LESSON 1 INTRODUCTION TO HIGH SPEED NETWORK

1.1 OBJECTIVES

The objective of this chapter is to introduce the reader about High Speed Network, performance attribute of high speed networks, what are the technologies of high peed networks and the cost of High Speed Networks.

1.2 INTORDUCTON TO HIGH SPEED NETWORKS

In future telecommunication systems, high-speed networks will be used to carry streams of real-time data between groups of personal C&C environments. In addition to the dramatic increase in network capacity that is needed to support this, future applications will need network support in several different areas:

- Mobile users will demand continuous network services, so the network should provide support for data streams between moving end-points.
- Communication in distributed campuses and offices will be clustered into groups of communicating users. The communication patterns within the groups will be one-to-many and many-to-many. Thus, the number of logical data streams grows network load as well, networks should provide efficient support for multicast communication.
- Future applications for personal C&C environments will be for live data, such as live audio and video, and will therefore have real-time requirements on delivery of data. Personal communication will be a dominating application, but applications like interactive TV and operation control also have real-time requirements.

1.2.1 Brief Networking History

- The World Wide Web (WWW): In spring 1989, CERN (the European Lab. For Particle Physics), Tim B. Lee proposed the idea of a distributed hypermedia technology to facilitate international exchange of researches findings using Internet. In 1991, a prototype WWW called *line-oriented browser* was developed at CERN and released to limited population. Later, the *graphically oriented browser,* Mosaic, developed at the NCSA at University of Illinois by Mark Andreson and others in 1992. 2 million copies of Mosaic have been delivered in a short period.
- Email triggered rapid growth of ARPAnet so the Web triggered explosive growth in internet.
- ISDN and Frame Relay: ISDN includes the integrated services and digital networking. The most noteworthy technical achievement of the ISDN effort was the development of specifications for Frame Relay. Frame Relay is now considered a matured technology. It has a proven track record and is used in many Fortune 500 companies.

Banks are big users of Frame Relay services, especially their automated teller machine networks.

- B-ISDN: ISDN was designed to provide data rates in the range of Kbps to Mbps. To provide higher data rate—Mbps to Gbps—the original ISDN (also called Narrowband ISDN, or N-ISDN) was extended to Broadband ISDN (B-ISDN).
- ATM: Just as frame relay is a technology developed as part of the ISDN effort and now widely used in non-ISDN applications, so ATM is a technology developed as part of the B-ISDN effort and now widely used in non-B-ISDN applications. ATM is becoming popular because it supports transport of data, video and voice communications over a single infrastructure. While data is somewhat resilient, video and voice applications demand timely and sometimes high bandwidth. ATM brings together these features and offers high-speed communication. Most of today's LAN operate at 4Mbps or 10Mbps. High speed LANs can operate at 100Mbps. ATM provides LAN and WAN connectivity at 155Mbps. ATM standards are being developed in the marketplace. The cost of ATM has also started to decrease. With standardization and a better cost/benefit model, ATM should be one of the technologies of the future.
- Marketplace Examples: T-1 networking will continue to be a viable option for many years. While T-1's is not new technology, it is a proven high bandwidth solution for many applications, including imaging, data communication and video conferencing. Many government agencies and commercial companies have become big users of T-1's. It is safe to say that most states have successfully deployed T-1's as part of their communications infrastructure.
- ATM utilization is not as widely deployed as T-1 and Frame Relay communications. The military has begun implementation of ATM based projects. The US Army is deploying this technology at Fort Bragg, North Carolina, Fort Hood, Texas and Fort

Stewart, Georgia. Three defense agencies at Fort Belvoir, Virginia are also using ATM as their backbone. Case Western Reserve University in Cleveland, Ohio is using ATM technology all the way to the desktop so students can work on assignments and even take classes via video over the network. North Carolina has implemented an ATM backbone for education and government use across the state.Leading Vendors & Users: The regional telephone companies like Ameritech, BellSouth, Pacific Bell, etc., along with national communications providers like AT&T, MCI and US West are the major providers of both T-1 and Frame Relay networking services. ATM is still fairly new territory for vendors. There are a large number of hardware and software vendors providing ATM products today. Since there is not a clear standard for ATM, the major internetworking vendors seem to be the biggest players in this technology. Companies like Bay Networks, Cascade, Cisco, Digital Equipment Corporation, FORE Systems, General DataComm, Hewlett-Packard and IBM are just a few of the companies offering ATM products. Several telephone companies like MCI and US West are offering ATM solutions as well. Local and regional cable television providers have moved into this service area

- Public network infrastructure corresponds to B-ISDN and consists of public telecommunications networks that support public telephony, cable TV (CATV), and WAN services.
- ATM LAN: ATM switches can serve as a backbone network in a local setting to interconnect a variety of traditional LANs (e.g. Ethernet).
- ATM WAN includes enterprise networks operating over leased lines such as optical fiber and wireless links.

1.2.3 Technology

High-speed networks will be based on optical fiber technology. Advances in optical components will open up new possibilities for network architectures; in particular, optical switches and amplifiers, and multi-wavelength techniques will have important roles in the future. However, it does not seem that the optical fibers that will be deployed during the next ten years will have significantly higher transmission speed than today's fibers. Typically, a distribution channel may carry 2.5 - 5 Gb/s. If we further assume the rate of the access channels to be nominally 100 Mb/s, it means that the ratio of distribution channel capacity to access channel capacity will decrease in the future. As a consequence, high capacity in backbone networks is achieved by using many individual channels, much more than what is used today. This motivates research in multi-channel network solutions, such as multiplexing in the wavelength and space domains.

The relatively large number of channels also has consequences in the area of switch architectures, since switching equipment will need to handle large numbers of incoming and outgoing channels. This calls for research in the area of switch architecture, and investigations of alternative switching concepts, such as distributed switching.

There are several problem areas in high-speed networks:

Multicast distribution: What techniques should be used to efficiently implement logical channels for multicast switching in a network with many subscribers, in order to support applications ranging from selective distribution to dynamically reconfigurable group communication?

Distributed switching: How can signalling protocols, addressing, routing and clock distribution be provided for distributed switching? What network structures are suitable for interconnecting switching elements? What are the advantages and disadvantages with using cir-

cuits, cells, and packets as the basis for switching, and how can variable length packet switching be provided in cell-based networks?

Network resource management and routing: What traffic control policies are most suitable; can feed-back schemes react fast enough, and are reservation-based policies cost-efficient? How are resources managed to support real-time multicast transmission?

Fault-tolerance in high-speed networks: How are networks organized to be resilient to link and node failures? What are the best ways to reconfigure high-speed networks after faults, and how are faults detected and located in optical networks?

Impact of new optical components on protocols: What medium access and switching techniques are best suited for multi-wavelength distribution techniques? How are protocols best designed for cost-efficient use of optical switches and amplifiers?

End-system integration: Does the close integration of communication and computation impose new requirements on optical fiber access techniques, switching, and network protocols? Is it possible and advantageous to combine network resource management with management of resources in the end-equipment (e.g. CPU scheduling)? How are protocols designed to enable efficient end-system interfaces?

High-speed switching: How will switching fabrics and circuit architectures change when moving from current low and medium rate switching systems towards 10 Gbit/s and 40 Gbit/s per port switching systems? How to dimension and design such switching systems? What impact will C&C environment applications have on switch architectures and switch dimensioning? Due to high operational speed, what constraints will implementations pose, and how are such constraints coupled to system design of switching systems and architectures? What implantation techniques and architectures need to be used for optical interfaces to such switches as well as to the overall switch fabric itself?

1.3 Need of High Speed Network

Emergence of High-speed LANs: In the 1990s, 2 significant trends have altered the role of the PC and LAN:

(1) The speed and computing power improvement of PCs, and

(2) MIS (management information systems). Organisations have recognised the LAN as a viable and essential computing platform.

The following are examples of requirements that call for high-speed LANs:

• Centralized server farms: There is need for users, client systems to be able to access huge amounts of data from multiple centralized servers.

• Power work groups: These groups typically consist of a small number of cooperating users who need to draw massive data files across the network.

• High-speed local backbone.

1.4 PERFORMANCE ATTRIBUTES

There are following attributes that determines the performance f a network:

Distance

Since communication involves at least two users or processes, and as no two entities may occupy the exact same geographical point at the same time, distance inevitably becomes a fundamental attribute Without distance, be it as large as astronomical units or as small as those in microelectronics, no network is conceivable. Every research into new network design must take the issue of distance into consideration. Distance is a fundamental property of the physical world and, in a network, it behaves as a resistance to the transport of a message. From physics, since the speed of EM waves is finite, distance translates into physical

propagation delay, higher number of user traffic sources, greater number of networking resources, higher chance of resource contention among users, greater probability of noise, errors, faults, and security problems. Since network topology depends on the distance, the latter plays a key role in network performance.

Asynchronism

Clearly, in the most general case, the timing of the communication between two or more users is likely to be asynchronous, i e. irregular in time. Assuming A and B as two users, neither A nor B can know, a priori and with certainty, when the other will initiate communication. Nevertheless, either one of them must be prepared to respond to the other when contacted. Also, when A contacts B, it has no a priori certain knowledge of how quickly B will respond. Although synchronous networks are conceivable and many are in operational use, the synchronous design principle breaks down as networks grow to encompass vast geographical distances, greatly increased number of nodes and users, and very high-speeds. The traffic sources interactions with the network elements are inherently asynchronous. Therefore, source traffic control techniques to ensure QoS in future networks must take into consideration the synchronism. While the nodes of a network interact asynchronously among themselves, faults and errors in the interactions between nodes may also occur asynchronously. Hence, the design of the timing and control algorithms as well as security techniques and recovery procedures, in future network evolution must increasingly focus on the asynchronicity.

Traffic sources

Traffic sources are naturally indispensable in networks and estimates of expected traffic in the network must constitute an integral component of any network design In general, a network designer does not possess a precise picture of how the network under design will be utilized after its deployment. As a result, traffic estimation in the real world can pose a significant challenge. Traffic engineering must include both high-level issues including the distribution as a function of time of the (1) number of traffic sources, (2) their bit rates, (3) session durations, and (4) QoS requirements, as well as lower-level issues such as, for each traffic source, (a) the cell arrival distribution, (b) whether the traffic is constant bit rate, busty, fractal, or self-similar, (c) cell-level traffic mix of video, audio, and data, and (d) active and silence interval times within a frame. The arrival of a cell at a network node is, in general, asynchronous.

Faults and errors

The existence of faults and design errors in any human engineered system, including high-speed networks, is only logical. Failures in the processors, switch fabrics, and links are likely. For high-speed network designs of the future, although resistance to failures will be high, the loss of cells and the consequent damage from any failure is likely to be very costly. Wherever possible, automatic and distributed detection of failures, fault tolerance, and self-recovery must be incorporated into the design of future networks. With the asynchronous nature of the faults and errors; the task can be especially challenging.

Transmission medium

Electromagnetic waves transmission invariably requires a medium. Whether it is wires, optical fibers, or free space, the EM transmission medium imposes a finite speed of propagation, which translates into a physical propagation delay. Transmission media also influence the energy required for transport, probability of loss, and the available transmission capacity. Consider, for example, communication between the earth and the moon via two competing media—free space and an optical fiber. The EM propagation speed through free space is 3!108 m/s while that through the fiber is 2!108 m/s, thus a single bit by itself is transported faster through free space. However, dispersion and energy dissipation are generally higher in free space than in optical fibers, implying a much higher bandwidth rate accompanied by lower cell loss in the optical link.

Bandwidth/data rate

The ability of a link or medium to transport bits is referred to as bandwidth/data rate and serves as the key to the viability of a network Bandwidth may be viewed as the number of lanes of a wide freeway that permits multiple vehicles to travel along the same route, simultaneously.

Delay

In addition to the physical propagation delay imposed by distance, the limited bandwidth disallows all of the bits of a packet to travel in parallel As a result, some of the bits must travel in sequence, giving rise to transmission delays.

Size of the information packet

The information packet is the basic quanta of information transport For the purpose of transport across a link, a packet is viewed as an atomic unit, i.e. all of its bits must be sent as a whole. It cannot be fragmented. Although it is a design parameter, the choice of the packet size always exerts a strong influence on the network performance.

1.5 NETWORK BACKBONE

Back in the early days of networking, all networks were linear networks. This meant that a network consisted of a number of workstations connected by a single cable. At the time, this was fine, because networks tended to be small, and so cable problems were fairly easy to isolate. On larger networks, however, locating a loose connection or cable break could be a difficult and time-consuming process, often requiring special—and expensive—equipment to determine the type and location of the cable problem.

To solve the problem caused by a single network cable, wiring centers, or hubs, were developed. The first hubs were nothing more than signal repeaters. They took a data signal from one cable segment and repeated it over another, extending the length of the network, as shown in Figure 1-1. However, these simple repeaters made it easy to divide networks into segments, meaning physical and/or logical groups of network users and resources. Dividing networks into segments enabled network managers to keep network users as close as possible to the network resources they used the most. This allowed the network administrator to control and isolate network traffic to a certain extent.

Figure 1-1. A basic hub configuration. The hub regenerates, or *repeats,* the transmission signals coming from the workstations

The potential usefulness of hubs went well beyond that of the first simple network repeaters. Soon after the appearance of the first basic hubs, a second generation of hubs appeared that offered management features. These hubs could collect management information from each network connection, then convert that information into a standardized format (for example, Simple Network Management Protocol), then export that information to a number of management reporting systems. Often these hubs also had buses that alowed them to support multiple transport protocol on the same high-performance backplane. Hubs appeared that could support multiple logical network segments within the hub, meaning that a hub could now contain more than one segment and that each of those segments could be managed separately. This enabled network managers to reconfigure network segments, making moves, ads, and changes on the fly and even from a remote console.

With these sophisticated new features came new roles for network hubs. Soon, network managers were using hubs to concentrate, or collapse, network nodes into increasingly complex hierarchies of networks, often within a single chassis. Simple, unmanaged wiring centre hubs were used to connect end-user nodes to their login servers. These simple hubs were in turn connected to one another by more sophisticated hubs. These intermediate-level hubs were then connected by extremely feature-rich and highly manageable chassis-based hubs, some employing bridging or routing modules. (See Figure 1-2.)

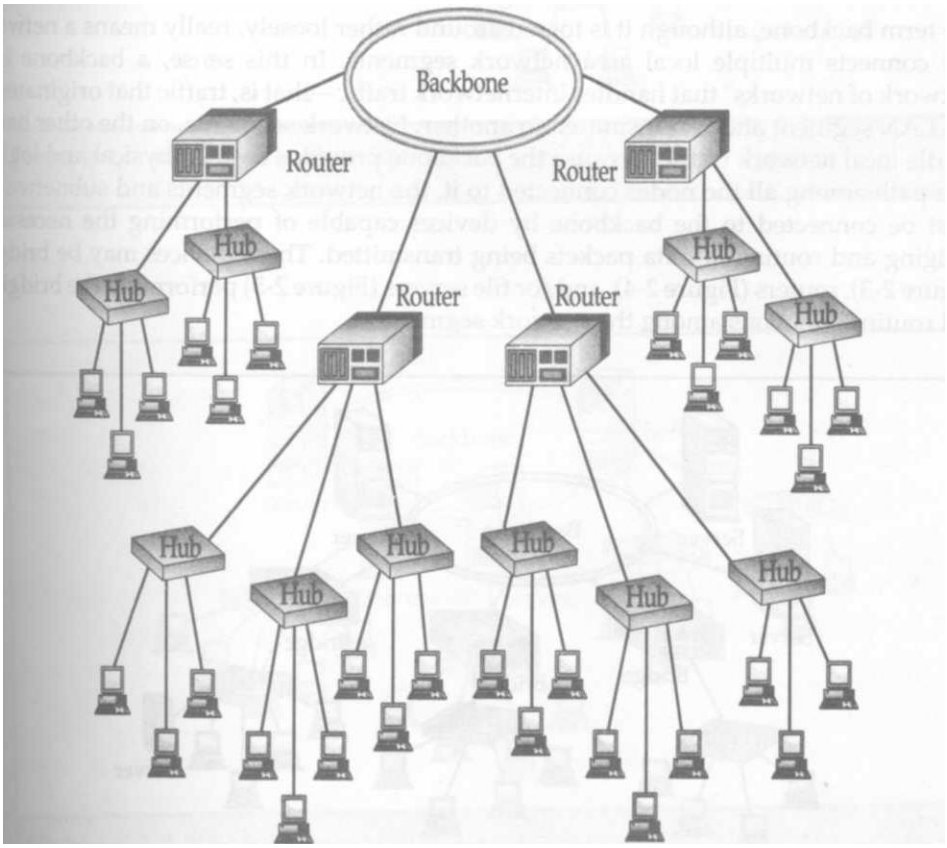


Figure 1.2

With this, the concept of structured wiring was born, and network managers began to realize its many benefits, including the following:

- Easier moves, adds, and changes
- Centralized management
- Improved scalability
- Better monitoring of network events and statistics
- Improved reporting
- Easier troubleshooting
- Improved fault tolerance
- Support for multiple transport protocols

1.6 COST OF HIGH SPEED NETWORKS

Now that you know where to expect slowdowns, and why they occur, it's nearly time to start talking about how to speed things up. However, before we do that we need to prepare you for sticker shock. High-speed networking is no cheap, it's not easy, and therefore it's not always worth it. Therefore, before you launch in' implementation project, I recommend that you do the following:

1. Determine exactly which segments will benefit from a high-speed protocol, detailed in the first four chapters of this book
2. Carefully estimate the costs of converting to a high-speed protocol

To help following worksheets outlines the major cost components of a high-speed network and describe how calculate the cash outlays associated with each. These worksheets appear at the end of chapter.

Hardware Cost

The cost of equipment is probably the first and most obvious expense related to installing a high-speed network. However, many costs directly and indirectly related to installing new hardware. The Hardware cost outlines the equipment and related costs that has to consider, along with the quantity and the cost of each.

Servers

When upgrading servers, be sure to contact your NOS vendor to find out exactly which network adapters and drivers are fully certified for the version of the NOS you now have. Remember, some high-speed networking technologies are relatively new, so the version of the network operating system you currently have installed may not support high-speed protocols. If it doesn't, a NOS upgrade—with all the attendant heartache—will be in order. Furthermore, a new version of the operating system—or even a high-speed NIC driver—may require other

hardware upgrades, such as increased memory or disk space. Be sure to ask your vendors about system hardware requirements, and figure any upgrades into your cost estimate.

Hubs

The number of hubs you require depends on two things:

- How many ports you will be converting to a high-speed protocol
- The port density of the hubs you want to purchase

If you are purchasing stackable hubs, the smallest unit you can buy is a hub. Note that some stackable hubs require separate terminators, so be sure to include the cost of these terminators in your estimate. If you have purchased or are planning to purchase a chassis-based unit, the smallest unit will be a chassis module. In either case, be sure to include the cost of the special cables and connectors required to attach the high-speed ports to your existing network.

Finally, remember to include the cost of any changes that you'll need to make in your wiring closet to accommodate the new hubs, such as additional racks and patch panels.

Routers

Your routers may require new physical interfaces, either internal or external, as well as software and firmware upgrades to work with a high-speed protocol. They may even have to be replaced altogether. In either case, some manual configuration will necessary, so be sure to include all the associated costs.

Switches

You may need to purchase or upgrade existing switches. Upgrading your switches may involve high-speed interfaces or firmware upgrades. Be sure to quiz your vendor to make sure you know everything involved in preparing your switches for high-speed networking.

Workstations

Implementing high-speed networking at the workstation involves many of the same considerations as upgrading servers. Contact your workstation vendor to make sure the network adapters you have chosen are compatible. Select network adapters with PCI buses if at all possible. Also, make sure the workstations are running a version of the operating system that supports the network adapter driver, and that they have sufficient memory and hard disk storage to accommodate the operating system and drivers. And remember, to get the desired performance, you may need to replace the workstation altogether.

Service Cost Worksheet

Service Cost worksheet when considering the cost of any outside service providers you may retain to help you with your high-spa networking implementation. This will include contract programmers to help you enhance applications and network integrators to help you upgrade your server's network operating system as well as your switches and routers. Don't forget PC maintenance companies that can help you upgrade your desktop workstations. Finally, you may need to hire cabling contractors to help upgrade racks, risers, and patch panels in your wiring closets.

Staffing and Staff Development Worksheet

Hiring and/or training a staff to install and maintain a high-speed network is a significant expense. High-speed networking technologies are relatively new; chances are your current staff hasn't been adequately trained in them. Therefore, before you dive into a high-speed network implementation, you'll need to make sure that your staff has acquired the necessary skills in both troubleshooting and management. This means they need to learn not only how to physically connect devices to the high-speed network but also to optimize drivers and operate management applications for the protocol.

Preparing your staff to handle these responsibilities includes sending them to courses and seminars, purchasing books and other reference materials, and possibly hiring temporary staff to keep your network running while your regular staff acquires expertise in high-speed networking.

Sometimes developing existing staff isn't enough. You may even have to hire additional staff that is already experienced in high-speed networking. If that's the case, be sure to include recruiting and hiring costs into your implementation budget. The Staffing and Staff Development worksheet can help you estimate your budget.

Time Estimation Worksheet

One of the hardest figures to estimate is that of time. Using the Time Estimate worksheet, come up with an estimated time to upgrade each server, hub, router, workstation, and switch; then, multiply that by the number of units of each piece of equipment you will convert.

Don't forget, the cost isn't limited to just time spent on the actual installation and configuration of your high-speed network. A major expense of converting your network protocol will be the costs of downtime, reduced productivity of everyone in your organization while systems are being optimized and the inevitable conversion problems are being solved, and reduced productivity of your staff while they become comfortable with the new equipment, software, and systems. Also, don't forget the opportunity cost associated with a network conversion: what won't get done well, or get done at all, while your staff is concentrating on implementing the high-speed protocol?

1.7 SUMMARY

- High-speed networks will be based on optical fiber technology
- There are several server components that affect network performance. Before contemplating upgrading to a higher-speed network, be sure to check out the processor speed, Disk subsystem, and Random Access Memory.

- The first hubs were nothing more than signal repeaters. They took a data signal from one cable segment and repeated it over another, extending the length of the network
- Performance problems are related to the connection between the host device and the network media, but aren't the result of insufficient bandwidth. Here are some of the more common performance-sapping problems that occur at the network connection.

1.8 KEYWORDS

- Network Interface Card(NIC)
- Management Information Systems(MIS)

1.9 Review Questions

Q1. What are high speed networks? Explain the areas of high speed networks.

Q2. How performance of high speed networks?

Q3. What is backbone? How it connect network components?

Q4. What are the cost factors on high speed networks?

1.10 Further Readings

- Tere Parnell - Building High Speed Networks, -- TMH
- Cooper E. - Broadband Network Technology -- Prentice Hall
- Tanenbaum - Computer Networks – PHL

LESSON 2 FIBER DISTRIBUTED DATA INTERFACE (FDDI)

2.1 OBJECTIVES

Main objective of this chapter is to introduce the reader about the Fiber Distributed Data Interface (FDDI).at the end of this chapter reader will able the answer about what is FDDI, how does it help in speed of data transmission from one station to another.

2.2 INTRODUCTION

FDDI stands for "Fiber Distributed Data Interface "FDDI is a group of networking specifications standardized by ANSI in the mid-1980s. An FDDI network supports data transfer speeds of 100 Mbps over a cable and uses a rotating token to define which system can send data at any given time. Fiber Distributed Data Interface, or FDDI, was the first 100Mbps transport protocol available for local area networks. When it first appeared, FDDI was expensive both to install and to manage. This was largely because the protocol ran only on fiber-optic cabling, which was scarce and difficult to install. Therefore, it was reserved almost exclusively for backbone use. Although most FDDI nodes are still mostly found on backbone segments, with the introduction of protocol's copper wire implementation—twisted-pair-physical media depend (TP-PMD)—along with a drop in the cost of installing optical fiber, this protocol finding its way to the desktops more often than ever before. Although, with the rec advent of other high-speed transport protocols that require less investment of time money to implement, FDDI may have missed its chance for widespread deployment, still may be the right choice for network's backbone or power workgroup.

FDDI networks are comprised of two physical paths, or "rings," that transfer data in opposite directions. The primary ring carries data between systems, while the secondary ring is used for redundancy. If a system on the network causes an interruption in the primary data path, the secondary ring is used until the primary ring is functional again. A variation of FDDI, called FDDI Full Duplex Technology (FFDT), uses the secondary ring as an additional primary channel. This type of FDDI network has no redundancy, but supports data transfer rates up to 200 Mbps.

2.3 DEFINITION

FDDI (Fiber Distributed Data Interface) is a set of ANSI and ISO standards for data transmission on fiber optic lines in a local area network (LAN) that can extend in range up to 200 km (124 miles). The FDDI protocol is based on the token ring protocol. In addition to being large geographically, an FDDI local area network can support thousands of users. FDDI is frequently used on the backbone for a wide area network (WAN).

2.3.1 FDDI Protocol

Fiber Distributed Data Interface (FDDI), developed by American National Standards Institute (ANSI) is a token passing ring network that operates at 100 Mb/s on optical fiber-medium. Its medium access control approach has close similarity with the IEEE 802.5 standard, but certain features have been added to it for higher reliability and better performance. Key features of FDDI are outlined in this section.

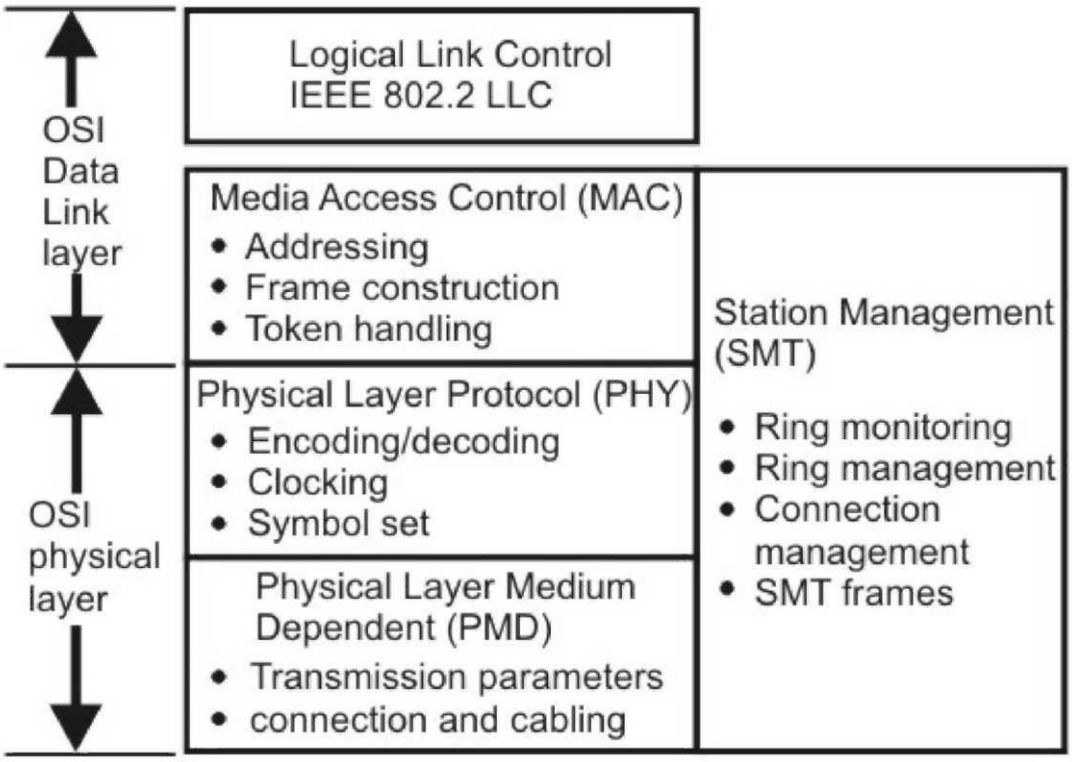


Figure 2.1 FDDI protocols

The FDDI standard divides transmission functions into 4 protocols: physical medium dependent (PMD), Physical (PHY), media access control (MAC) and Logical link control (LLC) as shown in Figure 2.1. These protocols correspond to the physical and data link layer of OSI reference model. Apart from these four protocols, one more protocol which span across both data link and physical layer (if considered of OSI), used for the station management.

2.3.2 Medium

As shown in Table 2.1, the standard physical medium is multi-mode 62.5/125 micron optical fiber cable using light emitting diode (LED) transmitting at 1300 nanometers, as the light source. FDDI can support up to 500 stations with a maximum distance of 2 Km between stations and maximum ring circumference of 200 Km. Single-mode 8-10/125 micron optical fiber cables has also been included in the standard for connecting a pair of stations separated by a distance in excess of 20 km.

Table 2.1 FDDI Physical layer specification

Trans. Medium	Optical Fiber 62.5/125 um	Twisted pair CAT5-UTP
Data Rate	100 Mbps	100Mbps
Signaling Technique	4B/5B/NRZ-I 125 Mbaud	MTL-3
Max. No. Repeaters	100	100
Max. distance	2Km	100m

The standard has also been extended to include copper media - Shielded Twisted Pair (STP) and some categories of Unshielded Twisted Pair (UTP) with a maximum distance of 100 m between stations. FDDI over copper is referred to as *Copper-Distributed Data Interface (CDDI).*

Optical fiber has several advantages over copper media. In particular, security, reliability, and performance are all enhanced with optical fiber media because fiber does not emit electrical signals. A physical medium that does emit electrical signals (copper) can be tapped and therefore vulnerable to unauthorized access to the data that is transmitted through the medium. In addition, fiber is immune to radio frequency interference (RFI) and electromagnetic interference (EMI). Fiber historically has supported much higher bandwidth (throughput potential) than copper, although recent technological advances have made copper capable of transmitting at 100 Mbps or more. Finally, FDDI allows 2 Km between stations using multimode fiber, and even longer distances using a single mode fiber. FDDI uses 4B/5B code for block coding. The 5-bit code is selected such that it has no more than one leading zero and no more than two trailing zeros and more than three consecutive 0's do not occur. Table 2.2 shows the encoded sequence for all the 4-bit data sequences. This is normally line coded with NRZ-I.

In case of failure of a node or a fiber link, the ring is restored by wrapping the primary ring to the secondary ring as shown in Fig. 2.2. The redundancy in the ring design provides a degree

Table 2.2 4B/5B encoding

Data Sequence	Encoded Sequence	Data Sequence	Encoded Sequence
0000	**11110**	**Q (Quiet)**	**00000**
0001	**01001**	**I (Idle)**	**11111**
0010	**10100**	**H (Halt)**	**00100**
0011	**10101**	**J (start delimiter)**	**11000**
0100	**01010**	**K (start delimiter)**	**10001**
0101	**01011**	**T (end delimiter)**	**01101**
0110	**01110**	**S (Set)**	**11001**
0111	**01111**	**R (Reset)**	**00111**
1000	**10010**		
1001	**10011**		
1010	**10110**		
1011	**10111**		
1100	**11010**		
1101	**11011**		
1110	**11100**		
1111	**11101**		

of fault tolerance, not found in other network standards. Further improvement in reliability and availability can be achieved by using *dual ring* of trees and *dual homing* mechanism.

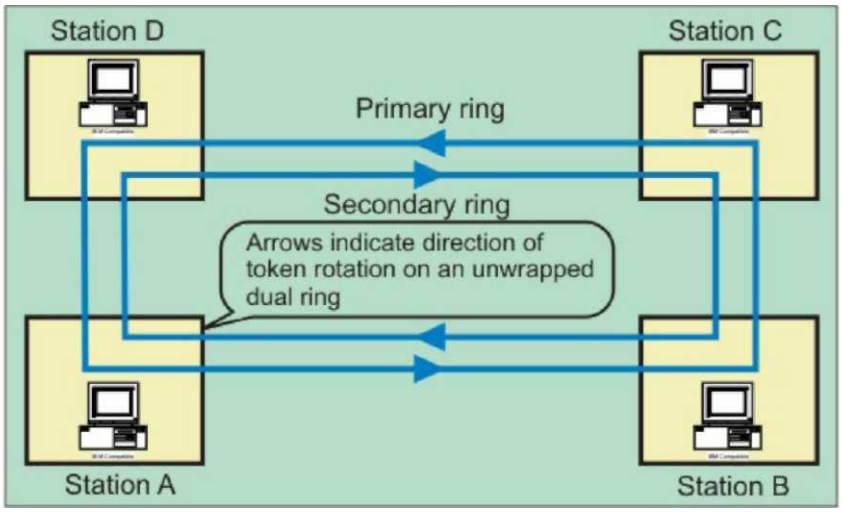


Figure 2.2 FDDI dual counter-rotating ring topology

2.3.3 Fault Tolerance

FDDI provides a number of fault-tolerant features. In particular, FDDI's dual-ring environment, the implementation of the optical bypass switch, and dual-homing support make FDDI a resilient media technology.

Advantages:

- FDDI supports real-time allocation of network bandwidth.
- This allows you to use a wide array of different types of traffic.
- FDDI has a dual ring that is fault-tolerant. The benefit here is that if a station on the ring fails or if the cable becomes damaged, the dual ring is automatically doubled back onto itself into a single ring.
- The FDDI compensates for wiring failures. The stations wrap within themselves when the wiring fails.
- Optical bypass switches are used that can help prevent ring segmentation. The failed stations are eliminated from the ring.

Disadvantages:

- There's a potential for multiple ring failures.
- As the network grows, this possibility grows larger and larger.
- The use of fiber optic cables is expensive.
- This has kept many companies from deploying FDDI in a widespread manner. Instead, they have been using copper wire and the similar method of CDDI.

2.4 ACCESS METHOD: TOKEN PASSING

FDDI is a protocol that employs a token-passing scheme to determine station access. Stations send and receive data packets when they receive the token—data

transmission is not managed by a central controlling station as it is with 100VG-AnyLAN. This token-passing scheme guarantees each station on the network a certain amount of bandwidth. If you take the number of nodes on the network and multiply it by the amount of time it takes each node to transmit a data packet, then you have the maximum amount of time it can take for any station to receive the token. This is what makes FDDI a deterministic network: each node will receive a minimum throughput that you can calculate.

While throughput on an FDDI network is fixed at any given time, it varies according to the number of nodes transmitting on the ring. Therefore, to join an FDDI segment, a must follow a fairly strict set of procedures, which are part of FDD's integrated management capabilities called *station management protocols,* or SMT. These ring initiation protocols first initialize and test the link from the new station to the ring. Next, the station initiates its connection to the ring using a distributed algorithm called the *claim token process. The* claim token process determines whether a token already exists and, if so, reconfigures token's path to include the new station. However, if no token is detected, the claim token protocol requires that all stations attempting to join the FDDI segment transmit sped packets, called *claim frames.* The stations use the claim frames to determine both

- An exact value for token rotation time
- Which station will initiate the new token

Once the token has been created and transmitted, it is used to arbitrate shared access for the stations using a timed token protocol. The first station to join the ring establish and tests its link with the ring, then generates a token that is passed from station to station. When a station receives the token, it can then transmit a fixed amount of

frames. To transmit information onto the ring, a station claims a token that is otherwise circling the ring.

FDDI has prioritization mechanisms implemented via bandwidth allocation. The first mechanism, called synchronous bandwidth allocation (SBA), enables managers to assign a fixed amount of bandwidth to a certain station or stations, thus giving them greater to the token. In SBA, bandwidth is allocated to stations as a percentage of *t*arget token rotation time (TTRT), which is the preset time it takes the token to make one rotation of the ring. Obviously, the total bandwidth allocated via SBA should not exceed 100 percent of the available bandwidth.

The second mechanism, referred to as the asynchronous class of service, takes the bandwidth that is not allocated via SBA and divides it equally among the stations on the ring. The asynchronous service works like this: each station keeps a token rotation timer that tells the station when next to expect the token. When the token next appears, the station compares the *target rotation time* (TRT), or expected time of arrival, to the TTRT, which is the preset time for the token to make one rotation. The TTRT is usually set at 8 MILLISECONDS. If the TRT is less than the TTRT, the station can grab the token and send asynchronous data frames. If the TRT is greater than the TTRT, the token is late, so stations that have only asynchronous class of service must defer to stations with synchronous bandwidth allocation.

During times of heavy traffic, the TRT can get so long that stations sending data packets with low priorities can be completely restricted from access to the ring for a time, eventually, however, all the stations with high-priority packets will send their data, lowering the TRT and thereby letting stations with low-priority transmissions have a chance to claim the token.

2.4.1 FDDI Dual-Ring Topology

Two of the main goals for FDDI during its development were speed and reliability. Optical fiber was selected as the network transmission medium because of its capability to transmit data at high speeds. The topology chosen was the ring topology, similar to token-ring networks. However, to provide enhanced reliability, a dual-ring topology was developed that uses two rings that transmit data in opposite directions(counter-rotating rings). Figure 2.3 shows the layout of a simple FDDI dual

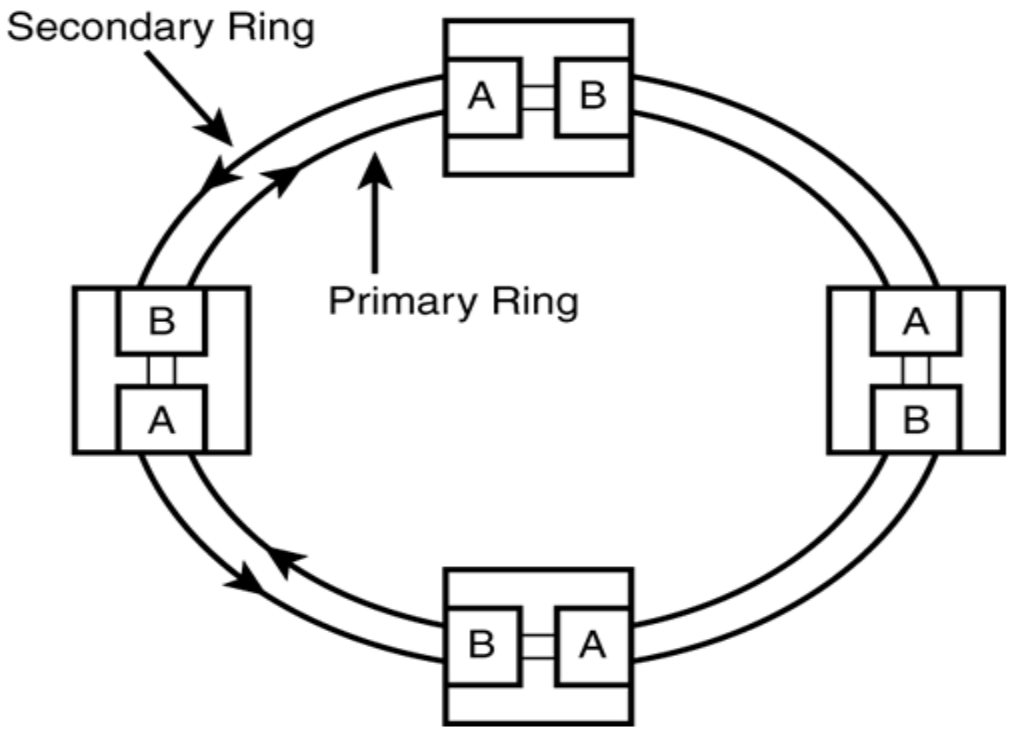


Figure 2.3 FDDI uses the dual counter-rotating ring topology.

2.4.2 FDDI Protocol Standards

FDDI is defined by standards from the American National Standard Institute (ANSI) and the International Organization for Standardization (ISO). The four key components of FDDI are as follows:

- Media Access Control (MAC) layer
- Physical (PHY) layer
- Physical Media Dependent (PMD) layer
- Station Management (SMT) protocol

2.4.3 Transmitting Data on an FDDI Ring

The method used by FDDI to transmit information across optical fiber is light. Two kinds of fiber optic cables can be used; they are classified as either *Single mode* or *Multimode.* Single-mode fiber uses a laser as its source of light and can be used over longer distances than Multimode fiber. Multimode fiber cables allow multiple rays of light, entering the cable at different angles, to carry signals through the cable and use a light-emitting diode (LED) as their light source.

2.4.4 Common Problems Using FDDI

Although FDDI rings perform some basic maintenance functions to help take care of problems, it is still necessary to monitor the LAN periodically to ensure that the network is operating optimally. Also, many problems can't be corrected by software, such as faulty network adapters or network cables. Tools you can use for monitoring and troubleshooting efforts include a cable tester (one intended for use with fiber-optic cable) and a standard LAN protocol analyser. Most FDDI vendors also will provide a station management application that can be used to examine ring functionality and gather statistics and error information. It is a good idea to get a thorough understanding of station management software so that you are better prepared when problems occur.

2.4.5 Ring Wrapping

This process allows for a malfunctioning node to be isolated from the other nodes in the ring. To restore the ring to normal functioning, it is necessary to track down the offending node and determine the cause of the failure. Station management software can provide the information you need to determine which station has left the ring. Perhaps the most common reason why ring wrapping occurs is a simple power failure. This can result from a faulty power supply in one of the attached workstations, or perhaps in a concentrator on the ring. It also can result from human error when someone who doesn't understand how the ring operates mistakenly powers down a station.

2.5 FDDI frame format

FDDI Frame Format:

The FDDI frame format is similar to the format of a Token Ring frame. This is one of the areas in which FDDI borrows heavily from earlier LAN technologies, such as Token Ring. FDDI frames can be as large as 4,500 bytes.

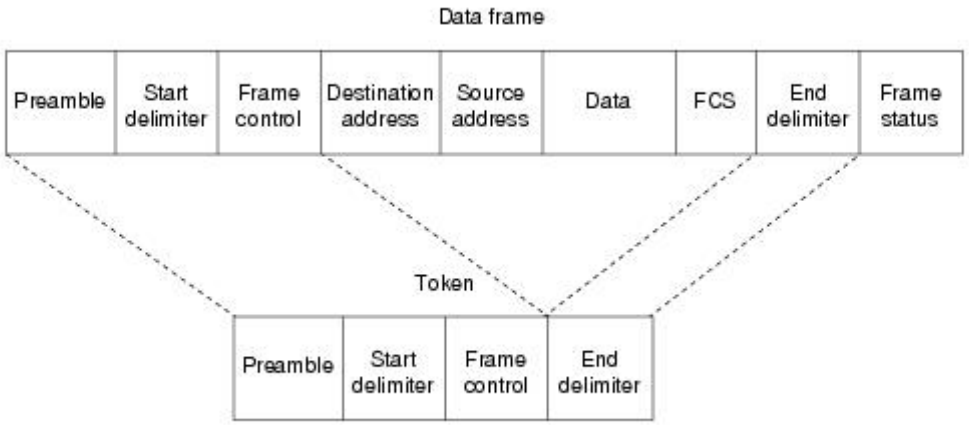


Figure 2.4 FDDI Frame Format

FDDI Frame Fields

The following descriptions summarize the FDDI data frame and token fields illustrated in the above figure.

- Preamble—Gives a unique sequence that prepares each station for an upcoming frame.
- Start delimiter—Indicates the beginning of a frame by employing a signaling pattern that differentiates it from the rest of the frame.
- Frame control—Indicates the size of the address fields and whether the frame contains asynchronous or synchronous data, among other control information.
- Destination address— Contains a unicast (singular), multicast (group), or broadcast (every station) address. As with Ethernet and Token Ring addresses, FDDI destination addresses are 6 bytes long.
- Source address— Identifies the single station that sent the frame. As with Ethernet and Token Ring addresses, FDDI source addresses are 6 bytes long.
- Data— Contains either information destined for an upper-layer protocol or control information.
- Frame check sequence (FCS)—Is filed by the source station with a calculated cyclic redundancy check value dependent on frame contents (as with Token Ring and Ethernet). The destination address recalculates the value to determine whether the frame was damaged in transit. If so, the frame is discarded.
- End delimiter—Contains unique symbols; cannot be data symbols that indicate the end of the frame.
- Frame status—Allows the source station to determine whether an error occurred; identifies whether the frame was recognized and copied by a receiving station

2.6 MEDIA ACCESS CONTROL

The FDDI media access control protocol is responsible for the following services.

(i) Fair and equal access to the ring by using a *timed token protocol.* To transmit on the ring, a station must first acquire the token. A station holds the token until it has transmitted all of its frames or until the transmission time for the appropriate service is over. Synchronous traffic is given a guaranteed bandwidth by ensuring that token rotation time does not exceed a preset value. FDDI implements these using three timers, *Token holding Timer* (THT), which determines how long a station may continue once it has captured a token. *Token Rotation Timer* (TRT) is reset every time a token is seen. When timer expires, it indicates that the token is lost and recovery is started. The *Valid Transmissions Timer* (VTT) is used to time out and recover from some transmit ring errors.

(ii) Construction of frames and tokens are done as per the format shown in Figure. The frame status (FS) byte is set by the destination and checked by the source station, which removes its frame from the ring and generates another token.

(iii) Transmitting, receiving, repeating and stripping frames and tokens from the ring, unlike IEEE 802.5, is possible for several frames on the ring simultaneously. Thus a station will transmit a token immediately after completion of its frame transmission. A station further down the ring is allowed to insert its own frame. This improves the potential throughput of the system. When the frame returns to the sending station, that station removes the frame from the ring by a process called *stripping.*

(iv) It also does *ring initialization, fault isolation* and error detection as we have discussed for IEEE 802.5.

2.7 SUMMARY

- FDDI stands for "Fiber Distributed Data Interface

- FDDI is a group of networking specifications standardized by ANSI in the mid-1980s
- FDDI, was the first 100Mbps transport protocol available for local area networks. When it first appeared, FDDI was expensive both to install and to manage. This was largely because the protocol ran only on fiber-optic cabling.
- FDDI network has no redundancy, but supports data transfer rates up to 200 Mbps.
- FDDI over copper is referred to as *Copper-Distributed Data Interface (CDDI).*
- **FDDI is a protocol that employs a token-passing scheme to determine station access.**

2.8 KEYWORDS

- ***FDDI- Fiber Distributed Data Interface***
- ***CDDI- Copper Distributed Data Interface.***
- ***SMT - Station Management Protocols***
- ***SBA- Synchronous Bandwidth Allocation***
- **TTRT- T*arget Token Rotation Time,***
- ***TRT - Target Rotation Time***
- FCS- Frame check sequence
- PHY- Physical (PHY) layer
- PDM- Physical Media Dependent
- SMT- Station Management (SMT) protocol

2.9 REVIEW QUESTIONS

Q1. Explain FDDI. How does it use to speed up data transmission in networking

Q2. Draw and explain FDDI frame format.

Q3. How data are transmitted in FDDI ring?

Q4. Explain media access mechanism in detail. How it is done in FDDI?

Q5. How token are passed in FDDI ring for data transmission? explain in detail.

2.10 FURTHER READINGS

- Douglas Comer, *Computer Networks and Internets with Internet Applications (third edition)*, Prentice Hall, Upper Saddle River, NJ, 2001, ISBN 0-13-091449-5.
- Deferring Real-Time Traffic for Improved Non-Real-Time Communication in FDDI Networks – Moncef Hamdaoui, Parameswaran Ramanathan – 1995
- Bounding Application-to-Application Delays for Multimedia Traffic in FDDI-Based Communication Systems – Fang Feng, Wei Zhao

LESSON 3 FAST ETHERNET

3.1 OBJECTIVE

The objectie of this chapter is to introduce the reader about fast ethernet which is a 100Base-T ethernat. After reading this chapter reader wil know the differnce of 10Base-T and 100Base-T and what are the variations of fast ethernet

3.2 INTRODUCTION

In computer networking, **Fast Ethernet** is a collective term for a number of Ethernet standards that carry traffic at the nominal rate of 100 Mbit/s, against the original Ethernet speed of 10 Mbit/s. Of the Fast Ethernet standards 100BASE-TX is by far the most common and is supported by the vast majority of Ethernet hardware currently produced. Fast Ethernet was introduced in 1995 and remained the fastest version of Ethernet for three years before being superseded by gigabit Ethernet

3.3. CLASSICAL ETHERNET

3.3.1. CSMA/CD

Classical Ethernet operates at 10Mbps over a bus topology LAN using the CSMA/CD medium access control protocol.

- Bus topology LAN: All stations attach directly to a linear transmission medium (bus). A transmission from any station propagates the length of the medium and can be received by all other stations. Each station has a unique address. Stations transmit data in frames where the destination address is included in the frame header.
- CSMA/CD: A station wishes to transmit first listens to the medium (carrier sense). If the medium is idle, the station may transmit. Collisions are detected and a procedure is used to deal with the situation if a collision occurs.

A more detailed CSMA/CD (carrier sense multiple access with collision detection) protocol is described as follows:

1. Station(s) listen before sending (carrier sense).
2. If the medium is quiet, station(s) start sending (multiple access).

3. If the medium is busy, stations(s) continue to listen until the channel is quiet, then transmit immediately.
4. Transmitting station can detect when the message collides with another station (collision detection).
5. Once a collision is detected, the station stops transmitting and sends a brief jamming signal to inform all stations.
6. After transmitting the jamming signal, the station waits a random amount of time and repeats step 1. The waiting follows the binary exponential backoff algorithm. The binary exponential backoff algorithm works as follows: A station initiates the retransmission. It also recognizes the mean waiting time. If a second attempt of retransmission is needed, the station doubles this mean waiting time and doubles it again on the third attempt, and so on. The aim of this algorithm is to minimize the probability of collision. After a number of unsuccessful attempts, the station gives up and reports an error.

The CSMA/CD protocol works well for light and bursty traffic. For the protocol to be effective, the message frame must be long enough so that the data time is very much longer than the propagation time.

3.3.2. Logical Link Control (LLC)

IEEE defined the LLC in its IEEE802.2 standard. The LLC protocol hides the differences between the various kinds of 802 networks by providing a single format and interface to the network layer. This format, interface, and protocol are all closely based on OSI. LLC forms the upper half of the data link layer, with the MAC sublayer below it. Figure 3.1 depicts the architecture.

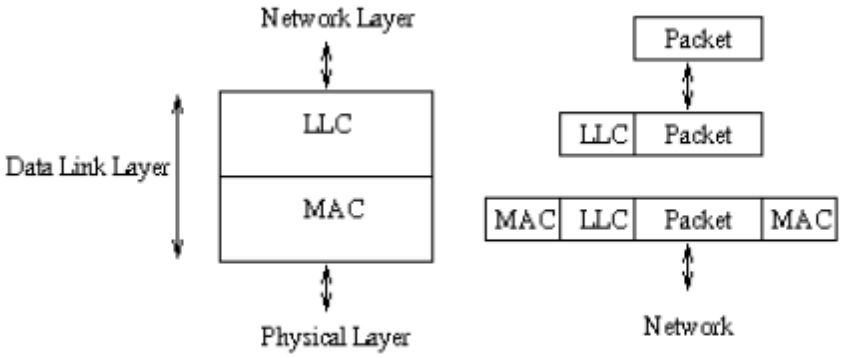


Figure 3.1 LLC and MAC protocols

Typical usage of LLC is as follows. The network layer on the sending machine passes a packet to LLC using the LLC access primitives. The LLC sublayer then adds an LLC header, containing sequence and acknowledgment numbers. the resulting structure is then inserted into the payload field of an 802.x frame and transmitted. At the receiver, the reverse process takes place. LLC provides three service options: unreliable datagram service, acknowledged datagram service, and reliable connection-oriented service. The LLC header is based on the older HDLC protocol.

3.4. FAST ETHERNET

3.4.1. Fast Ethernet Standards

Fast Ethernet (100BASE-T) refers to a set of IEEE 802.3 standards that provides a lowcost, Ethernet-compatible LAN operating at 100 Mbps. Figure 3.2 shows the terminology used.

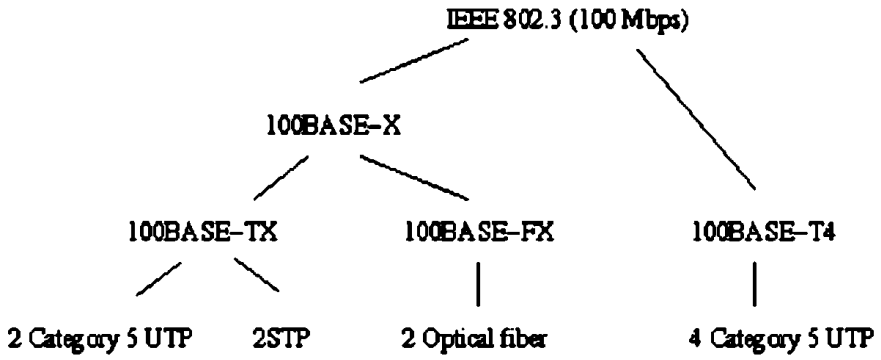


Figure 3.2 IEEE 802.3 100BASE-T options

100BASE-T is the high-speed extension of 10BASE-T. The next table compares the two technologies.

Features	**10BASE-T**	**100BASE-T**
Data rate	10 Mbps	100 Mbps
Standard	IEEE 802.3i	IEEE 802.3u
Topology	Star wired	Star wired
Access method	CSMA/CD	CSMA/CD
Frame size/format	1500 bytes/802.3	1500 bytes/802.3
Encoding	Manchester	4B/5B
Cabling required	UTP Cat. 3/4/5	UTP Cat. 5/STP type
No. of pairs	2	2
Signaling frequency	20 MHz	125 MHz
Distance	100m	100m

Where UTP represents unshielded twisted pairs, STP represents shielded twisted pairs.

Features of various types of 100BASE-T are compared in the next table:

Features	**100BASE-TX**	**100BASE-FX**	**100BASE-T4**
Transmission medium	2 pairs Cat. 5 UTP / 2 pairs STP	2 optical fiber	4 pairs Cat. 3/4/5 UTP
Signaling technique	4B5B, NRZ1	4B5B, NRZ1	8B6T, NRZ
Data rate	100 Mbps	100 Mbps	100 Mbps
Max. segment length	100m	100m	100m
Network span	200m	400m	200m
Physical topology	Star	Star	Star

3.4.2. Cable Standard

EIA/TIA (Electrical Industry Association/Telecommunication Industry Association) is a US standard body that defined cables into five categories:

- Category 1: Typically untwisted 22 AWG or 24 AWG wire with a wide range of impedance and attenuation values. Not for signaling speeds over 1 Mbps.
- Category 2: Similar to Category 1. It uses 22 or 24 AWG solid wire in twisted pairs.
- Category 3: It uses 24 AWG solid wire in twisted pairs and is useful for data transmission at speeds up to 16 Mbps.
- Category 4: It can use 22 or 24 AWG solid wire in twisted pairs. Category 4 has been superseded by Category 5 in most new installations.
- Category 5: This is 22 or 24 AWG unshielded twisted-pair cable and can handle data signaling under specific conditions at 100 Mbps.

3.4.3. Signaling Techniques

- Manchester encoding: A 1 data bit is sent as an uninverted clock pulse, and 0 data bit is sent as an inverted clock pulse. The Manchester encoding scheme is self-clocking, because it includes a signal transition in every bit period. But this makes the band rate twice as much as the bit rate of the signal; i.e., the efficiency of the Manchester encoding scheme is only 50%. For a 10 Mbps data rate, a baud rate of 20 Mbps is required.
- 4B5B NRZ1 encoding: The aim of the 4B5B and other encoding schemes used in Fast Ethernet LANs is to improve the efficiency while keeping the scheme selfchecking. In the 4B5B scheme, 4-bit code-words are encoded as 5-bit code-words. Of the 32 combinations of the 5-bit code-words, only 16 combinations needed to be used as valid code-words. To maintain the self-checking property, only those code words that have at least two transitions in their NRZ1 representation are used. For example, 0000 (a 4-bit code word) is represented by 11110 (a 5-bit code-word). 16 of the 32 (5-bit) code-words are used to represent data values. Of the remaining code-words, some are used as control codes. The 5-bit code-words that have three or more consecutive 0s are not used.
- 8B6T NRZ encoding: The nomenclature of this code implies that 8-bit binary codes are represented as 6-digit ternary codes. In a ternary code, each digit can take three distinct values, can be represented as 0, 1, and 2 (or +, - and 0). A 6-digit ternary code can represent 36 = 729 code combinations. A 8-bit code has 256 distinct code words. This gives a lot of freedom in choosing 256 code words that satisfy the self-checking property. _In the 100BASE-T4 scheme, three lines are used in each direction, each

link carrying a 33.3 Mbps data rate. Thus, the baud rate on each link becomes 6/8 * 33.3 = 25 Mbaud.

- 5B6B and 8B10B encoding: These encoding schemes use concepts similar to those described before. 5B6B is usd for the 100VG-AnyLAN specification. 8B10B is used in a 200 Mbaud standard called ESCON, developed by IBM.

3.5. CONFIGURATION

3.5.1. Repeaters and Switches

The 100BASE-T network is configured in a star-wire topology, with all stations connected directly to a central point called multiport repeater. The repeater has the following functions:

- A valid signal appearing on any single input is repeated on all output links.
- If two inputs occur at the same time, a jam signal is transmitted on all links.

Workstations belong to a single CSMA/CD network (such as stations linked to the same repeater) are said in the same collision domain. A bridge is used to link two repeaters to form a larger network. The bridge works in a store-and-forward fashion and therefore participates in two collision domains. The normal operation of an Ethernet is half duplex in which a station can either send or receive a frame but not both simultaneously. Full duplex operation enables the simultaneous transmit and receive, doubling the data rate of the system. To support duplex, the stations must have full duplex adapter cards. The Hub must be a switched hub or a bridge. In this case, each station may constitute a separate collision domain. In fact, there is no collision and the CSMA/CD algorithm is no longer needed. Switching hubs improve Ethernet performance by overcoming the basic limitations of the repeater architecture (i.e., if one is talking, everyone listens). In a switching system, messages pass only to those stations who need them. Other stations remain free to exchange additional traffic.

3.5.2. Mixed 10/100 Mbps Networks

If an organisation wants to introduce 100 Mbps network gradually, there are two products that can boost the system performance:

- 100 Mbps repeater: It provides full 100 Mbps access to a fixed number of users.
- 10 Mbps switch: It benefits all users without upgrading their existing computer hardware.

Figure 3.3 illustrates such an architecture. The 100 Mbps repeater creates a single 100 Mbps collision domain for all the high-performance workstations. The repeater also provides a 10 Mbps switched port bridging internally to the 100 Mbps service. The 10 Mbps switched port then connects to a 10 Mbps repeater, where a number of normal workstations are connected. In the figure, a thick line represents a link with 100 Mbps bandwidth capability, whereas a thin line represents a 10 Mbps bandwidth capability.

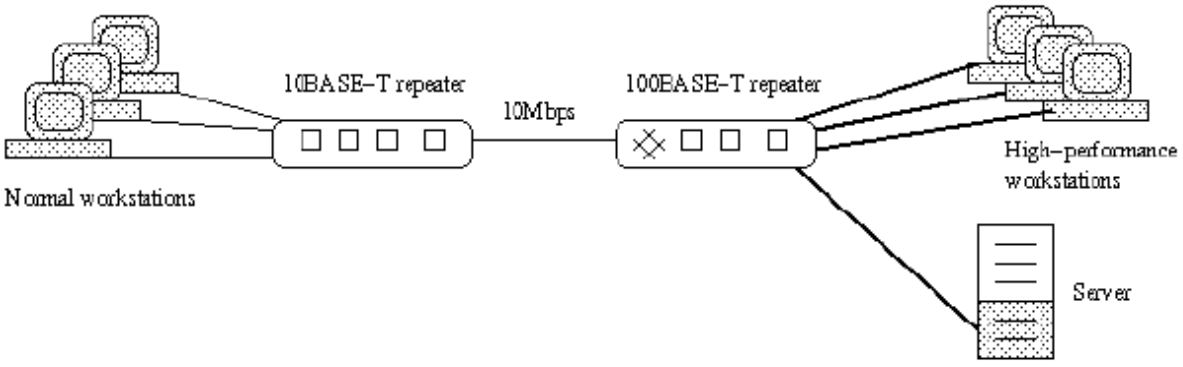


Figure 3.3 A mixed architecture

To extend the 100 Mbps network, a repeater with an expansion port can be used. In Figure 3.4, two repeaters with expansion port (100 Mbps switched) are used to create a two-level

hierarchy. Workstations connected to a different repeater belong to different collision domains.

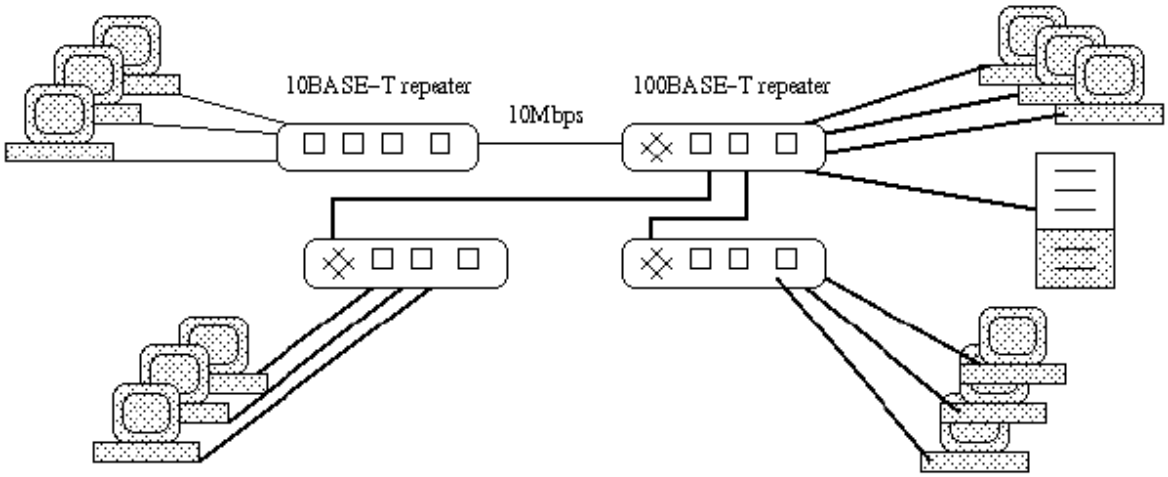


Figure 3.4 Expanding a 100 Mbps network

Another configuration is to use a 10 Mbps switch hub with a 100 Mbps switched port to connect a number of high-performance workstations and a powerful server. The server feeds 100 Mbps into the 100 BASE-T switch port and then the hub fans out to the workstations at 10 Mbps each. Another 10 Mbps repeater can be used to connect other normal workstations to the network. All workstations can benefit the speed boost of the server. Figure 3.5 shows such an example.

3.5.3. Pure 100 Mbps Networks

The minimum 100 Mbps network consists of a number of workstations connected to a 100 Mbps

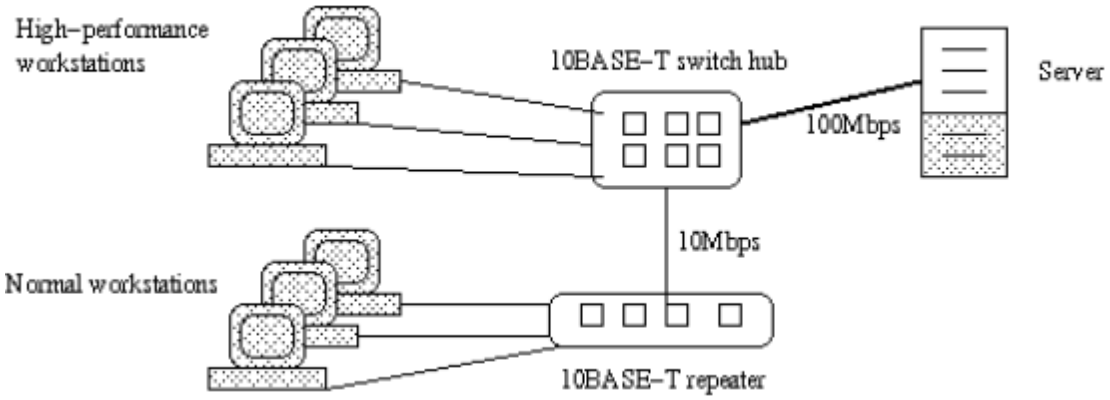


Figure 3.5 Use of a 10 Mbps switch

repeater (Figure 3.6(a)). As the network grows, the first repeater will eventually run out of ports. At this moment, three basic choices can be followed:

- Add a second repeater: As shown in Figure 3.6(b), workstations in this architecture belong to the same collision domain. Also, not all repeaters can be used in such atwo-repeater configuration.
- Use a larger, stackable repeater: A stackable repeater is an expandable repeater which can server a few hundred workstations. Figure 3.6 shows the architecture.
- Combine several repeaters with hierarchical expansion ports. To construct very large networks, Fast Ethernet uses switching. In a typical architecture, several repeaters connect to a switch, which in turn connects to a pool of network data servers. Figure 3.7 shows such an architecture.

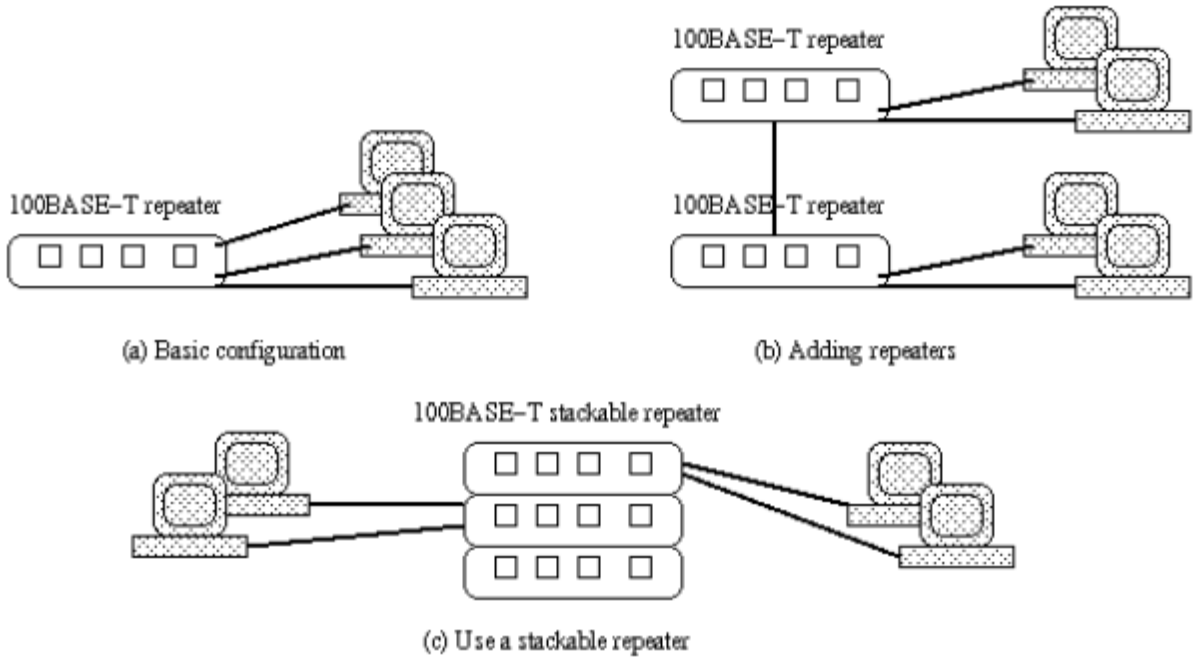


Figure 3.6 Pure Fast Ethernet architectures

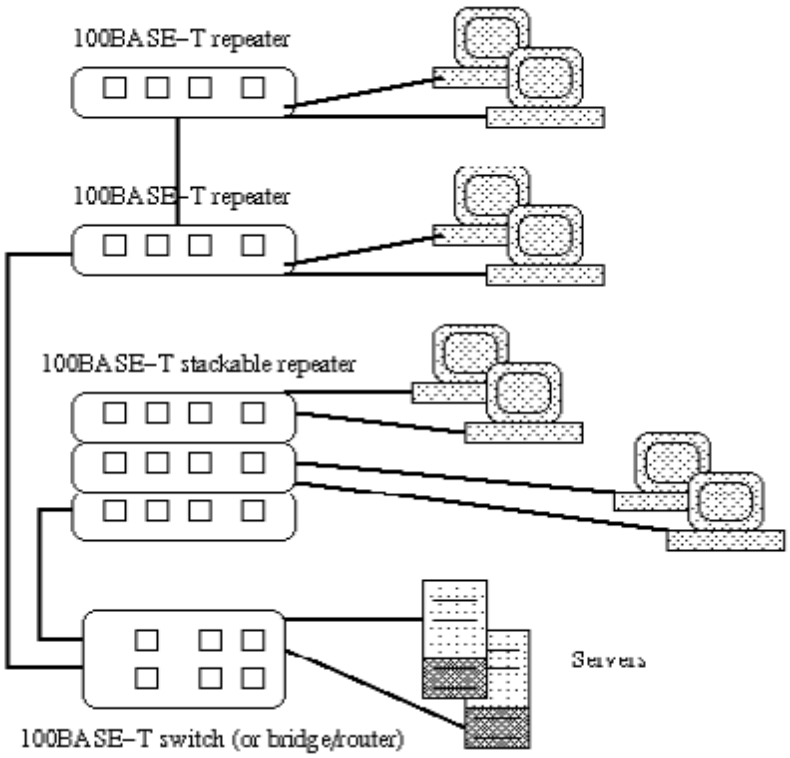


Figure 3.7: A large pure fast Ethernet Network

3.6 IEEE-802.3u

In CSMA/CD, it is important to ensure that the collision window is not a large fraction of overall packet size. When transmission speeds increase, this can only be achieved by either increasing average packet size or restricting maximum propagation delay (by constraining network size). The former option is undesirable since using different packet lengths would compromise compatibility with standard Ethernet so it is judged preferable to restrict the network size, by limiting cable runs to 450m (optical only) maximum. 802.3u specifies only a hub implementation of fast Ethernet, *100Base-T*, and there is no equivalent of 10Base5 or 10Base2. Restricting maximum propagation delay also allows the minimum packet size to remain unchanged. Recall that the minimum packet size in standard Ethernet is chosen so that any collision will always be detected before transmission is complete. If the transmission rate is a factor of ten faster, the cross-network propagation delay must be reduced by a factor of ten to compensate. Note, that these restrictions do not apply to switched full-duplex Ethernet, where collisions do not occur.

In 100Base-T, stations are attached to a central hub or switch using twisted pair or fibre optic point-to-point links. The commonest implementation is *100Base-TX* which uses two high-grade *Category 5 UTP* cables, one for each direction. Full duplex operation requires the use of a 100Base-T switch, as opposed to a simple hub, which merely mimics the operation of a single shared bus. Other implementations which are much less common are *100Base-T4* using 4 pairs of Category 3 UTP for half-duplex operation only and *100Base-FX*, which uses two multimode optic fibres, operating like 100Base-TX, but allowing stations to be up to 450m away from the hub in half-duplex mode. In full-duplex installations, 100Base-FX allows a station to be up to 2km from its switch. Another variant, not part of the original

802.3u, is 100Base-T2 which uses sophisticated equalisation and noise cancellation techniques to support multilevel signalling on two Category 3 twisted pairs.

The general operation of Fast Ethernet is similar to that of its slower analogue, although it operates ten times faster: for example, at 10Mbps, frames must be separated by an inter-frame gap of 9.6ms; in 802.3u, this is reduced to 960ns. The whole point of 100Base-T is to make migration from 10Base-T as easy as possible, so most 100Base-T hubs and switches can handle a mix of stations operating at both 10 and 100Mbps. Likewise, most 100Base-T station adapters (Ethernet cards) can be connected to a 10Base-T or 100Base-T port, and will adjust themselves to the correct speed automatically.

3.7 SUMMARY

100Base-T is probably best described as a high-performance economy model. It is probably the easiest of the high-speed network desktop protocols to implement. At $200 to $5,000 per managed port and falling dramatically, it nearly competes with 10Base-T cost. It also saves you money as well as time by building on proven technology and expertise. It's easy to integrate into an existing 10Base-T network. However, because it is modelled so closely after 10Base-T, it has some of the slower protocol's inherent obstacles limiting its scalability.

3.8 KEYWORDS

- MAC : Media Access Control
- *RS:* Reconciliation Sublayer
- *STP*: *Shielded Twisted Pair*
- MDI: *Medium Dependent Interface*.
- PCS: Physical Coding Sublayer.

- PMA: Physical Medium Attachment Sublayer.
- PMD: Physical Medium Dependent Sublayer.

3.9 REVIEW QUESTIONS

Q1. What do you mean by fast Ethernet?

Q2. Differentiate between 10Base-T Ethernet and 100Base-T Ethernet.

Q3. Define the function of physical layer in 802.3u.

Q4. What are the responsibilities of MAC layer in 802.3u?

Q5. Write note on followings

- 100Base-T4
- 100Base-T2
- 100Base-FX

3.10 FURTHER READINGS

- Pcweek Switched & Fast Ethernet Paperback by Robert A. Breyer (Author)
- Fast Ethernet Implementation and Migration Solutions (Mcgraw-Hill Series on Computer Communications)

LESSION 4 Gigabit Ethernet

4.1 OBJECTIVE

After reading this chapter the student will be able to discuss about the history if Ethernet, explain the gigabit Ethernet alliance (GEA), describe physical layer and list application of 10 gigabit Ethernet.

4.2 Introduction: Gigabit Ethernet

Ethernet is the world's most pervasive networking technology , since the 1970's. It is estimated that in 1996, 82% of all networking equipment shipped was Ethernet. In 1995 ,the Fast Ethernet Standard was approved by the IEEE. Fast Ethernet provided 10 times higher bandwidth, and other new features such as full-duplex operation, and auto-negotiation. This established Ethernet as a scalable technology. Now, with the emerging Gigabit Ethernet standard, it is expected to scale even further.

The Fast Ethernet standard was pushed by an industry consortium called the Fast Ethernet Alliance. A similar alliance, called the Gigabit Ethernet Alliance was formed by 11 companies in May 1996 , soon after IEEE announced the formation of the 802.3z Gigabit Ethernet Standards project. At last count, there were over 95 companies in the alliance from the networking, computer and integrated circuit industries.

A draft 802.3z standard was issued by IEEE in July 1997. The last technical changes are expected to be resolved by September. The standard is expected to be adopted by March 1998.

The new Gigabit Ethernet standards will be fully compatible with existing Ethernet installations. It will retain Carrier Sense Multiple Access/ Collision Detection (CSMA/CD) as the access method. It will support full-duplex as well as half duplex modes of operation. Initially, single-mode and multi mode fiber and short-haul coaxial cable will be supported. Standards for twisted pair cables are expected by 1999. The standard uses physical signalling technology used in Fiber Channel to support Gigabit rates over optical fibers.

Initially, Gigabit Ethernet is expected to be deployed as a backbone in existing networks. It can be used to aggregate traffic between clients and "server farms", and for connecting Fast Ethernet switches. It can also be used for connecting workstations and servers for high - bandwidth applications such as medical imaging or CAD.

History of Ethernet

Today, Ethernet is synonymous with the IEEE 802.3 standard for a "1-persistent CSMA/CD LAN". The 802.3 standard has an interesting history. The beginning, is generally considered to be the University of Hawaii ALOHA network. This system is the ancestor of all shared media networks. The original Ethernet, developed by Xerox was based on the ALOHA

system. It was a 2.94 Mbps CSMA/CD system and was used to connect over 100 personal workstations on a 1 Km cable. It was so successful, that Xerox, DEC and Intel came up with a 10 Mbps standard. The IEEE 802.3 standard was based on the 10 Mbps Ethernet.

CSMA/CD refers to the protocol used by stations sharing the medium, to arbitrate use of the medium. A sender has to "listen" to the medium. If no one else is transmitting, then the sender may transmit. If two senders start transmitting at the same time, then a *collision*is said to have occurred. Transmitting stations, therefore, have to listen to the medium for collisions while transmitting, and retransmit a packet after some time, if a collision occurs.

The original 802.3 standard was published in 1985. Originally two types of coaxial cables were used called *Thick Ethernet*and *Thin Ethernet*. Later unshielded copper twisted pair (UTP) , used for telephones, was added.

In 1980, when Xerox, DEC and Intel published the DIX Ethernet standard, 10 Mbps was a lot of bandwidth. Since then, as computing technology improved, network bandwidth requirements also increased. In 1995, IEEE adopted the 802.3u Fast Ethernet standard. Fast Ethernet is a 100 Mbps Ethernet standard. Fast Ethernet established Ethernet scalability. With Fast Ethernet came full-duplex Ethernet. Until, now, all Ethernets worked in half-duplex mode, that is, if there were only two station on a segment, both could not transmit simultaneously. With full-duplex operation, this was now possible.

The next step in the evolution of Ethernet is Gigabit Ethernet. The standard is being developed by the IEEE 802.3z committee.

1.2 The Gigabit Ethernet Alliance (GEA)

In March 1996, the IEEE 802.3 committee approved the 802.3z Gigabit Ethernet Standardization project. At that time as many as 54 companies expressed there intent to participate in the standardization project. The Gigabit Ethernet Alliance was formed in May 1996 by 11 companies : 3Com Corp., Bay Networks Inc., Cisco Systems Inc., Compaq Computer Corp., Granite Systems Inc., Intel Corporation, LSI Logic, Packet Engines Inc., Sun Microsystems Computer Company, UB Networks and VLSI Technology.

The Alliance represents a multi-vendor effort to provide open and inter-operable Gigabit Ethernet products. The objectives of the alliance are :

- supporting extension of existing Ethernet and Fast Ethernet technology in response to demand for higher network bandwidth.
- developing technical proposals for the inclusion in the standard
- establishment of inter-operability test procedures and processes

Currently membership of the alliance is over 95 companies. This indicates that the emerging standard will be backed by the industry. The alliance is pushing for speedy approval of the standard. So far, the standardization is proceeding without any delays, and is expected to be approved by March 1998.

Gigabit Ethernet (GigE) is becoming more and more popular. The accelerating growth of LAN traffic is pushing network administrators to look to higher-speed network technologies to solve the bandwidth crunch. These administrators—who typically have either Ethernet or FDDI backbones today—have several alternatives to choose from. Although each network faces different issues, **Gigabit Ethernet** meets several key criteria for choosing a high-speed network:

- Easy, straightforward migration to higher performance levels without disruption

- Low cost of ownership—including both purchase cost and support cost
- Capability to support new applications and data types
- Network design flexibility

One of the most important questions network administrators face is how to get higher bandwidth without disrupting the existing network. Gigabit Ethernet follows the same form, fit and function as its 10 Mbps and 100 Mbps Ethernet precursors, allowing a straightforward, incremental migration to higher-speed networking. All three Ethernet speeds use the same IEEE 802.3 frame format, full-duplex operation and flow control methods. In half-duplex mode, Gigabit Ethernet employs the same fundamental CSMA/CD access method to resolve contention for the shared media. And, Gigabit Ethernet uses the same management objects defined by the IEEE 802.3 group. Gigabit Ethernet is Ethernet, only faster.

Ethernet Frame Format

It is simple to connect existing lower-speed Ethernet devices to Gigabit Ethernet devices using LAN switches or routers to adapt one physical line speed to the other. Gigabit Ethernet uses the same variable-length (64- to 1514-byte packets) IEEE 802.3 frame format found in Ethernet and Fast Ethernet (Figure 4.1). Because the frame format and size are the same for all Ethernet technologies, no other network changes are necessary. This evolutionary upgrade path allows Gigabit Ethernet to be seamlessly integrated into existing Ethernet and Fast Ethernet networks.

Figure 4.1: Frame Format of Gigabit Ethernet

Full–and Half-Duplex Gigabit Ethernet Operation

Two nodes connected via a full-duplex, switched path can simultaneously send and receive packets. Gigabit Ethernet follows this standard to communicate in full-duplex mode

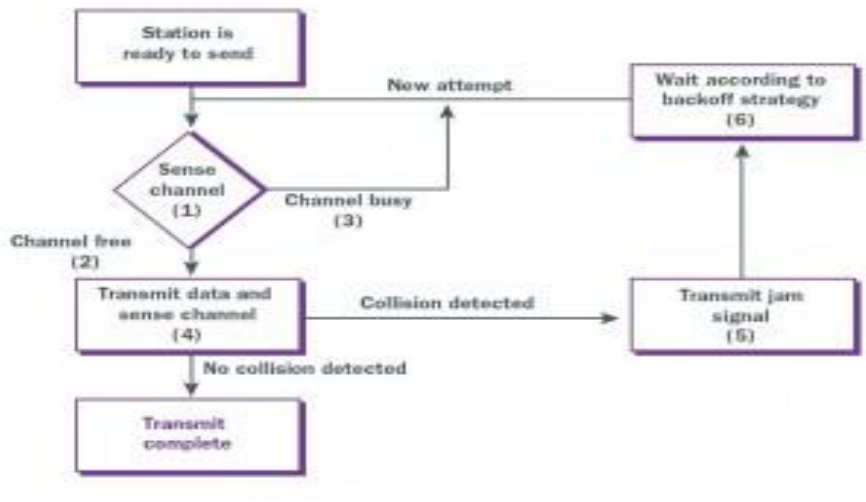


Figure 4.2: CDMA/CD access method

Gigabit Ethernet also employs standard Ethernet flow control methods to avoid congestion and overloading. When operating in half-duplex mode, Gigabit Ethernet adopts the same fundamental CSMA/CD access method to resolve contention for the shared media. The CSMA/CD method is illustrated in Figure 4.2.

The Gigabit Ethernet CSMA/CD method was enhanced in order to maintain a 200-meter collision diameter at gigabit speeds. Without this enhancement, minimum-sized Ethernet packets could complete transmission before the transmitting station senses a collision, thereby violating the CSMA/CD method. To resolve this issue, both the minimum CSMA/CD carrier time and the Ethernet slot time have been extended from their present value of 64 bytes to a new value of 512 bytes. (Note that the minimum packet length of 64

bytes has not been affected.) Packets smaller than 512 bytes have been augmented with a new carrier extension field following the CRC field. Packets longer than 512 bytes have not been extended. These changes, which can impact small-packet performance, have been offset by incorporating a new feature, called packet bursting, into the CSMA/CD algorithm. Packet bursting will allow servers, switches and other devices to send bursts of small packets in order to fully utilize available bandwidth.

Devices that operate in full-duplex mode (switches and buffered distributors) are not subject to the carrier extension, slot time extension or packet bursting changes. Full-duplex devices will continue to use the regular Ethernet 96-bits inter frame gap (IFG) and 64-byte minimum packet size.

Gigabit Ethernet Management Objects

As in the transition from Ethernet to Fast Ethernet, the fundamental management objects familiar to most network managers are carried forward with Gigabit Ethernet. For example, SNMP defines a standard method to collect device-level Ethernet information. SNMP uses management information base (MIB) structures to record key statistics such as collision count, packets transmitted or received, error rates and other device-level information. Additional information is collected by remote monitoring (RMON) agents to aggregate the statistics for presentation via a network management application. Because Gigabit Ethernet uses standard Ethernet frames, the same MIBs and RMON agents can be utilized to provide network management at gigabit speeds.

Gigabit Ethernet: Low Cost of Ownership

Cost of ownership is an important factor in evaluating any new networking technology. The overall cost of ownership includes not only the purchase price of equipment, but also the cost of training, maintenance and troubleshooting.

Competition and economies of scale have driven the purchase price of Ethernet connections down significantly. Though Fast Ethernet products have been shipping only since 1994, even these products have experienced significant price declines over the past two years. Gigabit Ethernet will follow the same price trends as Fast Ethernet. Products on the market today provide cost-effective connections for gigabit transmission rates. The IEEE's goal was to provide Gigabit Ethernet connections at two to three times the cost of a 100BASE-FX interface. As volume builds, reduced line width IC processes are implemented and lowcost opto-electronic devices are developed, the cost of Gigabit Ethernet interfaces will decline.

Switched Gigabit Ethernet connections are lower in cost than 622 Mbps ATM interfaces (assuming identical physical media interfaces), because of the relative simplicity of Ethernet and higher shipment volumes. Gigabit Ethernet repeater interfaces will be significantly lower in cost than 622 Mbps ATM connections, providing users with cost-effective alternatives for data center network backbone and server connections.

Over time, advances in silicon, including 0.35-micron CMOS ASIC technology, will provide even greater performance gains and cost reduction opportunities that will result in a new, even more cost-effective generation of Ethernet technology. Analysis indicates that 0.35-micron processes will achieve 1250 Mbps operation and economically fit one million gates on a single die. This is more than enough to fit a complete Ethernet switch, including

management, a significant amount of buffer memory, and an embedded 32-bit controller, on a single die—with obvious cost advantages.

Finally, because the installed base of users is already familiar with Ethernet technology, maintenance and troubleshooting tools, the support costs associated with Gigabit Ethernet will be far lower than other technologies. Gigabit Ethernet requires only incremental training of personnel and incremental purchase of maintenance and troubleshooting tools. In addition, deployment of Gigabit Ethernet is faster than alternative technologies. Once upgraded with training and tools, network support staff is able to confidently install, troubleshoot and support Gigabit Ethernet installations.

Gigabit Ethernet enables Support for New Applications and Data Types

The emergence of intranet applications portends a migration to new data types, including video and voice. In the past it was thought that video might require a different networking technology designed specifically for multimedia. But today it is possible to mix data and video over Ethernet through a combination of the following:

- Increased bandwidth provided by Fast Ethernet and Gigabit Ethernet, enhanced by LAN switching
- The emergence of new protocols, such as Resource Reservation Protocol (RSVP), that provide bandwidth reservation
- The emergence of new standards such as 802.1Q and 802.1p which will provide virtual LAN (VLAN) and explicit priority information for packets in the network
- The widespread use of advanced video compression

These technologies and protocols combine to make Gigabit Ethernet an extremely attractive solution for the delivery of video and multimedia traffic

Flexible Internetworking and Network Design

Network administrators today face a myriad of internetworking choices and network design options. They are combining routed and switched networks, and building intranets of increasing scale. Ethernet networks are shared (using repeaters) and switched based on bandwidth and cost requirements. The choice of a high-speed network, however, should not restrict the choice of internetworking or network topology.

Gigabit Ethernet can be switched, routed and shared. All of today's internetworking technologies, as well as such technologies such as IP-specific switching and Layer 3 switching, are fully compatible with Gigabit Ethernet, just as they are with Ethernet and Fast Ethernet. Gigabit Ethernet is available in a full duplex repeater (with the accompanying low cost per port) as well as on LAN switches and routers.

Gigabit Ethernet Technology

The simple migration and support offered by Ethernet, combined with the scalability and flexibility to handle new applications and data types, makes Gigabit Ethernet the strategic choice for high-speed, high-bandwidth networking.

Gigabit Ethernet is an extension to the highly successful 10 Mbps and 100 Mbps IEEE 802.3 Ethernet standards. Offering a raw data bandwidth of 1000 Mbps, Gigabit Ethernet maintains full compatibility with the huge installed base of Ethernet nodes.

The Gigabit Ethernet Standard—IEEE 802.3z

To recap the recent history of the Gigabit Ethernet standards process, in July, 1996, after months of initial feasibility studies, the IEEE 802.3 working group created the 802.3z Gigabit Ethernet task force. The key objectives of the 802.3z Gigabit Ethernet task force were to develop a Gigabit Ethernet standard that does the following:

• Allows half- and full-duplex operation at speeds of 1000 Mbps

• Uses the 802.3 Ethernet frame format

• Uses the CSMA/CD access method with support for one repeater per collision domain

• Addresses backward compatibility with 10BASE-T and 100BASE-T technologies

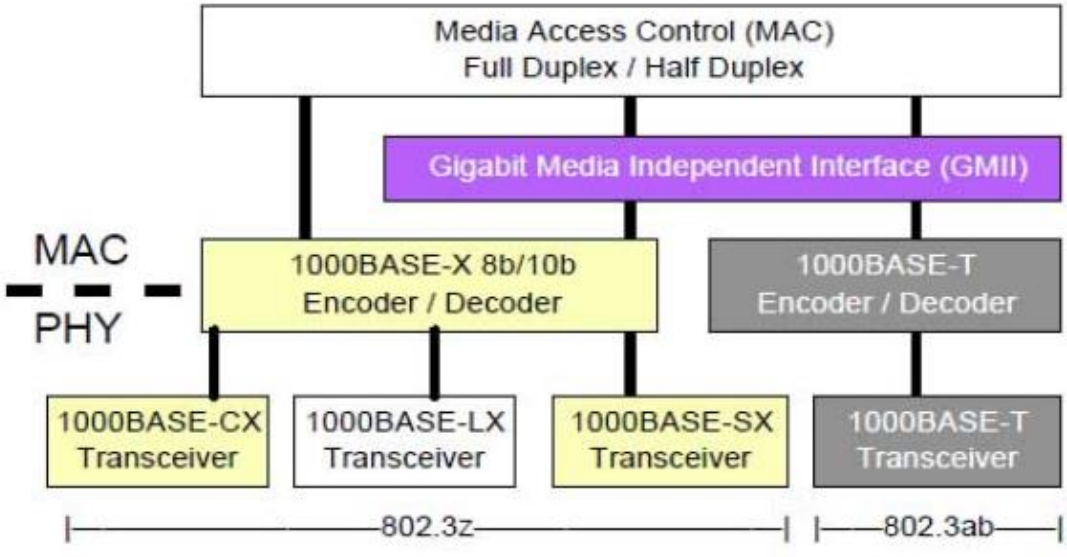


Figure 4.3: layered structure of Gigabit Ethernet

The task force identified three specific objectives for link distances: a multimode fiber-optic link with a maximum length of 550 meters; a single-mode fiber-optic link with a maximum length of 3 kilometers (later extended to 5 kilometers); and a copper based link with a maximum length of at least 25 meters. The IEEE is also actively investigating technology

that would support link distances of at least 100 meters over Category 5 unshielded twisted pair (UTP) wiring. This standards work will be completed this year. In addition, the task force decided to include a specification for an optional Gigabit Media Independent Interface (GMII) in the scope of its work.

4.3 PHYSICAL LAYER

The Physical Layer of Gigabit Ethernet uses a mixture of proven technologies from the original Ethernet and the ANSI X3T11 Fibre Channel Specification. Gigabit Ethernet is finally expected to support 4 physical media types . These will be defined in 802.3z (1000Base-X) and 802.3ab (1000Base-T).

1000Base-X

The 1000Base-X standard is based on the Fibre Channel Physical Layer. Fibre Channel is an interconnection technology for connecting workstations, supercomputers, storage devices and peripherals. Fibre Channel has a 4 layer architecture. The lowest two layers FC-0 (Interface and media) and FC-1 (Encode/Decode) are used in Gigabit Ethernet. Since Fibre Channel is a proven technology, re-using it will greatly reduce the Gigabit Ethernet standard development time.

Three types of media are include in the 1000Base-X standard :

- **1000Base-SX**850 nm laser on multi mode fiber.
- **1000Base-LX**1300 nm laser on single mode and multi mode fiber.
- **1000Base-CX**Short haul copper "twinax" STP (Shielded Twisted Pair) cable

The cabling distances to be supported are given in Table 1 :

Table 1. Cabling Types and Distances

Cable Type	Distance
Single-mode Fiber (9 micron)	3000 m using 1300 nm laser (LX)
Multi mode Fiber (62.5 micron)	300 m using 850 nm laser (SX) 550 m using 1300 nm laser (LX)
Multi mode Fiber (50 micron)	550 m using 850nm laser (SX) 550 m using 1300 nm laser (LX)
Short-haul Copper	25 m

1000Base-T

1000Base-T is a standard for Gigabit Ethernet over long haul copper UTP. The standards committee's goals are to allow up to 25-100 m over 4 pairs of Category 5 UTP. This standard is being developed by the 802.3ab task force and is expected to be completed by early 1999.

4.4 MAC LAYER

The MAC Layer of Gigabit Ethernet uses the same CSMA/CD protocol as Ethernet. The maximum length of a cable segment used to connect stations is limited by the CSMA/CD protocol. If two stations simultaneously detect an idle medium and start transmitting, a collision occurs.

Ethernet has a minimum frame size of 64 bytes. The reason for having a minimum size frame is to prevent a station from completing the transmission of a frame before the first bit has

reached the far end of the cable, where it may collide with another frame. Therefore, the minimum time to detect a collision is the time it takes for the signal to propagate from one end of the cable to the other. This minimum time is called the *Slot Time*. (A more useful metric is *Slot Size*, the number of bytes that can be transmitted in one Slot Time. In Ethernet, the slot size is 64 bytes, the minimum frame length.)

The maximum cable length permitted in Ethernet is 2.5 km (with a maximum of four repeaters on any path). As the bit rate increases, the sender transmits the frame faster. As a result, if the same frames sizes and cable lengths are maintained, then a station may transmit a frame too fast and not detect a collision at the other end of the cable. So, one of two things has to be done : (i) Keep the maximum cable length and increase the slot time (and therefore, minimum frame size) OR (ii) keep the slot time same and decrease the maximum cable length OR both. In Fast Ethernet, the maximum cable length is reduced to only 100 meters, leaving the minimum frame size and slot time intact.

Gigabit Ethernet maintains the minimum and maximum frame sizes of Ethernet. Since, Gigabit Ethernet is 10 times faster than Fast Ethernet, to maintain the same slot size, maximum cable length would have to be reduced to about 10 meters, which is not very useful. Instead, Gigabit Ethernet uses a bigger slot size of 512 bytes. To maintain compatibility with Ethernet, the minimum frame size is not increased, but the "carrier event" is extended. If the frame is shorter than 512 bytes, then it is padded with extension symbols. These are special symbols, which cannot occur in the payload. This process is called *Carrier Extension*.

Carrier Extension

Gigabit Ethernet should be inter-operable with existing 802.3 networks. Carrier Extension is a way of maintaining 802.3 minimum and maximum frame sizes with meaningful cabling distances.

For carrier extended frames, the non-data extension symbols are included in the "collision window", that is, the entire extended frame is considered for collision and dropped. However, the Frame Check Sequence (FCS) is calculated only on the original (without extension symbols) frame. The extension symbols are removed before the FCS is checked by the receiver. So the LLC (Logical Link Control) layer is not even aware of the carrier extension. Fig. 1 shows the ethernet frame format when Carrier Extension is used.

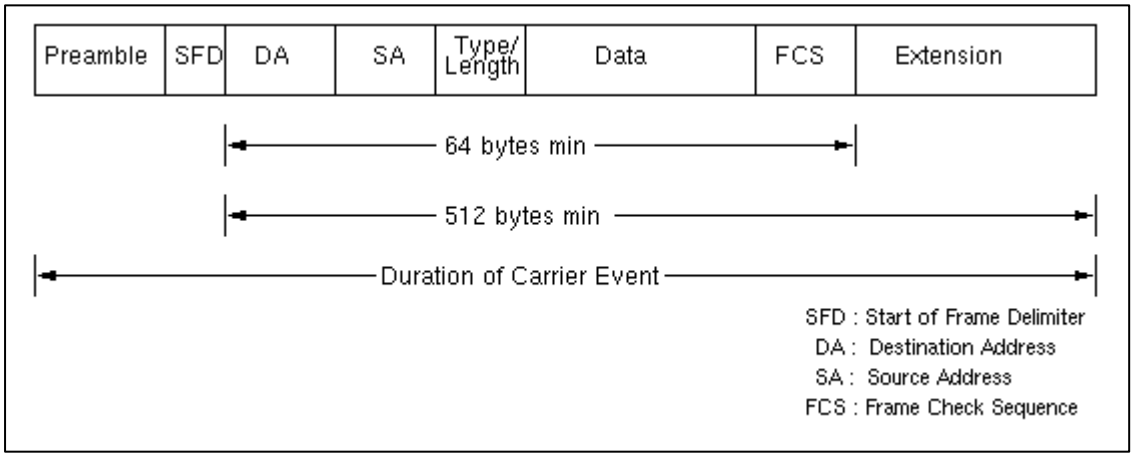


figure 4.4: Eternet Frame Format with Carrier Extension

Packet Bursting

Carrier Extension is a simple solution, but it wastes bandwidth. Up to 448 padding bytes may be sent for small packets. This results in low throughput. In fact, for a large number of small packets, the throughput is only marginally better than Fast Ethernet.

*Packet Bursting*is an extension of Carrier Extension. Packet Bursting is "Carrier Extension plus a burst of packets". When a station has a number of packets to transmit, the first packet

is padded to the slot time if necessary using carrier extension. Subsequent packets are transmitted back to back, with the minimum Inter-packet gap (IPG) until a burst timer (of 1500 bytes) expires. Packet Bursting substantially increases the throughput. Fig. 4.5. shows how Packet Bursting works.

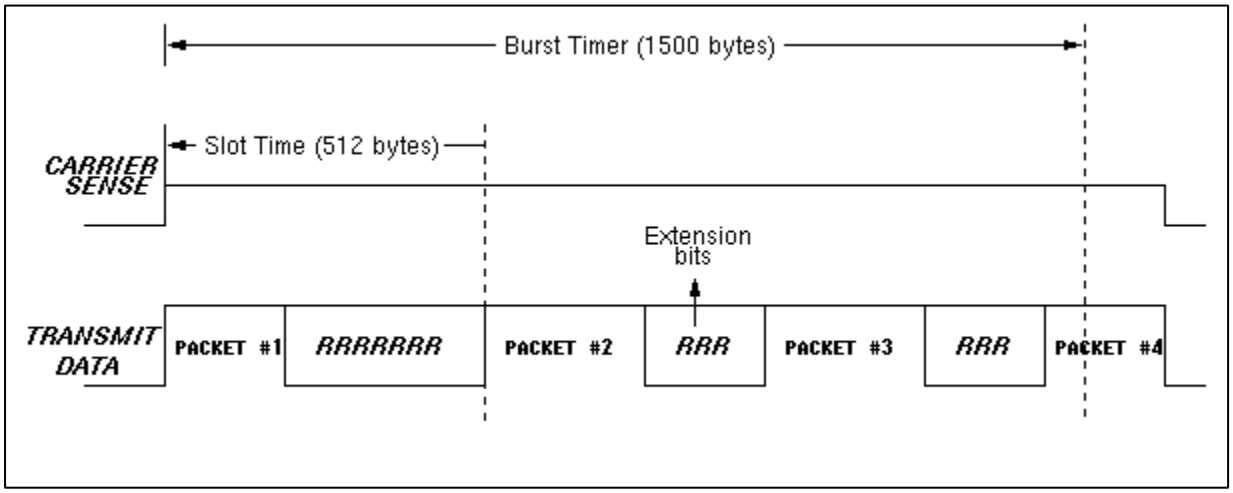


Figure 4.5: Packet Bursting

GMII (Gigabit Media Independent Interface)

The various layers of the Gigabit Ethernet protocol architecture are shown in Fig. 4.5. The GMII is the interface between the MAC layer and the Physical layer. It allows any physical layer to be used with the MAC layer. It is an extension of the MII (Media Independent Interface) used in Fast Ethernet. It uses the same management interface as MII. It supports 10, 100 and 1000 Mbps data rates. It provides separate 8-bit wide receive and transmit data paths, so it can support both full-duplex as well as half-duplex operation.

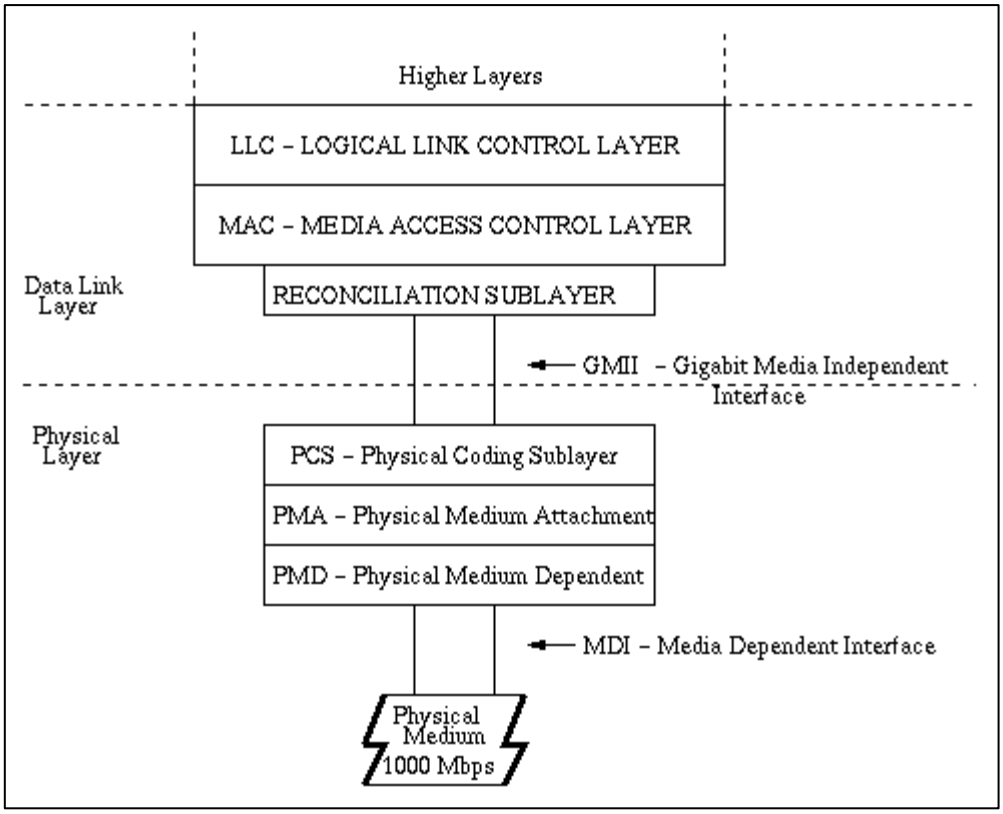


Figure 4.5: Gigabit Ethernet Protocol Architecture

The GMII provides 2 media status signals : one indicates presence of the carrier, and the other indicates absence of collision. The Reconciliation Sublayer (RS) maps these signals to Physical Signalling (PLS) primitives understood by the existing MAC sublayer. With the GMII, it is possible to connect various media types such as shielded and unshielded twisted pair, and single-mode and multi mode optical fibre, while using the same MAC controller.

The GMII is divided into three sublayers : PCS, PMA and PMD.

PCS (Physical Coding Sublayer)

This is the GMII sublayer which provides a uniform interface to the Reconciliation layer for all physical media. It uses 8B/10B coding like Fibre Channel. In this type of coding, groups of 8 bits are represented by 10 bit "code groups". Some code groups represent 8 bit data

symbols. Others are control symbols. The extension symbols used in Carrier Extension are an example of control symbols.

Carrier Sense and Collision Detect indications are generated by this sublayer. It also manages the auto-negotiation process by which the NIC (Network Interface) communicates with the network to determine the network speed (10,100 or 1000 Mbps) and mode of operation (half-duplex or full-duplex).

PMA (Physical Medium Attachment)

This sublayer provides a medium-independent means for the PCS to support various serial bit-oriented physical media. This layer serializes code groups for transmission and deserializes bits received from the medium into code groups.

PMD (Physical Medium Dependent)

This sublayer maps the physical medium to the PCS. This layer defines the physical layer signalling used for various media. The **MDI (Medium Dependent Interface)**, which is a part of PMD is the actual physical layer interface. This layer defines the actual physical attachment, such as connectors, for different media types.

Buffered Distributor

Ethernet today supports full-duplex media, physical layer as well MAC layer. However it still supports half-duplex operation to maintain compatibility. A new device has been proposed which provides hub functionality with full duplex mode of operation. It is called various names such as *Buffered Distributor, Full Duplex Repeater*and *Buffered Repeater.*The term "Buffered Distributor" is used for all these devices in the following discussion.

The basic principle is that CSMA/CD is used as the access method to the network and not to the link. A Buffered Distributor is a multi-port repeater with full-duplex links.

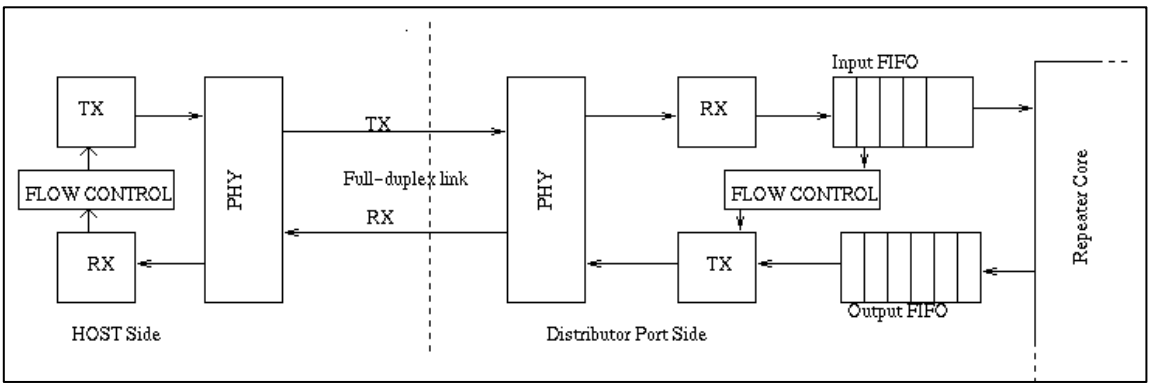


Figure 4.6: Buffered Distributor Architecture

Each port has an input FIFO queue and an output FIFO queue. A frame arriving to an input queue is forwarded to all output queues, except the one on the incoming port. Within the distributor, CSMA/CD arbitration is done to forward the frames to output queues.

Since collisions can no longer occur on links, the distance restrictions no longer apply. The only restriction on cabling distances is the characteristics of the physical medium, and not the CSMA/CD protocol.

Since the sender can flood the FIFO, frame based flow control is used between the port and the sending station. This is defined in the 802.3x standard and already used in Ethernet switches.

The motivation behind development of the Buffered Distributor is it's cost compared to a Gigabit switch and not a need to accommodate half duplex media. The Buffered Distributor provides full duplex connectivity, just like a switch, yet it is not so expensive, because it is just an extension of a repeater.

Topologies

This section discusses the various topologies in which Gigabit Ethernet may be used. Gigabit Ethernet is essentially a "campus technology", that is , for use as a backbone in a campus-wide network. It will be used between routers, switches and hubs. It can also be used to connect servers, server farms (a number of server machines bundled together), and powerful workstations.

Essentially, four types of hardware are needed to upgrade an exiting Ethernet/Fast Ethernet network to Gigabit Ethernet :

- Gigabit Ethernet Network Interface Cards (NICs)
- Aggregating switches that connect a number of Fast Ethernet segments to Gigabit Ethernet
- Gigabit Ethernet switches
- Gigabit Ethernet repeaters (or Buffered Distributors)

The five most likely upgrade scenarios are given below :

Upgrading server-switch connections

Most networks have centralized file servers and compute servers A server gets requests from a large number of clients. Therefore, it needs more bandwidth. Connecting servers to switches with Gigabit Ethernet will help achieve high speed access to servers. . This is perhaps the simplest way of taking advantage of Gigabit Ethernet.

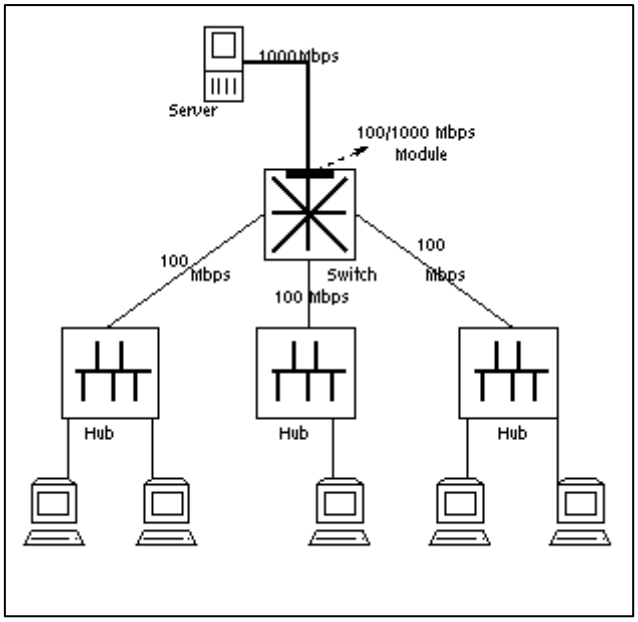
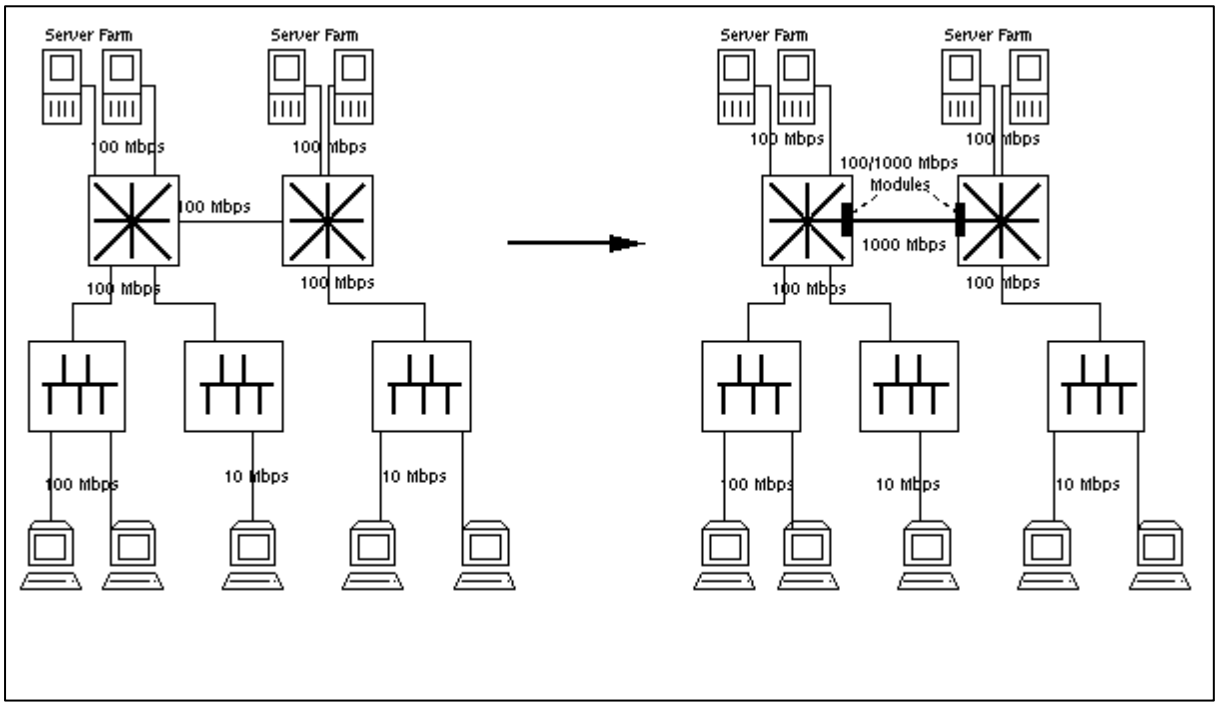


Figure 4.7: Server Switch Connection

Upgrading switch-switch connections

Another simple upgrade involves upgrading links between Fast Ethernet switches to Gigabit Ethernet links between 100/1000 Mbps switches.



Figuer 4.8: Upgrading Switch-Switch connections

Upgrading a Fast Ethernet backbone

A Fast Ethernet backbone switch aggregates multiple 10/100 Mbps switches. It can be upgraded to a Gigabit Ethernet switch which supports multiple 100/1000 Mbps switches as well as routers and hubs which have Gigabit Ethernet interfaces. Once the backbone has been upgraded, high performance servers can be connected directly to the backbone. This will substantially increase throughput for applications which require high bandwidth.

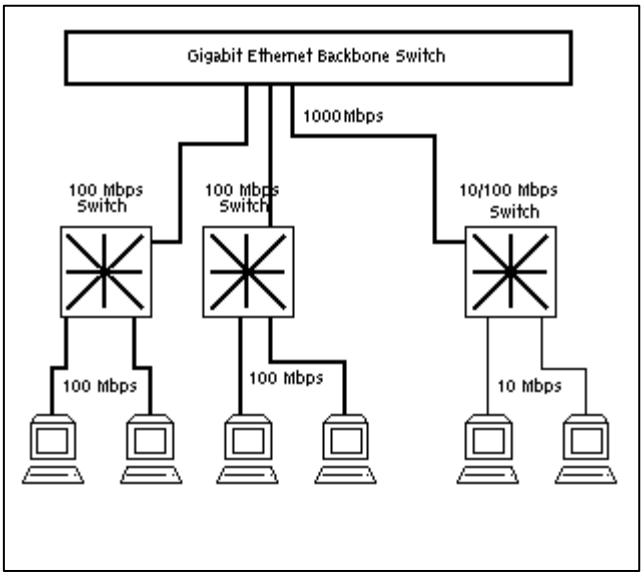


Figure 4.9: Upgrading the Backbone

Upgrading a Shared FDDI Backbone

Fiber Distributed Data Interface (FDDI) is a common campus or building backbone technology. An FDDI backbone can be upgraded by replacing FDDI concentrators or Ethernet-to-FDDI routers by a Gigabit Ethernet switch or repeater.

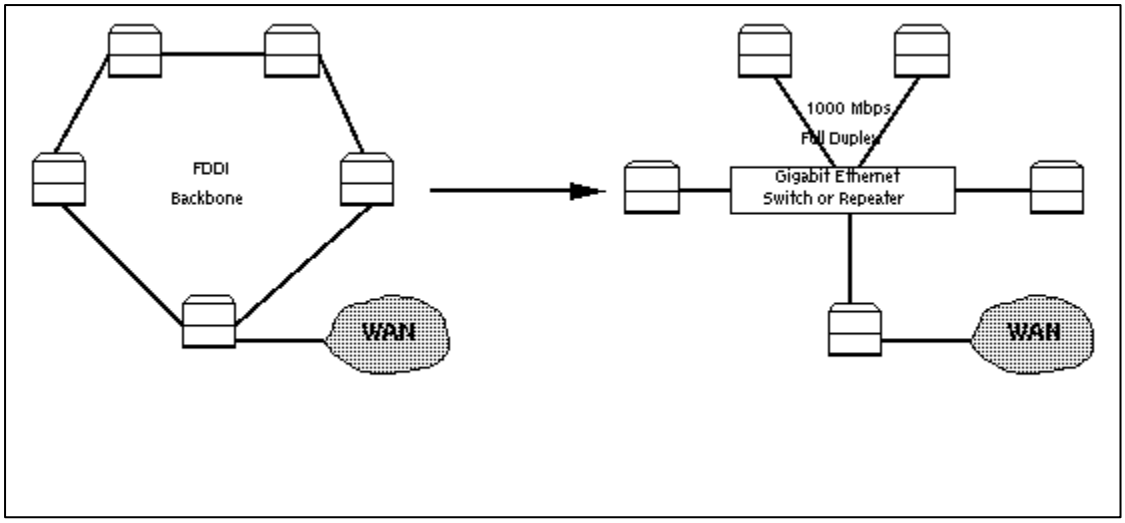


Figure 4.10: Upgrading a FDDI Backbone

Upgrading High Performance Workstations

As workstations get more and more powerful, higher bandwidth network connections are required for the workstations. Current high-end PCs have buses which can pump out more than 1000 Mbps. Gigabit Ethernet can be used to connect such high speed machines.

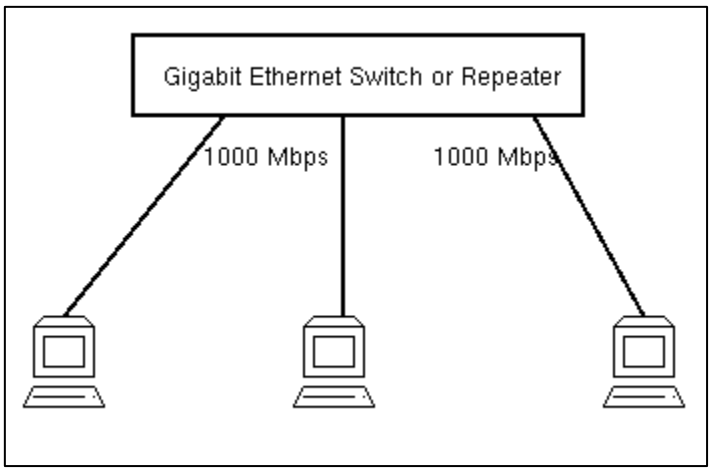


Figure 4.11: Upgrading High Performance Workstation

Gigabit Eithernet provides a wide pipe,it currently has no inherent mechanism for ensuring quality of bandwidth. However,it can use other technologies and quality of services protocol in conjunction with Gigabit Eithernet,although these present their own set of concerns:

- Resource Reservation Protocol (RSVP)
- Virtual LANs (vLANs)

4.5 RESOURCE RESERVATION PROTOCOL (RSVP)

The Resource Reservation Protocol (RSVP) guarantees that a certain amount of bandwidth is set aside for a designated application. Applications use RSVP to request from the network a specific amount of bandwidth for a specific "type of traffic. The job of RSVP is to set up bandwidth reservations over an established route. However, this protocol is still "emerging" - i.e., there is no standard implementation—which means that interoperability among different vendors' devices is not assured.

When implementing RSVP, the receiving device determines how much bandwidth it will reserve for a given type of traffic. It can change this level as necessary. However, you can centralize this control of bandwidth reservations—at least somewhat—by requiring each receiving station to specify its bandwidth reservations to one central device. This device then sets up the reservations for the entire network. However, there is a lot of overhead required in setting up and maintaining this centralized approach. For example every time a node wanted to reserve bandwidth for a new sending node, the central would have to send a notification to all senders. This would tie up a lot of band defeating the whole purpose of RSVP.

RSVP sets up bandwidth reservations using *packet classifiers* and *packet schedulers.* These are routines installed on the receiving devices that make service decisions about each packet received. The receiving device determines whether the requested bandwidth reservation is possible, based on bandwidth resources available and policy restriction, if any. If the

resources are available, the device, the packet classifier, and the packet scheduler transmit the packet. The packet scheduler makes forwarding decision based on the bandwidth reservations. The packet classifier routes the packet based on these forwarding decisions.

How do receiving devices know how much bandwidth to reserve for a given application ? They determine it by examining an applications *flow specification,* or flows flowspec. the flowspec is a pivotal part of the RSVP architecture. It specifies the bandwidth requirements of the application in question, as well as identifying the applicant's data stream as coming from that application. The packet scheduler uses the bandwidth requirement specified in an application's flowspec to establish the bandwidth allocation for that application.

Not all packets will be able to take advantage of the bandwidth requested by the flowspec. That's because each receiving device also has *filterspecs.* These are filters that designate which packets will receive priority treatment based on application or protocol header information. Packets that don't match the filter specs will not receive a guarantee of service. Instead, they will be transmitted on a best-effort basis only, even if they are addressed to the application's session.

4.6 VIRTUAL LAN

Setting up filters or constraints between different groups of users is awkward and time consuming with conventional bridges and routers. Network managers think in terms of workgroups, not the physical location of users. Therefore, they shouldn't have to set up a series of filtering statements based on physical ports. The connectionless nature of Gigabit Ethernet posed some problems for setting vLANs. After all, the nature of the protocol's access method was based on broadcast. However, two emerging standards, 802.1Q and 802.1p, are enabling virtual LAN capability for packet-based networks.

A virtual LAN is a list of device media access control (MAC) or network addresses that are independent of a physical port much like an access list used by some router vendors.

However, virtual LANs have network-wide significance. A device can access any other device on the same virtual LAN. Virtual LANs can define filters among themselves, just like routers can. Rather than configuring and reconfiguring routers every time end stations move, network managers can implement virtual LANs.

Devices on different media can be members of the same virtual LAN. Furthermore users can move end stations onto any segment within the virtual subnet without requiring address reconfiguration.

Virtual LANs enable network managers to group devices logically regardless of physical location and provide dedicated bandwidth and services to each.

ATM vs. Gigabit Ethernet

When ATM (Asynchronous Transfer Mode) was introduced, it offered 155 Mbps bandwidth, which was 1.5 times faster than Fast Ethernet. ATM was ideal for new applications demanding a lot of bandwidth, especially multimedia. Demand for ATM continues to grow for LAN's as well as WAN's.

On the one hand , proponents of ATM try to emulate Ethernet networks via LANE (LAN Emulation) and IPOA (IP over ATM). On the other, proponents of Ethernet/IP try to provide ATM functionality with RSVP(Resource Reservation Protocol) and RTSP (Real-time Streaming Transport Protocol). Evidently, both technologies have their desirable features, and advantages over the other. It appears that these seemingly divergent technologies are actually converging.

ATM was touted to be the seamless and scaleable networking solution - to be used in LANs, backbones and WANs alike. However, that did not happen. And Ethernet, which was for a long time restricted to LANs alone, evolved into a scalable technology.

As Gigabit Ethernet products enter the market, both sides are gearing up for the battle. Currently, most installed workstations and personal computers do not have the capacity to use these high bandwidth networks. So, the imminent battle is for the backbones, the network connections between switches and servers in a large network.

Gigabit Ethernet seems to be ready to succeed. It is backed by the industry in the form of the Gigabit Ethernet Alliance. The standardization is currently on schedule. Pre-standard products with claims of inter-operability with standardized products have already hit the market. Many Fast Ethernet pre-standard products were inter-operable with the standard. So it is expected that most pre-standard Gigabit Ethernet products will also be compatible with the standard. This is possible because many of the companies that have come out with products are also actively participating in the standardization process.

ATM still has some advantages over Gigabit Ethernet :

- ATM is already there. So it has a head start over Gigabit Ethernet. Current products may not support gigabit speeds,but faster versions are in the pipeline.
- ATM is better suited than Ethernet for applications such as video, because ATM has QOS (Quality of Service) and different services available such as CBR (constant bit rate) which are better for such applications. Though the IETF (Internet Engineering Task Force, the standards body for internet protocols) is working on RSVP which aims to provide QOS on Ethernet, RSVP has it's limitations. It is a "best effort"

protocol, that is , the network may acknowledge a QOS request but not deliver it. In ATM it is possible to guarantee QOS parameters such as maximum delay in delivery.

Gigabit Ethernet has its own strengths :

- The greatest strength is that it is Ethernet. Upgrading to Gigabit Ethernet is expected to be painless. All applications that work on Ethernet will work on Gigabit Ethernet. This is not the case with ATM. Running current applications on ATM requires some amount of translation between the application and the ATM layer, which means more overhead.
- Currently, the fastest ATM products available run at 622 Mbps. At 1000 Mbps, Gigabit Ethernet is almost twice as fast.

It is not clear whether any one technology will succeed over the other. It appears that sooner or later, ATM and Ethernet will complement each other and not compete.

4.7 SUMMARY

Gigabit Ethernet is the third generation Ethernet technology offering a speed of 1000 Mbps. It is fully compatible with existing Ethernets, and promises to offer seamless migration to higher speeds. Existing networks will be able to upgrade their performance without having to change existing wiring, protocols or applications. Gigabit Ethernet is expected to give existing high speed technologies such as ATM and FDDI a run for their money. The IEEE is working on a standard for Gigabit Ethernet, which is expected to be out by the beginning of 1998. A standard for using Gigabit Ethernet on twisted pair cable is expected by 1999.

4.8 KEYWORDS

IFG: Inter Frame Gap (IFG)

MIB: Management Information Base (MIB)

RSVP: Resource Reservation Protocol

IPOA: IP over ATM

QOS: Quality Of Services

4.9 REVIEW QUESTION

Q1. Ethernet is LAN technology discuss.

Q2. Draw and explain layered structure of Gigabit Ethernet.

Q3. What is the meaning of RSVP? How does it assure Quality of service?

Q4. How will you upgrade fast Ethernet backbone?

4.10 FURTHER READINGS

- Switched, Fast, and Gigabit Ethernet (3rd Edition) by Sean Riley
- gigabit Ethernet for Metro Area Networks by Paul Bedell

LESSON 5 Fibre Channel

5.1 OBJECTIVES

After reading this chapter the reader will be able to discuss fibre channel topologies, explain fibre channel layers: FC-0, FC-1, FC-2, FC-3 and FC-4.

5.2 INTRODUCTION

Fibre Channel didn't begin life as a network transport protocol. In fact, the specification was not originally intended to work as a network protocol at all. It was first designed and developed to interconnect high-speed peripherals—for example, a cluster of high-performance computers—to a shared mass storage device. The result of this intent is a generic architecture that is a combination of both network and channel connection technology

that concentrates on high-speed transmission and guaranteed delivery of data. This makes it suitable not only for connecting peripherals to hosts, but also for network connectivity over the short haul.

A channel connection provides either a direct or switched point-to-point connection between devices, such as a computer's processor and a peripheral device. The function of a channel isn't sophisticated in concept: it's supposed to transmit data as fast as possible from point A to point B. The destination address is not only predetermined, it is hard-wired—the data really can't go anywhere but its destination. Any error correction it performs is simple and is done in hardware. Routing and address resolution aren't don because they aren't required. Therefore, there's no address and error-correction| information that needs to be carried in the data packet, so the packet overhead is very low. Because of their point-to-point nature and consequently limited processing demands, channel connections are largely implemented in hardware.

Network connections, on the other hand, are multipoint connections that rely on addressing schemes to make sure that the data gets to the appropriate destination. Each data packet traveling along a network connection must contain an address, which each device on the network reads to determine whether the packet is intended for it. Furthermore, network connections routinely have relatively sophisticated error-detection and correction capabilities. Therefore, the packet must contain the address and error-correction information in its header, resulting in higher packet overhead than in a channel connection. However, network connections can support functions, such as routing, that a channel simply isn't designed to support.

Fibre Channel has features of both a channel and a network. It is a high-speed channel that connects devices to a network fabric. The network fabric describes the matrix of connections, from a single cable connecting two devices to a mesh created by a switch connecting many

devices. The fabric can be pretty much any combination of devices including hubs, loops, mainframes, and switches. A fabric can be created specifically to suit the application being supported. In any event, no matter how complex the network fabric is, as far as an individual Fibre Channel port is concerned, it manages a simple point-to-point connection between the workstation and the network fabric.

Fibre Channel's greatest asset is speed. Ethernet, at 10Mbps, transmit data far slower than computers can produce it. Therefore, the primary goal of Fibre channel is to provide computing devices with a throughput mechanism that is closer to the speed of their processors. And speed it has—it delivers banawidth from 133Mbps to 1.062Gbps over a Wide variety of cable types, including multimode fibre, coaxial cable, and shielded l-pair wire, and can support a network span of up to 10km. In fact, Fibre Channel is so efficient and fast that it can deliver speeds of 100MBps (i.e. megabytes per second) in both directions simultaneously, which means it is effectively a 200MBps full duplex channel. fibre Channel is extremely versatile because it is protocol independent. It is a generic transport mechanism, meaning that it can support command sets from several different other channel protocols, such as Small Computer System Interface (SCSI), Internet Protocol (IP), and High Performance Parallel Interface (HIPPI).

Its speed and guaranteed packet delivery make Fibre Channel a good protocol for connecting network devices in a relatively local environment, such as a workgroup or even a campus environment, as shown in Figure 5.1. But because of its high performance, wide network span, and support of various topologies such as point-to-point, arbitrated loop, and switched fabric, Fibre Channel has broader applications than as simply a networking protocol. It can connect LANs, midrange computers, mainframes, disk arrays, and server farms as part of one giant network.

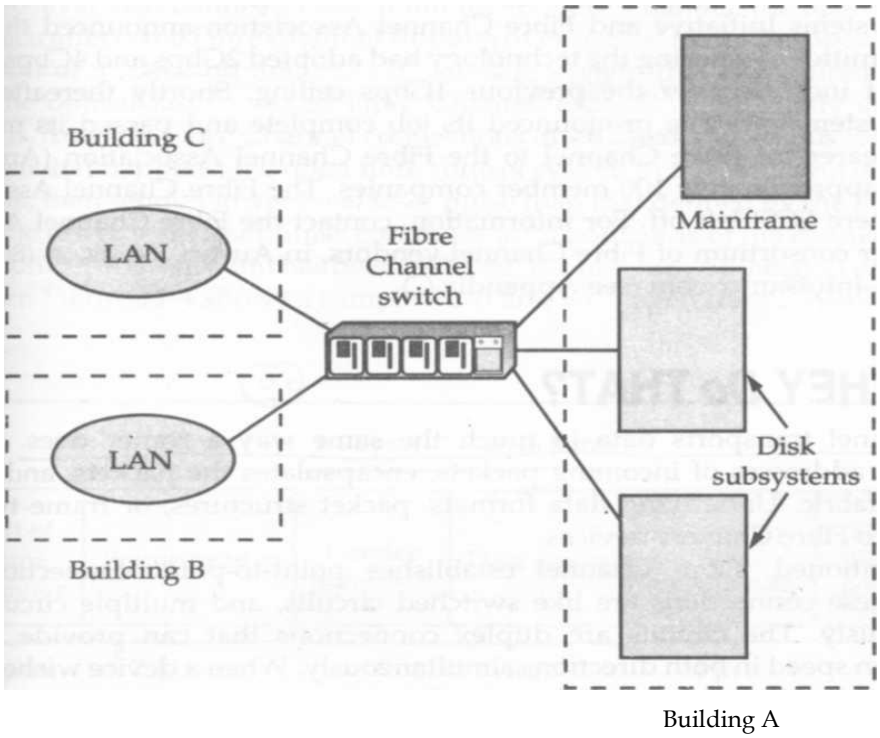


Figure 5.1: Fibre channel network in a campus environment

Fibre channel technology provides three different interconnect topologies to serve the combined needs of channel and network usages. These three topologies are:

- Fabric
- Point-to-point
- Arbitrated loop

Fabric Topology

A fabric topology permits dynamic interconnections between nodes through ports connected to the fabric. This is a standard network configuration. In the future, it may be possible to carry both network traffic and disk I/O over the same wire.

Point-to-Point Topology

In point-to-point host applications, two ports are used, connected to a link. The transmitter of each port is connected directly to the receiver of the opposite port. This topology limits the number of connections that can be made across the wire.

Arbitrated Loop Topology

The arbitrated loop topology is called fibre channel arbitrated loop (FC-AL). In this topology, each port arbitrates for access to the loop. Ports that "lose" the arbitration act as repeaters of all traffic on the loop. The loop contains a dedicated transmit channel and a dedicated receive channel that are clad together into one cable to form a loop out and back. This protocol allows up to 127 ports connected in a serial loop (one FL_Port and 126 NL_Ports). Silicon Graphics supports a maximum of 110 ports in a single rack.

Ports are called N_Ports (Node_Ports), NL_Ports (Node_Loop Ports), F_Ports (Fabric_Ports), or FL_Ports (Fabric_Loop Ports).

FC-AL ports represent a superset of functions and must work correctly with all three topologies (fabric, point-to-point, and arbitrated loop).

An NL_Port represents each disk in a disk array. Each NL_Port sees all messages and passes messages not addressed to that port. Ports passing messages are said to be in "repeat mode".

An FL_Port provides the connection from the fibre channel loop to the I/O board on the host, (in this case XIO or PCI). In the future, FL_Ports may provide a link from an arbitrated loop to a fabric.

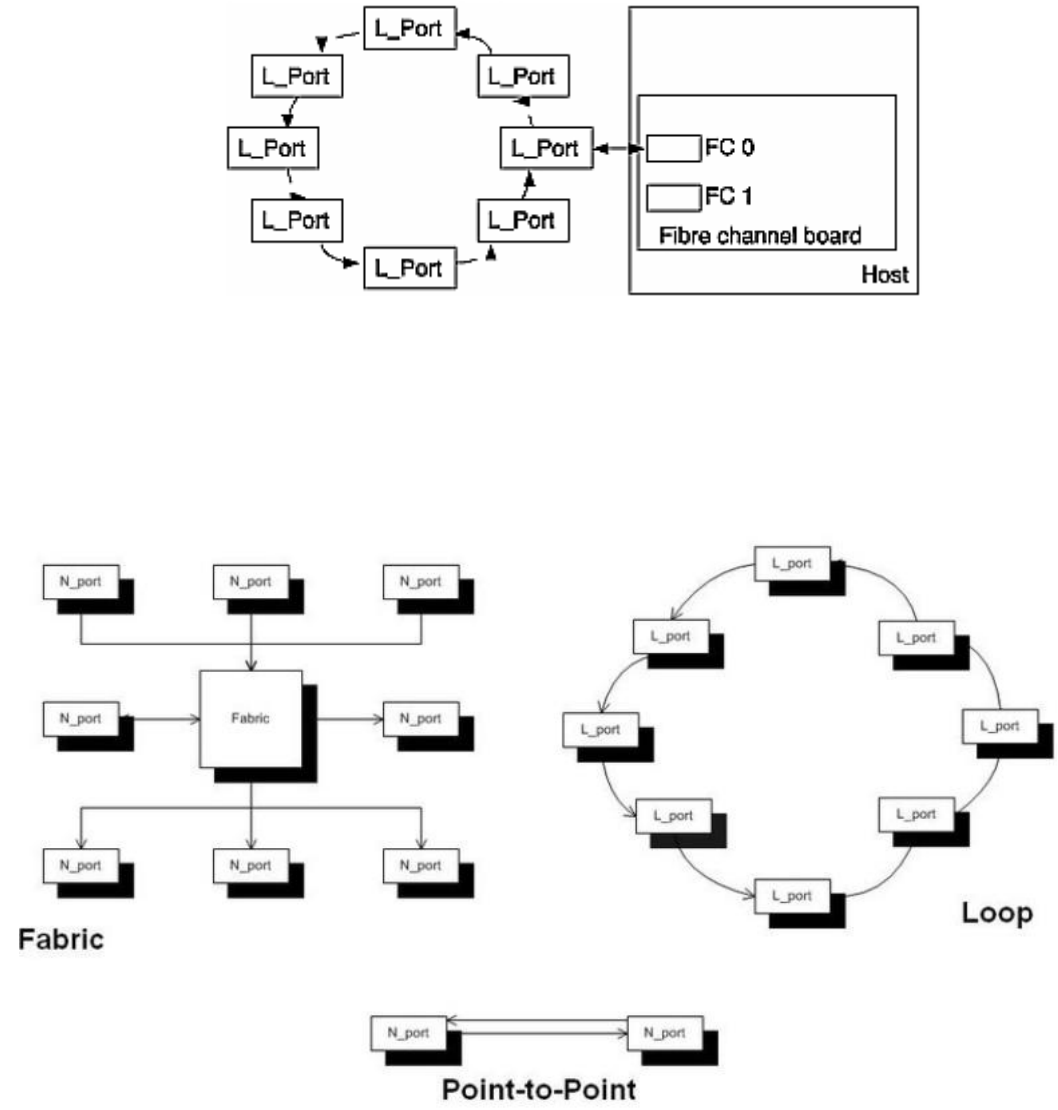


Figure 5.2 Fibre Channel Arbitrated Loop (FC-AL)

FC-AL is a shared-bandwidth distributed topology with arbitration fairness. The Silicon Graphics FC-AL option potentially supports 126 active L_Ports. Current support is limited to a maximum of 110 disks within a single rack. Communication on a loop is between two devices at a time. All nonactive FC-AL ports act as repeaters; all devices on the loop run at the same interface speed.

The arbitrated loop topology is a distributed topology in which each port includes the minimum necessary functions to establish a circuit. The ports are all L_Ports.

Disk L_Ports use the point-to-point protocol to create a point-to-point circuit between two L_Ports.

Port arbitration begins with one port replacing fill words between frames with an arbitration primitive called “ARBx.” If it goes completely around the loop and returns to the sending port, then that port has won arbitration.

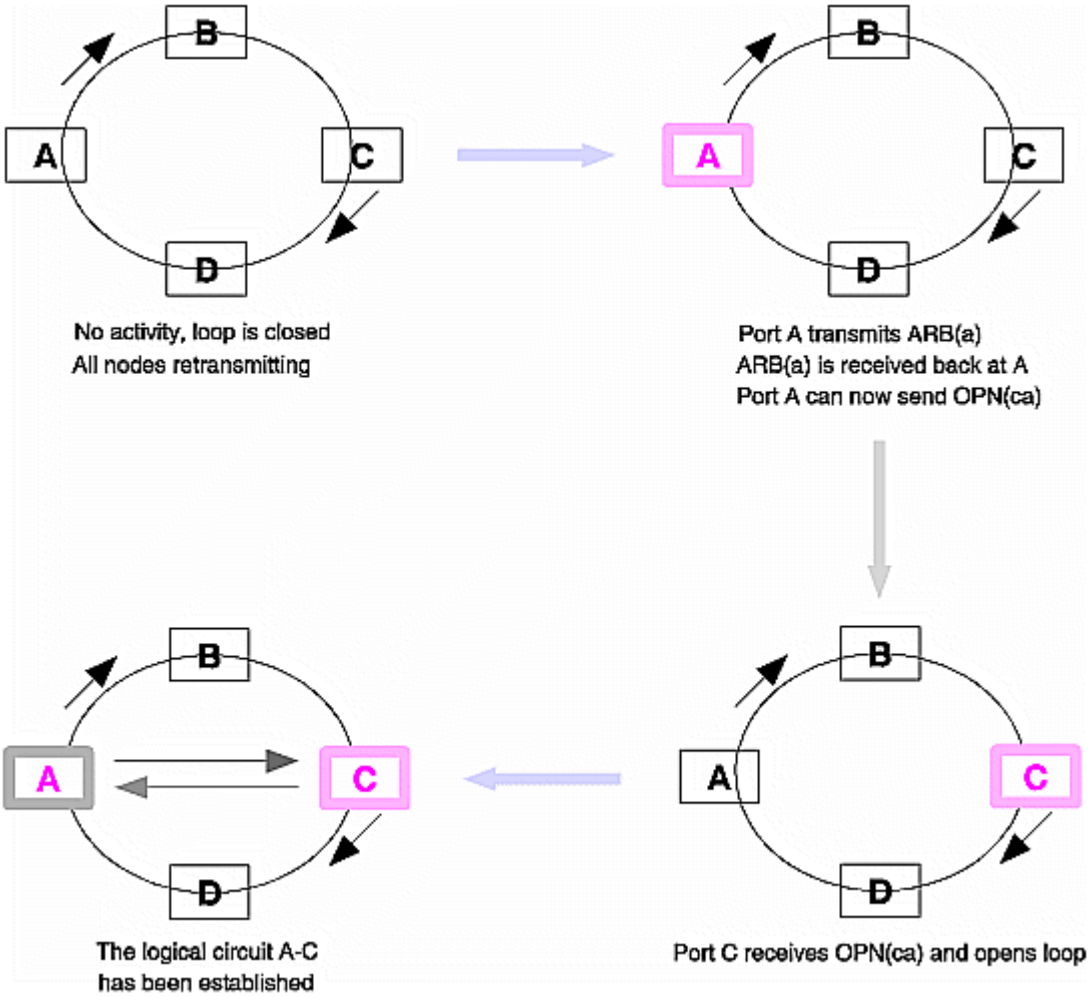


Figure 5.3 Port Arbitration Activity Example

A port that wins arbitration stops retransmitting received data and opens the loop between its receiver and transmitter by sending out the open primitive, OPNyx, where "y" represents the physical address of the destination port (AL_PD=Arbitrated Loop Physical Destination) and "x" represents the sending port. See Figure 5.3 for an example of loop functions.Closing a loop is done by transmitting the close primitive (CLS). The partner port finishes its work and retransmits the CLS, and then the loop is available for arbitration again.

The FC-AL loop interface is an implementation of SCSI-3 architecture that provides high-performance, greater connectivity, and high-availability features. FC-AL was designed as a low-cost method of connecting multiple hosts and storage devices without requiring the use of switches and fabrics. Note that multi-host configurations require special optional software and hardware.

Data Transfer in Fibre Channel

Data transfers in fibre channel use the following parameters

- DC-balanced 8b/10b signals with odd or even disparity
- variable-length data frames (maximum 2 KB)
- 32-bit CRC on frames

Combining channel and LAN characteristics, fibre channel uses buffers, one at the source node (port) and one at the destination. Each buffer can be any size.

Information can flow between two ports—actually between their buffers—in both directions simultaneously..

Data transfer takes place in units called *frames*, each a maximum of 2048 bytes of data. The frame contains the information to be transmitted, the address of the source and destination

ports, and link control information. Frames are of two types, Data frames and Link_Control frames. Data frames can be used as Link_Data frames and Device_Data frames. A set of related frames for one operation is called a *sequence*.

There are no limits on the size of a transfer. Frame sizes are transparent to upper-level software because the unit of transfer is a sequence.

5.3 LAYERED ARCHITECTURE

Fibre channel is a layered architecture with five layers: FC-0, FC-1, FC-2, FC-3, and FC-4. Figure- 5.3 diagrams the relationship between FC layers and OSI layers.

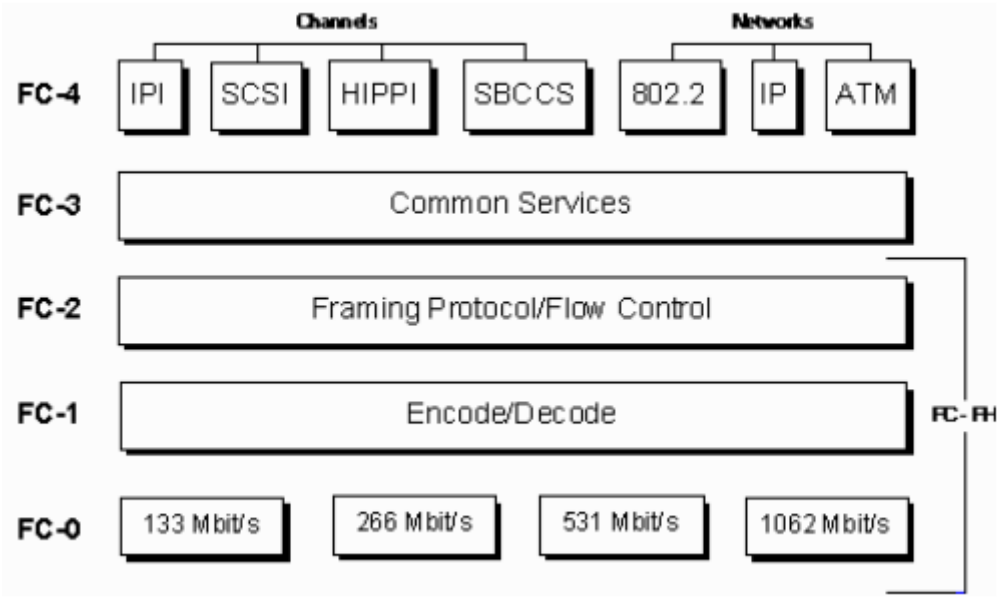


Figure 5.3 fibre channel layered architecture

5.4 LAYER 0 (FC-0 LAYER)

The lowest level (FC-0) defines the physical link in the system, including the fibre, connectors, optical and electrical parameters for a variety of data rates. Figure 5.4 shows the

schematic of the Fibre Channel optical link. The system bit error rate (BER) at the supported media and speeds is less than 10exp-12 . The physical level is designed for the use of large number of technologies to meet the widest range of system requirements. An end-to-end communicating route may consist of different link technologies to achieve the maximal performance and price efficiency.

Open Fibre Control

The FC-0 specifies a safety system - the Open Fibre Control system (OFC) - for SW laser data links, since the optical power levels exceed the limits defined by the laser safety standards. If an open fibre condition occurs in the link, the receiver of the Port the fibre is connected detects it and pulses its laser at a low duty cycle that meets the safety requirements. The receiver of the other port (at the other end of the fibre) detects this pulsing signal and also pulses it's transmitter at a low duty cycle. When the open fibre path is restored both ports receive the pulsing signals, and after a double handshaking procedure the connection is automatically restored within a few seconds.

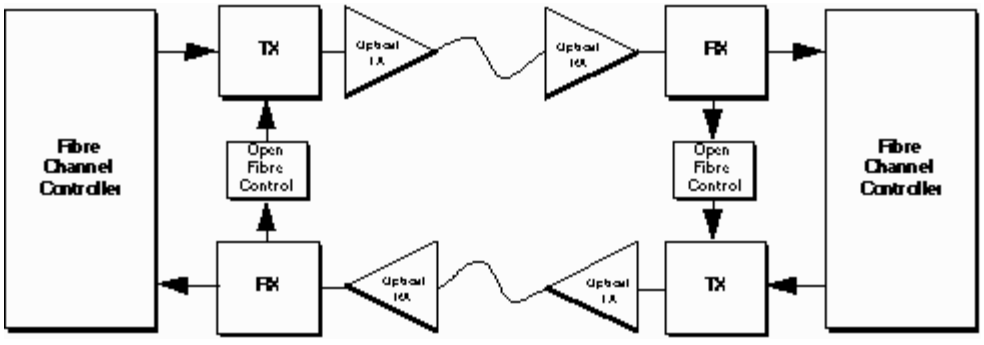


Figure 5.4 FC optical link

5.5 Layer 1 (FC-1 layer)

FC-1 defines the transmission protocol including serial encoding and decoding rules, special characters and error control. The information transmitted over a fibre is encoded 8 bits at a

time into a 10 bit Transmission Character. The primary rationale for use of a transmission code is to improve the transmission characteristic of information across a fibre. The transmission code must be DC balanced to support the electrical requirements of the receiving units. The Transmission Characters ensure, that short run lengths and enough transitions are present in the serial bit stream to make clock recovery possible.

FC-1 character conversion

An unencoded information byte is composed of eight information bits A,B,C,D,E,F,G,H and the control variable Z. This information is encoded by FC-1 into the bits a,b,c,d,e,i,f,g,h,j of a 10-bit Transmission Character. The control variable has either the value D (D-type) for Data characters or the value K (K-type) for special characters. Each valid Transmission Character has been given a name using the following convention: Zxx.y, where Z is the control variable of the unencoded FC-1 information byte, xx is the decimal value of the binary number composed of the bits E, D, C, B, and A, and y is the decimal value of the binary number composed of the bits H, G of the unencoded FC-1 information byte in that order. For example the name of the FC-1 Transmission Character composed of the hexadecimal "BC" special (K-type) code is K28.5.

The information received is recovered 10 bits at a time and those Transmission Characters used for data (D-type) are decoded into the one of the 256 8-bit combinations. Some of the remaining Transmission Characters (K-type) referred to as special characters, are used for protocol management functions. Codes detected at the receiver that are not D- or K- type are signalled as code violation errors.

Coding rules

Each data byte or special character has two (not necessarily different) transmission codes. The data bytes and special characters are encoded into these codes respectively, depending on the initial Running Disparity (RD). The RD is a binary parameter, which is calculated upon the balance of ones and zeros in the sub-blocks (the first six bits and the last four bits) of a transmission character. A new RD is calculated from the transmitted character at both the transmitter and the receiver. If the detected character has opposite RD the transmitter should have sent, (depending on the RD of the previous bit stream) the receiver indicates a disparity violation condition. A Transmission Word is composed of four contiguous transmission characters.

5.6 LAYER2 (FC-2 LAYER)

The Signaling Protocol (FC-2) level serves as the transport mechanism of Fibre Channel. The framing rules of the data to be transferred between ports, the different mechanisms for controlling the three service classes and the means of managing the sequence of a data transfer are defined by FC-2. To aid in the transport of data across the link, the following building blocks are defined by the standard:

- Ordered Set
- Frame
- Sequence
- Exchange
- Protocol

Ordered Set

The Ordered Sets are four byte transmission words containing data and special characters which have a special meaning. Ordered Sets provide the availability to obtain bit and word synchronization, which also establishes word boundary alignment. An Ordered Set always begins with the special character K28.5. Three major types of Ordered Sets are defined by the signaling protocol.

The Frame delimiters (the Start-of-Frame (SOF) and End-of-Frame (EOF) Ordered Sets) are Ordered Sets which immediately precede or follow the contents of a Frame. There are multiple SOF and EOF delimiters defined for the Fabric and N_Port Sequence control.

The two Primitive Signals: Idle and Receiver Ready (R_RDY) are Ordered Sets designated by the standard to have a special meaning. An Idle is a Primitive Signal transmitted on the link to indicate an operational Port facility ready for Frame transmission and reception. The R_RDY Primitive Signal indicates that the interface buffer is available for receiving further Frames.

A Primitive Sequence is an Ordered Set that is transmitted and repeated continuously to indicate specific conditions within a Port or conditions encountered by the receiver logic of a Port. When a Primitive Sequence is received and recognized, a corresponding Primitive Sequence or Idle is transmitted in response. Recognition of a Primitive Sequence requires consecutive detection of 3 instances of the same Ordered Set. The Primitive Sequences supported by the standard are Offline (OLS), Not Operational (NOS), Link Reset (LR) and Link Reset Response (LRR).

Frame

The basic building blocks of an FC connection are the Frames. The Frames contain the information to be transmitted (Payload), the address of the source and destination ports and link control information. Frames are broadly categorized as Data frames and Link_control frames. Data frames may be used as Link_Data frames and Device_Data frames, link control frames are classified as Acknowledge (ACK) and Link_Response (Busy and Reject) frames. The primary function of the Fabric is, to receive the Frames from the source port and route them to the destination port. It is the FC-2 layer's responsibility to break the data to be transmitted into Frame size, and reassemble the Frames.

Each Frame begins and ends with a Frame Delimiter (Figure 4) The Frame Header immediately follows the SOF delimiter. The Frame Header is used to control link applications, control device protocol transfers, and detect missing or out of order Frames. An optional header may contain further link control information. A maximum 2112 byte long field (payload) contains the information to be transferred from a source N_Port to a destination N_Port. The 4 bytes Cyclic Redundancy Check (CRC) precedes the EOF delimiter. The CRC is used to detect transmission errors.

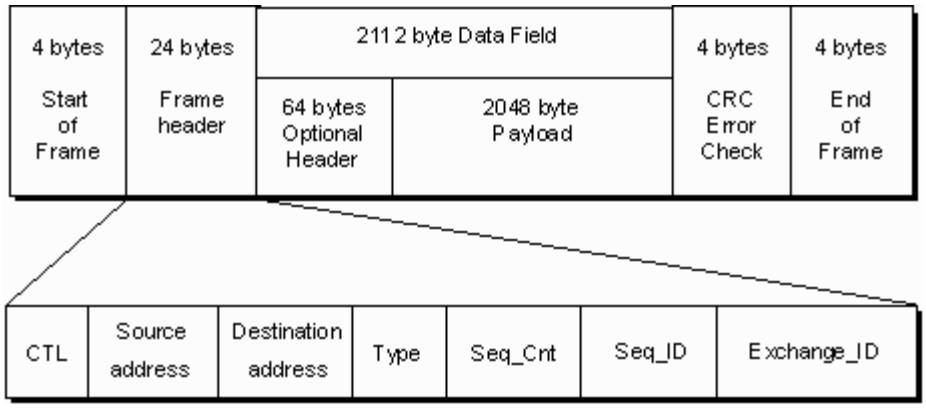


Figure 5.5 Frame Structure

Sequence

A Sequence is formed by a set of one or more related Frames transmitted unidirectionally from one N_Port to another. Each Frame within a sequence is uniquely numbered with a Sequence Count. Error recovery, controlled by an upper protocol layer is usually performed at Sequence boundaries.

Exchange

An Exchange is composed of one or more non concurrent sequences for a single operation. The Exchanges may be unidirectional or bidirectional between two N_Ports. Within a single Exchange, only one sequence may be active at any one time, but Sequences of different Exchanges may be concurrently active.

Protocol

The Protocols are related to the services offered by Fibre Channel. Protocols may be specific to higher-layer services, although Fibre Channel provides its own set of protocols to manage its operating environment for data transfer. The following Protocols are specified by the standard:

- Primitive Sequence Protocols are based on Primitive Sequences and specified for link failure.

- Fabric Login protocol: The interchanging of Service Parameters of an N_Port with the fabric.
- N_Port Login protocol: Before performing data transfer, the N_Port interchanges its Service Parameters with another N_Port.
- Data transfer protocol describes the methods of transferring Upper Layer Protocol (ULP) data using the Flow control management of Fibre Channel.

- N_Port Logout Protocol is performed when an N_Port requests removal of its Service Parameters from the other N_Port. This may be used to free up resources at the connected N_Port.

Flow control

Flow control is the FC-2 control process to pace the flow of Frames between N_Ports and between an N_Port and the Fabric to prevent overrun at the receiver. Flow control is dependent upon the service classes. Class 1 Frames use end-to-end flow control, class 3 uses only buffer-to-buffer, class 2 Frames use both types of flow control.

Flow control is managed by the Sequence Initiator (source) and Sequence Recipient (destination) Ports using Credit and Credit_CNT. Credit is the number of buffers allocated to a transmitting Port. The Credit_CNT represents the number of data frames which have not been acknowledged by the Sequence Recipient.

The end-to-end flow control process paces the flow of Frames between N_Ports. In this case the Sequence Recipient is responsible for acknowledging the received valid data Frames by ACK Frames. When the number of receive buffers are insufficient for the incoming Frame, a "Busy", when a Frame with error is received a "Reject" Frame will be sent to the Initiator Port. The Sequence Initiator is responsible for managing EE_Credit_CNT. The N_Port login is used to establish EE_Credit.

The buffer-to-buffer flow control is managed between an N_Port and an F_Port or between N_Ports in point-to-point topology. Each port is responsible for managing BB_Credit_CNT. BB_Credit is established during the Fabric Login The Sequence Recipient (destination) Port signals by sending a Receiver_Ready primitive signal to the transmitting Port whether it has free receive buffers for the incoming Frames..

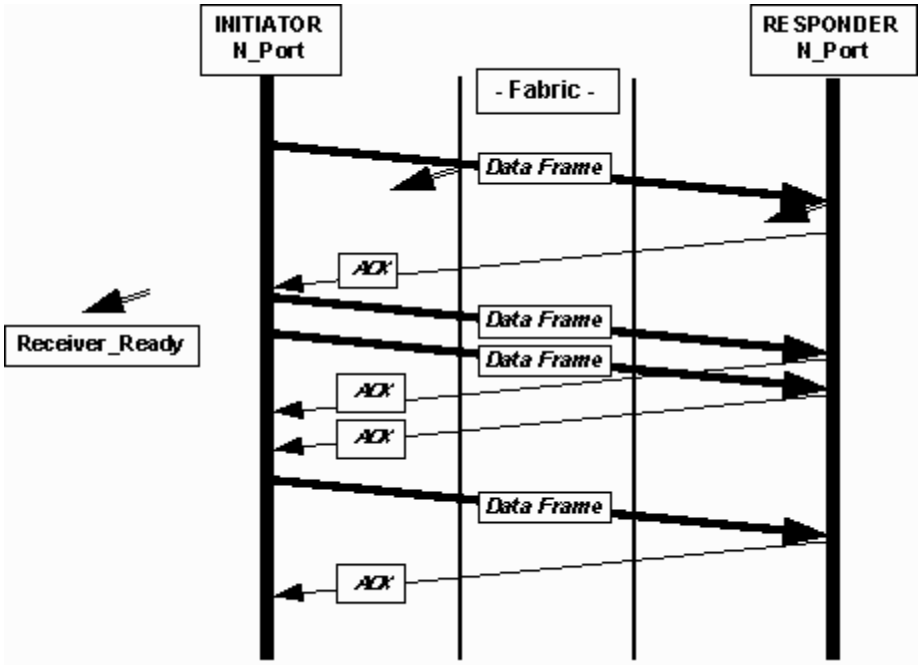


Figure 5.6 Class 1 Flow Control

Service Classes

To ensure efficient transmission of different types of traffic, FC defines three classes of service. Users select service classes based on the characteristics of their applications, like packet length and transmission duration, and allocate the services by the Fabric Login protocol.Class 1 is a service which provides dedicated connections, in effect providing the equivalent of a dedicated physical connection. Once established, a Class 1 connection is retained and guaranteed by the Fabric. This service guarantees the maximum bandwidth between two N_Ports, so this is the best for sustained, high throughput transactions. In Class 1, Frames are delivered to the destination Port in the same order as they are transmitted. Figure 5.6 shows the flow control management of a Class 1 connection.

Class 2 is a Frame-switched, connectionless service that allows bandwidth to be shared by multiplexing Frames from multiple sources onto the same channel or channels. The Fabric may not guarantee the order of the delivery and Frames may be delivered out of order. This

service class can be used, when the connection setup time is greater than the latency of a short message. Both Class 1 and Class 2 send acknowledgment Frames confirming Frame delivery. If delivery cannot be made due to congestion, a Busy frame is returned and the sender tries again. (Figure 5.7)

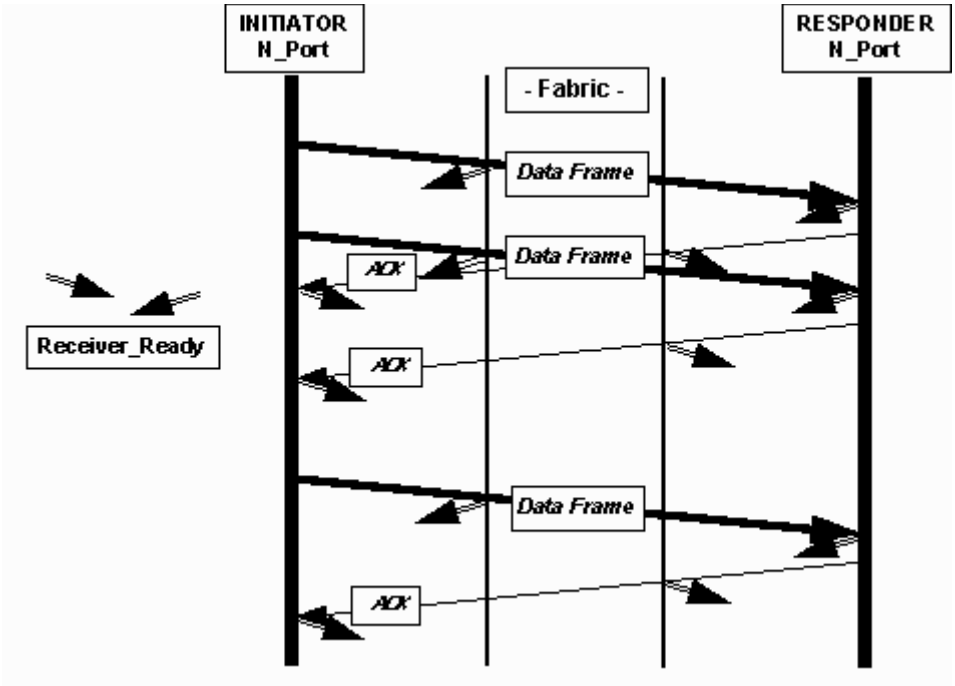


Figure 5.7 Class 2 Flow Control

Class 3 service is identical to Class 2, except that the Frame delivery is not confirmed. (Flow control is managed only on buffer level, see Figure 5.7) This type of transfer, known as datagram provides the quickest transmission by not sending confirmation. This service is useful for real- time broadcasts, where timeliness is key and information not received in time is valueless.

The FC standard also defines an optional service mode called intermix. Intermix is an option of Class 1 service, in which Class 1 Frames are guaranteed a special amount of bandwidth, but Class 2 and Class 3 Frames are multiplexed onto the channel, only when sufficient bandwidth is available to share the link.

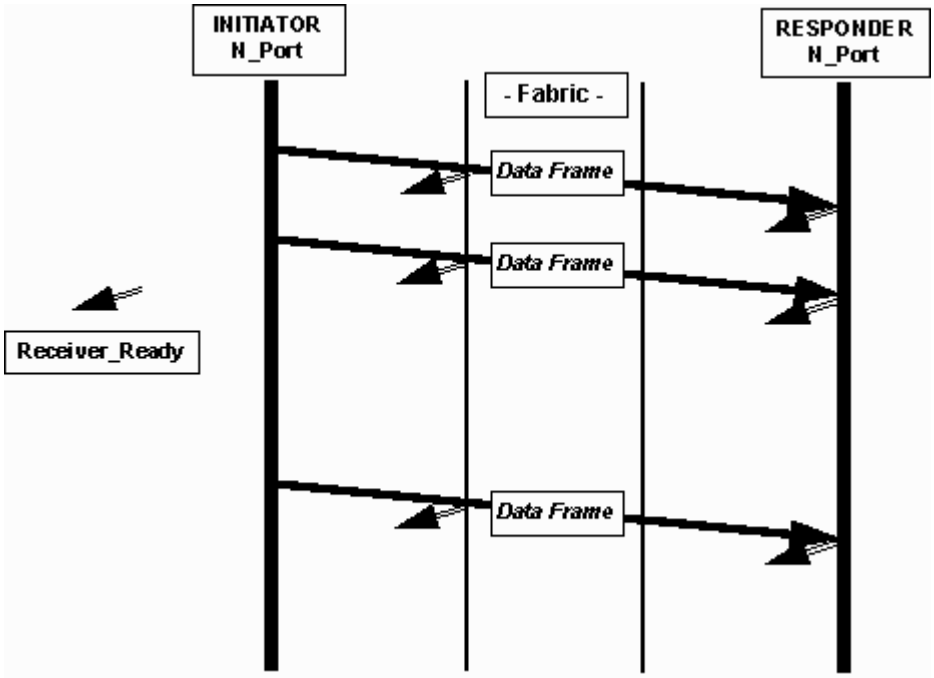


Figure 5.7 Class 3 Flow Control

5.7 LAYER 3 (FC-3 LAYER)

The FC-3 level of the FC standard is intended to provide the common services required for advanced features such as:

- Striping -To multiply bandwidth using multiple N_ports in parallel to transmit a single information unit across multiple links.
- Hunt groups - The ability for more than one Port to respond to the same alias address. This improves efficiency by decreasing the chance of reaching a busy N_Port.
- Multicast - Multicast delivers a single transmission to multiple destination ports. This includes sending to all N_Ports on a Fabric (broadcast) or to only a subset of the N_Ports on a Fabric.

5.8 LAYER 4 (FC-4 LAYER)

FC-4, the highest level in the FC structure defines the application interfaces that can execute over Fibre Channel. It specifies the mapping rules of upper layer protocols using the FC levels below. Fibre Channel is equally adept at transporting both network and channel information and allows both protocol types to be concurrently transported over the same physical interface.

The following network and channel protocols are currently specified or proposed as FC-4s:

- Small Computer System Interface (SCSI)
- Intelligent Peripheral Interface (IPI)
- High Performance Parallel Interface (HIPPI) Framing Protocol
- Internet Protocol (IP)
- ATM Adaptation Layer for computer data (AAL5)
- Link Encapsulation (FC-LE)
- Single Byte Command Code Set Mapping (SBCCS)
- IEEE 802.2

5.9 SUMMARY

- Fibre channel was first designed and developed to interconnect high-speed peripherals—for example, a cluster of high-performance computers—to a shared mass storage device
- A fabric topology permits dynamic interconnections between nodes through ports connected to the fabric
- Fibre channel is a layered architecture with five layers: FC-0, FC-1, FC-2, FC-3, and FC-4.

- FC-0 defines the physical link in the system, including the fibre, connectors, optical and electrical parameters for a variety of data rates
- FC-1 defines the transmission protocol including serial encoding and decoding rules, special characters and error control.
- FC-2 serves as the transport mechanism of Fibre Channel.
- FC-3 level of the FC standard is intended to provide the common services like striping, hunt groups, multicast.
- FC-4, the highest level in the FC structure defines the application interfaces that can execute over Fibre Channel

5.10 KEYWORDS

FC: Fibre Channel

SCSI: Small Computer System Interface.

IP: Internet Protocol.

HIPPI: High Performance Parallel Interface.

SOF: Start-of-Frame

EOF: End-of-Frame

5.11 REVIEW QUESTIONS

Q1. Explain the historical evolution of fibre channel

Q2. Discuss various topologies define in fibre channel

Q3. Explain character conversion of FC-1 of fibre channel.

Q4. Explain layered architecture of fibre channel in detail.

5.12 Further Readings

3T9.3 Task Group of ANSI: Fibre Channel Physical and Signaling Interface (FC-PH), Rev. 4.2 October 8, 1993

Fibre Channel Association: Fibre Channel: Connection to the Future, 1994, ISBN 1-878707- 19-1

Gary Kessler: Changing channels, LAN Magazine, December 1993, p69-78

LESSON 6 INTEGRATED SERVICES DIGITAL NETWORK

6.1 OBJECTIVE

Objective of this chapter is to introduce the reader about Integrated Services Digital Network (ISDN). After reading this chapter reader will be able to discuss about Integrated Services Digital Network what services it provide, ISDN Architecture and interfaces.

6.2 INTRODUCTION

ISDN referst to the set of communication protocols proposed by the telephone companies to permit telephone networks to carry data, voice and other source material. ISDN involves the digitization of the telephone network, which permits voice, data, text, graphics, music, video, and other source material to be transmitted over existing telephone wires.ISDN applications include high-speed image applications additional telephone lines in homes to serve the telecommuting industry, high-speed file transfer, and videoconferencing. ISDN is generally viewed as an alternative to frame relay and T1 wide area telephone service (WATS).

The digital connectivity has many benefits including,

1. ISDN provides a faster data transfer rate Telephone companies with the intention of creating a totally digital network for increased bandwidth speeds developed ISDN.
2. ISDN was developed to use the existing telephone wiring system to enhance WAN usage with out incurring major networking costs.
3. ISDN provides a faster data transfer rate Telephone companies with the intention of creating a totally digital network for increased bandwidth speeds developed ISDN.
4. ISDN was developed to use the existing telephone wiring system to enhance WAN usage with out incurring major networking costs.
5. ISDN provides access to digital video, circuit-switched data, and telephone network services by using the normal phone network that is circuit-switched.
6. ISDN offers much faster call setup than modem connections because it uses out-of-band (D channel) signaling. For example ISDN calls can be setup in less than one second.
7. ISDN can provide a clear data path over which to negotiate PPP links.
8. ISDN offers Dial on Demand Routing, which means you only pay for the time that you use the link.

ISDN Devices:

ISDN devices include terminals,

-Terminal adapters (TAs),

-Network-termination devices,

-Line-termination equipment,

-Exchange-termination equipment.

ISDN terminals come in two types. Specialized ISDN terminals are referred to as terminal equipment type 1 (TE1). Non-ISDN terminals, such as DTE, are referred to as terminal equipment type 2 (TE2). TE1s connect to the ISDN network through a four-wire, twisted-pair digital link. TE2s connect to the ISDN network through a TA.

Beyond the TE1 and TE2 devices, the next connection point in the ISDN network is the network termination type 1 (NT1) or network termination type 2 (NT2) device. These are network-termination devices that connect the four-wire subscriber wiring to the conventional two-wire local loop. In North America, the NT1 is a customer premises equipment (CPE) device. In most other parts of the world, the NT1 is part of the network provided by the carrier. The NT2 is a more complicated device that typically is found in digital private branch exchanges (PBXs) and that performs Layer 2 and 3 protocol functions and concentration services. An NT1/2 device also exists as a single device that combines the functions of an NT1 and an NT2.

ISDN specifies a number of reference points that define logical interfaces between functional groups, such as TAs and NT1s. ISDN reference points include the following:

- **R** - The reference point between non-ISDN equipment and a TA.
- **S** - The reference point between user terminals and the NT2.
- **T** - The reference point between NT1 and NT2 devices.
- **U** - The reference point between NT1 devices and line-termination equipment in the carrier network. The U reference point is relevant only in North America, where the NT1 function is not provided by the carrier network.

Figure 1 illustrates a sample ISDN configuration and shows three devices attached to an ISDN switch at the central office. Two of these devices are ISDN-compatible, so they can be attached through an S reference point to NT2 devices. The third device (a standard, non-

ISDN telephone) attaches through the reference point to a TA. Any of these devices also could attach to an NT1/2 device, which would replace both the NT1 and the NT2. In addition, although they are not shown, similar user stations are attached to the far-right ISDN switch.

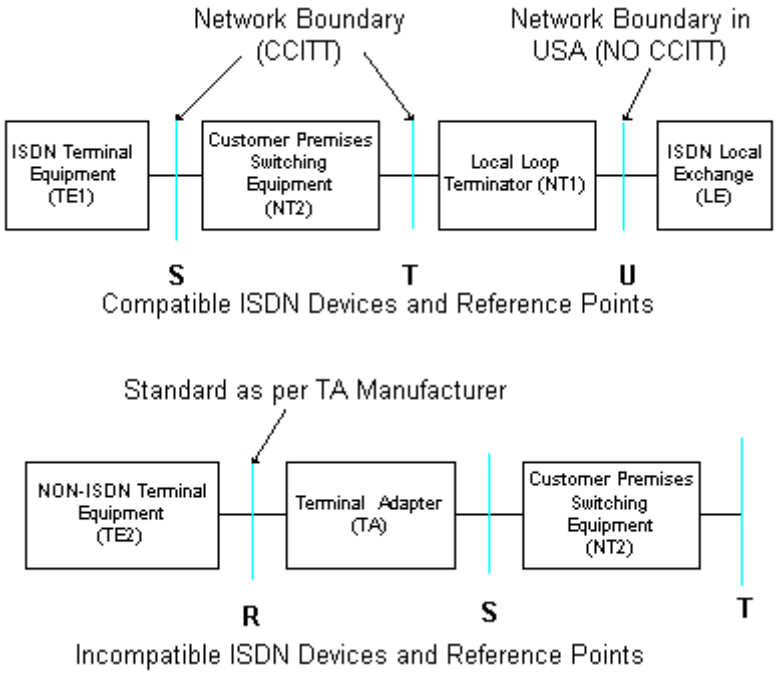


Figure 1: Sample ISDN Configuration Illustrates Relationships Between Devices and Reference Points

6.3 ISDN SERVICES:

There are two types of services associated with ISDN:

- BRI
- PRI

ISDN BRI sercive:

The ISDN Basic Rate Interface (BRI) service offers two B channels and one D channel (2B+D). BRI B-channel service operates at 64 kbps and is meant to carry user data; BRI D-

channel service operates at 16 kbps and is meant to carry control and signaling information, although it can support user data transmission under certain circumstances.

ISDN PRI Service:

ISDN Primary Rate Interface (PRI) service offers 23 B channels and 1 D channel in North America and Japan, yielding a total bit rate of 1.544 Mbps (the PRI D channel runs at 64 kbps). ISDN PRI in Europe, Australia, and other parts of the world provides 30 B channels plus one 64-kbps D channel and a total interface rate of 2.048 Mbps.

6.4 ISDN System Architecture:

LAYER 1:

ISDN physical layer (Layer 1) frame formats differ depending on whether the frame is outbound (from terminal to network) or inbound (from network to terminal). (See figure 2)

The frames are 48 bits long, of which 36 bits represent data. The bits of an ISDN physical layer frame are used as follows:

- **F** - Provides synchronization
- **L** - Adjusts the average bit value
- **E** - Ensures contention resolution when several terminals on a passive bus contend for a channel
- **A** - Activates devices
- **S** - Is unassigned
- **B1**, **B2**, and **D** - Handle user data

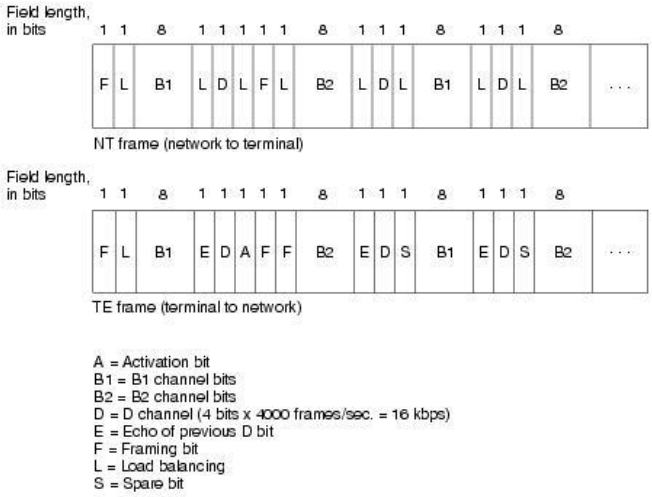


Figure 2: ISDN Physical Layer Frame Formats Differ Depending on Their Direction

Multiple ISDN user devices can be physically attached to one circuit. In this configuration, collisions can result if two terminals transmit simultaneously. Therefore, ISDN provides features to determine link contention. When an NT receives a D bit from the TE, it echoes back the bit in the next E-bit position. The TE expects the next E bit to be the same as its last transmitted D bit.

LAYER 2:

Layer 2 of the ISDN signaling protocol is Link Access Procedure, D channel (LAPD). LAPD is similar to High-Level Data Link Control (HDLC) and Link Access Procedure, Balanced (LAPB). As the expansion of the LAPD acronym indicates, this layer is used across the D channel to ensure that control and signaling information flows and is received properly. The LAPD frame format (see Figure 3: LAPD Frame Format Is Similar to That of HDLC and LAPB)is very similar to that of HDLC; like HDLC, LAPD uses supervisory, information, and unnumbered frames. The LAPD protocol is formally specified in ITU-T Q.920 and ITU-T

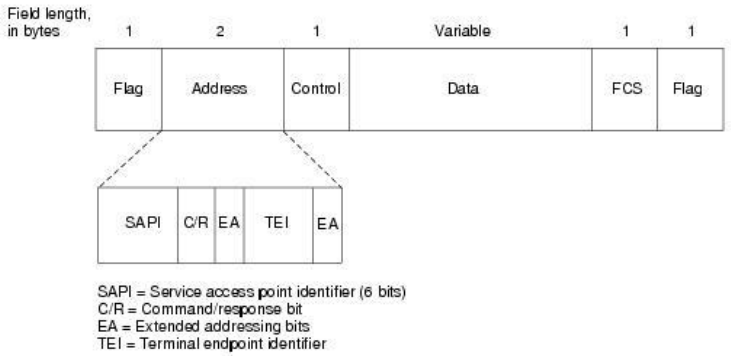


Figure 3: LAPD Frame Format Is Similar to That of HDLC and LAPB

The LAPD Flag and Control fields are identical to those of HDLC. The LAPD Address field can be either 1 or 2 bytes long. If the extended address bit of the first byte is set, the address is 1 byte; if it is not set, the address is 2 bytes. The first Address-field byte contains the service access point identifier (SAPI), which identifies the portal at which LAPD services are provided to Layer 3. The C/R bit indicates whether the frame contains a command or a response. The Terminal Endpoint Identifier (TEI) field identifies either a single terminal or multiple terminals. A TEI of all ones indicates a broadcast.

LAYER 3:

Two Layer 3 specifications are used for ISDN signaling: ITU-T (formerly CCITT) I.450 (also known as ITU-T Q.930) and ITU-T I.451 (also known as ITU-T Q.931). Together, these protocols support user-to-user, circuit-switched, and packet-switched connections. A variety of call-establishment, call-termination, information, and miscellaneous messages are specified, including SETUP, CONNECT, RELEASE, USER INFORMATION, CANCEL, STATUS, and DISCONNECT. These messages are functionally similar to those provided by the X.25 protocol. See figure 4.

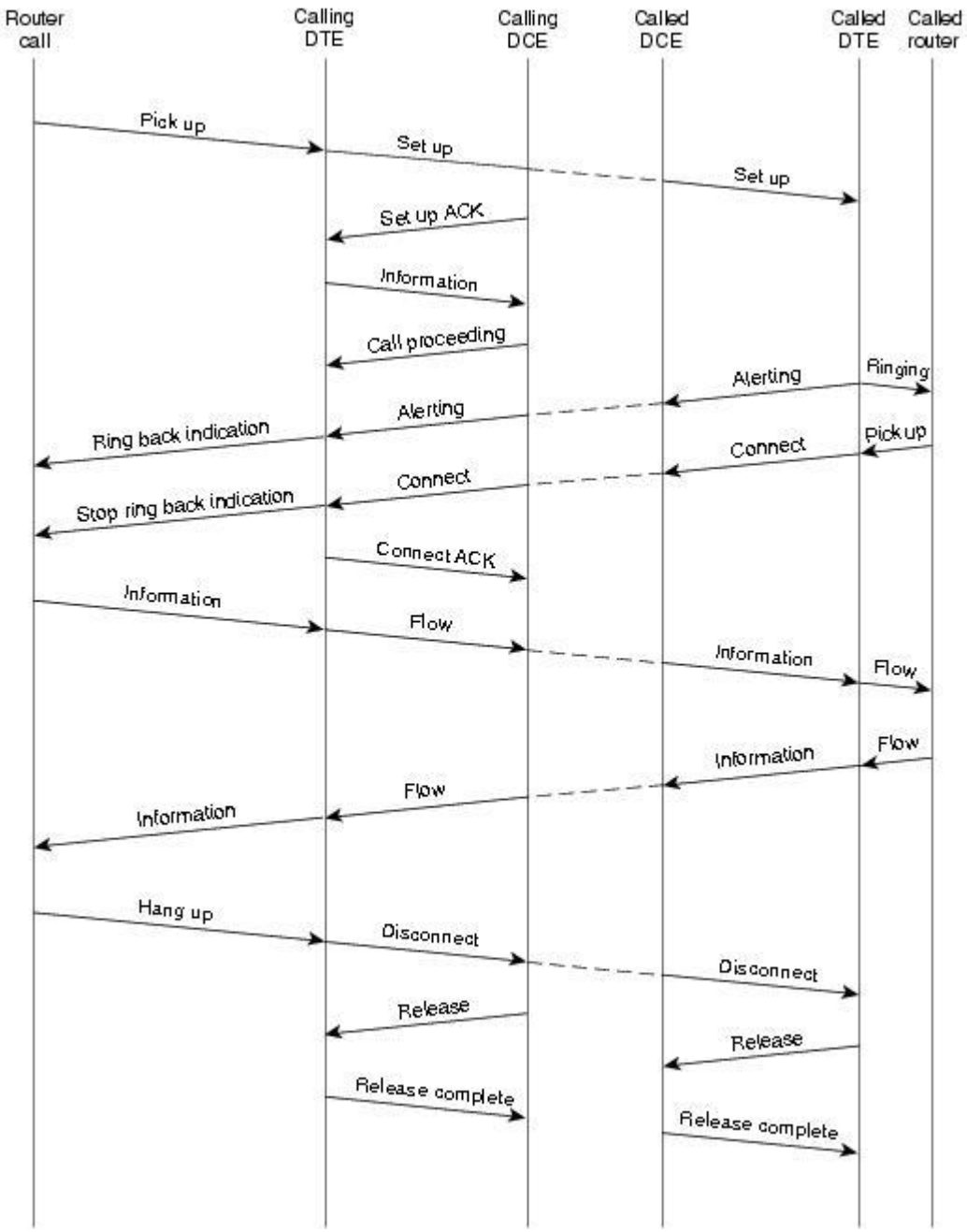


Figure 4: An ISDN Circuit-Switched Call Moves Through Various Stages to Its Destination

6.5 BROADBAND ISDN

The original specifications for the integrated services digital network (ISDN), were based around voice and non-voice telephone-type services: telephony, data, telex, facsimile, as it

was hoped that the ISDN would evolve from the (then) emerging digital telephone networks. Indeed, this is one of the reasons that the fundamental element of an ISDN link is the 64 Kb/s B-Channels. However, the planning for ISDN was started around 1976, and as technology evolved, so did the requirements of the users that wanted to use this technology. In 1988, the CCITT released a document that described a new set of Broadband ISDN (B-ISDN) services. To distinguish this new concept from the original ISDN service, we now refer to the latter as Narrowband ISDN (N-ISDN).

Broadband ISDN services

The need for a Broadband ISDN service sprung from the growing needs of the customers. The planned Broadband ISDN services can broadly be categorized as follows:

Interactive services. These are services allowing information flow between two end users of the network, or between the user and the service provider. Such services can be subdivided:

Conversational services. These are basically end-to-end, real-time communications, between users or between a user and a service provider, e.g. telephone-like services. Indeed, B-ISDN will support N-ISDN type services. (Note also that the user-to-user signaling, user-to-network signaling, and inter-echange signaling are also provided but outside our scope.) Also the additional bandwidth offered will allow such services as video telephony, video conferencing and high volume, high speed data transfer.

Messaging services. This differs from conversational services in that it is mainly a store-and-forward type of service. Applications could include voice and video mail, as well as multi-media mail and traditional electronic mail.

Retrieval services. This service provides access to (public) information stores, and information is sent to the user on demand only. This includes things like tele-shopping, videotex services, still and moving pictures, telesoftware and entertainment.

Distribution services. These are mainly broadcast services, are intended for mainly one way interaction from a service provider to a user:

No user control of presentation. This would be for instance, a TV broadcast, where the user can choose simply either to view or not. It is expected that cable TV companies will become interested in Broadband ISDN as a carrier for the high definition TV (HDTV) services that are forseen for the future.

User controlled presentation. This would apply to broadcast information that the user can partially control, in that the user can decide which part of it he/she accesses, e.g. teletext and news retrieval services.

Protocol Reference Model

The network is described in terms of a protocol reference model (PRM) (Figure 5). Not all of the PRM is fully defined. The main aspects of the model are that it can be viewed in terms of the three planes -- user plane, control plane and management plane -- and in terms of the 3 layers -- ATM adaptation layer, ATM layer and the Physical layer.

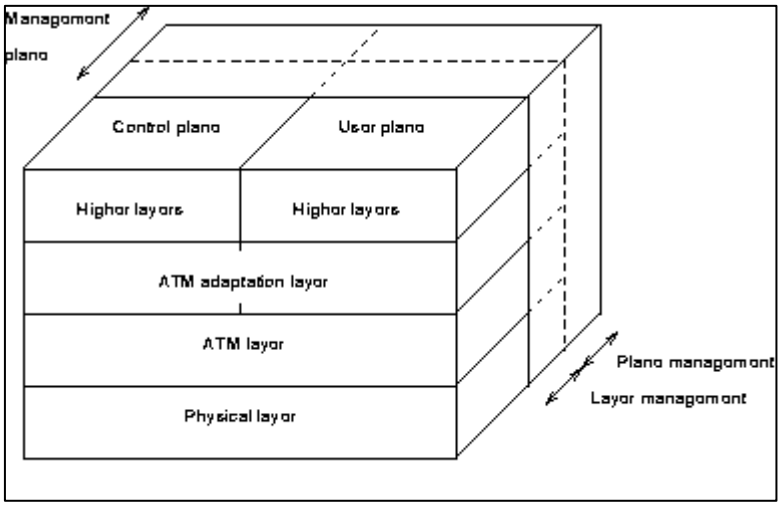


Figure 5: Broadband ISDN protocol reference model

The functions of the layers are as follows:

- ATM adaptation layer (AAL). This layer is responsible for mapping the service offered by ATM to the service expected by the higher layers. It has two sublayers.
  - Convergence sub layer (CS). Responsible for presenting the ATM service to the higher layers. The functionality of this sub layer is very much dependent on the higher layer service.
  - Segmentation and reassembly (SAR). This layer is responsible for, at the transmitter, splitting the higher level PDU into 48 octet chunks, and at the receiving side, to reassemble the 48 octet chunks back into the original PDU.
- ATM Layer. This layer is independent of the physical medium over which transmission is to take place. It has four functions:
  - Generic flow control (GFC) function. This can be used to alleviate short term overload conditions above the ATM layer, as it is accessible by the user.
  - Cell header generation and extraction. At the transmitter, adds header information to a cell and at the receiver removes it.
  - Cell multiplex and demultiplex. At the transmitter, multiplex cells into one continuous stream and at the header demultiplex the cells according to VPI and VCI values.
- Physical layer. This consists of two sublayers:
  - Transport Convergence (TC). This sub layer has five functions:
    - Cell rate decoupling. Insertion and extraction of idle cells.
    - Header error control (HEC) generation and verification. In the transmitter, generation of the HEC, and in the receiver checking of the HEC. The HEC that is used can detect and correct a 1 bit error and can further detect certain multiple bit errors.

- Cell delineation. In the receiver, detection of cell boundaries.
- Transmission frame adaptation. Adapts cell flow according to the payload of the Physical level frame being used, e.g. for SDH.
- Transmission frame generation and recovery. At the transmitter, generates Physical level frames, and at the receiver, extracts the ATM cells from the Physical level frame.

- Physical medium (PM). This contains two sublayers:
  - Bit timing. Insertion and extraction of bit timing information and generation and reception of waveforms.
  - Physical medium. Bit transmission, bit alignment and optical ↔electrical conversion, if required. (The physical medium need not be optical, at least for transmission rates of 155Mb/s and lower.)

This (sub-) layering of the PRM is depicted in Figure 6.

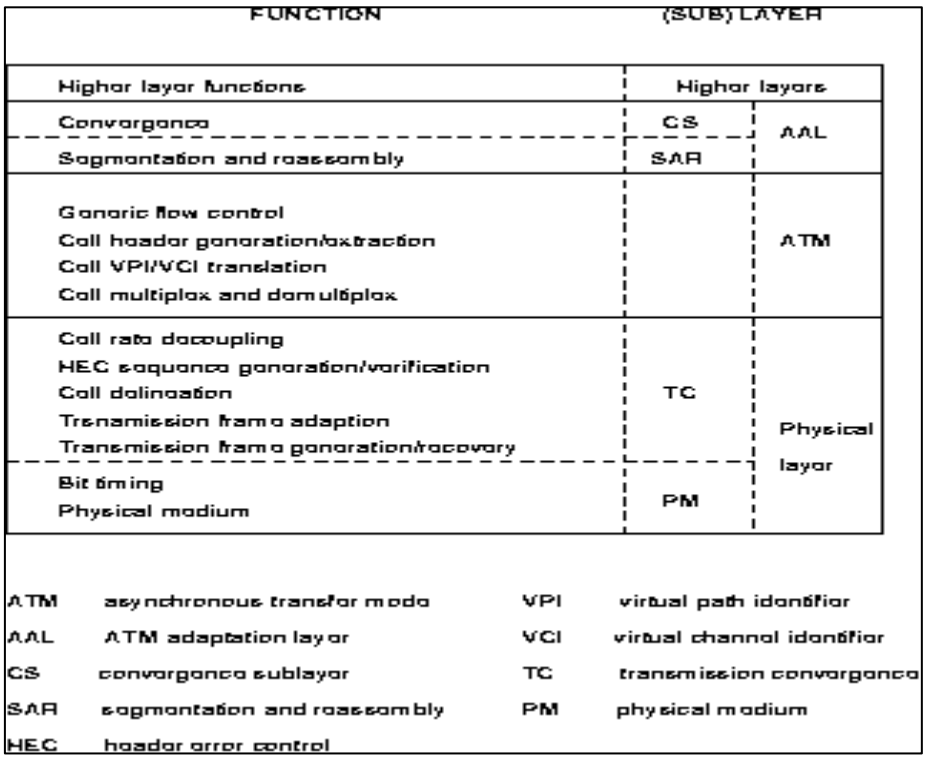


Figure 6: Broadband ISDN layer functionality

The management plane consists of two functions to perform layer management and plane management. The plane management is not layered as the other layers are. This is because it relies needs information on all aspects of the the system to provide management facilities for the systems as a whole. The layer management provides information and control facilities for the protocol entities that exists in each individual layer. This includes operation and maintenance (OAM) functions for each layer.

The control plane is responsible for the supervision of connections, including call set-up, call release and maintenance.

The user plane provides for the transfer of user information. It also includes mechanisms to perform error recovery, flow control etc.

Broadband ISDN intends to offer many Mb/s to the user, but intends to remain backwards compatible with Narrowband ISDN. Indeed, the Narrowband services will eventually need to be offered over the global Broadband network to come. To this extent the user interface to Broadband ISDN is very similar to that for Narrowband ISDN. Figure 7 shows the position of the user to network interface (UNI), as well as the internal network to network interface (NNI) for BISDN.

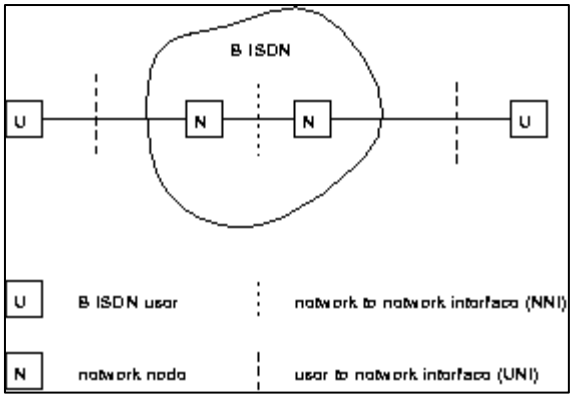


Figure 7: Broadband ISDN user and network interfaces

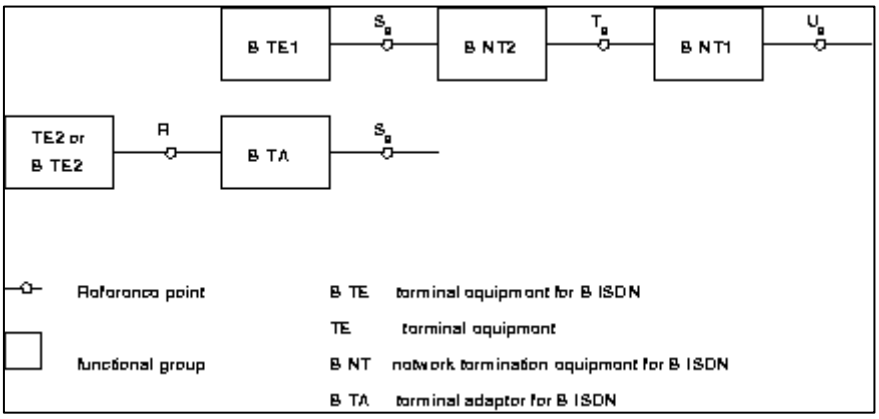


Figure 8: Broadband ISDN UNI configuration reference points

Note that in Figure 8 , it is expected that Narrowband ISDN (or even other PSTN) equipment will be able to connect to the Broadband network via a suitable terminal adaptor. The various functional groups are now described:

- B-NT1. This group contains functions that are considered to be part of OSI layer 1. It represents the physical connection point to the network, i.e. the socket on the wall. It includes functions such as:
  - Line transmission termination. Provision of the physical connection.
  - Transmission interface handling. The interface to the transmission channel, be it electrical or optical.

- Operation and Maintenance (OAM). This is not normally associated with the socket in the wall. However, it is expected that for B-ISDN, more sophisticated management capabilities will be required than at present.

- B-NT2. This group contains OSI layer 1 and higher OSI layer functions:
  - Adaptation functions. For different physical media and network topologies.
  - Multiplexing and demultiplexing. The user data may be sent and received on several VCCs and VPCs.
  - Buffering. User data may be sent and/or received at varying rates with respect to the B-ISDN user and the network.
  - Signaling. VCCs/VPCs must be established, controlled and released.
  - Interface. Interaction with the B-ISDN user.

- B-TE1. Equipment requiring B-ISDN access.
- B-TA. Equipment allowing connection of other B-ISDN, N-ISDN and non-ISDN equipment.
- B-TE2. B-ISDN with special interface needs or N-ISDN equipment.
- TE2. Non-ISDN equipment.

Note that these are logical units. The physical implementation may be quite different. For instance, it may be common to find the following in the same physical unit, depending on need: B-NT1 and B-NT2; B-TE1 and B-NT2; B-TA and B-TE2 etc.

Further, the way in which the terminal equipment is connected to the user-to-network interface via B-NT1/B-NT2 is not restricted with respect to local topologies (Figure 9).

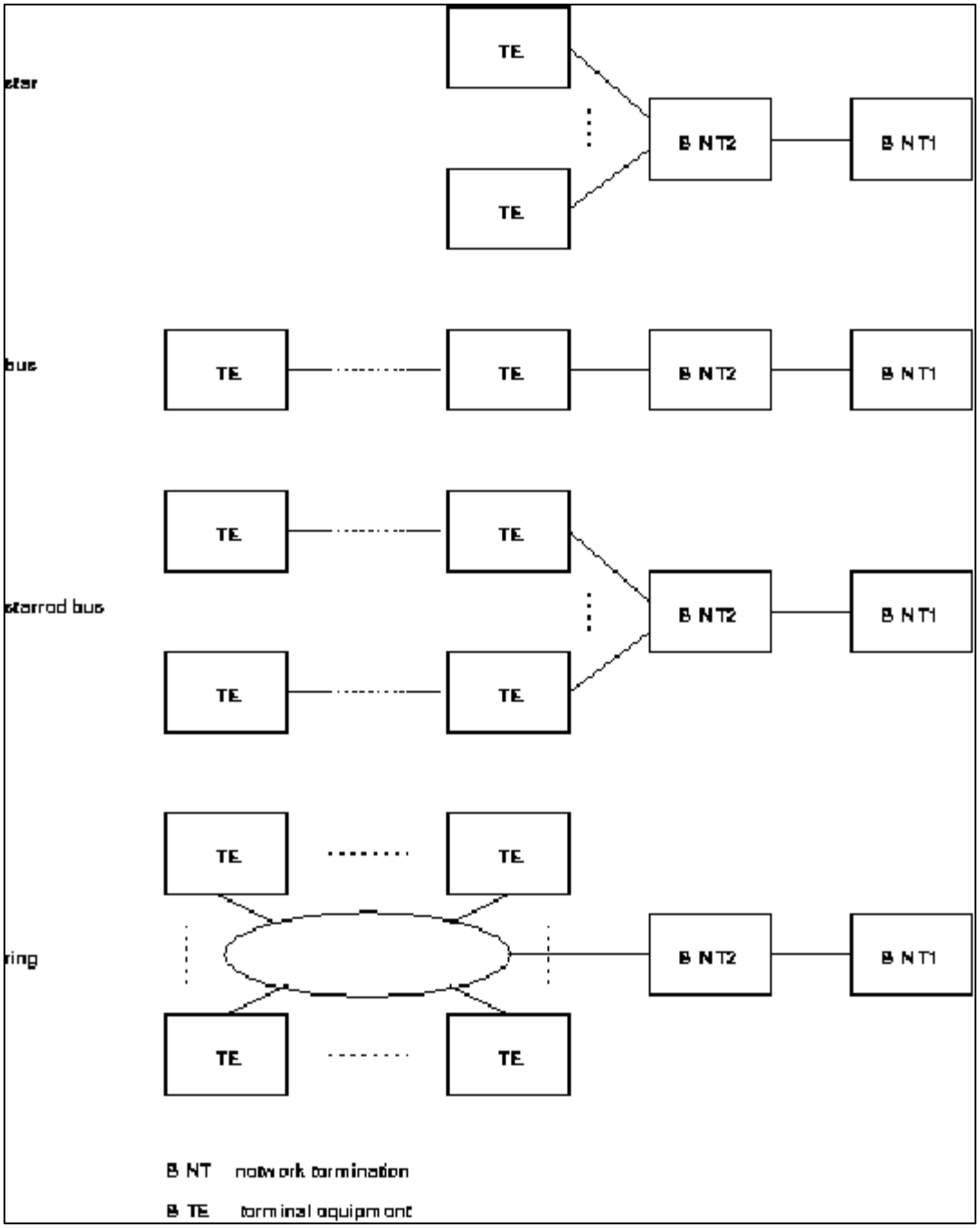


Figure 9: Broadband ISDN multiple interface configurations

The B-NT2 equipment is considered to be the customer premisis equipment (CPN) (Figure 10). This could in real terms be an private branch exchange (PBX) or other local switch.

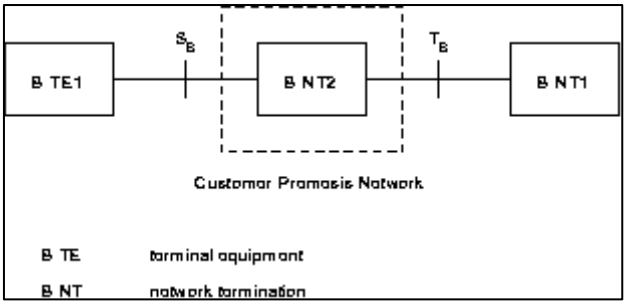


Figure 10: B-ISDN customer premisis network configuration

The discussion above has mentioned the OSI reference model. This was developed in collaboration between the ISO and the (then) CCITT. It seems surprising therefore that there is no defined relationship between the B-ISDN PRM and the OSI reference model. Figure 11 is the author's view of the relationship between the two.

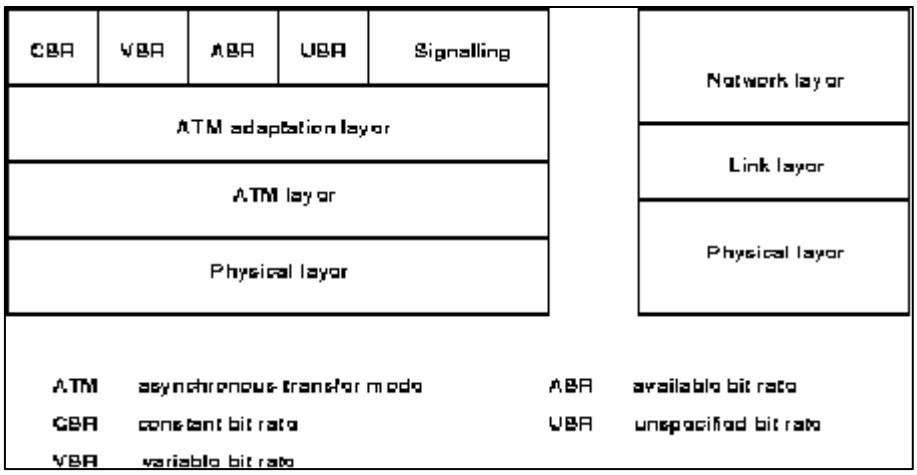


Figure 11: Broadband ISDN PRM compared with OSI model

As there is unlikely to be a user interface directly to the AAL, included in this figure are the interfaces to the service classes defined by the ATM Forum for the UNI:

- **Constant bit rate (CBR)**. The CBR service offers a very simple, reliable guaranteed channel that effectively acts as circuit emulation. The QoS of this service must be maintained throughout the lifetime of a CBR connection, as the data rate is expected to be constant. It is intended for use by applications with stringent real-time constraints on delay and jitter, e.g.\ real-time video.

- **Variable bit rate (VBR).** This service is also intended for use by by real-time applications. However, it differs from CBR in that it does not expect the data rate to be constant, i.e. the sources may use variable bit rate coding for efficiency and also be statistically multiplexed.
- **Available bit rate (ABR)**. This service class offers the B-ISDN user some degree of fairness, and also control of loss or delay with respect to QoS, but is intended for non real-time applications. It is likely that ABR QoS statements will specify that there are minimum acceptable parameters, but that if better QoS should become available then it will be used. ABR is intended for use by unit-oriented applications such as database access and electronic mail.
- **Unspecified bit rate (UBR)**. UBR is intended for applications that send data very sporadically and the use of CBR, VBR or ABR would be wasteful of resources. In fact, this service class is effectively a best-effort approach which is similar to today's IP. Applications that use this service would have non-real time requirements and not be too sensitive to loss, e.g. file transfers.

6.6 OPERATION AND MANTAINANCE

OAM functions in the network are performed on five OAM hierarchical levels associated with the ATM and physical layers of the protocol reference model. The functions result in corresponding bidirectional information flows F1, F2, F3, F4 and F5 referred to as OAM flows. Not all of these flows need to be present. The OAM functions of a missing level are performed at the next higher level. The levels are as follows:

- *Virtual channel level:* Extends between network elements performing VCC termination functions and is shown extending through one or more VPCs

- *Virtual path level:* Extends between network elements performing VPC termination functions and is shown extending through one or more transmission paths.
- *Transmission path level:* Extends between network elements assembling/disassembling the payload of a transmission system and associating it with its OAM functions. Cell delineation and Header Error Control (HEC) functions are required at the endpoints of each transmission path. The transmission path is connected through one or more digital sections.
- *Digital section level:* Extends between section endpoints and comprises a maintenance entity.
- *Regenerator section level:* A regenerator section is a portion of a digital section and as such is a maintenance sub-entity.

This layered concept and the requirements of independence of the layers from each other lead to the following principles. OAM functions related to OAM levels are independent from the OAM functions of other levels and have to be provided at each level. Each level, where OAM functions are required, is able to carry out its own processing to obtain quality and status information.

6.7 SUMMARY:

ISDN is comprised of digital telephony and data-transport services offered by regional telephone carriers. ISDN involves the digitization of the telephone network to transmit voice, data, text, graphics, music, video, and other source material over existing telephone wires.

ISDN devices include the following:

- Terminals
- Terminal adapters (TAs)
- Network-termination devices
- Line-termination equipment
- Exchange-termination equipment

The ISDN specification references specific connection points that define logical interfaces between devices.

ISDN uses the following two types of services:

- Basic Rate Interface (BRI, which offers two B channels and one D channel (2B+D)
- Primary Rate Interface (PRI), which offers 23 B channels and 1 D channel in North America and Japan, and 30 B channels and 1 D channel in Europe and Australia

ISDN runs on the bottom three layers of the OSI reference model, and each layer uses a different specification to transmit data.

6.8 KEYWORDS

- Integrted services data Network (ISDN)
- Broadband Integrated services Data Network (B-ISDN)
- Terminal Endpoint Identifier (TEI)
- service access point identifier (SAPI)
- Terminal adapters (TAs)
- Header Error Control (HEC)
- Primary Rate Interface (PRI)
- Basic Rate Interface (BRI)
- Operation and Maintenance (OAM)
- High-Level Data Link Control (HDLC)

6.9 REVIEW QUESTIONS

Q1. What are the two speeds of ISDN PRI services?

Q2. Explian two types of services associated with ISDn

Q3. Explian B-ISDN services

Q4. Discribe operation and Maintenance of B-ISDN network.

Q5. Write in detail about various B-ISDN Services.

6.10 FURTHER READING

- http://docstore.mik.ua/univercd/cc/td/doc/product/wanbu/bpx8600/8_4/over84/sysbrdtk.htm
- https://fmfi-uk.hq.sk/Informatika/Distribuovane%20Systemy/knihy/ICN/ch1s6.

LESSON 7 ASYNCHRONOUS TRANSFER MODE

7.1 Objectives

The objective of this chapter is to introduce students about Asynchronous Transfer Mode (ATM**)** cell relay protocol designed by the ATM Forum and adopted by the ITU-T. The combination of ATM and SONET will allow high-speed interconnection of all the world's networks. In fact, ATM can be thought of as the "highway" of the information superhighway.

7.2 INTRODUCTION TO ATM

Asynchronous Transfer Mode (ATM) is the cell relay protocol designed by the ATM Forum and adopted by the ITU-T. The combination of ATM and SONET will allow high-speed

interconnection of all the world's networks. In fact, ATM can be thought of as the "highway" of the information superhighway.

Design Goals

Among the challenges faced by the designers of ATM, six stand out.

1. Foremost is the need for a transmission system to optimize the use of high-data-rate transmission media, in particular optical fiber. In addition to offering large band-widths, newer transmission media and equipment are dramatically less susceptible to noise degradation. A technology is needed to take advantage of both factors and thereby maximize data rates.
2. The system must interface with existing systems and provide wide-area interconnectivity between them without lowering their effectiveness or requiring their replacement.
3. The design must be implemented inexpensively so that cost would not be a barrier to adoption. If ATM is to become the backbone of international communications, as intended, it must be available at low cost to every user who wants it.
4. The new system must be able to work with and support the existing telecommunications hierarchies (local loops, local providers, long-distance carriers, and so on).
5. The new system must be connection-oriented to ensure accurate and predictable delivery.
6. Last but not least, one objective is to move as many of the functions to hardware as possible (for speed) and eliminate as many software functions as possible (again for speed).

7.3 OVERVIEW OF ATM

ATM is a cell-switching and multiplexing technology that combines the benefits of circuit switching (guaranteed capacity and constant transmission delay) with those of packet switching (flexibility and efficiency for intermittent traffic). It provides scalable bandwidth from a few megabits per second (Mbps) to many gigabits per second (Gbps). Because of its asynchronous nature, ATM is more efficient than synchronous technologies, such as time-division multiplexing (TDM). With TDM, each user is assigned to a time slot, and no other station can send in that time slot. If a station has much data to send, it can send only when its time slot comes up, even if all other time slots are empty. However, if a station has nothing to transmit when its time slot comes up, the time slot is sent empty and is wasted. Because ATM is asynchronous, time slots are available on demand with information identifying the source of the transmission contained in the header of each ATM cell.

ATM transfers information in fixed-size units called cells. Each cell consists of 53 octets, or bytes. The first 5 bytes contain cell-header information, and the remaining 48 contain the payload (user information). Small, fixed-length cells are well suited to transferring voice and video traffic because such traffic is intolerant of delays that result from having to wait for a large data packet to download, among other things. Figure 7.1 illustrates the basic format of an ATM cell.

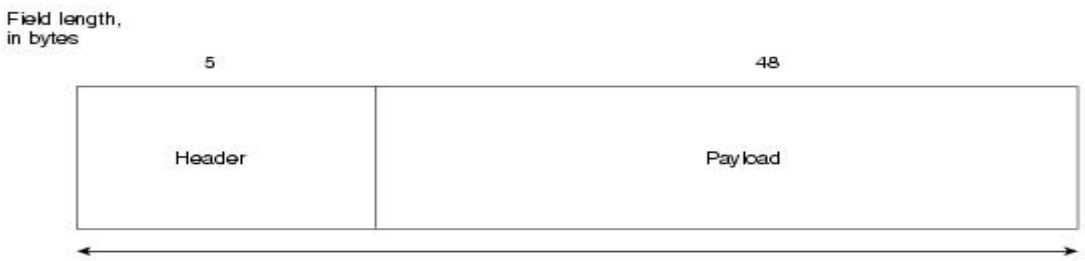


Figure 7.1: An ATM Cell Consists of a Header and Payload Data

7.4 ATM REFERENCE MODEL

ATM is almost similar to cell relay and packets witching using X.25and frame relay. Like packet switching and frame relay, ATM involves the transfer of data in discrete pieces also, like packet switching and frame relay, ATM allows multiple logical connections to multiplex over a single physical interface in the case of ATM the information flow on each logical connection is organized into fixed-size packets, called cells. ATM is a streamlined protocol with minimal error and flow control capabilities: this reduces the overhead of processing ATM cells and reduces the number of overhead bits required with each cell, thus enabling ATM to operate at high data rates. The use of fixed-size cells simplifies the processing required at each ATM node, again supporting the use of ATM at high data rates. The ATM architecture uses a logical model to describe the functionality that it supports. ATM functionality corresponds to the physical layer and part of the data link layer of the OSI reference model. The protocol reference model shown makes reference to three separate planes:

- **User plane** provides for user information transfer, along with associated controls (e.g. flow control, error control).
- **Control plane** performs call control and connection control functions.
- **Management plane** includes plane management, which performs management function related to a system as a whole and provides coordination between all the planes, and layer management which performs management functions relating to resource and parameters residing in its protocol entities .

The ATM reference model is composed of the following ATM layers:

• **Physical layer**—Analogous to the physical layer of the OSI reference model, the ATM physical layer manages the medium-dependent transmission.

• **ATM layer**—Combined with the ATM adaptation layer, the ATM layer is roughly analogous to the data link layer of the OSI reference model. The ATM layer is responsible for the simultaneous sharing of virtual circuits over a physical link (cell multiplexing) and passing cells through the ATM network (cell relay). To do this, it uses the VPI and VCI information in the header of each ATM cell.

• **ATM adaptation layer (AAL)**—Combined with the ATM layer, the AAL is roughly analogous to the data link layer of the OSI model. The AAL is responsible for isolating higher-layer protocols from the details of the ATM processes. The adaptation layer prepares user data for conversion into cells and segments the data into 48-byte cell payloads.

Finally, the higher layers residing above the AAL accept user data, arrange it into packets, and hand it to the AAL. Figure 7.2 illustrates the ATM reference model.

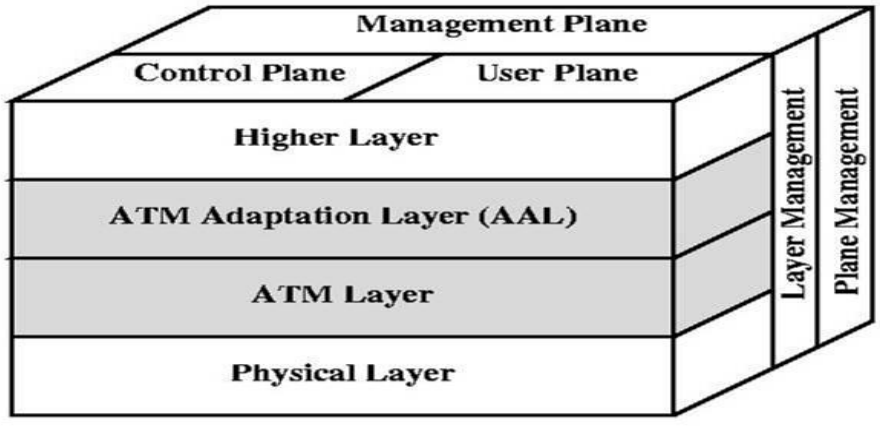


Figure 7.2 ATM Reference model

ATM Adaptation Layers (AAL) The use of Asynchronous Transfer Mode (ATM) technology and services creates the need for an adaptation layer in order to support information transfer protocols, which are not based on ATM. This adaptation layer defines how to segment and reassemble higher-layer packets into ATM cells, and how to handle various transmission aspects in the ATM layer.

Application Adaptation Layer

The **application adaptation layer (AAL)** was designed to enable two ATM concepts. First ATM must accept any type of payload both data frames and streams of bits. A data frame can come from an upper-layer protocol that creates a clearly defined frame to be sent to a carrier network such as ATM. A good example is the Internet. ATM also carry multimedia payload. It can accept continuous bit streams and break them into chunks to be encapsulated into a cell at the ATM layer. AAL uses two sublayers accomplish these tasks.

Whether the data are a data frame or a stream of bits, the payload must be segmented into 48-byte segments to be carried by a cell. At the destination, these need to be reassembled to recreate the original payload. The AAL defines a sublayer called a **segmentation and reassembly (SAR)** sublayer, to do so. Segmentation is at the source; reassembly, at the destination.

Before data are segmented by SAR, they must be prepared to guarantee the integrity of the data. This is done by a sublayer called the **convergence sublayer (CS).**

ATM defines four versions of the AAL: **AAL1, AAL2, AAL3/4, and AAL5.** T**he** versions today are AAL1 and AAL5.

AAL1 AAL1 supports applications that transfer information at constant bit rate such as video and voice. It allows ATM to connect existing digital telephone such as voice channels

and T lines. Figure 7.3 shows how a bit stream of data is chopped into 47-byte chunks and encapsulated in cells.

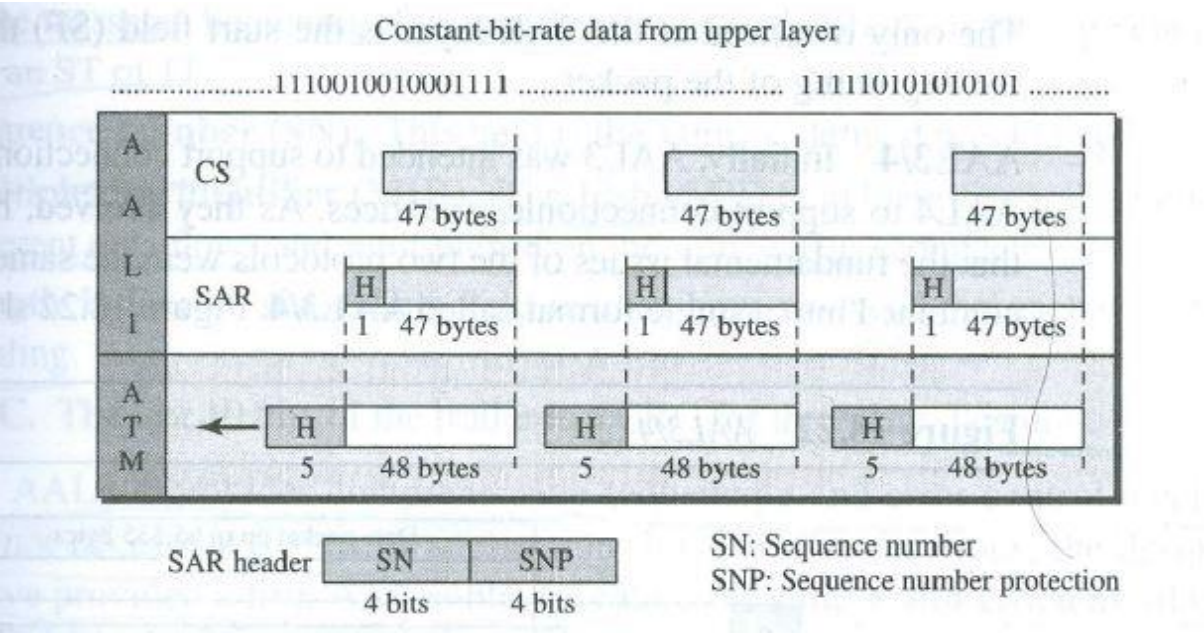


Figure 7.3: AAL1

The CS sublayer divides the bit stream into 47-byte segments and passes them to the SAR sublayer below. Note that the CS sublayer does not add a header.

The SAR sublayer adds 1 byte of header and passes the 48-byte segment to the ATM layer. The header has two fields:

- **Sequence number (SN).** This 4-bit field defines a sequence number to order the bits. The first bit is sometimes used for timing, which leaves 3 bits for sequencing (modulo 8).
- **Sequence number protection (SNP).** The second 4-bit field protects the first field. The first 3 bits automatically correct the SN field. The last bit is a parity bit that detects error over all 8 bits.

AAL2 Originally AAL2 was intended to support a variable-data-rate bit stream but it has been redesigned. It is now used for low-bit rate traffic and short-frame traffic such as audio (compressed or uncompressed), video, or fax. A good example of AAL2 use is in mobile telephony. AAL2 allows the multiplexing of short frames into one cell.

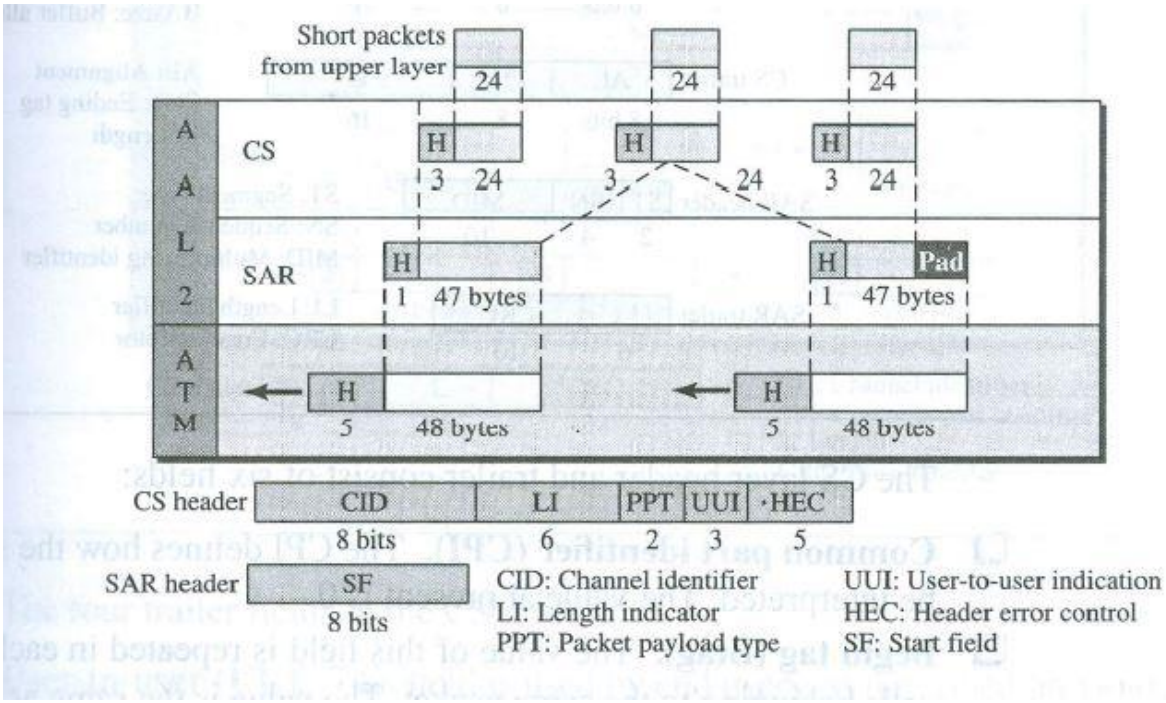


Figure 7.4: AAL2

Figure 7.4 shows the process of encapsulating a short frame from the same source (the same user of mobile phone) or from several source(several user of mobile telephones) into one cell.

The CS layer overhead consists of five fields:

- **Channel identifier (CID).** The 8-bit CID field defines the channel (user) of the short packet.
- **Length indicator (LI).** The 6-bit LI field indicates how much of the final packet data is.

□ **Packet payload type (PPT).** The PPT field defines the type of packet.

□ **User-to-User Indicator(UUI).** The UUI field can be used by end-to-end users.

□ **Header Error Control (HEC).** the last 5 bits is use to correct errors in the header.

The only overhead at the SAR is the start field(SF) and define the offset from the beginning of the packet.

AAL3/4 Initially, AAL3 was intended to support connection-oriented data services and AAL4 to support connectionless services. As they evolved, however, it become evident that the fundamental issues of the two protocols were the same. They have therefor combined into a single format called AAL 3/4. Figure 7.5 shows the AAL 3/4 sublayer

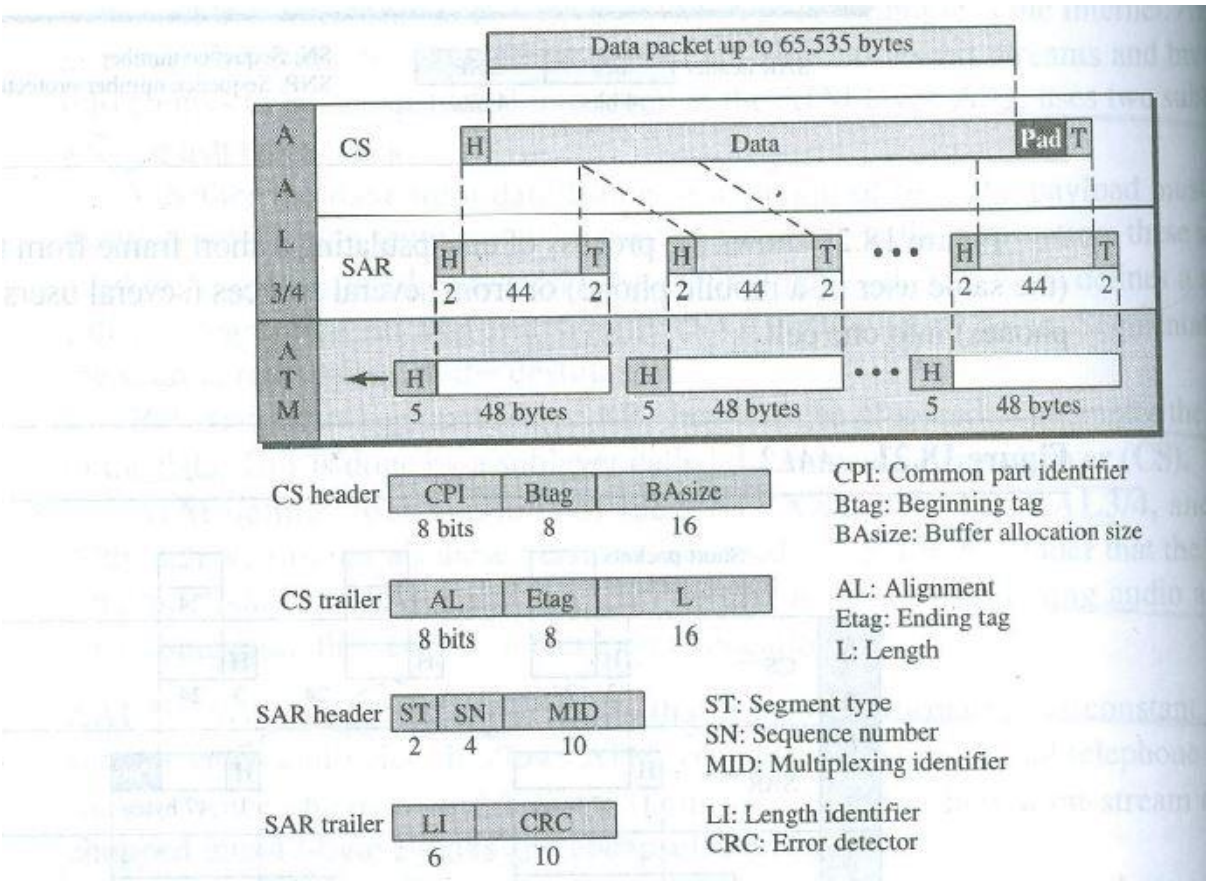


Figure 7.5: AAL 3/5

The CS layer header and trailer consist of six fields:

- **Common part identifier (CPI).** The CPI defines how the subsequent fields are to be interpreted. The value at present is 0.
- **Begin tag (Btag).** The value of this field is repeated in each cell to identify in all cells belonging to the same packet. The value is the same as the Etag.
- **Buffer allocation size (BAsize).** The 2-byte BA field tells the receive what size buffer is needed for the coming data.
- **Alignment (AL).** The 1-byte AL field is included to make the rest |of the trailer 4 bytes long.
- **Ending tag (Etag).** The 1-byte ET field serves as an ending flag, Its value is same as that of the beginning tag.
- **Length (L).** The 2-byte L field indicates the length of the data unit.

The SAR header and trailer consist of five fields:

- **Segment type (ST).** The 2-bit ST identifier specifies the position of the segment in the message: beginning (00), middle (01), or end (10). A single-segment message has an ST of 11.
- **Sequence number (SN).** This field is the same as defined previously.
- **Multiplexing identifier (MID).** The 10-bit MID field identifies cells coming from different data flows and multiplexed on the same connection
- **Length indicator (LI).** This field defines how much of the packet is data, not padding.
- **CRC.** The last 10 bits of the trailer is a CRC for the entire data unit.

AAL5 AAL3/4 provides comprehensive sequencing and error control mechanisms that are not necessary for every application. For these applications, the designers of ATM have provided a fifth AAL sublayer, called the **simple and efficient adaptation layer (SEAL).**

AAL5 assumes that all cells belonging to a single message travel sequentially and that control functions are included in the upper layers of the sending applications. Figure 7.6 shows the AAL5 sublayer.

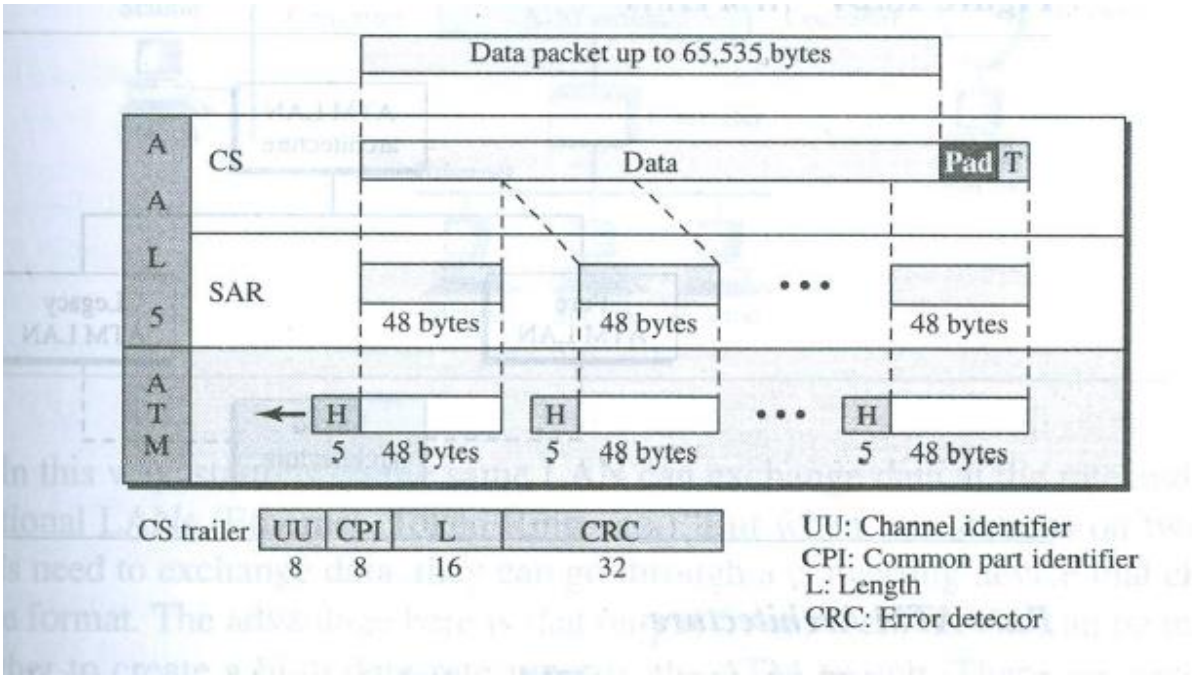


Figure 7.6: AAL5

The four trailer fields in the CS layer are

- **User-to-user (UU).** This field is used by end users, as described previously.
- **Common part identifier (CPI).** This field is the same as defined previously.
- **Length (L).** The 2-byte L field indicates the length of the original data.
- **CRC.** The last 4 bytes is for error control on the entire data unit.

7.5 ATM Switching Operation

ATM interconnection is capable of addressing almost 2^{12} VPs of up to almost 2^{16} VCs each (in practice some of the VP and VC numbers are reserved) (Figure 7.7)

Figure 7.7: ATM virtual connection

A **Virtual Path (VP)** denotes the transport of ATM cells belonging to virtual channels which share a common identifier, called the Virtual Path Identifier (VPI), which is also encoded in the cell header. A virtual path, in other words, is a grouping of virtual channels which connect the same end-points. This two layer approach results in improved network performance. Once a virtual path is set up, the addition/removal of virtual channels is straightforward

Virtual Channel (VC) denotes the transport of ATM cells which have the same unique identifier, called the Virtual Channel Identifier (VCI). This identifier is encoded in the cell header. A virtual channel represents the basic means of communication between two end-points, and is analogous to an X.25 virtual circuit.

7.6 ATM Cell Header Format

An ATM cell consists of a 5 byte header and a 48 byte payload. The payload size of 48 bytes was a compromise between the needs of voice telephony and packet networks, obtained by a simple averaging of the US proposal of 64 bytes and European proposal of 32, said by some to be motivated by a European desire not to need echo-cancellers on national trunks.

ATM defines two different cell formats: NNI (Network-network interface) and UNI (User-network interface). Most ATM links use UNI cell format.

Diagram of the UNI ATM Cell

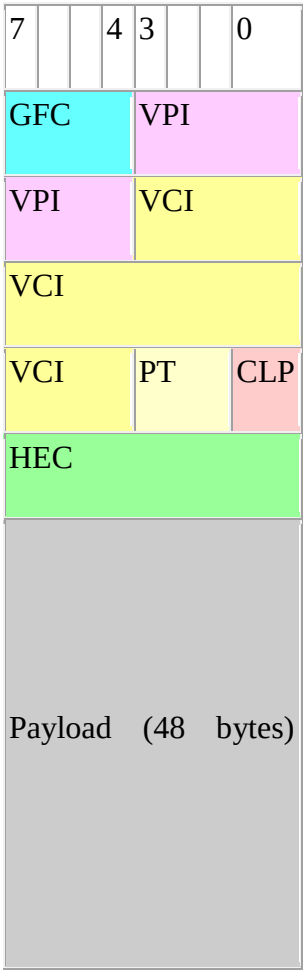


Diagram of the NNI ATM Cell

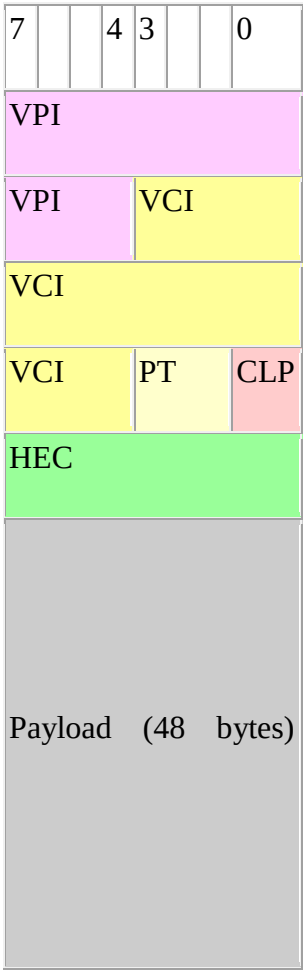


Figure 7.8: Format of UNI and NNI cell

GFC = Generic Flow Control (4 bits) (default: 4-zero bits)

VPI = Virtual Path Identifier (8 bits UNI) or (12 bits NNI)

VCI = Virtual channel identifier (16 bits)

PT = Payload Type (3 bits)

CLP = Cell Loss Priority (1-bit)

HEC = Header Error Correction (8-bit CRC, polynomial = $X^8 + X^2 + X + 1$)

The PT field is used to designate various special kinds of cells for Operation and Management (OAM) purposes, and to delineate packet boundaries in some AALs.

Several of ATM's link protocols use the HEC field to drive a CRC-Based Framing algorithm, which allows the position of the ATM cells to be found with no overhead required beyond what is otherwise needed for header protection. The 8-bit CRC is used to correct single-bit header errors and detect multi-bit header errors. When multi-bit header errors are detected, the current and subsequent cells are dropped until a cell with no header errors is found.

In a UNI cell the GFC field is reserved for a local flow control/submultiplexing system between users. This was intended to allow several terminals to share a single network connection, in the same way that two ISDN phones can share a single basic rate ISDN connection. All four GFC bits must be zero by default.The NNI cell format is almost identical to the UNI format, except that the 4-bit GFC field is re-allocated to the VPI field, extending the VPI to 12 bits. Thus, a single NNI

7.7ATM ADDRESSING

ATM Cell Addressing

Each ATM cell contains a two-part address, VPI/VCI, in the cell header. This address uniquely identifies an individual ATM virtual connection on a physical interface. Virtual Channel Indicator (VCI) bits are used to identify the individual circuit or connection. Multiple virtual circuits that traverse the same physical layer connection between nodes are

grouped together in a virtual path (Figure 7.9). The virtual path address is given by the Virtual Path Indicator (VPI) bits. The Virtual Path can be viewed as a trunk that carries multiple circuits all routed the same between switches.

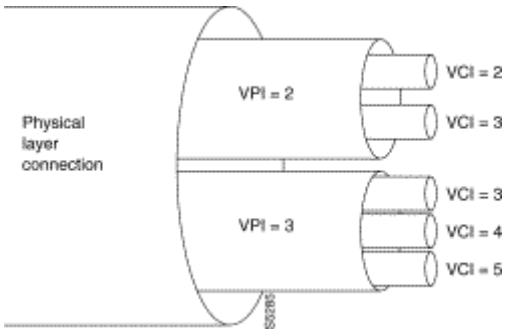


Figure 7.9: Virtual Paths and Virtual Channels

The VPI and VCI addresses may be translated at each ATM switch in the network connection route. They are unique only for a given physical link. Therefore, they may be reused in other parts of the network as long as care is taken to avoid conflicts. Figure 7.10 illustrates switching using VP only, which may be done at tandem switches while Figure 7.11 illustrates switching on VC as well as VP.

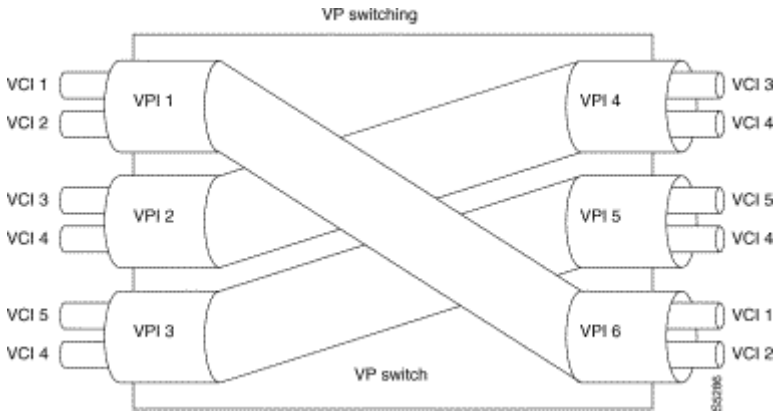


Figure 7.10: VP-only Switching

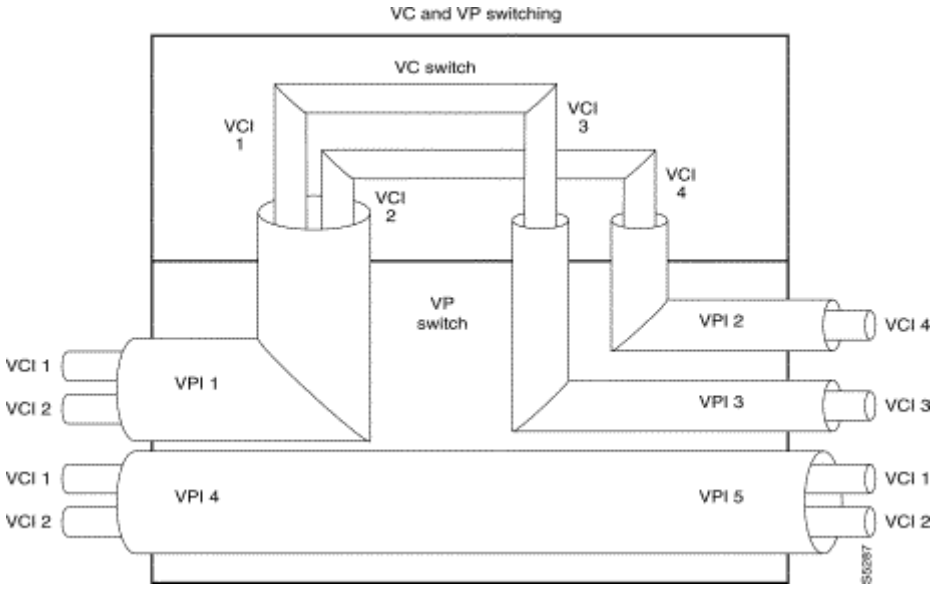


Figure 7.11: VP and VC Switching

The VCI field is 16 bits wide with UNI and NNI header types described earlier. This allows for a total possible 65, 535 unique circuit numbers. The UNI header reserves 8 bits for VPI (256 unique paths) while the NNI reserves 12 bits (4,096 unique paths) as it is likely that more virtual paths will be routed between networks than between a user and the network. The STI header reserves 8 bits for VCI and 10 bits for VPI addresses.

ATM addresses are needed to support the use of switched virtual connections (SVCs) through an ATM network.

At the simplest level, ATM addresses are 20 bytes long and have three distinct parts:

- **Network prefix**

the first 13 bytes identify the location of a specific switch in the network. The use of this

portion of the address can vary considerably depending on its address format. Each of the three standards ATM addressing schemes provides information about ATM switch locations differently. These schemes include the data country/region code (DCC) format, the international code designator (ICD) format, and the E.164 format proposed by the ITU-T for international telephone numbering use in broadband ISDN networks.

- **Adapter media access control address**

The next 6 bytes identify a physical endpoint, such as a specific ATM adapter card, using a media access control layer address that is physically assigned to the ATM hardware by its manufacturer. The use and assignment of media access control addresses for ATM hardware are identical to how this addressing works for Ethernet, Token Ring, and other IEEE 802.*x* technologies.

- **Selector (SEL)**

The last byte is used to select a logical connection endpoint on the physical ATM adapter. Although all ATM addresses fit this basic three-part structure, there are significant differences in the exact format of the first 13 bytes, depending on the addressing format or whether the ATM network is for public or private use.

All of the three ATM address formats that are currently in widespread use (DCC, ICD, and E.164) include the following characteristics:

- They comply with the Network Service Access Point (NSAP) addressing plan as proposed by the Open Standards Interconnection (OSI) protocol suite of the International Standards Organization (ISO).

- Each can be used to establish and interconnect privately built ATM networks that support switched virtual connections (SVCs).

7.8 ATM IN DATA LINK LAYER

It is now time to begin our journey up through the ATM protocol layers of Fig. 1 -30. The ATM physical layer covers roughly the OSI physical and data link layers, with the physical medium dependent sublayer being functionally like the OSI physical layer and the transmission convergence (TC) sublayer having data link functionality. There are no physical layer characteristics specific to ATM. Instead, ATM cells are carried by SONET, FDDI, and other transmission systems. Therefore we will concentrate here on the data link functionality of the TC sublayer, but we will discuss some aspects of the interface with the lower sublayer later on.

When an application program produces a message to be sent, that message works its way down the ATM protocol stack, having headers and trailers added and undergoing segmentation into cells. Eventually, the cells reach the TC sublayer for transmission. Let us see what happens to them on the way out the door.

Cell Transmission

The first step is header checksumming. Each cell contains a 5-byte header consisting of 4 bytes of virtual circuit and control information followed by a 1-byte checksum. Although the contents of the header are not relevant to the TC sublayer, curious readers wishing a sneak preview should turn to Fig. 5-62. The checksum only covers the first four header bytes, not the payload field. It consists of the remainder after the 32 header bits have been divided by

the polynomial $x^8 + x^2 + x + 1$. To this the constant 01010101 is added, to provide robustness in the face of headers containing mostly 0 bits.

The decision to checksum only the header was made to reduce the probability of cells being delivered incorrectly due to a header error, but to avoid paying the price of checksumming the much larger payload field. It is up to higher layers to perform this function, if they so desire. For many real-time applications, such as voice and video, losing a few bits once in a while is acceptable (although for some compression schemes, all frames are equal but some frames are more equal). Because it covers only the header, the 8-bit checksum field is called the **HEC (Header Error Control).**

A factor that played a major role in this checksumming scheme is the fact that ATM was designed for use over fiber, and fiber is highly reliable. Furthermore, a major study of the U.S. telephone network has shown that during normal operation 99.64 percent of all errors on fiber optic lines are single-bit errors (AT&T and Bellcore, 1989). The HEC scheme corrects all single-bit errors and detects many multibit errors as well. If we assume that the probability **of** a single-bit **error** is 10^{-8}, then the probability of a cell containing a detectable multibit header error is about $1(T^{13}$. The probability of a cell slipping through with an undetected header error is about 10 , which means that at OC-3 speed, one bad cell header will get through every 90.000 years. Although this may sound like a long time, once the earth has, say, 1 billion ATM telephones, each used 10 percent of the time, over 1000 bad cell headers per year will go undetected.

For applications that need reliable transmission in the data link layer, Shac-ham and McKenney (1990) have developed a scheme in which a sequence of consecutive cells are EXCLUSIVE ORed together. The result, an entire cell, is appended to the sequence. If one cell is lost or badly garbled, it can be reconstructed from the available information.

Once the HEC has been generated and inserted into the cell header, the cell is ready for transmission. Transmission media come in two categories: asynchronous and synchronous. When an asynchronous medium is used, a cell can be sent whenever it is ready to go. No timing restrictions exist.

With a synchronous medium, cells must be transmitted according to a predefined timing pattern. If no data cell is available when needed, the TC sublayer must invent one. These are called idle cells.

Another kind of non-data cell is the OAM (Operation And Maintenance) cell. OAM cells are also used by the ATM switches for exchanging control and other information necessary for keeping the system running. OAM cells also have some other special functions. Por example, the 155.52-Mbps OC-3 speed matches the gross data rate of SONET, but an STM-1 frame has a total of 10 columns of overhead out of 270, so the SONET payload is only 260/270 x 155.52 Mbps or 149.76 Mbps. To keep from swamping SONET, an ATM source using SONET would normally put out an OAM cell as every 27th cell, to slow the data rate down to 26/27 of 155.52 Mbps and thus match SONET exactly. The job of matching the ATM output rate to the rate of the underlying transmission system is an important task of the TC sublayer.

On the receiver's side, idle cells are processed in the TC sublayer, but OAM cells are given to the ATM layer. OAM cells are distinguished from data cells by having the first three header bytes be all zeros, something not allowed for data cells. The fourth byte describes the nature of the OAM cell.

Another important task of the TC sublayer is generating the framing information for the underlying transmission system, if any. For example, an ATM video camera might just produce a sequence of cells on the wire, but it might also produce SONET frames with the ATM cells embedded inside the SONET payload. In the latter case, the TC sublayer would

generate the SONET framing and pack the ATM cells inside, not entirely a trivial business since a SONET payload does not hold an integral number of 53-byte cells.

Although the telephone companies clearly intend to use SONET as the underlying transmission system for ATM, mappings from ATM onto the payload fields of other systems have also been defined, and new ones are being worked on. In particular, mappings onto Tl, T3, and FDDI also exist.

Cell Reception

On output, the job of the TC sublayer is to take a sequence of cells, add a HEC to each one, convert the result to a bit stream, and match the bit stream to the speed of the underlying physical transmission system by inserting OAM cells as filler. On input, the TC sublayer does exactly the reverse. It takes an incoming bit stream, locates the cell boundaries, verifies the headers (discarding cells with invalid headers), processes the OAM cells, and passes the data cells up to the ATM layer.

The hardest part is locating the cell boundaries in the incoming bit stream. At the bit level, a cell is just a sequence of 53 x 8 = 424 bits. No 01111110 flag bytes are present to mark the start and end of a cell, as they are in HDLC. In fact, there are no markers at all. How can cell boundaries be recognized under these circumstances?

In some case, the underlying physical layer provides help. With SONET, for example, cells can be aligned with the synchronous payload envelope, so the SPE pointer in the SONET header points to the start of the first full cell. However, sometimes the physical layer provides no assistance in framing. What then?

The trick is to use the HEC. As the bits come in, the TC sublayer maintains a 40-bit shift register, with bits entering on the left and exiting on the right. The TC sublayer then inspects the 40 bits to see if it is potentially a valid cell header. If it is, the rightmost 8 bits will be

valid HEC over the leftmost 32 bits. If this condition does not hold, the buffer does not hold a valid cell, in which case all the bits in the buffer are shifted right one bit, causing one bit to fall off the end, and a new input bit is inserted at the left end. This process is repeated until a valid HEC is located. At that point, the cell boundary is known because the shift register contains a valid header.

The trouble with this heuristic is that the HEC is only 8 bits wide. For any given shift register, even one containing random bits, the probability of finding a valid HEC is 1/256, a moderately large value. Used by itself, this procedure would incorrectly detect cell headers far too often.

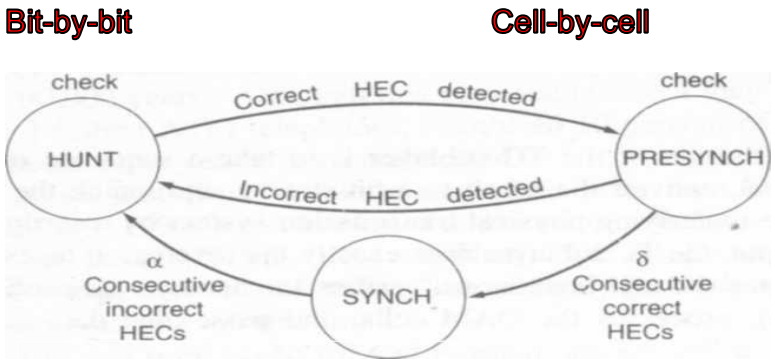


Figure 7.12: Finite state machine of recognition algorithm

To improve the accuracy of the recognition algorithm, the finite state machine of Figure 7.12 is used. Three states are used: *HUNT*, *PRESYNCH*, and *SYNCH*. In the *HUNT* state, the TC sublayer is shifting bits into the shift registers one at a time looking for a valid HEC. As soon as one is found, the finite state machine switches to *PRESYNCH* state, meaning that it

has tentatively located a cell boundary. It now shifts in the next 424 bits (53 bytes) without examining them. If its guess about the cell boundary was correct, the shift register should now contain another valid cell header, so it once again runs the HEC algorithm. If the HEC is incorrect, the TC goes back to the *HUNT* state and continues to search bit-by-bit for a header whose HEC is correct.

On the other hand, if the second HEC is also correct, the TC may be onto something, so it shifts in another 424 bits and tries again. It continues inspecting headers in this fashion until it has found 8 correct headers in a row, at which time it assumes that it is synchronized and moves into the *SYNCH* state to start normal operation. Note that the probability of getting into *SYNCH* state by accident with a purely random bit stream is $2\sim^{86}$, which can be made arbitrarily small by choosing a large enough 8. The price paid for a large 8, however, is a longer time to synchronize.

In addition to resynchronizing after losing synchronization (or at startup), the TC sublayer needs a heuristic to determine when it has lost synchronization, for example after a bit has been inserted or deleted from the bit stream. It would be unwise to give up if just one HEC was incorrect, since most errors are bit inversions, not insertions or deletions. The wisest course here is just to discard the cell with the bad header and hope the next one is good. However, if a HECs in a row are bad, the TC sublayer has to conclude that it has lost synchronization and must return to the *HUNT* state.

Although unlikely, it is conceivable that a malicious user could try to spoof the TC sublayer by inserting a data pattern into the payload field of many consecutive cells that imitates the HEC algorithm. Then, if synchronization were ever lost, it might be regained in the wrong place. To make this trick much harder, the payload bits are scrambled on transmission and descrambled on reception.

Before leaving the TC sub layer, one comment is in order. The mechanism chosen for cell delineation requires the TC sub layer to understand and use the header of the ATM layer above it. Having one layer make use of the header of a higher layer is in complete violation of the basic rules of protocol engineering. The idea of having layered protocols is to make each layer be independent of the

7.9 ATM IN TRANSPORT LAYER

ATM has a fixed transmission unit, called the ``Cell''. Each Cell is 53 octets, of which 5 octets are used for the ATM header and the remaining 48 octets can be used for data transport (this part is call the *Payload* or Service Data Unit (SDU)). The header can be this small, because instead of the destination ATM address, only the information that is needed locally for data forwarding is included in each Cell. This information consists of a Virtual Path Identifier (VPI) and Virtual Channel Identifier (VCI) and other administrative data. More on the use of these Virtual Channels and Paths in Section .

When an ATM connection is set up, the calling party specifies what kind of traffic the connection intends to use and what QoS it expects from the network. This information is put in the traffic descriptor and the QoS parameter respectively and then both are passed as part of the connection setup request to the network, along with other parameters like the destination address. The parameters are carried in Information Elements in the signalling messages. The Information Elements are:

- The Service Category (a.k.a. Transfer Mode): This is a property of the Virtual Channel Connection (VCC) that is being set up. Currently defined service categories are: Constant Bit-Rate (CBR), (non) real-time Variable Bit-Rate (rt-VBR, nrt-VBR), Available Bit-Rate (ABR) and Unspecified Bit-Rate (UBR). The Service

Category defines which QoS parameters should be specified for the connection (e.g. Cell Delay (CD), Cell Delay Variation (CDV) and Cell-Loss Ratio (CLR)). For CBR connections, all three above parameters are performance objectives, but for ABR, only CLR is needed.

- Traffic Descriptor: Peak Cell Rate (PCR), Sustainable Cell Rate (SCR), Maximum Burst Size (MBS). The traffic descriptor is used to allocate resources on the network.
- The QoS Class gives the values for the relevant QoS parameters for the chosen Service Category.

After the connection with a certain Service Category has been setup, the ATM client and the ATM network have a traffic contract associated with the connection. If the users of the connection attempt to use too much resources of the network, i.e. exceed the traffic contract parameters, the network may drop cells or use other means within the capabilities of the network to keep the contract. Traffic parameters can be re-negotiated while the connection is still in use. For more information about the information elements used in the signalling messages, please see the UNI specifications

7.10 SUMMARY

1. Asynchronous Transfer Mode (ATM**)** cell relay protocol designed by the ATM Forum and adopted by the ITU-T.
2. ATM is a cell-switching and multiplexing technology that combines the benefits of circuit switching (guaranteed capacity and constant transmission delay) with those of packet switching

3. ATM is a cell-switching and multiplexing technology that combines the benefits of circuit switching (guaranteed capacity and constant transmission delay) with those of packet switching
4. The ATM reference model is composed three layers namely physical layer, ATM layer, ATM Adaptation layer.
5. ATM defines four versions of the AAL: AAL1, AAL2, AAL3/4, and AAL5.

7.11 KEYWORDS

- ATM: Asymmetric Transfer mode.
- AAL: ATM Adaptation layer.
- transmission convergence (TC)
- international code designator (ICD) format
- **Common part identifier (CPI).**
- Generic Flow Control(GFC)
- Virtual Path (VP)

7.12 REVIEW QUESTIONS

Q1. Explain Asynchronous Transfer mode.

Q2. Draw and explain ATM reference model.

Q3. Explain ATM adaptation layers in detail. What is the format of header of each layer?

Q4.How switching is done in ATM?

Q5. Define the address format of ATM cell address.

7.13 FURTHER READINGS

1 H. Dutton and P. Lenhard, "Asnchronous Transfer Mode (ATM) Technical Overview", 2nd Ed., Prentice Hall, 1995

2 T. M. Chen, and S. S. Liu, "ATM Switching Systems", Artech House, INC., 1995

LESSON 8 Switched Multimegabit Data Services (SMDS)

8.1 OBJECTIVE

Such a system is called a public network. It is analogous to, and often a part of, the public telephone system. We already briefly looked at one new service, DQDB, In the following sections we will study four other example services, SMDS.

8.2 INTRODUCTION:

Switched Multimegabit Data Service (SMDS) is a packet-switched datagram service designed for very high-speed wide-area data communications. SMDS offers data throughputs that will initially be in the 1- to 34-Mbps range and is being deployed in public networks by the carriers in response to two trends. The first trend is the proliferation of distributed processing and other applications that require high-performance networking. The second trend is the decreasing cost and high-bandwidth potential of fiber media, making support of such applications over a wide-area network (WAN) viable.

SMDS is described in a series of specifications produced by Bell Communications Research (Bellcore) and adopted by the telecommunications equipment providers and carriers. One of these specifications describes the *SMDS Interface Protocol* (SIP), which is the protocol between a user device (referred to as *customer premises equipment*, or *CPE*), and SMDS network equipment.

The SIP is based on an IEEE standard protocol for metropolitan-area networks (MANs): that is, the *IEEE 802.6* Distributed Queue Dual Bus (DQDB) standard. Using this protocol, CPE such as routers can be attached to an SMDS network and use SMDS service for high-speed internetworking.

Technology Basics

Figure 8.1 shows an internetworking scenario using SMDS. In this figure, access to SMDS is provided over either a 1.544-Mbps (*DS-1,* or *Digital Signal 1*) or 44.736-Mbps (*DS-3, or Digital Signal 3*) transmission facility. Although SMDS is usually described as a fiber-based service, DS-1 access can be provided over either fiber or copper-based media with sufficiently good error characteristics. The demarcation point between the carrier's SMDS network and the customer's equipment is referred to as the *subscriber network interface* (SNI).

SMDS data units are capable of containing up to 9,188 octets (bytes) of user information. SMDS is therefore capable of encapsulating entire *IEEE 802.3*, *IEEE 802.4*, *IEEE 802.5*, and *FDDI* frames. The large packet size is consistent with the high-performance objectives of the service.

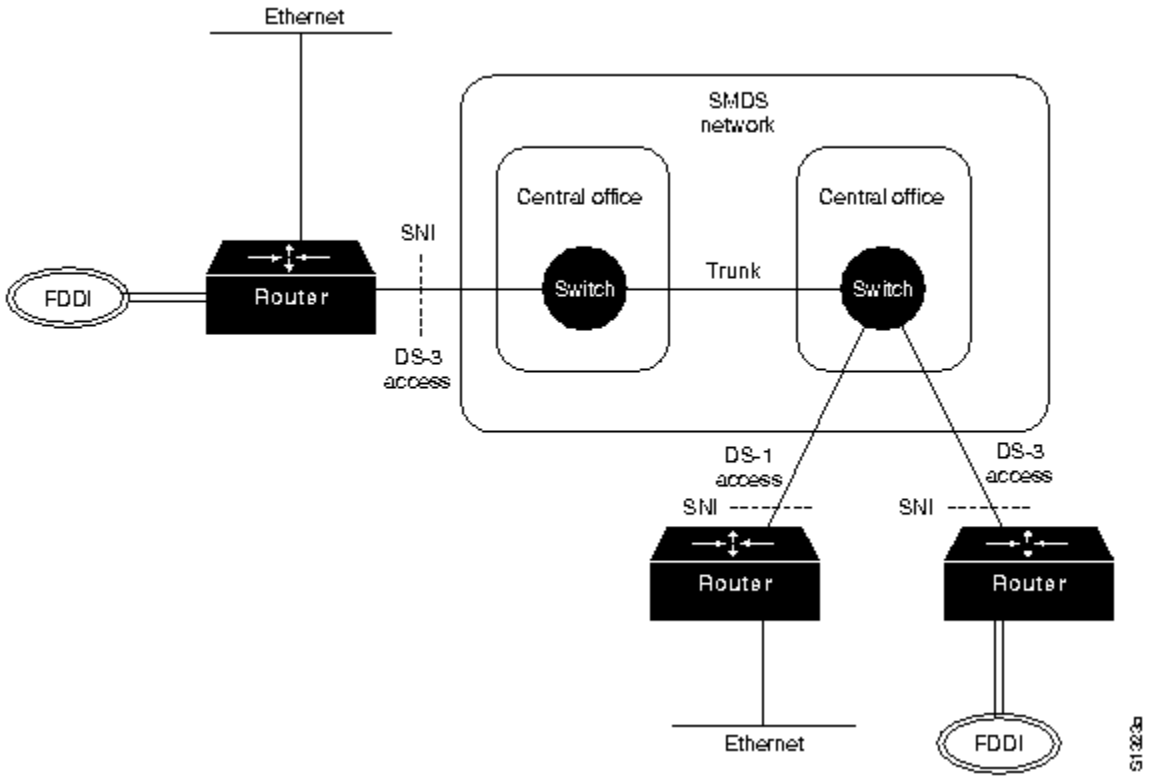


Figure 8-1 7SMDS Internetworking Scenario

Addressing

Like other datagram protocols, SMDS data units carry both a source and a destination address. The recipient of a data unit can use the source address to return data to the sender and for functions such as address resolution (discovering the mapping between higher-layer addresses and SMDS addresses). SMDS addresses are 10-digit addresses that resemble conventional telephone numbers.

In addition, SMDS supports group addresses that allow a single data unit to be sent and then delivered by the network to multiple recipients. Group addressing is analogous to multicasting on local-area networks (LANs) and is a valuable feature in internetworking applications where it is widely used for routing, address resolution, and dynamic discovery of network resources

(such as file servers).

SMDS offers several other addressing features. Source addresses are validated by the network to ensure that the address in question is legitimately assigned to the SNI from which it originated. Thus, users are protected against *address spoofing*---that is, a sender pretending to be another user. Source and destination address screening is also possible. Source address screening acts on addresses as data units are leaving the network, while destination address screening acts on addresses as data units are entering the network. If the address is disallowed, the data unit is not delivered. With address screening, a subscriber can establish a private virtual network that excludes unwanted traffic. This provides the subscriber with an initial security screen and promotes efficiency because devices attached to SMDS do not have to waste resources handling unwanted traffic.

Access Classes

To accommodate a range of traffic requirements and equipment capabilities, SMDS supports a variety of access classes. Different access classes determine the various maximum sustained information transfer rates as well as the degree of burstiness allowed when sending packets into the SMDS network.

On DS-3-rate interfaces, access classes are implemented through credit management algorithms, which track credit balances for each customer interface. Credit is allocated on a periodic basis, up to some maximum. Then, the credit balance is decremented as packets are sent to the network.

The operation of the credit management scheme essentially constrains the customer's equipment to some sustained or average rate of data transfer. This average rate of transfer is less than the full information carrying bandwidth of the DS-3 access facility. Five access

classes, corresponding to sustained information rates of 4, 10, 16, 25, and 34 Mbps, are supported for DS-3 access interface. The credit management scheme is not applied to DS1rate access interfaces.

8.3 DATA COMUNICATION SERVICES

Switched Multimegabit Data Service (SMDS) is a high-speed, packet-switched, datagram-based WAN networking technology used for communication over public data networks (PDNs). SMDS can use fiber- or copper-based media; it supports speeds of 1.544 Mbps over Digital Signal level 1 (DS-1) transmission facilities, or 44.736 Mbps over Digital Signal level 3 (DS-3) transmission facilities. In addition, SMDS data units are large enough to encapsulate entire IEEE 802.3, IEEE 802.5, and Fiber Distributed Data Interface (FDDI) frames. This chapter summarizes the operational elements of the SMDS environment and outlines the underlying protocol. A discussion of related technologies, such as Distributed Queue Dual Bus (DQDB) is also provided. The chapter closes with discussions of SMDS access classes and cell formats.

SMDS Network Components

SMDS networks consist of several underlying devices to provide high-speed data service. These include customer premises equipment (CPE), carrier equipment, and the subscriber network interface (SNI). CPE is terminal equipment typically owned and maintained by the customer. CPE includes end devices, such as terminals and personal computers, and intermediate nodes, such as routers, modems, and multiplexers. Intermediate nodes, however, sometimes are provided by the SMDS carrier. Carrier equipment generally consists of high-speed WAN switches that must conform to certain network equipment specifications, such as those outlined by Bell Communications Research (Bellcore). These specifications

define network operations, the interface between a local carrier network and a long-distance carrier network, and the interface between two switches inside a single carrier network.

The SNI is the interface between CPE and carrier equipment. This interface is the point at which the customer network ends and the carrier network begins. The function of the SNI is to render the technology and operation of the carrier SMDS network transparent to the customer. Figure 8.2 illustrates the relationships among these three components of an SMDS network.

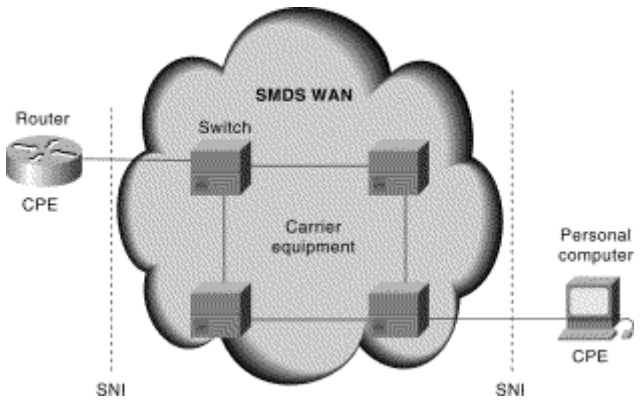


Figure 8.2: The SNI Provides an Interface between the CPE and the Carrier Equipment in SMDS

SMDS Interface Protocol

The *SMDS Interface Protocol (SIP)* is used for communications between CPE and SMDS carrier equipment. SIP provides connectionless service across the subscriber network interface (SNI), allowing the CPE to access the SMDS network. SIP is based on the IEEE 802.6 Distributed Queue Dual Bus (DQDB) standard for cell relay across metropolitan-area networks (MANs). The DQDB was chosen as the basis for SIP because it is an open standard that supports all the SMDS service features. In addition, DQDB was designed for compatibility

with current carrier transmission standards, and it is aligned with emerging standards for Broadb and ISDN (BISDN), which will allow it to interoperate with broadband video and voice services. Figure 8.3 illustrates where SIP is used in an SMDS network.

SIP Levels

SIP consists of three levels. SIP Level 3 operates at the Media Access Control (MAC) sublayer of the data link layer of the OSI reference model. SIP Level 2 operates at the MAC sublayer of the data link layer. SIP Level 1 operates at the physical layer of the OSI reference model. Figure 8.4 illustrates how SIP maps to the OSI reference model, including the IEEE data link sublayers.

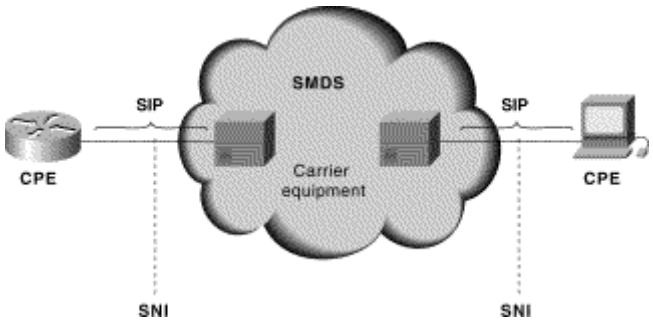


Figure 8.3: SIP Provides Connectionless Service Between the CPE and Carrier Equipment

SIP Level 3 begins operation when user information is passed to it in the form of SMDS service data units (SDUs). SMDS SDUs then are encapsulated in a SIP Level 3 header and trailer. The resulting frame is called a Level 3 protocol data unit (PDU). SIP Level 3 PDUs then are passed to SIP Level 2.

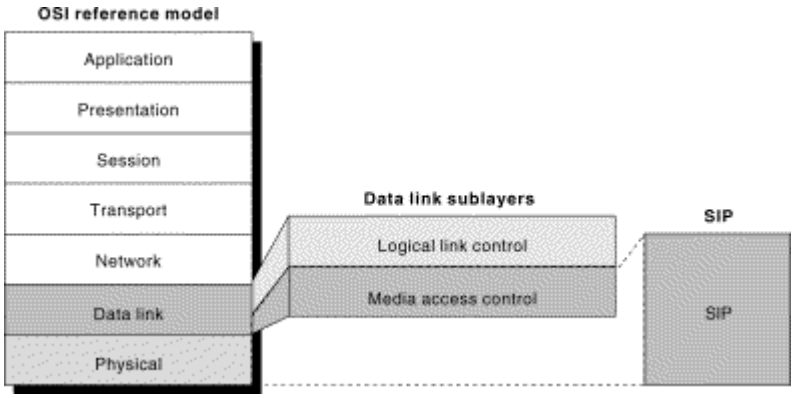


Figure 8.4 : SIP Provides Services Associated with the Physical and Data Link Layers of the OSI Model

SIP Level 2, which operates at the Media Access Control (MAC) sublayer of the data link layer, begins operating when it receives SIP Level 3 PDUs. The PDUs then are segmented into uniformly sized (53-octet) Level 2 PDUs, called cells. The cells are passed to SIP Level 1 for placement on the physical medium.

SIP Level 1 operates at the physical layer and provides the physical-link protocol that operates at DS-1 or DS-3 rates between CPE devices and the network. SIP Level 1 consists of the transmission system and Physical Layer Convergency Protocol (PLCP) sublayers. The transmission system sublayer defines the characteristics and method of attachment to a DS-1 or DS-3 transmission link. The PLCP specifies how SIP Level 2 cells are to be arranged relative to the DS-1 or DS-3 frame. PLCP also defines other management information.

Distributed Queue Dual Bus

The *Distributed Queue Dual Bus (DQDB)* is a data link layer communication protocol

designed for use in metropolitan-area networks (MANs). DQDB specifies a network topology composed of two unidirectional logical buses that interconnect multiple systems. It is defined in the IEEE 802.6 DQDB standard.

An access DQDB describes just the operation of the DQDB protocol (in SMDS, SIP) across a user-network interface (in SMDS, across the SNI). Such operation is distinguished from the operation of a DQDB protocol in any other environment (for example, between carrier equipment within the SMDS PDN).

The access DQDB is composed of the basic SMDS network components:

- **Carrier equipment—**A switch in the SMDS network operates as one station on the bus.
- **CPE—** One or more CPE devices operate as stations on the bus.
- **SNI**— the SNI acts as the interface between the CPE and the carrier equipment.

Figure 8.5 depicts a basic access DQDB, with two CPE devices and one switch (carrier equipment) attached to the dual bus.

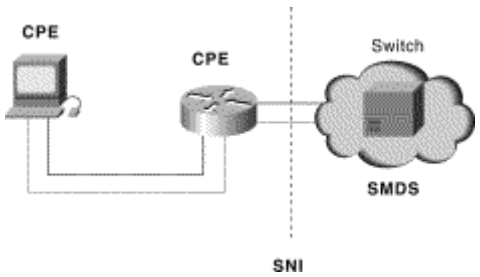


Figure 8.5: A Basic Access DQDB May Consist of an End Node, a Router, and a Switch

An SMDS access DQDB typically is arranged in a single-CPE configuration or a multi-CPE configuration.

A single-CPE access DQDB configuration consists of one switch in the carrier SMDS network and one CPE station at the subscriber site. Single-CPE DQDB configurations create a two-node DQDB sub network. Communication occurs only between the switch and the one CPE device across the SNI. No contention is on the bus because no other CPE devices attempt to access it.

A multi-CPE configuration consists of one switch in the carrier SMDS network and a number of interconnected CPE devices at the subscriber site (all belonging to the same subscriber). In multi-CPE configurations, local communication between CPE devices is possible. Some local communication will be visible to the switch serving the SNI, and some will not.

Contention for the bus by multiple devices requires the use of the DQDB distributed queuing algorithm, which makes implementing a multi-CPE configuration more complicated than implementing a single-CPE configuration.

SMDS Access Classes

SMDS access classes enable SMDS networks to accommodate a broad range of traffic requirements and equipment capabilities. Access classes constrain CPE devices to a sustained or average rate of data transfer by establishing a maximum sustained information transfer rate and a maximum allowed degree of traffic burstiness. (Burstiness, in this context, is the propensity of a network to experience sudden increases in bandwidth demand.) SMDS access classes sometimes are implemented using a credit-management scheme. In this case, a credit-management algorithm creates and tracks a credit balance for each customer interface. As packets are sent into the network, the credit balance is

decremented. New credits are allocated periodically, up to an established maximum. Credit management is used only on DS-3-rate SMDS interfaces, not on DS-1-rate interfaces.

Five access classes are supported for DS-3-rate access (corresponding to sustained information rates). Data rates supported are 4, 10, 16, 25, and 34 Mbps.

SMDS Addressing Overview

SMDS *protocol data units (PDUs)* carry both a source and a destination address. SMDS addresses are 10-digit values resembling conventional telephone numbers.

The SMDS addressing implementation offers group addressing and security features.

SMDS group addresses allow a single address to refer to multiple CPE stations, which specify the group address in the Destination Address field of the PDU. The network makes multiple copies of the PDU, which are delivered to all members of the group. Group addresses reduce the amount of network resources required for distributing routing information, resolving addresses, and dynamically discovering network resources. SMDS group addressing is analogous to multicasting on LANs.

SMDS implements two security features: source address validation and address screening. *Source address validation* ensures that the PDU source address is legitimately assigned to the SNI from which it originated. Source address validation prevents address spoofing, in which illegal traffic assumes the source address of a legitimate device. *Address screening* allows a subscriber to establish a private virtual network that excludes unwanted traffic. If an address is disallowed, the data unit is not delivered.

8.4 SMDS PACKET FORMAT

SMDS Reference: SIP Level 3 PDU Format

Figure 8.6 illustrates the format of the SMDS Interface Protocol (SIP) Level 3 protocol data unit (PDU).

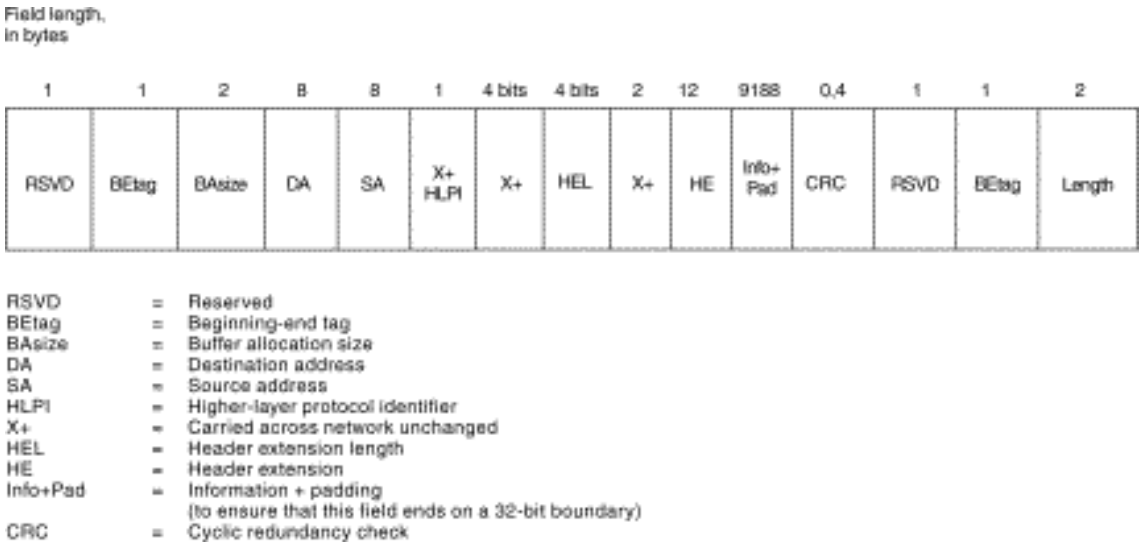


Figure 8.6: A SIP Level 3 Protocol Data Unit Consists of 15 Fields

The following descriptions briefly summarize the function of the SIP Level 3 PDU fields illustrated in Figure 8.6.

- **X+—**ensures that the SIP PDU format aligns with the DQDB protocol format. SMDS does not process or change the values in these fields, which may be used by systems connected to the SMDS network.

- **RSVD**—Consists of zeros.

- **BEtag**—Forms an association between the first and last segments of a segmented SIP Level 3 PDU. Both fields contain identical values and are used to detect a condition in which the last segment of one PDU and the first segment of the next PDU are both lost, which results in the receipt of an invalid Level 3 PDU.
- **BAsize**—Contains the buffer allocation size.
- **Destination address (DA)**—Consists of two parts:
    - **Address type**—Occupies the 4 most significant bits of the field. The Address Type can be either 1100 or 1110. The former indicates a 60-bit individual address, while the latter indicates a 60-bit group address.
    - **Address**—Gives the individual or group SMDS address for the destination. SMDS address formats are consistent with the North American Numbering Plan (NANP).

  The 4 most significant bits of the Destination Address subfield contain the value 0001 (the internationally defined country code for North America). The next 40 bits contain the binary-encoded value of the 10-digit SMDS address. The final 16 (least significant) bits are populated with ones for padding.

- **Source address (SA)**—Consists of two parts:
    - **Address type**—Occupies the 4 most significant bits of the field. The Source Address Type field can indicate only an individual address.
    - **Address**—Occupies the individual SMDS address of the source. This field follows the same format as the Address subfield of the Destination Address field.

- **Higher layer protocol identifier (HLPI)**—Indicates the type of protocol encapsulated in the Information field. The value is not important to SMDS, but it can be used by certain systems connected to the network.
- **Header extension length (HEL)**—Indicates the number of 32-bit words in the Header Extension (HE) field. Currently, the field size for SMDS is fixed at 12 bytes. (Thus, the HEL value is always 0011.)
- **Header extension (HE)**—Contains the SMDS version number. This field also conveys the carrier-selection value, which is used to select the particular interexchange carrier to carry SMDS traffic from one local carrier network to another.
- **Information and Padding (Info + Pad)**—Contains an encapsulated SMDS service data unit (SDU) and padding that ensures that the field ends on a 32-bit boundary.
- **Cyclic redundancy check (CRC)**—Contains a value used for error checking.
- **Length**—Indicates the length of the PDU.

SMDS Reference: SIP Level 2 Cell Format

Figure 8.7 illustrates the format of the SMDS Interface Protocol (SIP) Level 2 cell format.

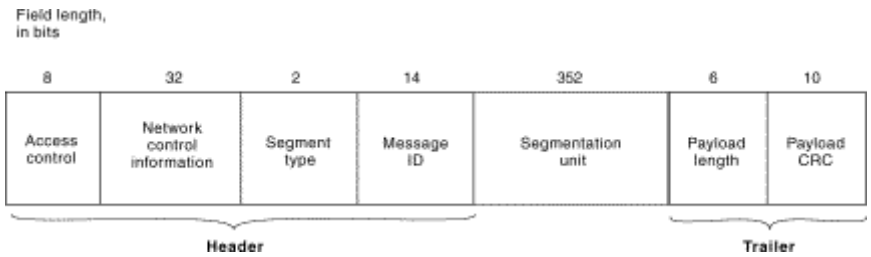


Figure 8.7: Seven Fields Comprise the SMDS SIP Level 2 Cell

The following descriptions briefly summarize the functions of the SIP Level 2 PDU fields illustrated in above figure:

- **Access control**—Contains different values, depending on the direction of information flow. If the cell was sent from a switch to a CPE device, only the indication of whether the Level 3 protocol data unit (PDU) contains information is important. If the cell was sent from a CPE device to a switch, and if the CPE configuration is multi-CPE, this field can carry request bits that indicate bids for cells on the bus going from the switch to the CPE device.
- **Network control information**—Contains a value indicating whether the PDU contains information.
- **Segment type**—Indicates whether the cell is the first, the last, or a middle cell from a segmented Level 3 PDU. Four possible segment type values exist:
    - **00**—Continuation of message
    - **01**—End of message
    - **10**—Beginning of message
    - **11**—Single-segment message
- **Message ID**—Associates Level 2 cells with a Level 3 PDU. The message ID is the same for all the segments of a given Level 3 PDU. In a multi-CPE configuration, Level 3 PDUs originating from different CPE devices must have a different message ID. This allows the SMDS network receiving interleaved cells from different Level 3 PDUs to associate each Level 2 cell with the correct Level 3 PDU.
- **Segmentation unit**—contains the data portion of the cell. If the Level 2 cell is empty, this field is populated with zeros.
- **Payload length**—indicates how many bytes of a Level 3 PDU actually are contained in the Segmentation Unit field. If the Level 2 cell is empty, this field is populated with zeros.

- **Payload cyclic redundancy check (CRC)**—Contains a CRC value used to detect errors in the following fields:
    - Segment Type
    - Message ID
    - Segmentation Unit
    - Payload Length
    - Payload CRC

  The Payload CRC value does not cover the Access Control or the Network Control Information fields.

8.5 SUMMARY

SMDS is a high-speed, packet-switched, datagram-based WAN networking technology used for communication over public data networks (PDNs). SMDS can use fiber- or copper-based media. It supports speeds of 1.544 Mbps over DS-1 transmission facilities, or 44.736 Mbps over DS-3 transmission facilities.

The following devices comprise SMDS networks:

- Customer premises equipment (CPE)
- Carrier equipment
- Subscriber network interface (SNI)

The SNI is the interface between the CPE and carrier equipment; it transparently enables data transmission between the two networks.

- SMDS uses SIP to communicate between CPE and the carrier site using the DQDB standard for cell relay across MANs.

- SIP consist of the following three levels:
    - SIP Level 3, which operates at the MAC sublayer of the data link layer of the OSI reference model
    - SIP Level 2, which also operates at the MAC sublayer of the data link layer of the OSI reference model
    - SIP Level 1, which operates at the physical layer of the OSI reference model
- SMDS PDUs carry both a source and a destination address, and offer both group addressing and security features.

8.6 REVIEW QUESTIONS

Q1. What is SMDS? Define what data services it provides.

Q2. Define SMDS protocol interface?

Q3. Define Packet format of each level of SMDS Interface Protocol?

8.7 KEYWORD

- SMDS- Switched Multimegabit data services.
- DQDB- Distributed Queue Dual Bus
- SIP- *SMDS Interface Protocol*
- PDU- protocol data unit
- Customer premises equipment (CPE)
- Subscriber network interface (SNI)
- Payload cyclic redundancy check **(CRC)**—Contains a CRC value used to detect
- Higher layer protocol identifier (HLPI)—
- Carrier equipment—A switch in the SMDS network operates as one station on the bus.

- CPE— one or more CPE devices operate as stations on the bus.
- SNI—The SNI acts as the interface between the CPE and the carrier equipment
- Public data networks (PDNs).

8.8 FURTHER READING

- McDysan, David E.; Darren L. Spohn (1999). *ATM Theory and Applications*. Montreal: McGraw-Hill.
- Bellcore. *Generic System Requirements in Support of a Switched Multi-Megabit Data Service.* Technical Advisory, TA-TSY-000772; October 1989.
- Bellcore. *Local Access System Generic Requirements, Objectives, and Interface Support of Switched Multi-Megabit Data Service.* Technical Advisory TA-TSY-000773, Issue 1; December 1985.
- Bellcore. *Switched Mutli-Megabit Data Service (SMDS) Operations Technology Network Element Generic Requirements.* Technical Advisory TA-TSY-000774.
- Telcordia. *Generic Requirements for SONET.*

LESSON 9 FRAME RELAY

9.1 OBJECTIVES

The objective of this chapter is to introduce the reader about Frame Relay. Chapter includes the packet format, network implementation of frame relay and architecture of frame relay.

9.2 INTRODUCTION

Frame Relay is an example of a packet-switched technology. Packet-switched networks enable end stations to dynamically share the network medium and the available bandwidth. The following two techniques are used in packet-switching technology:

- Variable-length packets
- Statistical multiplexing

Variable-length packets are used for more efficient and flexible data transfers. These packets are switched between the various segments in the network until the destination is reached.

Statistical multiplexing techniques control network access in a packet-switched network. The advantage of this technique is that it accommodates more flexibility and more efficient use of bandwidth. Most of today's popular LANs, such as Ethernet and Token Ring, are packet-switched networks.

Frame Relay often is described as a streamlined version of X.25, offering fewer of the robust capabilities, such as windowing and retransmission of last data that are offered in X.25. This is because Frame Relay typically operates over WAN facilities that offer more reliable connection services and a higher degree of reliability than the facilities available during the late 1970s and early 1980s that served as the common platforms for X.25 WANs. As mentioned earlier, Frame Relay is strictly a Layer 2 protocol suite, whereas X.25 provides services at Layer 3 (the network layer) as well. This enables Frame Relay to offer higher performance and greater transmission efficiency than X.25, and makes Frame Relay suitable for current WAN applications, such as LAN interconnection.

Frame Relay Standardization

Initial proposals for the standardization of Frame Relay were presented to the Consultative Committee on International Telephone and Telegraph (CCITT) in 1984. Because of lack of interoperability and lack of complete standardization, however, Frame Relay did not experience significant deployment during the late 1980s.

A major development in Frame Relay's history occurred in 1990 when Cisco, Digital Equipment Corporation (DEC), Northern Telecom, and StrataCom formed a consortium to focus on Frame Relay technology development. This consortium developed a specification that conformed to the basic Frame Relay protocol that was being discussed in CCITT, but it

extended the protocol with features that provide additional capabilities for complex internetworking environments. These Frame Relay extensions are referred to collectively as the Local Management Interface (LMI).

Since the consortium's specification was developed and published, many vendors have announced their support of this extended Frame Relay definition. ANSI and CCITT have subsequently standardized their own variations of the original LMI specification, and these standardized specifications now are more commonly used than the original version.

Internationally, Frame Relay was standardized by the International Telecommunication Union-Telecommunications Standards Section (ITU-T). In the United States, Frame Relay is an American National Standards Institute (ANSI) standard.

9.3 FRAME RELAY ARCHITECTURE

Frame Relay Devices

Devices attached to a Frame Relay WAN fall into the following two general categories:

- Data terminal equipment (DTE)
- Data circuit-terminating equipment (DCE)

DTEs generally are considered to be terminating equipment for a specific network and typically are located on the premises of a customer. In fact, they may be owned by the customer. Examples of DTE devices are terminals, personal computers, routers, and bridges.

DCEs are carrier-owned internetworking devices. The purpose of DCE equipment is to provide clocking and switching services in a network, which are the devices that actually transmit data through the WAN. In most cases, these are packet switches.

Figure 9.1: DCEs Generally Reside Within Carrier-Operated WANs shows the relationship between the two categories of devices.

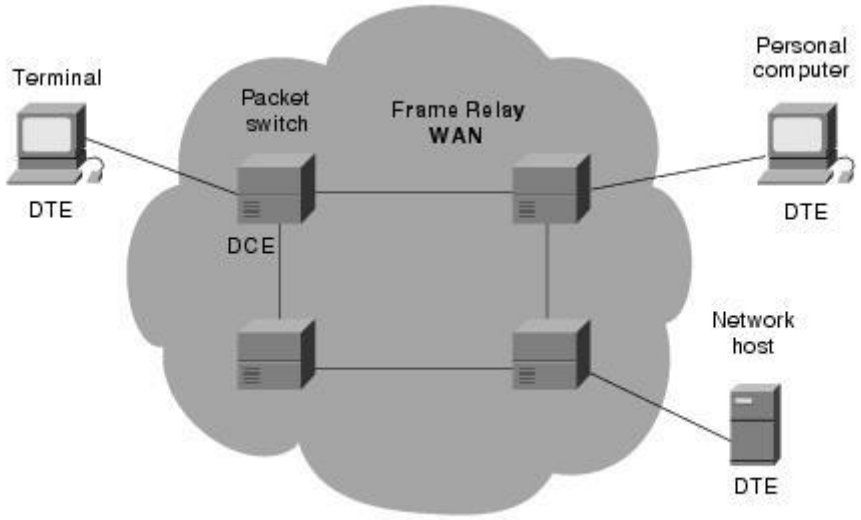


Figure 9.1: DCEs Generally Reside Within Carrier-Operated WANs

The connection between a DTE device and a DCE device consists of both a physical layer component and a link layer component. The physical component defines the mechanical, electrical, functional, and procedural specifications for the connection between the devices. One of the most commonly used physical layer interface specifications is the recommended standard (RS)-232 specifications. The link layer component defines the protocol that establishes the connection between the DTE device, such as a router, and the DCE device, such as a switch. This article examines a commonly utilized protocol specification used in WAN networking: the Frame Relay protocol.

Frame Relay Virtual Circuits

Frame Relay provides connection-oriented data link layer communication. This means that a defined communication exists between each pair of devices and that these connections are associated with a connection identifier. This service is implemented by using a Frame Relay virtual circuit, which is a logical connection created between two data terminal equipment (DTE) devices across a Frame Relay packet-switched network (PSN).

Virtual circuits provide a bidirectional communication path from one DTE device to another and are uniquely identified by a data-link connection identifier (DLCI). A number of virtual circuits can be multiplexed into a single physical circuit for transmission across the network. This capability often can reduce the equipment and network complexity required to connect multiple DTE devices.

A virtual circuit can pass through any number of intermediate DCE devices (switches) located within the Frame Relay PSN.

Frame Relay virtual circuits fall into two categories: switched virtual circuits (SVCs) and permanent virtual circuits (PVCs).

Switched Virtual Circuits

Switched virtual circuits (SVCs) are temporary connections used in situations requiring only sporadic data transfer between DTE devices across the Frame Relay network. A communication session across an SVC consists of the following four operational states:

- **Call setup** - The virtual circuit between two Frame Relay DTE devices is established.
- **Data transfer** - Data is transmitted between the DTE devices over the virtual circuit.
- **Idle** - The connection between DTE devices is still active, but no data is transferred. If an SVC remains in an idle state for a defined period of time, the call can be terminated.
- **Call termination** - The virtual circuit between DTE devices is terminated.

After the virtual circuit is terminated, the DTE devices must establish a new SVC if there is additional data to be exchanged. It is expected that SVCs will be established, maintained, and terminated using the same signaling protocols used in ISDN.

Few manufacturers of Frame Relay DCE equipment support switched virtual circuit connections. Therefore, their actual deployment is minimal in today's Frame Relay networks.

Previously not widely supported by Frame Relay equipment, SVCs are now the norm. Companies have found that SVCs save money in the end because the circuit is not open all the time.

Permanent Virtual Circuits

Permanent virtual circuits (PVCs) are permanently established connections that are used for frequent and consistent data transfers between DTE devices across the Frame Relay network. Communication across PVC does not require the call setup and termination states that are used with SVCs. PVCs always operate in one of the following two operational states:

- **Data transfer** - Data is transmitted between the DTE devices over the virtual circuit.
- **Idle** - The connection between DTE devices is active, but no data is transferred. Unlike SVCs, PVCs will not be terminated under any circumstances when in an idle state.

DTE devices can begin transferring data whenever they are ready because the circuit is permanently established.

Data-Link Connection Identifier

Frame Relay virtual circuits are identified by data-link connection identifiers (DLCIs). DLCI values typically are assigned by the Frame Relay service provider (for example, the telephone company).

Frame Relay DLCIs have local significance, which means that their values are unique in the LAN, but not necessarily in the Frame Relay WAN.

Figure illustrates how two different DTE devices can be assigned the same DLCI value within one Frame Relay WAN.

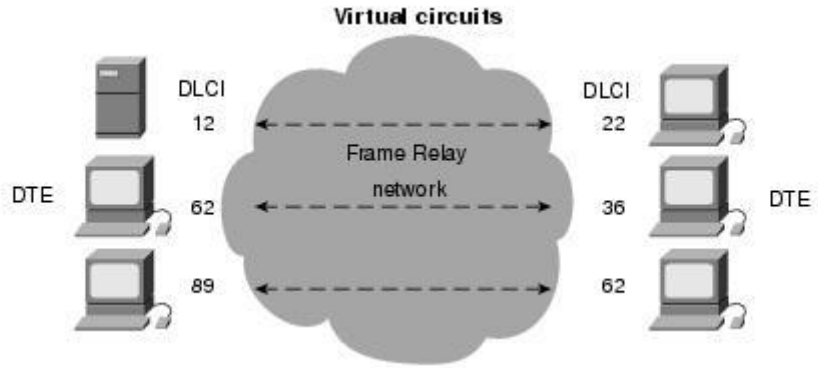


Figure 9.2: A Single Frame Relay Virtual Circuit Can Be Assigned Different DLCIs on Each End of a VC

Congestion-Control Mechanisms

Frame Relay reduces network overhead by implementing simple congestion-notification mechanisms rather than explicit, per-virtual-circuit flow control. Frame Relay typically is implemented on reliable network media, so data integrity is not sacrificed because flow control can be left to higher-layer protocols. Frame Relay implements two congestion-notification mechanisms:

- Forward-explicit congestion notification (FECN)
- Backward-explicit congestion notification (BECN)

FECN and BECN each is controlled by a single bit contained in the Frame Relay frame header. The Frame Relay frame header also contains a Discard Eligibility (DE) bit, which is used to identify less important traffic that can be dropped during periods of congestion.

The FECN bit is part of the Address field in the Frame Relay frame header. The FECN mechanism is initiated when a DTE device sends Frame Relay frames into the network. If the

network is congested, DCE devices (switches) set the value of the frames' FECN bit to 1. When the frames reach the destination DTE device, the Address field (with the FECN bit set) indicates that the frame experienced congestion in the path from source to destination. The DTE device can relay this information to a higher-layer protocol for processing. Depending on the implementation, flow control may be initiated, or the indication may be ignored.

The BECN bit is part of the Address field in the Frame Relay frame header. DCE devices set the value of the BECN bit to 1 in frames traveling in the opposite direction of frames with their FECN bit set. This informs the receiving DTE device that a particular path through the network is congested. The DTE device then can relay this information to a higher-layer protocol for processing. Depending on the implementation, flow-control may be initiated, or the indication may be ignored.

Frame Relay Discard Eligibility

The Discard Eligibility (DE) bit is used to indicate that a frame has lower importance than other frames. The DE bit is part of the Address field in the Frame Relay frame header.

DTE devices can set the value of the DE bit of a frame to 1 to indicate that the frame has lower importance than other frames. When the network becomes congested, DCE devices will discard frames with the DE bit set before discarding those that do not. This reduces the likelihood of critical data being dropped by Frame Relay DCE devices during periods of congestion.

Frame Relay Error Checking

Frame Relay uses a common error-checking mechanism known as the cyclic redundancy check (CRC). The CRC compares two calculated values to determine whether errors occurred during the transmission from source to destination. Frame Relay reduces network overhead by implementing error checking rather than error correction. Frame Relay typically is

implemented on reliable network media, so data integrity is not sacrificed because error correction can be left to higher-layer protocols running on top of Frame Relay.

Frame Relay Local Management Interface

The Local Management Interface (LMI) is a set of enhancements to the basic Frame Relay specification. The LMI was developed in 1990 by Cisco Systems, StrataCom, Northern Telecom, and Digital Equipment Corporation. It offers a number of features (called extensions) for managing complex internetworks. Key Frame Relay LMI extensions include global addressing, virtual circuit status messages, and multicasting.

The LMI global addressing extension gives Frame Relay data-link connection identifier (DLCI) values global rather than local significance. DLCI values become DTE addresses that are unique in the Frame Relay WAN. The global addressing extension adds functionality and manageability to Frame Relay internetworks. Individual network interfaces and the end nodes attached to them, for example, can be identified by using standard address-resolution and discovery techniques. In addition, the entire Frame Relay network appears to be a typical LAN to routers on its periphery.

LMI virtual circuit status messages provide communication and synchronization between Frame Relay DTE and DCE devices. These messages are used to periodically report on the status of PVCs, which prevents data from being sent into black holes (that is, over PVCs that no longer exist).

The LMI multicasting extension allows multicast groups to be assigned. Multicasting saves bandwidth by allowing routing updates and address-resolution messages to be sent only to specific groups of routers. The extension also transmits reports on the status of multicast groups in update messages.

9.4 Frame Relay Frame Formats

To understand much of the functionality of Frame Relay, it is helpful to understand the structure of the Frame Relay frame. Figure 9.3 depicts the basic format of the Frame Relay frame, and Figure 9.4 illustrates the LMI version of the Frame Relay frame.

Flags indicate the beginning and end of the frame. Three primary components make up the Frame Relay frame: the header and address area, the user-data portion, and the frame check sequence (FCS). The address area, which is 2 bytes in length, is comprised of 10 bits representing the actual circuit identifier and 6 bits of fields related to congestion management. This identifier commonly is referred to as the data-link connection identifier (DLCI). Each of these is discussed in the descriptions that follow.

Standard Frame Relay Frame

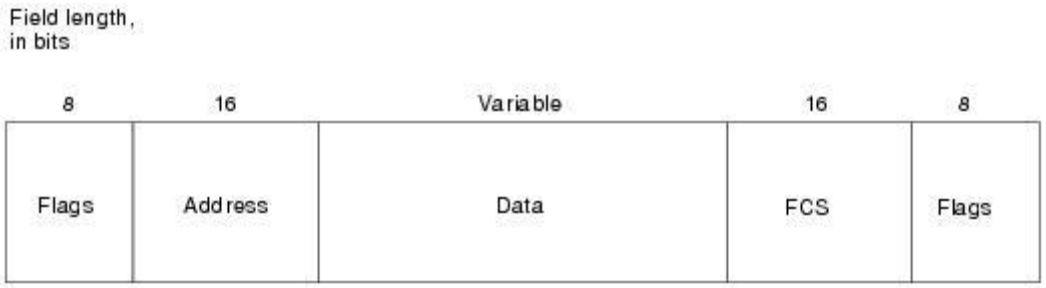


Figure 9.3: Five Fields Comprise the Frame Relay Frame

The following descriptions summarize the basic Frame Relay frame fields illustrated in Figure 9.3.

- **Flags** - Delimits the beginning and end of the frame. The value of this field is always the same and is represented either as the hexadecimal number 7E or as the binary number 01111110.

- **Address** - Contains the following information:
  - **DLCI** - The 10-bit DLCI is the essence of the Frame Relay header. This value represents the virtual connection between the DTE device and the switch. Each virtual connection that is multiplexed onto the physical channel will be represented by a unique DLCI. The DLCI values have local significance only, which means that they are unique only to the physical channel on which they reside. Therefore, devices at opposite ends of a connection can use different DLCI values to refer to the same virtual connection.
  - **Extended Address (EA)** - The EA is used to indicate whether the byte in which the EA value is 1 is the last addressing field. If the value is 1, then the current byte is determined to be the last DLCI octet. Although current Frame Relay implementations all use a two-octet DLCI, this capability does allow longer DLCIs to be used in the future. The eighth bit of each byte of the Address field is used to indicate the EA.
  - **C/R** - The C/R is the bit that follows the most significant DLCI byte in the Address field. The C/R bit is not currently defined.
  - **Congestion Control** - This consists of the 3 bits that control the Frame Relay congestion-notification mechanisms. These are the FECN, BECN, and DE bits, which are the last 3 bits in the Address field.

    Forward-explicit congestion notification (FECN) is a single-bit field that can be set to a value of 1 by a switch to indicate to an end DTE device, such as a router, that congestion was experienced in the direction of the frame transmission from source to destination. The primary benefit of the use of the FECN and BECN fields is the capability of higher-layer protocols to react intelligently to these congestion indicators. Today, DECnet and OSI are the only higher-layer protocols that implement these capabilities.

Backward-explicit congestion notification (BECN) is a single-bit field that, when set to a value of 1 by a switch, indicates that congestion was experienced in the network in the direction opposite of the frame transmission from source to destination.

Discard eligibility (DE) is set by the DTE device, such as a router, to indicate that the marked frame is of lesser importance relative to other frames being transmitted. Frames that are marked as "discard eligible" should be discarded before other frames in a congested network. This allows for a basic prioritization mechanism in Frame Relay networks.

- **Data** - Contains encapsulated upper-layer data. Each frame in this variable-length field includes a user data or payload field that will vary in length up to 16,000 octets. This field serves to transport the higher-layer protocol packet (PDU) through a Frame Relay network.
- **Frame Check Sequence** - Ensures the integrity of transmitted data. This value is computed by the source device and verified by the receiver to ensure integrity of transmission.

LMI Frame Format

Frame Relay frames that conform to the LMI specifications consist of the fields illustrated in

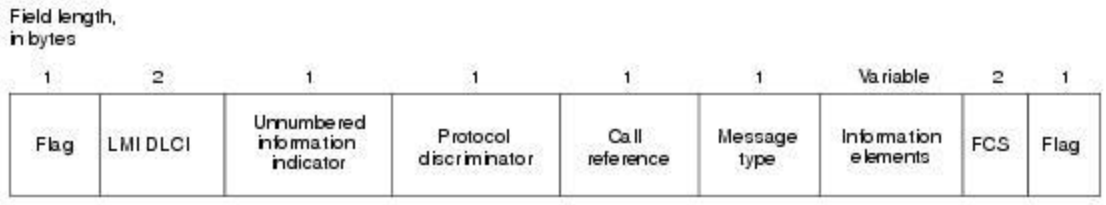


Figure 9.4: Nine Fields Comprise the Frame Relay That Conforms to the LMI Format

The following descriptions summarize the fields illustrated in Figure 9.4.

- **Flag** - Delimits the beginning and end of the frame.
- **LMI DLCI** - Identifies the frame as an LMI frame instead of a basic Frame Relay frame. The LMI-specific DLCI value defined in the LMI consortium specification is DLCI = 1023.
- **Unnumbered Information Indicator** - Sets the poll/final bit to zero.
- **Protocol Discriminator** - Always contains a value indicating that the frame is an LMI frame.
- **Call Reference** - Always contains zeros. This field currently is not used for any purpose.
- **Message Type** - Labels the frame as one of the following message types:
  - **Status-inquiry message** - Allows a user device to inquire about the status of the network.
  - **Status message** - Responds to status-inquiry messages. Status messages include keepalives and PVC status messages.
- **Information Elements** - Contains a variable number of individual information elements (IEs). IEs consist of the following fields:
  - **IE Identifier** - Uniquely identifies the IE.
  - **IE Length** - Indicates the length of the IE.
  - **Data** - Consists of 1 or more bytes containing encapsulated upper-layer data.
- **Frame Check Sequence (FCS)** - Ensures the integrity of transmitted data.

9.5 FRAME RELAY LAYERS

Frame relay reduces the complexity of the physical network without disrupting higher-level network functions. Frame Relay functions using only the bottom two layers of the OSI

model, as compared to X.25 which includes the Network layer (see Figure 9.5). By reducing the amount of processing required, and by efficiently using high-speed digital transmission lines, frame relay can improve performance and response times for most applications.

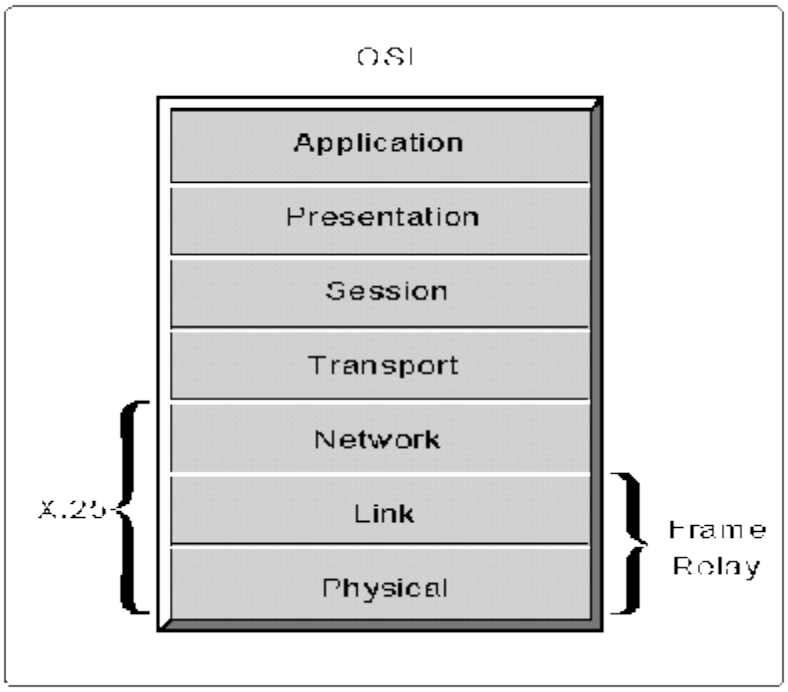


Figure 9.5: Frame relay uses only the bottom two layers of the OSI model.

Increased Interoperability via International Standards Frame relay's simplified link layer protocol can be implemented over existing technology. Access devices often require only software changes or simple hardware modifications to support the interface standard. Existing packet switching equipment and T1/E1 multiplexers often can be upgraded to support frame relay over existing backbone networks.

Frame relay is an accepted interface standard that vendors and service providers are adhering to and implementing. There is exceptionally good interoperability between the various standards. Further, most equipment vendors and service providers have pledged their support for frame relay development and standards. The simplicity of the frame relay protocol

accommodates quick and easy interoperability testing procedures between devices from different vendors. This interoperability testing is currently in progress among vendors, as are certification processes for carriers providing frame relay services.

9.6 FRAME RELAY NETWORK IMPLEMENTATION

A common private Frame Relay network implementation is to equip a T1 multiplexer with both Frame Relay and non-Frame Relay interfaces. Frame Relay traffic is forwarded out the Frame Relay interface and onto the data network. Non-Frame Relay traffic is forwarded to the appropriate application or service, such as a private branch exchange (PBX) for telephone service or to a video-teleconferencing application.

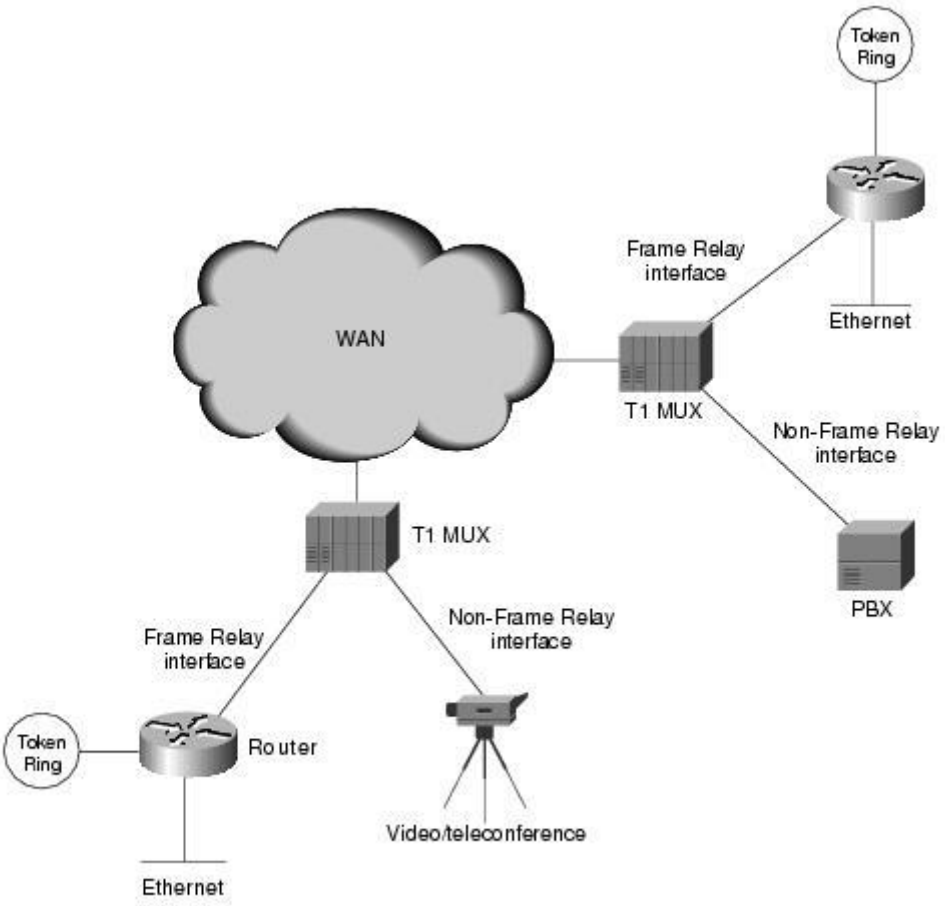


Figure 9.6: A simple Frame Relay network

A typical Frame Relay network consists of a number of DTE devices, such as routers, connected to remote ports on multiplexer equipment via traditional point-to-point services such as T1, fractional T1, or 56-Kb circuits. An example of a simple Frame Relay network is shown in figure 9.6.

The majority of Frame Relay networks deployed today are provisioned by service providers that intend to offer transmission services to customers. This is often referred to as a public Frame Relay service. Frame Relay is implemented in both public carrier-provided networks and in private enterprise networks. The following section examines the two methodologies for deploying Frame Relay.

Public Carrier-Provided Networks

In public carrier-provided Frame Relay networks, the Frame Relay switching equipment is located in the central offices of a telecommunications carrier. Subscribers are charged based on their network use but are relieved from administering and maintaining the Frame Relay network equipment and service.

Generally, the DCE equipment also is owned by the telecommunications provider. DTE equipment either will be customer-owned or perhaps will be owned by the telecommunications provider as a service to the customer.

The majority of today's Frame Relay networks are public carrier-provided networks.

Private Enterprise Networks

More frequently, organizations worldwide are deploying private Frame Relay networks. In private Frame Relay networks, the administration and maintenance of the network are the responsibilities of the enterprise (a private company). All the equipment, including the switching equipment, is owned by the customer.

9.7 SUMMARY

Frame Relay is a networking protocol that works at the bottom two levels of the OSI reference model: the physical and data link layers. It is an example of packet-switching technology, which enables end stations to dynamically share network resources.

Frame Relay devices fall into the following two general categories:

- Data terminal equipment (DTEs), which include terminals, personal computers, routers, and bridges
- Data circuit-terminating equipment (DCEs), which transmit the data through the network and are often carrier-owned devices (although, increasingly, enterprises are buying their own DCEs and implementing them in their networks)

Frame Relay networks transfer data using one of the following two connection types:

- Switched virtual circuits (SVCs), which are temporary connections that are created for each data transfer and then are terminated when the data transfer is complete (not a widely used connection)
- Permanent virtual circuits (PVCs), which are permanent connections

The DLCI is a value assigned to each virtual circuit and DTE device connection point in the Frame Relay WAN. Two different connections can be assigned the same value within the same Frame Relay WAN-one on each side of the virtual connection.

In 1990, Cisco Systems, StrataCom, Northern Telecom, and Digital Equipment Corporation developed a set of Frame Relay enhancements called the Local Management Interface (LMI). The LMI enhancements offer a number of features (referred to as extensions) for managing complex internetworks, including the following:

- Global addressing
- Virtual circuit status messages
- Multicasting

9.8 KEYWORDS

- Consultative Committee on International Telephone and Telegraph (CCITT)
- Local Management Interface (LMI).
- International Telecommunication Union-Telecommunications Standards Section (ITU-T).
- Data terminal equipment (DTEs)
- Backward-explicit congestion notification (BECN)
- Forward-eplicit congestion notification(FECN)
- Discard Eligibility(DS)

9.9 REVIEW QUESTIONS

Q1. What kind of technology is Frame Relay?

Q2. Name the two kinds of packet-switching techniques discussed in this article, and briefly describe each.

Q3. Describe the difference between SVCs and PVCs.

Q4. What is the data-link connection identifier (DLCI)?

Q5. Describe how LMI Frame Relay differs from basic Frame Relay.

9.10 FURTHER READINGS

- Frame relay networks: specifications and implementations, by Uyless Black. (McGraw-Hill, 1994, ISBN 0-07-005558-0).
- A comprehensive and up-to-date treatment of everything you wanted to know about frame relay networks.
- ISDN and Broadband ISDN with Frame Relay and ATM, Third Edition, by William Stallings.
- ISDN, SONET, Frame Relay and ATM. However, it does not cover Frame Relay Forum and ATM Forum agreements.
- Emerging Communications Technologies, by Uyless Black.
- SNMP, A Guide to Network Management, by Dr. Sidnie Feit (McGraw-Hill).
- The Basics Book of Frame Relay
- The Guide to Frame Relay and Packet Networking by Nathan J. Muller and Robert Davidson.

LESSON 10 DIGITAL SUBSCRIBER LINE

10.1 OBJECTIVES

The subject of interest in this chapter is the use of Digital Subscriber Line (DSL) technology to increase the rate and improve the quality of data communications over copper cable. It is an important topic both within the context of data communications today and into the future. All, or almost all, aspects of this subject will be explored. However, it seems rather forbidding just to

jump into this topic. Rather, it is more appropriate to take a step back and talk about the nature of communications first, in order to introduce some needed terminology. Such a step back will also provide us with a broader perspective on the subject of DSL technology as a transmission facilitator.

10.2 INTRODUCTION

The widespread use of the Internet and especially the World Wide Web have opened up a need for high bandwidth network services that can be brought directly to subscriber's homes. These services would provide the needed bandwidth to surf the web at lightning fast speeds and allow new technologies such as video conferencing and video on demand. Currently, Digital Subscriber Line (DSL) and Cable modem technologies look to be the most cost effective and practical methods of delivering broadband network services to the masses.

10.3 DSL TECHNOLOGIES

Digital Subscriber Line

A Digital Subscriber Line makes use of the current copper infrastructure to supply broadband services.

A DSL requires two modems, one at the phone companies end and one at the subscribers end. The use of the term modem is not entirely correct because technically a DSL modem does not do modulation /demodulation as in a modem that uses the normal telephone network. DSL's also have the added benefit of transmitting telephone services on the same set of wire as data services. DSL's common in many flavor, and are sometimes referred to as xDSL, the x standing for the specific type.

For years it has been believed that the upper limit for transmitting data on analog phone lines was 56 kb/s. This limit is set using the maximum possible bandwidth and no compression. The reason for this limit is that POTS or Plain Old Telephone Service uses the lower 4 Khz only. The limit imposed by the POTS lines does not take advantage of all the bandwidth available on copper, which is on the order of 1 Mhz. The xDSL technologies take advantage of this difference and uses the upper frequencies for data services. Previously this was not possible because of the interference that the data services would cause in the POTS band. Advances in digital signal processing have eliminated the near-end crosstalk that results from the use of the upper bandwidth for data. The new DSP technologies allow data and POTS to be transmitted on the same set of copper wires without interfering with each other. DSL technologies were initially tested for use with video on demand (VOD) and interactive television (ITV) services. Lack of a "killer application" for these services and competition from the cable TV industry in these areas forced the telephone companies to look for a different application for their technologies. With the popularity of the World Wide Web and telecommuting on the rise the DSL technologies moved to providing network and phone services to the home. Other areas where DSL technologies are targeted for are Intranet access, LAN to LAN connections, Frame Relay, ATM Network access, and leased line provisioning.

10.4 ADSL

Asymmetric Digital Subscriber Line

The most promising of the DSL technologies is ADSL or Asymmetric Digital Subscriber Line. ADSL looks to make the most impact in residential access and the SOHO (Small Office Home

Office) market. Just like the name implies ADSL is asymmetric, meaning that the downstream bandwidth is higher than the upstream bandwidth. Downstream refers to traffic in the direction towards the subscriber, and upstream refers to data sent from the subscriber back to the network. This is done because of the kinds traffic that ADSL is designed to carry. Asymmetry is used to increase the downstream bandwidth. This works because all of the downstream signals can be of the same amplitude thus eliminating crosstalk between downstream channels. Upstream signals would have to put up with more interference because the amplitude of the upstream signals would be of smaller amplitude because they are originating from different distances. The asymmetric nature of ADSL lends itself well to applications like the web and client server applications.

To achieve the asymmetry ADSL divides its bandwidth into four classes of transport.

- higher bandwidth simplex channel
- lower bandwidth duplex channel
- duplex control channel
- POTS channel

Transmission on the high bandwidth simplex channel and the lower bandwidth duplex channel do not interfere in any way with the POTS channel. So ADSL can carry both data a POTS on the same medium, which makes it ideal for residential and small office use.

ADSL bandwidth is currently standardized by ANSI (American National Standards Institute). Tables 1and 2 detail the four transport classes which are based on multiples of T-1 (1.5 Mb/s) downstream bandwidth. There are also three more classes that are based on the European E-1 (2.0 Mb/s) standard which is shown in the second chart. These classes are all based on the

maximum bandwidth available on each channel. The actual rates depend on factors such as wire gauge, local loop length, and line condition. In this case, the local loop length is the distance from the central office to the subscriber.

Table 1: ADSL Transport Classes (T-1 based multiples)

Class	1	2	3	4
Downstream simplex Channel	6.144 Mb/s	4.608 Mb/s	3.072 Mb/s	1.536 Mb/s
Upstream duplex channel	640 kb/s 576 kb/s of usable Bandwidth	608 kb/s 544 kb/s of usable bandwidth	608 kb/s 544 kb/s of usable bandwidth	176 kb/s 160 kb/s of usable bandwidth
Control channel	64 kb/s	64 kb/s	64 kb/s	16 kb/s
POTS channel	64 kb/s	64 kb/s	64 kb/s	64 kb/s

Table 2: ADSL Transport Classes (E-1 based multiples)

Class	2M1	2M2	2M3
Downstream simplex channel	6.144 Mb/s	4.096 Mb/s	2.048 Mb/s
Upstream duplex channel	640 kb/ s	608 kb/s	176 kb/s
Control channel	64 kb/s	64 kb/s	16 kb/s
POTS channel	64 kb/s	64 kb/s	64 kb/s

10.4.1 Competing Standards

One issue yet to be resolved with ADSL is the debate between CAP (Carrierless Amplitude Modulation) and DMT (Discrete Multitone) line code standards. DMT has been standardized by ANSI, but currently products using CAP have been released by various companies. Different companies support different standards and neither of them has become a de facto standard. CAP

technologies have been quicker to get to market but DMT is gaining. The main drawback with DMT is that it has been expensive to deploy up until recently. Both methods have their advantages. CAP is a single carrier modulation technique that uses three frequency ranges. CAP uses 900 Mhz for downstream data, 75 Mhz for upstream data , and 4 Khz for POTS service. CAP takes the data channels and treats them like one big pipe on which to send data. DMT is

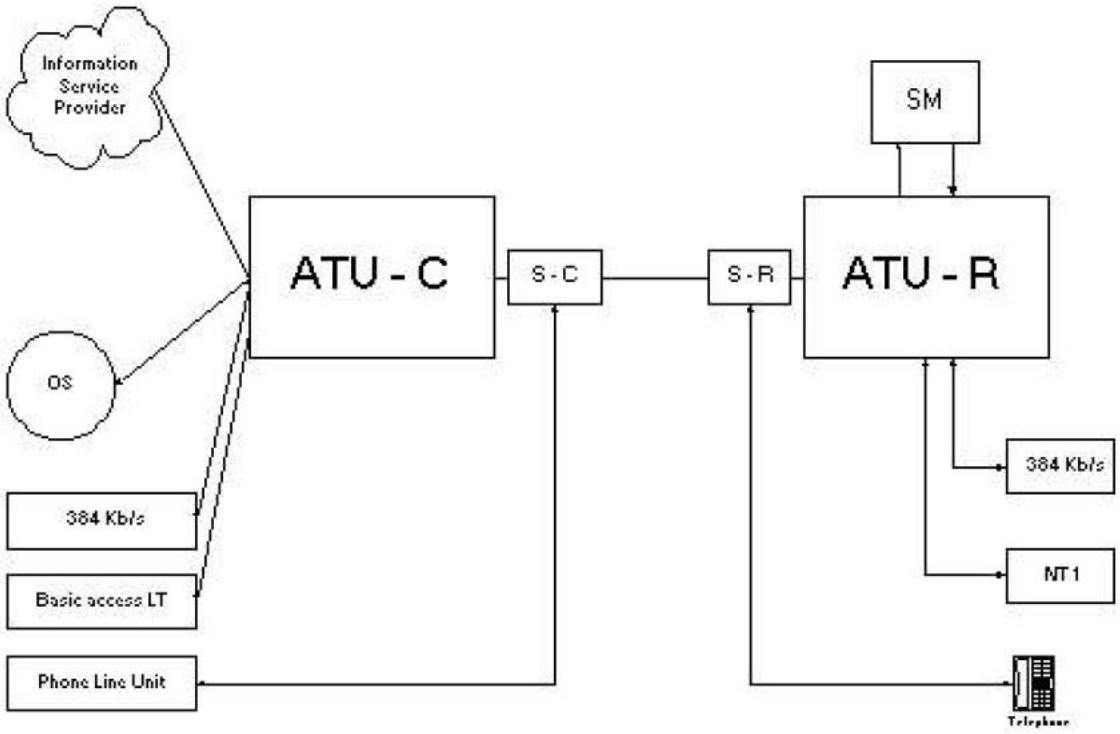


Figure 1: Basic ADSL system

different in that it breaks the data transmission channels into 256 subchannels and then selects the best ones on which to send its data. DMT fits better in to a RADSL or Rate Adaptive DSL, scheme due to the fact that it has that ability to select channels that have lower levels of interference on them. CAP generally provides 1.5 Mb/s downstream and 64 kb/s up stream. In contrast DMT transmits 6 Mb/s downstream and 640 kb/s upstream. DMT is not without disadvantages, the DMT equipment requires more power and therefore operates at higher

temperature which limits the number of ADSL / DMT modems that can be stacked together at a central office.

The basic ADSL model is given in figure 1. The -C and -R designations are given to the terminal equipment on the Central Office and the remote end respectively. The S-C and the S-R units split the POTS signal in and out of the ADSL signal.

10.4.2 Trends

Of all the DSL technologies ADSL in being tested the most right now. Widespread deployment is not far off.

To help this along companies are making complete DMT based ADSL transceivers on a chips at lower costs than the chip sets previously used. Some of those chips also incorporate Ethernet serial transceivers in them to make it easier to interface with current LAN technologies.

There is also some debate between HDSL or High-data-rate DSL and ADSL. HDSL provides a symmetric data pathway at 1.544 Mb/s which matches the speed of today's T-1's. HDSL has been around longer and is currently being used to effectively provision local access to T-1's. HDSL has it place in thatmarket while ADSL can provide a better service to homes and small business that use the web and client server technologies.

10.5 HDSL

High-data-rate Digital Subscriber Line

The most common DSL deployed today is HDSL. HDSL is mostly used to provision other services by telephone companies. HDSL symmetrically delivers 1.544 Mb/s over two sets of

copper twisted pair. Which is the same rate as a T-1 type connection. This allows telco's (short for telephone companies) to use HDSL to deliver T-1 services. HDSL's operating range is about 12,000 feet, and it is possible to extend that by using repeaters along the line to the customer. HDSL is mostly used to deploy PBX network connections , interexchange POP's (Point Of Presence), and directly connecting servers to the Internet.

10.6 SDSL

Single-line Digital Subscriber Line also known as Symmetric Digital Subscriber Line Similar to HDSL, SDSL delivers the same 1.544 Mb/s, but it does it on a single set of twisted pair of copper. This limits SDSL's reach to 10,000 feet. SDSL could take hold in niche markets like residential video conferencing or connecting LAN's over short distances.

10.7 VDSL

Very-high-rate Digital Subscriber Line

VDSL technology operates on a single set of copper twisted pair, and delivers data in the range of 13 Mb/s to 52 Mb/s. This high bandwidth does not come without a price, the range of VDSL is limited to between 1,000 and 4,500 feet. The VDSL standard is still in the works but there are already applications for the technology. One use for it is in getting high data rate services from the telephone companies central office to the subscriber via a FTTN (Fiber To The Neighborhood) network.. FTTN encompasses the Fiber to the Curb technologies and uses VDSL as the customers connection to the telephone companies fiber based network.

VDSL would be used to connect premises distribution networks to the Optical Network Unit or ONU. The optical network unit is in turn connected via fiber optical line to the telco's central office.

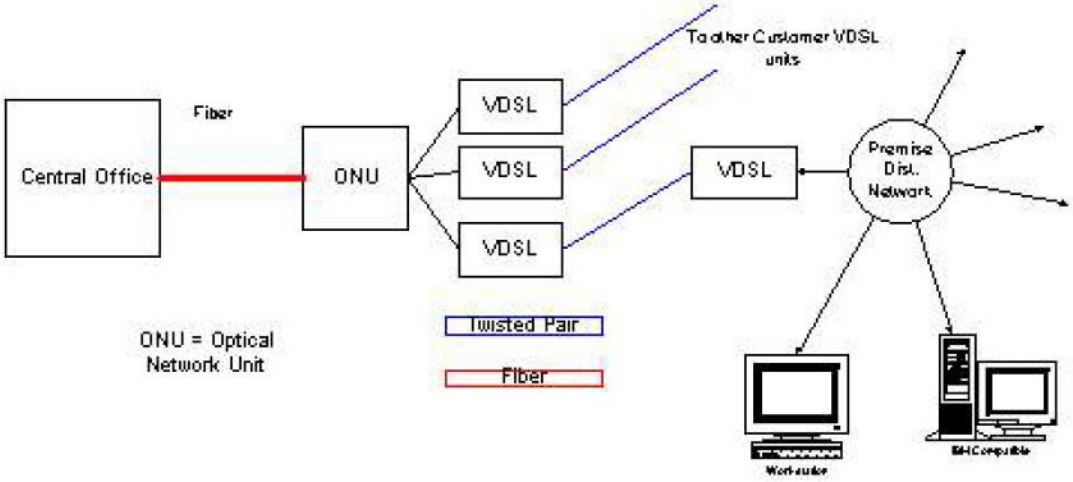


Figure 2: A typical use of VDSL

Figure 2 shows how VDSL might be used to bring data services from an ONU to the customer's premises distribution network.

Since VDSL is still in discussion right now there are no solid standards but some reachable goals have been set in an early draft by the ADSL Forum. Data rates have been projected as multiples of SONET and SDH. These speeds have been chosen because of proposed use of VDSL as a solution for delivering fiber type bandwidth to customers over copper.

Table 3: Proposed Data rates for VDSL

Data rate	Range in feet	Range in meters
12.96 - 13.8 Mb/s	4500 ft	1500 m
25.92 - 27.6 Mb/s	3000 ft	1000 m
51.84 - 55.2 Mb/s	1000 ft	300 m

10.8 RADSL

Rate Adaptive Digital Subscriber Line

RADSL is derived from ADSL technologies with some added features. RADSL automatically adjusts line speed based on the condition of the line. In areas where there is a large variance in the distance between the central office and the subscribers RADSL helps to provide a more consistent service for it's subscribers by taking the uncertainties of line conditions out of the equation when setting up a DSL connection. RADSL can adjust line speed based on the gauge of the wire, the distance between subscriber and the central office, and the condition of the line. It also takes care of fluctuations that the weather can induce into the line.

10.9 Comparison of DSL Technologies

Table 4: DSL Comparison

DSL Upstream	**Upstream** Bandwidth	Downstream Bandwidth	Range	Media	Symmetry
ADSL	16 kb/s to 640 kb/s	1.5 Mb/s to 9 Mb/s	18, 000 feet Single	twisted pair	Asymmetric
HDSL	1.544 Mb/s	1.544 Mb/s	12, 000 feet	Two twisted pairs	Symmetric
SDSL	1.544 Mb/s	1.544 Mb/s	10,000 feet Single	twisted pair	Symmetric
RADSL	varies with in ADSL range	varies with in ADSL range	18,000 feet	Single twisted pair	Asymmetric
VDSL	13 Mb/s - 52 Mb/s	1.6 Mb/s to 2.3 Mb/s	1,000 - 4,500 feet	Single twisted pair	Both

10.10 SUMMARY

Spurned by Internet growth and the convergence of television with the personal computer, high bandwidth data services are soon to in homes everywhere. DSL and cable modems look to be an economically promising alternative to today's current 56 kb/s modems that promise to bring much higher data rates to the home. Two industries race for the business of the data hunger

consumer, the telephone companies and cable companies. Cable modems look to make use of the high bandwidth coax and hybrid fiber networks already used to provided television service, while the phone companies use their existing copper and more recent transition to digital networks as a way to leverage their current infrastructure to meet the high speed demands of their customers. In the near future broadband services while may become as commonplace and necessary as today's telephone service.

10.11 KEYWORDS

ADSL - Asymmetric Digital Subscriber Line. DSL with higher bandwidth in one direction than the other.

ANSI - American National Standards Institution.

ATM - Asynchronous Transfer Mode. Digital switched network, transfers in fixed length 53 byte cells.

ATU-C - ADSL Terminal Unit Central Office. Terminal nearer the central office or remote network node.

ATU-R - ADSL Terminal Unit Remote. ADSL terminal nearer to the subscriber.

CAP - Carrierless Amplitude Modulation/Phase Modulation. A possible technology used in ADSL.

Current standards put

emphasis on DMT technology.

DMT - Discrete Multitone. A version of multi-carrier modulation that allows allocation of payload data bits and transmitter power among more than one subchannel depending on loss, and interference among each subchannel. A candidate technology for ADSL.

Drop Wire - Last part of the loop connecting distribution cable to the customer premises.

DS1 - Digital Signal 1 : 1.544 Mb/s with a payload of 1.536 Mb/s bi-directional.

DS2 - Digital Signal 2 : 6.312 Mb/s , which can transport 4 DS1's asynchronously.

ECH - Echo Canceler with hybrid. used in DSL and HDSL systems.

FTTC - Fiber To The Curb. Fiber optic lines that run to a remote electronics node close to the subscribers location.

FTTH - Fiber To The Home. Fiber optic line that run to the subscribers home.

POTS - Plain Old Telephone Service. Telephone service as we know it today.

RBOCS - Regional Bell Operating Companies. Local exchange carriers.

RFI - Radio frequency interference.

VDSL - Very-high-rate Digital Subscriber Line. Capable of transporting 50 Mb/s payload or greater.

10.12 REVIEW QUESTIONS

Q1. What is Digital subscriber Line?

Q2. Define Asynchronous Digital Subscriber Line

Q3. Explain High-Data-Rate DSl in detial.

Q4. Write short note on Quantization.

Q5. What is very high data rate digital subscriber line?

10.13 FUTURE READINGS

- Marlis Humphrey, John Freeman, Paradyne Corp., "How xDSL supports Broadband Services to the Home", IEEE Network Jan./Feb 1997 Notes.
- http://www.adsl.com/adsl_reference_model.html
- Philip Kyees, Ronald C McConnell, Kamran Sistanizadeh, "ADSL: A New Twisted-Pair Access to the Information Highway", IEEE Communications Magazine, April 1995
- Angela Littwin, "ADSL: Ready for Prime Time?", Telecommunications, Dec. 1996
- Arielle Emmett, "Motorola plays ADSL chip, CopperGold bets on single-chip transceiver", America's Network, Dec. 1, 1996, http://www.americasnetwork.com/
- "Data Over Cable Interface Specifications - Cable Modem Termination System Network side Interface Specification", http://www.cablelabs.com

LESSON 11 MANAGEMENT SYSTEM FOR HIGH SPEED NETWORKS

11.1 OBJECTIVE

Objective of this chapter is to introduce reader about the management of High speed Network. Management of traffic management, application management and device management.

11.2 INTRODUCTION TO MANAGEMENT SYSTEM TO HSN

An essential part of planning your network infrastructure is designing and implementing methods for managing it. "Network management" is certainly a broad term, which can and does encompass everything from determining the amount of traffic on a particular segment to monitoring how an application uses memory on a network client to tracking whether certain devices are functioning properly.

In this chapter we will take a look at three very broad areas of network management and the various management systems used for each. The areas of management are as follows:

- Traffic management
- Application management
- Device management

11.3 TRAFFIC MANAGEMENT

When most people talk about "network management," they are really talking about *traffic management.* The purpose of traffic management is to determine how much of what type of data traffic is traveling over your network infrastructure at any given time. Upon determining this, you can decide whether your network infrastructure can handle that amount of traffic, and make changes in the bandwidth or access method or protocol type to ensure optimum throughput.

Traffic management systems fall generally into three categories:

- Packet generators
- Network analyzers
- Application testers

Packet Generators

Packet generators operate at the Network and Data-Link layers of the OSI model. They load a network with "synthetic traffic." Synthetic traffic is composed of packets that the packet generator created (rather than packets from real network production traffic). These packets are sent in huge volume from a central location on the network. The goal of packet generators is to “stress test” a network, or finds its maximum traffic limits. Packet generators are able to flood a network, while normal production traffic is rarely able to do so.

The downside of packet generators is that the traffic patterns they produce aren't real. Ordinary network traffic doesn't come as fast, as constant, or as uniformly shaped and directed as the traffic produced by packet generators. Therefore, while packet generators are great for determining the maximum theoretical limits of traffic that your network can handle, they don't tell you much about how your network will perform under day-to-day traffic loads. A well-known packet generator is SmartBits from Netcom Systems.

Network Analyzers

Network analyzers monitor actual network traffic. They allow you to take a sample of traffic and analyze its content and composition. For example, if you are getting a slow response time on a particular segment of your network, you could place a network analyzer on that segment to monitor it for several hours—or even several days. The network analyzer would record the type of traffic (10Base-T, token ring, etc.), the network protocol (IP, IPX, router protocols, and others), how much of the available bandwidth is being used, and information about collisions, incomplete packets, and packet retransmissions. This information can help you locate "chattering" cards (cards that are malfunctioning and flooding the network with bad packets), overloaded segments, and “top talkers"—users who take up most of the network's traffic resources, Some examples of network analyzers are Internet Advisor from Hewlett-Packard and Sniffer from Network General.

Application Testers

Application tester’s model and measure node-to-node traffic patterns from the Application layer. They simulate real-world traffic associated with using specific types of distributed network applications such as databases and document preparation packages. The packages allow you to test the impact that adding a given application would have on your network so

that you can plan infrastructure changes accordingly. Some examples of application testers are Chariot from Ganymede Software, Inc., and Dynameasure from Bluecurve, Inc.

11.4 APPLICATION MANAGEMENT

While traffic management devices are great for helping you locate problems and Bottlenecks in the infrastructure, they don't tell you anything about the resources that are, being used by network applications, nor the resulting availability and response time of those applications. This is the venue of application management systems, which have arrived fairly recently on the scene.

In fact, application management systems are so new that at the time of this writing, there are no standards for application management. Right now, application vendors are Ring proprietary technology to implement management and monitoring systems. For example, Oracle Corporation has enabled SNMP support for its database, and has developed its own proprietary management console. This works well in a single-vendor situation, but in reality, most enterprise environments will have multiple mission-critical applications from a variety of vendors. Each of these applications should be monitored—and a separate management system for each critical application simply isn't workable. The ideal application management product should be able to monitor a wide range of applications with an open and scalable architecture.

Therefore, the Desktop Management Task Force (DMTF) is working on developing such standards. Currently the Distributed Application Performance (DAP) Working Group of the DMTF has been charged with defining a standard model for the behavior of distributed applications. This model will define the runtime behavior for applications for *each unit of work. A* unit of work could be a transaction, a database query, or even an input/output event.

The stated goals of the DAP working group are as follows:

- Extend the core/meta schema, if necessary, to model runtime performance and incorporate quality of service objects
- Relate these models to the common schemas of the other management domains: Systems/Devices, Network, Applications, Service, Support, and User
- Evolve standard methods to populate the model developed independent of operating system or performance agent availability

11.5 DEVICE MANAGEMENT

One of the most widespread needs in network management is to monitor devices to determine whether they are functioning properly. This is called *device management.*

Most device management packages gather information from devices by using *agents* to collect information from devices all over the network, then send it back to a centralized *management console.* At the management console, the network manager can use the information to create reports on the health of the network.

Device management systems also use agents to send alerts about trouble on the network such as device failures or performance problems. Agents reside throughout the network on servers, routers, switches, and client computers. Agents gather information about the device on which they are installed and send it back to the management console. Management console functions include network topology mapping, event reporting, traffic monitoring, network diagnostic functions, trend analysis, and report generation.

The most widely used management applications that employ agents use the *Simple Network Management Protocol* (SNMP). SNMP is a management protocol for TCP/IP networks. Due most likely to the UNIX roots of TCP/IP networks, most SNMP management systems are UNIX based, although Windows NT systems are becoming more popular. In any event,

SNMP was developed by the Internet community, and it defines how agents and management consoles interact through request and response messaging.

Simple Network Management Protocol (SNMP)

SNMP is a communication protocol for gathering information from network devices. Each network device hosts an agent that gathers information about the performance of me device, then sends that information to a management console. Each piece of information gathered about a device is defined by a *managed object.* A managed object is a logical representation of a real physical entity on the network. Each managed object gathers some information from the device hosting it. The managed object then maintains that information for use by the management system.

All of these managed objects are organized as a management information base (MIB). *MIB* is just a hierarchical database of managed objects. There are both generic and vendor-specific MIBs. Individual vendors must design managed objects for each piece of information they want to monitor on their devices. The vendors must provide these managed objects in the form of a MIB that adheres to SNMP standards.

Figure 19-1 shows the components of a typical SNMP management environment and how these components work together. The components are as follows

- Devices Network devices such as hubs, switches, and routers hosting agents that gather information about the devices and send it to the management console.
- Proxy agent In devices that aren't able to host an agent, a proxy agent can be run by another device on behalf of the SNMP-incapable device.
- Management systems The application that assembles the information gathered by agents.

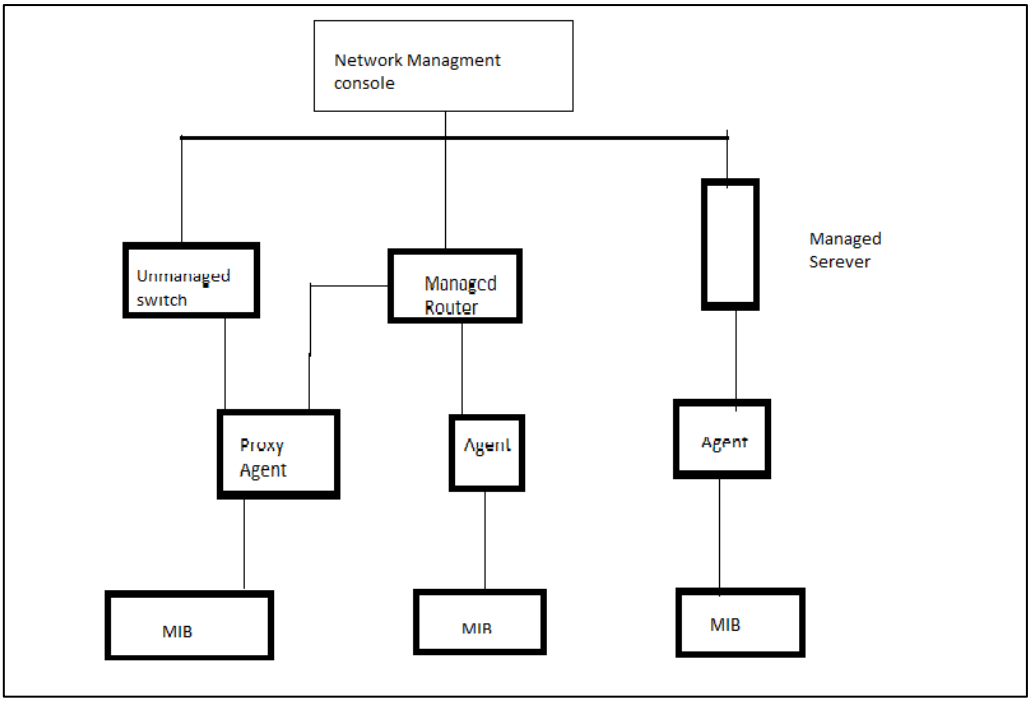


Figure 19.1

- SNMP The protocol that provides both the query language for gathering information from—as well as the transport protocol for sending the information to—the agents that runs in the devices.

Most SNMP management systems discover the topology of the network automatically, then display the network topology in the form of a graphic. Administrators can choose a particular segment to view in more detail. They can then select a particular device to monitor, and view all information collected from it.

The agents running on the managed devices constantly collect information. Each agent collects only that information it is specifically charged with collecting. For example, an agent on a switch may count each packet received on each port. Collecting this information and sending it to the management system is the sole job of the SNMP agent.

Back at the central collection point, the management system continuously contacts the various agents to obtain the information they have collected. SNMP uses UDP, which is part of the TCP/IP suite, to carry its messages across the network.

Web-Based Network Management

Web-based network management lets network managers monitor their networks using Web technologies like browsers and HTTP. It lets administrators monitor devices remotely, in real time, and collect information about systems. It also lets them monitor troubleshoot systems and create reports about network problems or trends. Web-based network management uses the TCP/IP protocols, which are the coins of the realm on the Internet and corporate intranets. Therefore, the same Web management tools can be used to manage a switch in the same room or anywhere in the world via the worldwide Internet. This trend is becoming so popular that many vendors of network management systems have already added Web browser interfaces to their management systems.

Web-Based Enterprise Management (WBEM)

In July of 1996, five companies formed the Web-Based Enterprise Management (WBEM) consortium with the express goal of integrating disparate management protocols, management information, and management systems. The five companies are

- BMC Software
- Cisco Systems
- Microsoft
- Intel

While WBEM was founded by these five companies, its initiative is now supported by over 50 vendors. So, what is the WBEM initiative all about? While SNMP is still the most widely used management protocol, WBEM would like to see new Internet standards such **as** Hypermedia Management Protocol (HMMP) and Common Information Model (CIM) take the forefront in network management. The HMMP advantage says the WBEM Consortium, is

that it lets network management system vendors build systems that use all the same protocols and provide access through the same user interface. As well, the WBEM Consortium is interested in developing a common data structure to facilitate easier, more uniform access to corporate data.

HYPERMEDIA MANAGEMENT PROTOCOL (HMMP) Hypermedia Management Protocol (HMMP) is an Internet protocol used to access and publish data across the Internet, as well as to exchange management and control messages between HMMP nodes. As a client/server protocol, it is similar to Hypertext Transfer Protocol (HTTP). HMMP clients request data from HMMP servers. However, unlike HTTP, HMMP clients can be management processes running in standard Web browsers or management and monitoring processes that are hosted by managed devices. When HMMP servers receive requests from clients, they either respond to those requests with information or pass the request on to another server. In the latter instance, the requesting server is now a client of the server to which it is passing the request. As you can see, HMMP doesn't require that every client know how to manage every device it needs to use directly. Instead, clients can make requests to servers that manage groups of devices. The devices it manages may in fact use proprietary management mechanisms and protocols unknown to the client. This dqesn't affect the client making the request, because the servers always respond to the clients' requests for information using HMMP.

COMMON INFORMATION MODEL (CIM) Common Information Model (CIM) is also part of the WBEM initiative. CIM is a data structure that uses HMMP to access and manipulate data held by HMMP management hosts. The CIM structure is a set of standard classes and types that represent hardware devices and other such managed objects in the network management environment.

11.6 KEYWORDS

- Management information base (MIB)
- The purpose of traffic management is to determine how much of what type of data traffic is traveling over your network infrastructure at any given time
- Device management packages gather information from devices by using *agents* to collect information from devices all over the network, then send it back to a centralized *management console*

11.6 REVIEW QUESTIONS

Q1. How management of High Speed network is done?

Q2. Write note on traffic management.

Q3. How application management is done in High Speed Network?

Q4. How devices are managed in High Speed Network?

11.7 FURTHER READINGS

- Tere Parnell - Building High Speed Networks, -- TMH
- Cooper E. - Broadband Network Technology -- Prentice Hall
- Tanenbaum - Computer Networks -- PHL
- Green P.E. - Fiber Optics Networks - Prentice Hall.
- Goralski W.J. -- Introduction to ATM Networking -- MGH.

LESSON 12 VIDEO COMPRESSIONS

12.1 OBJECTIVE

Data compression is required storage and transmission of data. The chapter contains various video compression techniques.

12.2 INTRODUCTION TO VIDEO COMPRESSION

Video applications require some form of data compression to achieve reasonable precondition for storage and transmission. The digital video compression is one of the main issue in digital video coding, enabling efficient distribution and interchange of visual information.

Video codecs are devices that are used to compress and decompress as well as to encode and decode video streams. The most complex part of a codec is the compress/decompress function. Codecs can do their work by hardware but also by software with fast processors. The main goal of coding is the bit-rate reduction for storage and transmission of the video source while retaining video quality as good as possible. There are a number of international standards and also many proprietary techniques for digital video compression. The basic idea

behind video compression is to remove spatial redundancy within a video frame and temporal redundancy between adjacent video frames.

There are two main types of compression techniques, lossless and lousy. In the lossless compression a frame can be decompressed into the original exactly. The compression ratio of lossless methods (Huffman, Arithmetic, LZW, and RLE) is not high enough for digital video communication. In the lousy compression we create compressed data that can be decompressed into images that look similar to the original (as human eye sees them) but are different in digital form.

Video compression techniques

The human eye is more sensitive to changes in brightness than to chromaticity changes. Therefore the image data is first divided into one luminance and two chrominance components, and the chrominance components are subsampled relative to the luminance component. After this step the usual lossy compression method used in digital video compression is based on Discrete Cosine Transform (DCT) and quantization. This technique reduces the high spatial frequency components from the image since the human viewer is more sensitive to the reconstruction errors of low frequency components. The purpose of the quantization step is to represent the DCT-coefficients with the precision what is needed to achieve the required image quality. The zig-zag step arranges the high frequency coefficients to the end of the stream and since most of them have become zero after the quantization, run length encoding (RLE) is used for further compression. The upper left corner coefficient represents the mean value of the block and is encoded using the difference from the previous block (DPCM). The final step in the compression process is to minimize the entropy using Huffman or arithmetic coding. The encoded frame is often called I-frame (intra frame)

because the encoding process uses no information from other frames. The block diagram of the encoding process is in Figure 12.1.

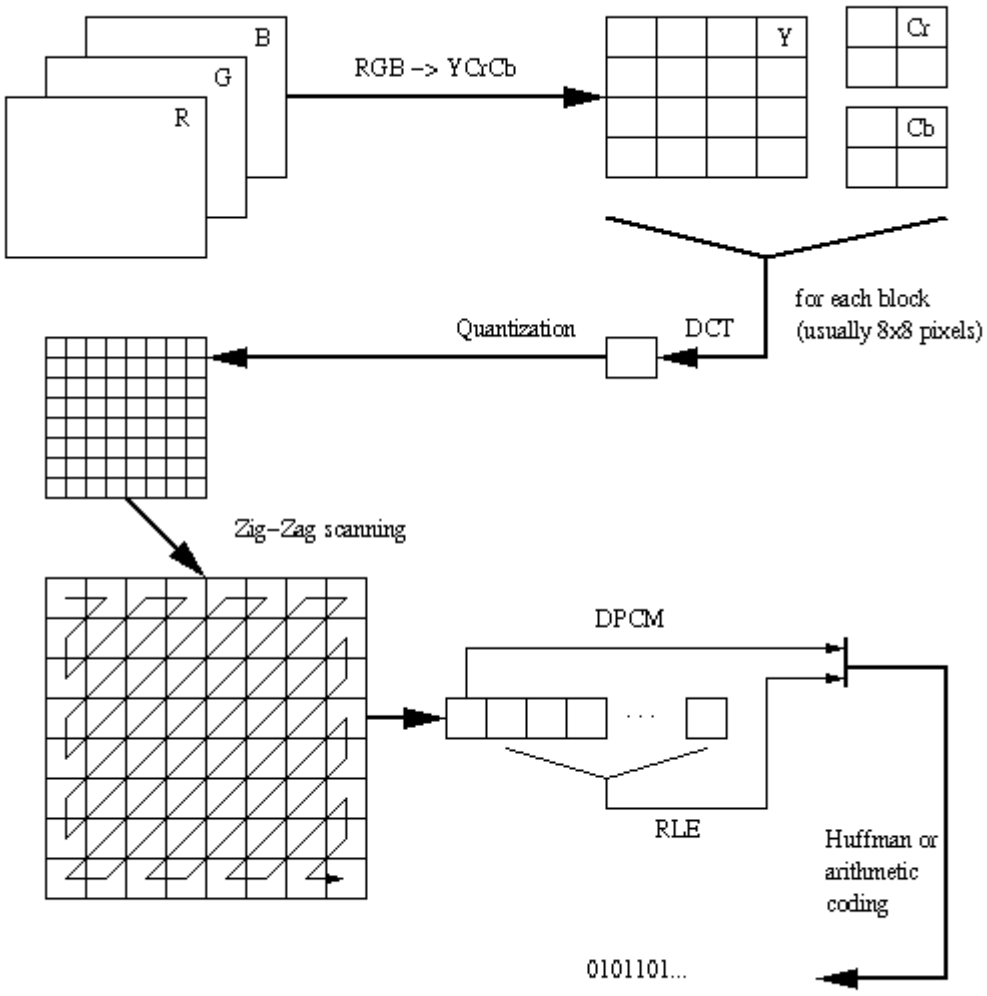


Figure 12.1 Block diagram of video compression

In addition to the previous compression technique, the temporal redundancy between frames can be utilized for further compression. The basic method is to calculate the prediction error between corresponding blocks in the current and previous frames. The error values are then send to the compression process. Compressed frames generated using prediction are usually

called P-frames. When using both previous and future frames as reference, the frame is called B-frame (bidirectional frame).

Motion compensated prediction is an efficient tool to reduce temporal redundancy between frames. The concept of motion compensation contains the motion estimation between video frames (Figure 12.2). The motion is described by a small number of motion vectors which gives the translation of a block of pixels between frames. The motion vectors and compressed prediction errors are then transmitted.

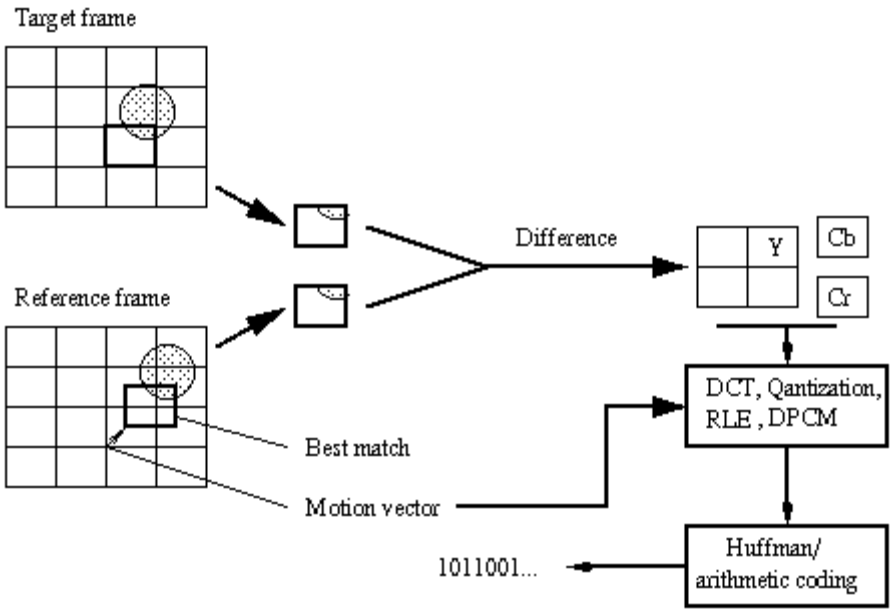


Figure 12.2 Motion compensation

At higher level of compression the block based DCT-transform introduces some blocking artefacts. Newer methods no longer require this subdivision of the image and permit deep levels of compression without these artefacts. Wavelet transformation, a powerful tool for compressing information, represents pictures, as waves that can be described mathematically

in terms of frequency, energy and time. The advanced mathematics underlying wavelet transformation is very complex.

The fractal compression has gained some global interest but offers so far no benefits over DCT based methods.

H.261 video codec

H.261 is a video coding standard published by ITU in 1990. It is designed for ISDN network data rates which are multiples of 64 kbit/s. It is the most widely used international video compression standard in video conferencing today. A functional block diagram of the video codec is shown in Figure 12.3

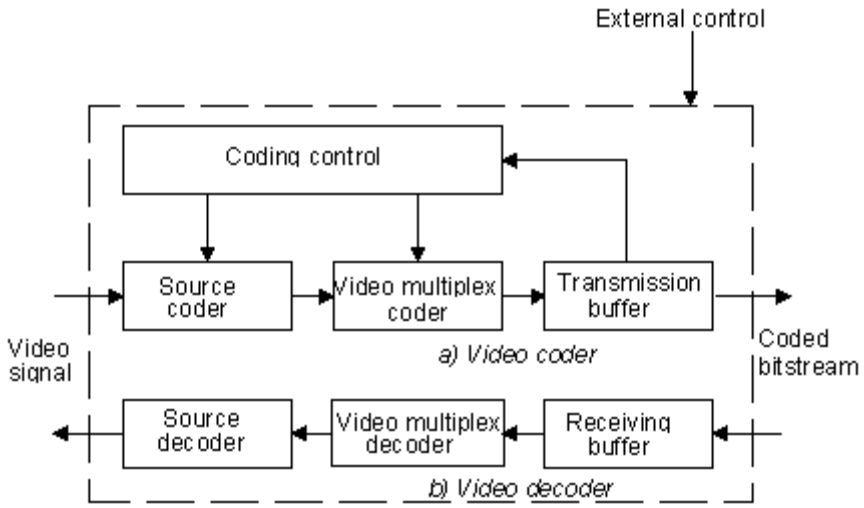


Figure 12.3: A block diagram of the H.261 video codec

The main element in the encoding process is a source coder which compresses the incoming video signal. The source coder operates on non-interlaced pictures occurring 30000/1001 times per second. Pictures are coded as values of the luminance and two colour difference components (Y, Cb, Cr). These components are defined in ITU-R BT.601 recommendation.

The Cb and Cr matrices are a quarter the size of the Y matrix. To solve the incompatibility between different TV standards (PAL, SECAM and NTSC), the CIF and QCIF picture structures were adopted. The QCIF format, which employs half the CIF spatial resolution, must be supported by all H.261 compatible codecs. QCIF is intended for videophone applications where head-and-shoulders pictures are sent. The full CIF format is used when several people must be viewed in a conference room. for example.

H.261 Picture Formats

Picture format	**Lumin. pixels**	**Lumin. lines**	
QCIF	176	144	mandatory
CIF	352	288	Optional

Each picture is divided into groups of blocks (GOB) in the encoding process. The CIF picture has 12 GOBs and QCIF has three.

GOBs in CIF		**GOBs in QCIF**
1	2	1
3	4	2
5	6	3
7	8	
9	10	
11	12	

Each GOB is divided into 33 macro blocks. The macro block header defines location of a macro block within the GOB, type of coding, possible motion vectors, and which blocks will actually be coded within the macro block.

macro blocks in GOB										
1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	21	22
23	24	25	26	27	28	29	30	31	32	33

Each macro block is further divided into six blocks. Four of them represent the luminance and two the chrominance components of the signal. Each block has 8x8 pixels, so the colour resolution is half of the luminance resolution in both dimensions.

Luminance (Y)		**Chrominance/blue (Cb)**	**Chrominance/red (Cr)**
1	2	5	6
3	4		

Following is a short description of the coder's main functions:

Prediction

The prediction is inter-picture and may be augmented by motion compensation. The coding mode in which prediction is applied is called "inter". The coding mode is called "intra" if no prediction is applied. The intra coding mode comprise the picture level, and blocks (8x8 pixels) are encoded only with reference to themselves and are sent directly to the block coding process.

The inter coding uses another picture frame as a reference. The prediction error is calculated between corresponding macro blocks in the current and previous frames. A criteria if a block is transmitted or not, is not defined in the standard.

Motion compensation

Motion compensation support in a H.261 encoder is optional, but it is required from the decoder. In motion compensation the previous frame is searched for the best reference macroblock. The prediction error and motion vectors are sent to the block transformation process. Neither the search area nor the method to compute motion vectors are included in the standard.

Block coding

In block coding both intra coded and inter coded frames are coded as 8x8 blocks using two dimensional DCT, which gives an 8x8 coefficient matrix for each block.

Quantization

After DCT the coefficients are quantized. This step produces coefficients values, which are much smaller in magnitude than the original values and most of them become zero. The purpose of this step is to represent the coefficients with no greater precision than is necessary to achieve the required image quality.

Entropy coding

Extra lossless compression is achieved by RLE-coding coefficients and then further entropy coding these with Huffman-technique. This produces a variable length code.

Multiplexing

The video multiplex coder combines the compressed data with various side information, which indicates various operation modes, into hierarchical bit stream that can be interpreted by decoders.

Transmission buffer

The transmission buffer is employed to smooth the varying bit rate from the source coder to adapt it to the communication channel.

The coded bit stream is composed of two types of frames: intra coded I-frames and inter coded P-frames. An example of an encoded sequence might be: **I P P P I P P P ....** Where there are three predicted P-frames after each I-frame.

H.263

The ITU-T H.263 recommendation is based on H.261 and is targeted for compressing the moving picture at low bitrates.

The coding algorithm of H.263 has some improvements and changes over that used by H.261 and it can often achieve the same quality as H.261 with less than half the number of bits in the coded stream. That is why it is replacing H.261 in many applications in the Internet, where bandwidth is often a critical resource.

H.263 codec has the following new features over H.261:

- Half pixel precision is used for the motion compensation, as opposed to H.261 where full pixel precision and loop filter are used.
- Some parts of the data stream structure are now optional, so the codec can be configured for a lower bitrate or better error recovery.
- There are four negotiable coding options included: Unrestricted motion vectors mode, syntax-based arithmetic coding, advanced prediction mode and PB-frames mode.
- The H.263 supports new picture resolutions.

H.263 Picture Formats

Picture format	Lumin. pixels	Lumin. lines	
SQCIF	128	96	mandatory
QCIF	176	144	mandatory
CIF	352	288	optional
4CIF	704	576	optional
16CIF	1408	1152	optional

The PB-frame consists of two frames coded as one unit. The P-frame is predicted from the previous P-frame and the B-frame which is predicted from both the previous and the current P-frame

12.4 VIDEO COMPRESSION: MPEG

The acronym MPEG stands for Moving Picture Expert Group, which worked to generate the specifications under ISO, the International Organization for Standardization and IEC, the International Electro technical Commission. What is commonly referred to as "MPEG video" actually consists at the present time of two finalized standards, MPEG-11 and MPEG-22, with a third standard, MPEG-4, was finalized in 1998 for Very Low Bitrate Audio-Visual Coding. The MPEG-1 and MPEG-2 standards are similar in basic concepts. They both are based on motion compensated block-based transform coding techniques, while MPEG-4 deviates from these more traditional approaches in its usage of software image construct descriptors, for target bit-rates in the very low range, < 64Kb/sec. Because MPEG-1 and MPEG-2 are finalized standards and are both presently being utilized in a large number of

applications, this paper concentrates on compression techniques relating only to these two standards. Note that there is no reference to MPEG-3. This is because it was originally anticipated that this standard would refer to HDTV applications, but it was found that minor extensions to the MPEG-2 standard would suffice for this higher bit-rate, higher resolution application, so work on a separate MPEG-3 standard was abandoned.

The current thrust is MPEG-7 "Multimedia Content Description Interface" whose completion is scheduled for July 2001. Work on the new standard MPEG-21 "Multimedia Framework" has started in June 2000 and has already produced a Draft Technical Report and two Calls for Proposals.

MPEG-1 was finalized in 1991, and was originally optimized to work at video resolutions of 352x240 pixels at 30 frames/sec (NTSC based) or 352x288 pixels at 25 frames/sec (PAL based), commonly referred to as Source Input Format (SIF) video. It is often mistakenly thought that the MPEG-1 resolution is limited to the above sizes, but it in fact may go as high as 4095x4095 at 60 frames/sec. The bit-rate is optimized for applications of around 1.5 Mb/sec, but again can be used at higher rates if required. MPEG-1 is defined for progressive frames only, and has no direct provision for interlaced video applications, such as in broadcast television applications.

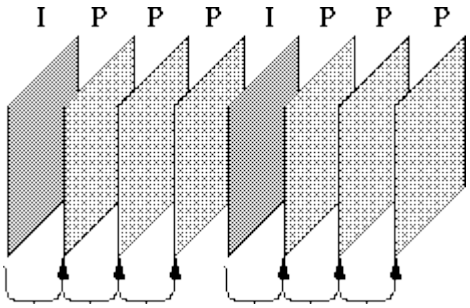
MPEG-2 was finalized in 1994, and addressed issues directly related to digital television broadcasting, such as the efficient coding of field-interlaced video and scalability. Also, the target bit-rate was raised to between 4 and 9 Mb/sec, resulting in potentially very high quality video. MPEG-2 consists of profiles and levels. The profile defines the bit stream scalability and the colorspace resolution, while the level defines the image resolution and the maximum bit-rate per profile. Probably the most common descriptor in use currently is Main Profile, Main Level (MP@ML) which refers to 720x480 resolution video at 30 frames/sec, at bit-

rates up to 15 Mb/sec for NTSC video. Another example is the HDTV resolution of 1920x1080 pixels at 30 frame/sec, at a bit-rate of up to 80 Mb/sec. This is an example of the Main Profile, High Level (MP@HL) descriptor. A complete table of the various legal combinations can be found in reference2.

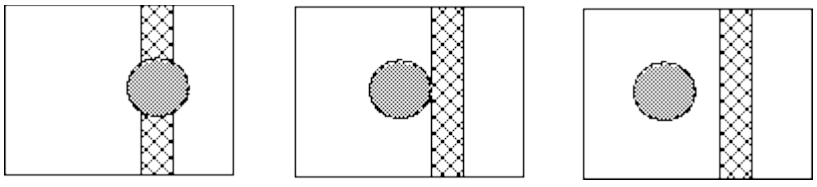
MPEG Video

MPEG compression is essentially a attempts to overcome some shortcomings of H.261 and JPEG:

- Recall H.261 dependencies:



  - The Problem here is that many macro blocks need information is **not** in the reference frame.
  - For example:

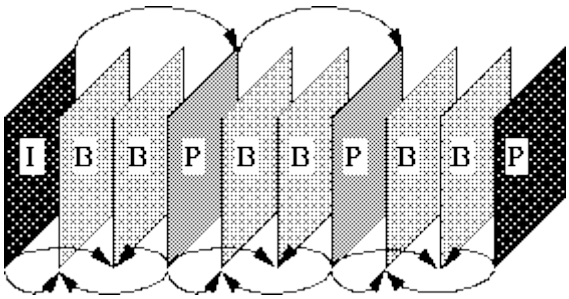


  - The ***MPEG solution*** is to add a third frame type which is a bidirectional frame, or *B-frame*

- B-frames search for macro block in *past* and *future* frames.

- Typical pattern is IBBPBBPBB IBBPBBPBB IBBPBBPBB

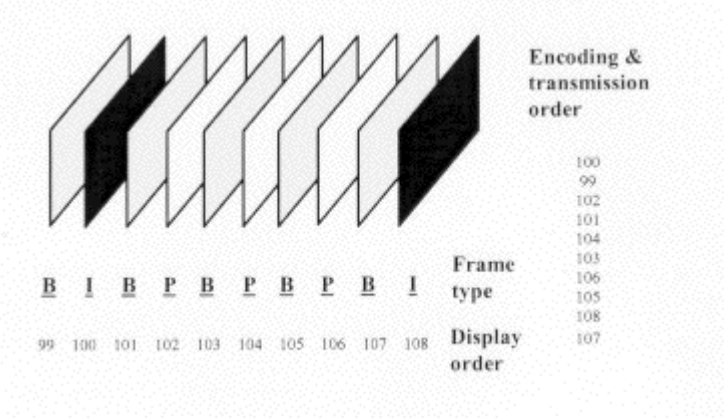
Actual pattern is up to encoder, and need not be regular.



B-Frames

The MPEG encoder also has the option of using forward/backward interpolated prediction. These frames are commonly referred to as bi-directional interpolated prediction frames, or B frames for short. As an example of the usage of I, P, and B frames, consider a group of pictures that lasts for 6 frames, and is given as I,B,P,B,P,B,I,B,P,B,P,B, As in the previous I and P only example, I frames are coded spatially only and the P frames are forward predicted based on previous I and P frames. The B frames however, are coded based on a forward prediction from a previous I or P frame, as well as a backward prediction from a succeeding I or P frame. As such, the example sequence is processed by the encoder such that the first B frame is predicted from the first I frame and first P frame, the second B frame is predicted from the second and third P frames, and the third B frame is predicted from the third P frame and the first I frame of the next group of pictures. From this example, it can be seen that backward prediction requires that the future frames that are to be used for backward prediction be encoded and transmitted first, out of order. There is no defined limit to the number of consecutive B frames that may be used in a group of pictures, and of course the optimal number is application dependent. Most broadcast quality applications however, have

tended to use 2 consecutive B frames (I,B,B,P,B,B,P) as the ideal trade-off between compression efficiency and video quality.



B-Frame Encoding

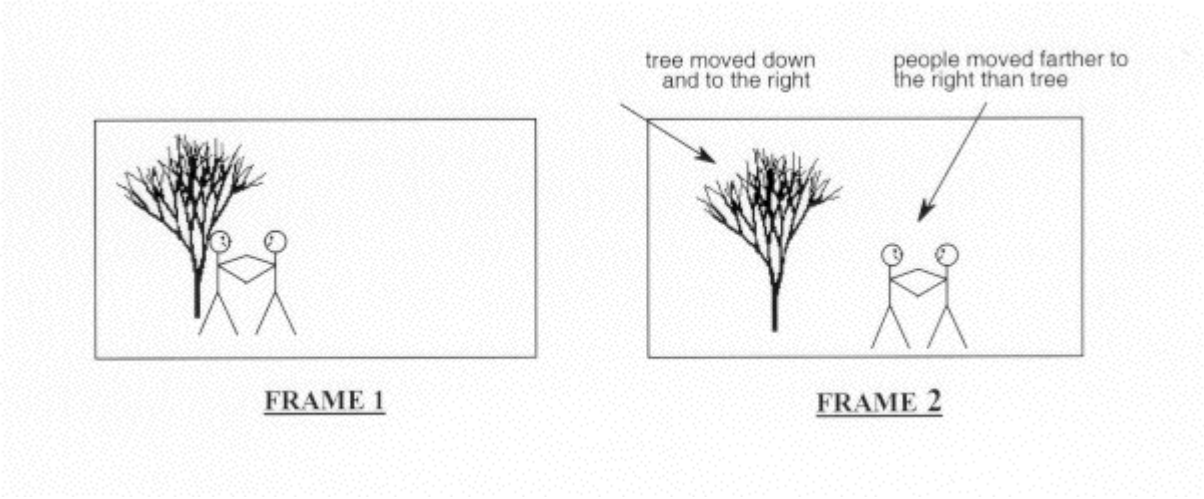
The main advantage of the usage of B frames is coding efficiency. In most cases, B frames will result in less bits being coded overall. Quality can also be improved in the case of moving objects that reveal hidden areas within a video sequence. Backward prediction in this case allows the encoder to make more intelligent decisions on how to encode the video within these areas. Also, since B frames are not used to predict future frames, errors generated will not be propagated further within the sequence.

One disadvantage is that the frame reconstruction memory buffers within the encoder and decoder must be doubled in size to accommodate the 2 anchor frames. This is almost never an issue for the relatively expensive encoder, and in these days of inexpensive DRAM it has become much less of an issue for the decoder as well. Another disadvantage is that there will necessarily be a delay throughout the system as the frames are delivered out of order as was shown in Figure . Most one-way systems can tolerate these delays, as they are more objectionable in applications such as video conferencing systems.

Motion Estimation

The temporal prediction technique used in MPEG video is based on motion estimation. The basic premise of motion estimation is that in most cases, consecutive video frames will be similar except for changes induced by objects moving within the frames. In the trivial case of zero motion between frames (and no other differences caused by noise, etc.), it is easy for the encoder to efficiently predict the current frame as a duplicate of the prediction frame. When this is done, the only information necessary to transmit to the decoder becomes the syntactic overhead necessary to reconstruct the picture from the original reference frame. When there is motion in the images, the situation is not as simple.

Figure shows an example of a frame with 2 stick figures and a tree. The second half of this figure is an example of a possible next frame, where panning has resulted in the tree moving down and to the right, and the figures have moved farther to the right because of their own movement outside of the panning. The problem for motion estimation to solve is how to adequately represent the changes, or differences, between these two video frames.



12.5 SUMMARY

The methods used in the standards are quite similar and represent the current development in compression techniques. The video compression algorithms have traditionally developed for constant bit-rate channels. Heterogenous packet networks like the Internet require however some new approaches. Packet losses are not rare and loss patterns may be bursty. Compression schemes that rely on low error rate do not operate well under these circumstances.

With algorithms used in MPEG and H.261/ H.263, the packet loss causes a significant degrade in quality due to their method of removing temporal redundancy. The prediction is based on the decoded signal and the model assumes that decoder shares an identical state. Both MPEG and H.261 rely on intra frames to eventually resynchronize, but at low bitrates the resynchronization intervals can be too wide and decoded bit stream may virtually never be error free.

The solution is to reduce the resynchronization interval. This scheme is used in M-JPEG where each frame is coded independently. However, this approach results in low compression because of redundant information. Another solution is to forego motion compensation, use only intra frames and make heavy use of conditional replenishment. In this model only the blocks that change are transmitted, and at some low rate all the blocks in the image are transmitted to guarantee that lost blocks are eventually retransmitted. The MPEG-2 and H.261/263 standards are flexible enough to make codecs to utilize aggressive conditional replenishment and intra coding with a fully compliant syntax of the standards .

12.6 KEYWORDS

- Video codecs are devices that are used to compress and decompress as well as to encode and decode video streams. The most complex part of a codec is the compress/decompress function.
- Differential Pulse Code Modulation (DPCM) on DC component
- Discrete Cosine Transform (DCT)
- run length encoding (RLE)

12.7 REVIEW QUESTIONS

Q1. What is image compression? Why compression is needed?

Q2. Define Discrete cosine Transformation.

Q3. How data is compressed usins run length encoding?

Q4. What is quantization? What are the types of quantization?

12.8 FURTHER READING/INFORMATION

- G.K. Wallace, *The JPEG Still Picture Compression Standard*
- CCITT, *Recommendation H.261*
- D. Le Gall, *MPEG: A Video Compression Standard for Multimedia Applications*
- K. Patel, et. al., *Performance of a Software MPEG Video Decoder*
- P. Cosman, et. al., *Using Vector Quantization for Image Processing*