN.

MIB – See "Management Information Base".

Microsegmentation – The process of dividing up LAN segments to contain fewer users on a shared media LAN, increasing performance by reducing congestion. It is generally implemented with LAN switches.

Mid-Level Networks – The transit networks that make up the second level of the Internet hierarchy. They connect the sub-networks to the backbone networks. Also known as regionals.

MLL – See "Maximum Lobe Length".

MPOA – See "Multi-Protocol Over ATM".

MTU – See "Maximum Transfer Unit".

Multicast – A form of broadcast in which a packet is delivered to a pre-defined subset of all possible destinations. A specific multicast destination address is used.

Multilink PPP – A form of PPP which uses inverse multiplexing of multiple wide area circuits to achieve a higher-bandwidth virtual connection.

Multimode – A form of fiber optic cabling in which light is able to follow multiple paths as it traverses the cable. Less expensive, and with a lower maximum rate and distance, than single mode fiber optic cable.

Multiplex – To transmit two or more messages or message streams on a single channel, typically through the use of frequency-division multiplexing, time division multiplexing, or statistical time division multiplexing.

Multiplexer – A device used for division of a transmission facility into two or more subchannels, either by splitting the frequency band into narrower bands or by allotting a common channel to several different transmitting devices one at a time. Also known as a mux.

Multi-Protocol Over ATM (MPOA) – A protocol developed by the ATM Forum which provides a standard method for the routing of multiple protocols across an ATM network. The first version of MPOA supports only IP traffic.

N

NAT – See "Network Address Translation".

NDIS (Network Driver Interface Specification) – Developed by Microsoft for writing hardware-independent drivers. NDIS allows multiple protocol stacks (e.g., TCP/IP and NetWare) to share a single network interface module and the software which supports it.

NEBS (Network Equipment-Building System) – Bellcore has devised a three-tier system of criteria for NEBS compliance to ensure that the telecommunications equipment that various operating companies purchase is suitable for their needs, and to reduce the time and expense for manufacturers. The main purpose of this three-tier system is to identify criteria levels and the impact of any non-conforming result. The levels cover safety, environmental, and equipment operability under increasingly rigorous conditions.

NetWare – A protocol suite developed by Novell Corporation. The second most widely used protocol in LANs, after TCP/IP.

Network Address Translation (NAT) – A process by which addresses (typically IP addresses) are translated from one set of addresses to another. Typically used to allow a large address space to be used within a campus network when only a very limited address space is available for that organization's connection to the Internet.

Network Segment – A portion of a network set apart from other network sections by a bridge, router, or switch. Each network segment supports a single medium access protocol.

NHRP (Next Hop Resolution Protocol) – An IAB protocol which provides a cut-through service between end stations in an ATM network.

NIC (Network Interface Card) – A physical plug-in module which goes into a workstation or server and provides the connection to a network.

Non-Real Time Variable Bit Rate (nrt-VBR) – A form of ATM transmission in which clock frequency can vary, but mean variation of delay between cells is guaranteed. A typical use is transmission of stored video.

O

OC-3 – A standard ATM / SONET rate and framing specification; approximately 155 Mbps.

OC-12 – A standard ATM / SONET rate and framing specification; approximately 622 Mbps.

ODI (Open Datalink Interface) – The Novell standard for hardware-independent drivers. ODI can simultaneously support multiple protocol stacks.

Open Shortest Path First (OSPF) – An IAB protocol which is used by IP routers to determine the optimal path along which to move a packet. Like other routing protocols, OSPF requires regular exchange of information among the routers, from which each router calculates the optimal path toward any given subnet. OSPF is a relatively advanced "vector" protocol.

Open Systems Interconnection (OSI) – See "International Organization for Standardization".

Optical Bypass – A capability in FDDI for enhanced failure resistance. A DAS station, such as a concentrator, generates a DC voltage to an attached mechanical optical bypass unit, through which pass all optical signals between the station and the ring. If the station fails, the voltage drops, and the optical bypass unit defaults to a state in which the ring optically passes straight through the bypass unit and cuts out the station.

OSPF – See "Open Shortest Path First".

P

Packet – (1) A variable-length layer-three protocol entity containing address and control information, plus data. Examples include IP and IPX packets. (2) A variable-length layer-two protocol entity containing address and other control information, plus data. Examples include Ethernet and token ring packets. These are also referred to as "frames," and in this book the term "packet" generally refers to a layer-three entity.

Packet Filtering – The ability of a bridge, router, or gateway to limit propagation of packets between two or more interconnected networks.

Packet Switching – A communications method in which variable-length packets are individually routed between hosts.

Partial Packet Discard (PPD) – A process of intelligent cell discard that occurs in an ATM switch when its buffer capacity is exceeded. PPD discards traffic for whole upper-layer PDUs when congestion is encountered. This is done by identifying which cells have been segmented from an individual frame (or packet) and discarding those cells associated with that frame.

PCR – See "Peak Cell Rate".

PDU – See "Protocol Data Unit".

Peak Cell Rate (PCR) – The maximum rate at which ATM cells can be transmitted across a virtual circuit, specified in cells per second and defined by the interval between the transmission of the last bit of one cell and the first bit of the next.

Permanent Virtual Circuit (PVC) – A connection in a connection-oriented network which is established through configuration, rather than dynamically.

Phase Jitter – The result of repeaters regenerating a signal which has experienced envelope delay in transmission through electronics and cable. Phase jitter is removed by processing the data stream through a buffer and reclocking it.

PNNI (Private Network-to-Node Interface) – An advanced, dynamic routing protocol that operates between ATM switches. It is based on link-state protocols, such as OSPF, with extensions that enable switches to advertise their own capabilities, such as capacity and delay.

Point-to-Point Protocol (PPP) – The successor to SLIP, PPP is a layer-two protocol which provides router-to-router and computer-to-network connections across a wide area circuit, generally in a TCP/IP network. See "SLIP."

Port Mirroring – A capability, typically in a switch, which allows a network manager to replicate the real-time data flow from one port at another port. Typically, the second port is attached to a protocol analyzer.

PPD – See "Partial Packet Discard".

PPP – See "Point to Point Protocol".

pps – Packets per second.

Primary Rate Interface (PRI) – An ISDN subscriber interface which operates over a copper or fiber cable connection, providing one control (D) channel at 64 Kbps, and 23 (North America) or 30 (international) bearer (B) channels, at 64 Kbps each. The B channels are sometimes combined to provide various transmission rates. PRI is the interface commonly provided to business ISDN subscribers.

Protocol – A formal description of messages to be exchanged and rules to be followed for two or more systems to exchange information.

Protocol Converter – A device for translating the protocol of one network or device to the corresponding protocol of another network or device. A protocol converter enables equipment with different conventions to communicate with one another.

Protocol Data Unit (PDU) – A defined data unit passed from one protocol layer to another. Each protocol layer encapsulates the PDU from the layer above within the information which it adds.

Protocol Independent Multicast (PIM) Dense Mode – An IAB multicast protocol similar to DVMRP in that it uses Reverse Path Forwarding but does not require any particular unicast protocol. It is useful when multicast senders/receivers are in close proximity to one another, there are few senders and many receivers, the volume of multicast traffic is high, and the stream of multicast traffic is constant.

Protocol Independent Multicast (PIM) Sparse Mode – An IAB multicast protocol that works by defining a rendezvous point that is common to both sender and receiver. Sender and receiver initiate communication at the rendezvous point, and when flow begins it occurs over an optimized path. This is useful when there are few receivers in a group, senders and receivers are separated by WAN links, and the traffic is intermittent.

Proxy – The mechanism whereby one system acts for another system in responding to protocol requests.

Proxy ARP – The technique in which one device, usually a router, answers and issues ARP requests for another device.

PVC – See "Permanent Virtual Circuit".

Q

Quality of Service (QoS) – The requirements a network must provide to an individual flow of information (e.g., a voice call, an interactive video conference, a data file transfer) so that the information is optimally delivered. Elements of QoS can include maximum transit delay, maximum variability of delay, level of data path security, and prioritization with regard to other traffic.

R

RADIUS (Remote Access Dial-In User Service) – An IAB UDP-based protocol used for carrying authentication, authorization, accounting, and security information between a client and a server. Developed to better manage large serial line and modem pools, RADIUS leverages a single user database containing user ID/password and user authorized server types. The client/server model supports security via PAP, CHAP, UNIX login, and other authentication schemes, such as challenge/response systems.

Random Early Discard (RED) – A process of intelligent cell discard that occurs within an ATM switch when its buffer capacity is exceeded. RED discards cells in a round-robin fashion among affected connections.

Real Time Variable Bit Rate (rt-VBR) – A form of ATM transmission in which clock frequency can vary, but maximum delay and maximum variation of delay between cells are guaranteed. A typical use is real-time videoconferencing.

Remote Monitor (RMON) – An IAB specification for a set of MIBs which are used to communicate statistical, event, and other management control information between a managed device and a network management station.

Repeater – A device which propagates electrical signals from one segment to another without routing, buffering, or filtering.

Request for Comment (RFC) – A document written and registered within a process of dialogue managed by the IAB; the collective substance of the most recent RFPs on various topics relating to the Internet forms the body of Internet standards.

RIF – See "Routing Information Field".

Ring – A LAN topology in which each device is connected to two other workstations, with the connections forming a ring. Data is sent from device to device around the ring in a single direction. Each device acts as a repeater by resending messages to other devices. Examples include token ring and FDDI.

Ring Error Monitor for Token Ring – A ring resident function which maintains statistical records of error conditions on the ring operation.

Ring In and Ring Out (RI / RO) – The token ring connectors on the MAU that connect it to other MAUs. Unlike lobe ports, Ring In / Ring Out ports support a "wrap" capability; if an RI / RO cable is disconnected, the ring wraps back on itself, maintaining viability.

RIP – See "Routing Information Protocol".

Riser Cabling – That portion of a building's cabling system which extends from the main distribution frame to the wiring closets. For data, this is often fiber optic cable. For voice, it is fiber optic cable if the PBX is distributed, and twisted pair copper cable otherwise.

RJ-11 – A standard connector commonly used to terminate voice connections.

RJ-45 – A standard connector commonly used to terminate data connections.

RMON – See "Remote Monitor".

Round Trip Delay – A measure of the delay in a network from request sent to reply received.

Route – The path that network traffic takes from its source to its destination.

Router – A layer-three device responsible for making decisions regarding which of several paths network traffic will follow. To do this, it uses a routing protocol to gain information about the network, and algorithms to choose the best route based on several criteria (known as routing metrics). Routers interconnect subnets.

Routing – The process of delivering a message across a network or networks via the most appropriate path.

Routing Domain – A set of routers exchanging routing information within administrative boundaries.

Routing Information Field (RIF) – A field in a token ring or FDDI frame header which provides information used by source-routing bridges to move the frame through a network. The RIF specifies a series of interleaved ring numbers and bridge numbers.

Routing Information Protocol (RIP) – An IAB protocol which is used by IP routers to determine the optimal path along which to move a packet. Like other routing protocols, RIP requires regular exchange of information among the routers, from which each router calculates the optimal path toward any given subnet. RIP is a relatively simple "link-state" protocol.

rt-VBR – See "Real Time Variable Bit Rate".

RSVP – See "Internet ReSerVation Protocol".

S

SAP – See "Service Advertising Protocol".

SAR – See "Segmentation and Reassembly".

SAS – See "Single Attached Station".

Segment – An electrically continuous piece of a bus-based LAN, typically Ethernet. Segments can be joined together using repeaters, switches, bridges, or routers.

Segmentation – Increasing the available bandwidth per device by dividing a network with bridges, switches, or routers to decrease the number of nodes on a segment.

Segmentation and Reassembly (SAR) – A process that occurs within an ATM access device, such as a LAN switch, or sometimes in a LAN switch. In a SAR process, information carried in data frames, such as Ethernet, or voice frames, such as a DS-0 channel, is divided into cells. The SAR is responsible for mapping data from the AAL Convergence Sublayer into the cell payloads of an ATM cell stream.

Sequenced Packet Exchange (SPX) – The layer-four protocol used in Novell's NetWare protocol suite. SPX provides a connection-oriented transport-layer service.

Service Advertising Protocol (SAP) – A protocol used in Novell's NetWare protocol suite which allows servers to inform workstations of their availability, through periodic broadcast packets.

Shielded Twisted Pair (STP) – Copper cable that includes one or more sets of cable pairs which have been molded into an insulating material and covered by a braided shielding conductor. STP offers better noise protection than unshielded twisted pair (UTP) but is much more expensive and more difficult to use. Commonly associated with early token ring networks.

Signaling – Communications between devices to set up calls and tear them down.

Simple Network Management Protocol (SNMP) – An IAB protocol designed to manage networking devices. With SNMP a management station can configure a supported device (SET); can request that the device send statistical, status, and configuration information (GET); and can receive unsolicited alarms from a device (TRAP).

Single Attached Station (SAS) – A form of FDDI connection in which a single ring is supported. Typically used for connecting workstations and servers to a concentrator.

Single Mode – A form of fiber optic cabling in which light follows a single path as it traverses the cable. More expensive, and with a higher maximum rate and distance, than multimode fiber optic cable.

SNA (Systems Network Architecture) – An important protocol suite developed by IBM Corporation beginning in the 1970s for use in both local and wide area communications. Pioneered many modern communications techniques. Many SNA networks are still in place at large organizations, although they are generally converting to TCP/IP.

SNMP – See "Simple Network Management Protocol".

SONET (Synchronous Optical Network) – A set of standards for data communication over fiber optic cable at speeds of 51.84 Mbps and above.

Source Route Bridge – A bridge which is capable of processing the Routing Information Field in a token ring or FDDI frame to determine whether or not to forward that particular frame.

Source Routing (SRB) – A protocol in which the end stations determine the path that frames will follow between them. An end station sends a preliminary route-finding broadcast frame, which turns into many frames, each following a separate route, and each accumulating a statement of the path it has followed. The one that arrives first is assumed to have followed the fastest path, and its path is then specified in all subsequent frames. Source routing is used in some, but not all, token ring and FDDI networks.

Source Route Transparent (SRT) – A protocol which is used in some token ring networks, which uses source routing for frames that need it, and uses transparent bridging for other frames. A variant (SRTB) translates from one type of frame to the other, so that end stations with disparate configurations can communicate.

Spanning Tree – A protocol specified in the IEEE 802.1D standard which allows a network to have a topology that contains physical loops. Spanning Tree operates in bridges and switches. It opens certain paths to create a tree topology, thereby preventing packets from looping endlessly on the network.

Spanning Tree Domain – A portion of a network in which a single Spanning Tree operates.

SPX – See "Sequenced Packet Exchange".

S/T Interface – A physical interface in an ISDN Basic Rate service which uses two copper pairs.

Standby Monitor – Any 802.5 token ring adapter currently attached (active) to the ring which is not the active monitor. One standby monitor assumes the role of the active monitor if it is no longer present on the ring.

Star – A network topology in which each node is connected to a central point.

Station Cabling – See "Horizontal Cabling".

Statistical Time Division Multiplexing (STDM) – Also known as statistical multiplexing. A form of time division multiplexing in which a given data stream can obtain more or less bandwidth dynamically, based on its needs and on the demands of other data streams. Widely used in devices such as routers, LAN switches, and frame relay switches.

Store and Forward – A method of switching in which a message is received as a whole, buffered, and then resent. All routers and virtually all current switches work in this manner. See "Cut-through".

STP – See "Shielded Twisted Pair". Not related to the popular engine-cleansing fuel additive.

Subnet – A portion of a network in which all stations share a common subnet address.

Subnet Address – The subnet portion of an IP address.

Subnet Mask – See "Address Mask".

Sustainable Cell Rate – The maximum throughput bursty traffic can achieve within a given virtual circuit without risking cell loss.

SVC – See "Switched Virtual Circuit".

Switched Virtual Circuit (SVC) – A connection in a connection-oriented network which is established dynamically, rather than through network configuration. An SVC is set up through a protocol which operates between a switch and an end station, and between switches.

Synchronous – Signals that are sourced from the same timing reference and have the same frequency. For example, in high-speed wide area digital communications, the network commonly provides a reference clocking source to which each subscriber's equipment synchronizes its transmissions.

Synchronous Transfer Mode – B-ISDN communications method that transmits a group of different data streams synchronized to a single reference clock.

T

T1 – See "DS-1".

T3 – See "DS-3".

Tagging – See "Frame Tagging".

TAXI – An early standard for ATM transmission at 100 Mbps. Not commonly used now.

TCP – see "Transmission Control Protocol".

TCP/IP – The various protocols which support the Internet and many private networks. An instance of the advantages of synergistic cooperation over central planning.

TELNET – The protocol within the TCP/IP protocol suite which provides a terminal emulation function.

Time Division Multiplexing (TDM) – A method of multiplexing in which multiple information streams "take turns" with a single communications channel. Each stream is allocated a specified percentage of the common channel.

TLS – See "Transparent LAN Service".

Token – A unique packet that is passed around a token ring or FDDI LAN continuously. When a device wishes to transmit, it waits until it receives the token, attaches its message to the token, and transmits it. The device then removes its message from the ring when the token and message return to it.

Token ring – A network architecture standardized in IEEE 802.5 in which the devices on a ring transmit data while they are in possession of a token which passes from node to node continuously. Token ring operates at 4 or 16 Mbps.

Topology – Can be either physical or logical. Physical topology describes the physical connections of a network and the geometric arrangement of links and nodes that make up that network. Logical topology describes the possible logical connections between nodes, and indicates which pairs of nodes are able to communicate.

TOS – See "Type of Service".

TP/PMD – See "Twisted Pair / Physical Medium Dependent".

Transmission Control Protocol (TCP) – A layer-four protocol in the set of protocols which supports the Internet and many private networks. TCP provides a guaranteed transport service.

Transparent LAN Service (TLS) – A service provided by a common carrier in which multiple end user locations are interconnected, using layer-two or layer-three processes, in such a way that the entire network appears to be located on a single site. Commonly implemented today using an ATM service, with LAN switches equipped with ATM uplinks at the customers' sites.

Tree – A LAN topology in which there is only one route between any two of the nodes on the network. The pattern of connections resembles a tree.

Twisted Pair – Insulated copper wires twisted together with the twists or lays varied in length to reduce potential signal interference between the pairs. They are usually bundled together and wrapped in a cable sheath. New data grade Unshielded Twisted Pair (Category 5) is specified for 100 Mbps transmission.

Twisted Pair / Physical Medium Dependent (TP/PMD) – A physical-level specification for FDDI which allows it to operate over unshielded twisted pair and shielded twisted pair copper cable. Sometimes referred to as "CDDI".

Type of Service (TOS) – A field within an IP header which can be used by the device originating the packet, or by an intermediate networking device, to signal a request for a specific QoS level.

U

U Interface – A physical interface in an ISDN Basic Rate service which uses a single copper pair.

UBR – See "Unspecified Bit Rate".

UDP – See "User Datagram Protocol".

UNI (User-to-Network Interface) – An interface point between ATM end users and a private ATM switch, or between a private ATM switch and the public carrier ATM network; defined by physical and protocol specifications per ATM Forum UNI documents. The standard adopted by the ATM Forum to define connections between users or end stations and a local switch.

Unicast – A frame which is sent from one station to another. A unicast contains the specific MAC addresses of the source and destination devices.

Unspecified Bit Rate (UBR) – A form of ATM transmission in which an information stream is supported on whatever bandwidth is available after other connection types have been satisfied. No congestion control is provided. UBR is commonly used to support information streams originating in LAN switches with ATM uplinks.

URL (Universal Resource Locator) – The English equivalent of an IP address and path that describes the location of an HTML (or other type) document on the World Wide Web. The first part of the URL describes the protocol to be used, the second is the DNS location of the server where the document is located, and the last is the path to the document.

User Datagram Protocol – A layer-four protocol in the TCP/IP protocol suite which serves as a connectionless alternative to TCP. Among other functions, UDP is used by SNMP.

V

VCI – See "Virtual Channel Identifier".

Virtual Channel – A single connection across a UNI or NNI allowing the switching of various ATM cells in a virtual path to different destinations.

Virtual Channel Identifier (VCI) – Identifier in an ATM cell of local significance across UNI or NNI which distinguishes data of one virtual channel from the data of another.

Virtual Circuit – A link that behaves like a dedicated point-to-point line or a system that delivers packets in sequence, as would happen on an actual point-to-point network.

Virtual Path – Contains virtual circuits that are to be switched together to a common destination such as an inter-exchange carrier.

Virtual Path Identifier (VPI) – The field in the ATM cell header that labels (identifies) a particular virtual path.

Virtual Router Redundancy Protocol (VRRP) – A non-Cisco-driven IAB protocol that allows several routers on a multi-access link to utilize the same virtual IP address. One router will be elected as a master, with the other router(s) acting as backup(s) in case of master router failure. Host systems may be configured with a single default gateway, rather than running an active routing protocol. See also Hot Standby Router Protocol.

VLAN (Virtual Local Area Network) – In a switched network, a logical collection of devices, such as all the workstations and servers with a particular IP subnet address, which are grouped into a broadcast domain.

VPI – See "Virtual Path Identifier".

W

Web – See "World Wide Web".

Wide Area Network (WAN) – A network which covers a larger geographical area than a single end user location, and in which telecommunications links are implemented, normally leased from service provider(s).

World Wide Web (Web) – The set of HTTP servers, and clients which access them, which are interconnected via the Internet.

X

X.25 – An ITU standard for the interface to a public packet-switching network. Generally connection-oriented.

Index

-48 VDC 140

1000Base-CX 99, 145

1000Base-LX 99, 145

1000Base-SX 99, 145

1000Base-T 99, 145

100BaseFX 132, 145

100BaseTX 132, 145

10Base2 16, 132, 145

10Base5 16, 132, 145

10BaseFL 61, 132, 145

10BaseT 19, 132, 145

115 VAC 140

230 VAC 140

25M ATM 26

802.1d Spanning Tree 17

802.1p 129

802.1Q 129, 133

802.x 145

A

AAL 83, 145

ABR 81, 135, 145

Access Control Method 145

Active Monitor 146

address mask 146

address services 107

Agent 146

ANSI 146

ANSI DS-1.617 Annex D 138

any-to-any 2

any-to-any MAC-layer translation 133

API 146

AppleTalk 54

ARL 146

ARP 146

ARP/InARP 138

ARQ 148

ASIC 25, 146

ATM 47, 75, 127, 147

ATM 25.6 132

ATM Adaptation Layer 5 89

ATM DS-1 132

ATM DS-3 132

ATM E1 132

ATM E3 132

ATM edge switch 111

ATM Forum 22, 93, 114, 147

ATM Forum Traffic Management 4.0 136

ATM LAN Emulation 133

ATM OC-12 141

ATM OC-12 / STM-4 132

ATM OC-3 141

ATM OC-3c / STM-1 132
ATM switching 127
ATM-ARP 147
AUI 147
authentication 3, 127, 148
authentication services 127
authorization 148
autonomous system 148
AutoTracker 136

B

B-ISDN 150
backbone 148
backplane 148
bandwidth 148
Baseband 149
Bellcore 149
BGP4 63, 149
BGP4 routing for IP 141
BISDN 2, 74
BootP 149
BPDU 138
BRI 149
broadband 149
broadband LAN 150
broadcast 17, 150
broadcast domain 150
broadcast storm 150
broadcast suppression 129
broadcast throttling 133
BUS 91, 150

C

CAC 151
cache 39, 151
campus network 151
CBR 80, 135, 151
CE 135, 151
cell 151
cell discard 151
checksum 152
CIDR 152
Circuit Emulation 114
circuit switching 152
Classical IP 89
Classical IP Over ATM 134
CLP 151
CLR 152
coax 152
collapsed backbone 152
collision 153
configurable DLCMI support 138
congestion control 153
Connection Admission Control 153

copper 99

COPS 153

CRC 153

CSMA/CD 151

CSU 154

custom-defined field value 137

D

DAP 108

DAS 154

DCE 154

DDNS 128

DECNet 54

DHCP 50, 107, 155

DHCP / BootP relay 134

DHCP relay 128

DHCP server 128

DIBOC 135

distributed processing architecture 140

DLCI 138, 154

DNS 108, 155

DS-0 156

DS-1 82, 156

DS-1 AAL1 133

DS-3 82, 117, 156

DSU 155

DTE 155

Dual Generic Cell Rate Algorithms 136

DVMRP 134, 155

Dynamic LAN Emulation 133

E

E-1 156

E-3 156

E1 AAL1 133

E3 117

EFCI 157

EGP 157

ELAN 90, 157

EMI 157

end-to-end ATM 75

EPD 88, 157

Ethernet 15, 100, 127

Explicit Forward Congestion Indication 136

explicit rate and relative rate flow control 136

F

fabric blocking 86

Fast Ethernet 21, 100, 129

FDDI 21, 127

FDDI LANs 136

FDM 158

Fibre Channel 99

firewall services 106
firewalls 3, 127
forwarding table 39
FPGA 158
frame relay 112
Frame Relay Forum 22, 75, 124
FRF.5 160
FRF.5 / FRF.8 interworking 129
FRF.8 160
FRF.9 160
FRF.9 compression 129
FTP 66, 122

G

GARP 57
Gigabit 3
Gigabit Ethernet 3, 127
Gigabit switching 130

H

head-of-line blocking 87
HEC 161
high-density token ring 141
hot-swappable switching modules 140
HRE 134
HRE II 134
HSRP 161
HSTRA 46
HTML 161
HTTP 162
hubs 24, 162

I

IAB 8, 119, 162
IAB RFC 1757 121
ICMP 162
IDF 162
IEEE 22, 124
IEEE 802.10 163
IEEE 802.1D 162
IEEE 802.1p 162
IEEE 802.1Q 162
IEEE 802.2 162
IEEE 802.3 163
IEEE 802.3z 99
IEEE 802.5 163
IETF 8, 92, 124, 163
IETF RFC 1483 89
IETF RFC 1490 113
IETF RFC 1577 89
IGMP 134, 163
IGP 163
IGRP 63
IISP 84, 135, 163

ILMI 163
integral DSU 133
integrated security 130
intelligent fabric 2
Internet 116
Inverse multiplexing 2
IP 61, 134, 164
IP firewall 129
IP multicast switching 129, 134
IP routing capability 116
IP security firewall 116
IP subnet 107
IP subnet address 137
IP VLAN 51
IPX 134, 164
IPX and SPX spoofing 138
IPX network number 137
IPX VLAN 51
ISDN 2, 73, 116
ISDN bandwidth 138
ISDN Basic Rate 133
ISDN dial backup 133, 138
ISO 164
ITU 22, 81, 124, 166
ITU Q.933 Annex A 138

J

JPEG 113

L

LAN 2, 95, 135
LAN Emulation 95, 135
LAN switching 130
LANE 93, 141, 147
LANE 1.0 91
LANE 2.0 91
LANs 127
layer-three multicast routing 129
layer-three switching 65, 127
layer-two multicast groups 129
layer-two switching 127
LDAP 108, 128, 167
leased lines 112
LEC 90, 167
LECS 91, 166
LES 91, 167
link aggregation 133
LLC 61, 167
LMI revision 1.0 138
lobe 42
lobe port 167

M

MAC 15, 137
MAC Address 168
MAC layer 3, 100, 168
MAC-layer bridge 17, 168
MAC-Layer protocol 168
MAC-Layer switching 168
MAN 168
MAU 168
MDF 168
MIB 119
mobility services 106
MPEG II 86
MPLS 68
MPOA 3, 93, 141, 170
multi-layer path trace 129
multicast 17
multicast group 137
Multilink PPP 141, 170
multimode 170
multimode fiber optic 99
multiplexer 170

N

NAT 107, 171
NDIS 170
NEBS 170
NetBIOS 44
NetWare 61, 171
network address translation 128
network interface cards 24
network management 130
network modeling 129
NHRP 66, 92, 135, 171
NIC 171
NLSP 63
nrt-VBR 81, 135, 171

O

OC-1 81
OC-12 82, 171
OC-12 ATM 25
OC-3 81, 117, 171
OC-3 ATM 25
OC-48 82
ODI 171
Omni-3wx 138
Omni-5wx 138

Omni-9wx 138

OmniStack 1000 139

OmniStack 1900 139

OmniStack 2000 139

OmniStack 3000 139

OmniStack 4000 139

OmniStack 5000 139

OmniSwitch 127

Optimized Device Switching 133

OSI 172

OSPF 63, 84, 134, 171

P

passive backplane 140

PBX 12

PCR 173

PDU 174

PNNI 84, 135, 173

policy-based configuration 129

port mirroring 139, 173

PPD 88, 173

PPP 117

PRI 141, 173

prioritization 81

prioritized ELANs 129

PVCs 79, 135, 173

Q

Q.922 Annex A framing 138

QoS 3, 128

queues 86

R

RADIUS 128

RED 88, 175

redundancy 140

RFC 8

RFC 1293 138

RFC 1483 133

RFC 1490 138

RFC 1577 134

RFC 1755 90

RIF 45, 176

RIP 61, 113

RIP / SAP spoofing 129

RIP II 63, 134

RISC 25

RJ-11 176

RJ-45 176

RMON 120

routed frame relay 138

routed IP 138
routers 21, 111
routing protocols 63
routing table 63
RS-232 / V.24 / V.28 133
RS-422 / RS-449 133
RS-530 133
RSVP 109, 129, 164
rt-VBR 80, 135

S

S/T interface 133, 179
SAP 113, 177
SAR 89
SAS 178
security firewalls 24
single mode fiber optic 99
SNA 44, 124, 178
SNMP 119, 178
SONET 81, 178
Source Route Bridging 133, 179
Spanning Tree Bridging 133, 179
SPX 178
SRB 179
SRT 45, 179
SRTB 45
standards 22
STDM 180
STP 178
subnet 122
SVC 79, 180
switched campus network 111
Switched Network Services 128

T

tagging protocol 51
TCP 62, 182
TCP/IP 44
TDM 181
TELNET 181
TLS 114, 181
token ring 2, 127
token ring switch 122
TP/PMD 182
traffic management 89
Transparent Bridging 133
Transparent LAN Services 114
twinax 99

U

UBR 81, 135, 183

UDP 119, 183

UNI 83, 135, 183

unicast 183

unshielded twisted pair 99

URL 183

user authentication 24, 129

User Datagram Protocol 183

UTP 99

V

V.35 133

VCI 83, 184

VINES 137

virtual circuits 79, 184

virtual LANs 113

Virtual Path 184

VPI 184

VRRP 184

W

WAN 75

WAN access switching 130

WAN integration 103

Web 185

wide area ATM 114

wide area networking 3, 111

X

X-Cell 134

X-Vision 139

X.21 133

X.25 112, 185

XOS 128

Xylan 3

Z

zero hop routing 92

Acknowledgements

Design, layout, and illustrations for Switching Book II were created by Xylan's Creative Services team – Phil Krahn, Chuck Downs, Clarce Estrada, and Ed Youngblood.

It was edited, over and over, by Jan Rosser.

Additional contributions (large and small, but mostly large) were made by:

Anthony Alegrete (lots!), Sangeeta Anand, James Chong, Jeff Hayes, Keith Higgins, Roger Hockaday, Mike Le Blanc, Albert Lew, Dave Rodewald, Sean Thomas, Ides Vanneuville, Kevin Walsh, Bob Wyan, and the rest of the Xylan team.

Most of the concept and writing was by Douglas Hill.

Reply

If you have any corrections, comments or suggestions for Xylan's Switching Book II we would appreciate your input. Please fill out the following and send it to:

Xylan

26707 West Agoura Road • Calabasas, CA 91302

or you can fax it to: 818-878-4714

Attention: Switching Book

Please rate the book on the following on a scale from 1 to 5 (five being best)

CLARITY	1	2	3	4	5
CONTENT	1	2	3	4	5
ILLUSTRATIONS	1	2	3	4	5

What do you like most about the Switching Book II?

What do you like least?

If you have any suggestions or note any areas which are in need of further editing, please write them down here so we can make the proper improvements in the next edition of this book:

Notes

Notes