Optical Communication Networks

EE654

Lectuer - 5

Spring 2016

Client Layers of the Optical Layer

The predominant client layers in backbone networks today are SONET/SDH, Ethernet, and the Optical Transport Network (OTN). These protocols would correspond to the physical layer in the OSI hierarchy (see Figure 1.6). SONET/SDH as part of the first generation of optical networks was the earliest to be deployed in backbone networks and has been very successful over the years. It is particularly adept at supporting constant bit rate (CBR) connections, and it multiplexes these connections into higher speed optical connections by using time division multiplexing. Originally designed for low speed voice and CBR connections, up to 51 Mb/s, it now supports data network, packet traffic that can have link transmission rates in the tens of gigabits per second. An important feature of SONET/SDH is that it provides carrier grade service of high availability.

SONET/SDH can transport packets for data networks due to data link layer protocols that adapt packet traffic to its connections. Generic Framing Procedure (GFP) is an adaptation method that works for a variety of data networks, including IP, Ethernet, and Fibre Channel.

OTN builds upon the concepts of SONET/SDH and has been designed to carry all types of data traffic including SONET/SDH traffic. It has been enhanced to operate at very high transmission rates, and it has a complete and flexible set of operation and management features.

Ethernet started as a local-area network (LAN) using a coaxial cable. Today, it is carried over all communication physical media including twisted pair, wireless, and fiber optic cables. It offers a wide range of data rates: 10 Mb/s, 100 Mb/s, 1 Gb/s, and 10 Gb/s. Ethernet spans the data link and physical layers.

IP will also be covered, even though it is not a proper client layer of the optical layer. It is at the network layer and is not carried directly over optical paths. However, it is the predominant packet transport technology for many applications including the Internet, and much of the traffic carried by optical networks is IP traffic. Thus, it is important to understand IP since optical networks should efficiently support its traffic. In addition, some of the ideas of the IP protocol have been applied to design optical networks.

IP uses *connectionless* routing, where packets are forwarded based only on the packets' destinations. It has been enhanced with the *multiprotocol label switching* (MPLS), protocol which is a *connection-oriented* routing mechanism. In connection-oriented routing, streams of packets are organized into flows, and routing is done per flow. Flows are identified by labels, and these labels are carried by packets to identify their flow and to facilitate packet forwarding along routes.

In the metro network, there are several types of client layers such as Gigabit Ethernet, 10-Gigabit Ethernet, Fibre Channel, Resilient Packet Ring (RPR) as well as SONET/SDH. Fibre Channel is used in the so-called storage-area networks to interconnect computers and their peripherals. RPR is at the data link layer and is not a proper client layer of the optical layer. However, for data packet traffic, it is an alternative to SONET/SDH, and like SONET/SDH, it uses a ring network topology to provide high availability of service.

Table 6.1 Transmission rates for asynchronous and plesiochronous signals, adapted from [SS96].

Level	North America	Europe	Japan
0	0.064 Mb/s	0.064 Mb/s	0.064 Mb/s
1	1.544 Mb/s	2.048 Mb/s	1.544 Mb/s
2	6.312 Mb/s	8.448 Mb/s	6.312 Mb/s
3	44.736 Mb/s	34.368 Mb/s	32.064 Mb/s
4	139.264 Mb/s	139.264 Mb/s	97.728 Mb/s

5.1 SONET and SDH

SONET (Synchronous Optical Network) is the current transmission and multiplexing standard for high-speed signals within the carrier infrastructure in North America. A closely related standard, SDH (Synchronous Digital Hierarchy), has been adopted in Europe and Japan and for most submarine links.

In order to understand the factors underlying the evolution and standardization of SONET and SDH, we need to look back in time and understand how multiplexing was done in the public network. Prior to SONET and SDH, the existing infrastructure was based on the *plesiochronous digital hierarchy* (PDH), dating back to the mid-1960s. (North American operators refer to PDH as the *asynchronous* digital hierarchy.) At that time the primary focus was on multiplexing digital voice circuits. An analog voice circuit with a bandwidth of 4 kHz could be sampled at 8 kHz and quantized at 8 bits per sample, leading to a bit rate of 64 kb/s for a digital voice circuit. This became the widely accepted standard. Higher-speed streams were defined as multiples of this basic 64 kb/s stream. Different sets of standards emerged in different parts of the world for these higher-speed streams, as shown in Table 6.1. In North America, the 64 kb/s signal is called DS0 (digital signal-0), the 1.544 Mb/s signal is DS1, the 44.736 Mb/s is DS3, and so on. In Europe, the hierarchy is labeled E0, E1, E2, E3, and so on, with the E0 rate being the same as the DS0 rate. These rates are widely prevalent today in carrier networks and are offered as leased line services by carriers to customers, more often than not to carry data rather than voice traffic.

PDH suffered from several problems, which led carriers and vendors alike to seek a new transmission and multiplexing standard in the late 1980s. This resulted in the SONET/SDH standards, which solved many problems associated with PDH. We explain some of the benefits of SONET/SDH below and contrast it with PDH.

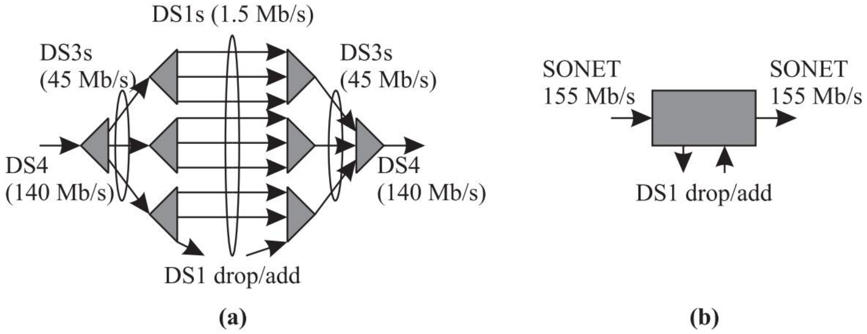


Figure 6.1 Comparison of asynchronous and synchronous multiplexing. (a) In the asynchronous case, demultiplexers must be stacked up to extract a lower-speed stream from a multiplexed stream. (b) In the synchronous case, this can be done in a single step using relatively simple circuitry.

1. **Multiplexing simplification:** In asynchronous multiplexing, each terminal in the network runs its own clock, and while we can specify a nominal clock rate for the signal, there can be significant differences in the actual rates between different clocks. For example, in a DS3 signal, a 20 ppm (parts per million) variation in clock rate between different clocks, which is not uncommon, can produce a difference in bit rate of 1.8 kb/s between two signals. So when lower-speed streams are multiplexed by interleaving their bits, extra bits may need to be stuffed in the multiplexed stream to account for differences between the clock rates of the individual streams. As a result, the bit rates in the asynchronous hierarchy are not integral multiples of the basic 64 kb/s rate, but rather slightly higher to account for this bit stuffing. For instance, a DS1 signal is designed to carry 24 64 kb/s signals, but its bit rate (1.544 Mb/s) is slightly higher than 24 × 64 kb/s.

   With asynchronous multiplexing, it is very difficult to pick out a low-bit-rate stream, say, at 64 kb/s, from a higher-speed stream passing through, say, a DS3 stream, without completely demultiplexing the higher-speed stream down to its individual component streams. This results in the need for "multiplexer mountains," or stacked-up multiplexers, each time a low-bit-rate stream needs to be extracted, as shown in Figure 6.1. This is a relatively expensive proposition and also compromises network reliability because of the large amount of electronics needed overall.

   The synchronous multiplexing structure of SONET/SDH provides significant reduction in the cost of multiplexing and demultiplexing. All the clocks in the network are perfectly synchronized to a single master clock, and as a consequence, the rates defined in SONET/SDH are integral multiples of the basic rate and no bit stuffing is needed when multiplexing streams together. As a result, a lower-speed signal can be extracted from a multiplexed SONET/SDH stream in a single step by locating the appropriate positions of the corresponding bits in the multiplexed signal. This makes the design of SONET multiplexers and demultiplexers much easier than their asynchronous equivalents. We will explore this in more detail in Section 6.1.1.

2. **Management:** The SONET and SDH standards incorporate extensive management information for managing the network, including extensive performance monitoring, identification of connectivity and traffic type, identification and reporting of failures, and a data communication channel for transporting management information between the nodes. This is mostly lacking in the PDH standards.
3. **Interoperability:** Although PDH defined multiplexing methods, it did not define a standard format on the transmission link. Thus different vendors used different line coding, optical interfaces, and so forth to optimize their products, which made it very difficult to connect one vendor's equipment to another's via a transmission link. SONET and SDH avoid this problem by defining standard optical interfaces that enable interoperability between equipment from different vendors on the link.
4. **Network availability:** The SONET and SDH standards have evolved to incorporate specific network topologies and specific protection techniques and associated protocols to provide high-availability services. As a consequence, the service restoration time after a failure with SONET and SDH is much smaller—less than 60 ms—than the restoration time in PDH networks, which typically took several seconds to minutes.

5.1.1 SONET/SDH Layers

The SONET layer consists of four sublayers—the *path, line, section,* and *physical* layers. Figure 6.4 shows the top three layers. Each layer, except for the physical layer, has a set of associated overhead bytes that are used for several purposes. These overhead bytes are added whenever the layer is introduced and removed whenever the layer is terminated in a network element. The functions of these layers will become clearer when we discuss the frame structure and overheads associated with each layer in the next section.

The path layer in SONET (and SDH) is responsible for end-to-end connections between nodes and is terminated only at the ends of a SONET connection. It is possible that intermediate nodes may do performance monitoring of the path layer

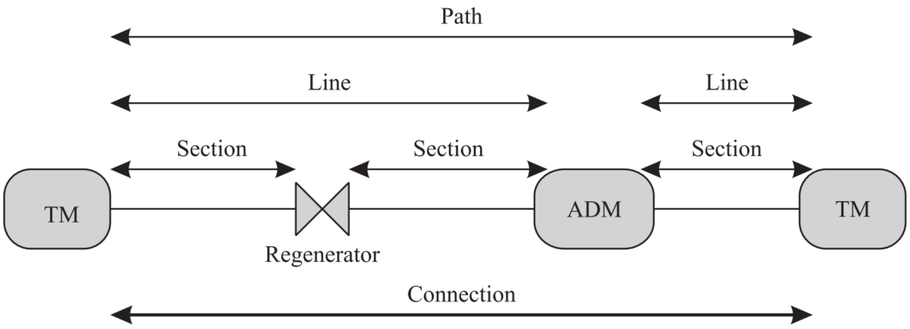


Figure 6.4 SONET/SDH layers showing terminations of the path, line, and section layers for a sample connection passing through terminal multiplexers (TMs) and add/drop multiplexers (ADMs). The physical layer is not shown.

signals, but the path overhead itself is inserted at the source node of the connection and terminated at the destination node.

Each connection traverses a set of links and intermediate nodes in the network. The line layer (*multiplex section* layer in SDH) multiplexes a number of path-layer connections onto a single link between two nodes. Thus the line layer is terminated at each intermediate line terminal multiplexer (TM) or add/drop multiplexer (ADM) along the route of a SONET connection. The line layer is also responsible for performing certain types of protection switching to restore service in the event of a line failure.

Each link consists of a number of sections, corresponding to link segments between regenerators. The section layer (*regenerator-section* layer in SDH) is terminated at each regenerator in the network.

Finally, the physical layer is responsible for actual transmission of bits across the fiber.

5.1.2 SONET frame Structure

Figure 6.5 shows the structure of an STS-1 frame. A frame is 125 μs in duration (which corresponds to a rate of 8000 frames/s), regardless of the bit rate of the SONET signal. This time is set by the 8 kHz sampling rate of a voice circuit. The frame is a specific sequence of 810 bytes, including specific bytes allocated to carry overhead information and other bytes carrying the payload. We can visualize this frame as consisting of 9 rows and 90 columns, with each cell holding an 8-bit byte.

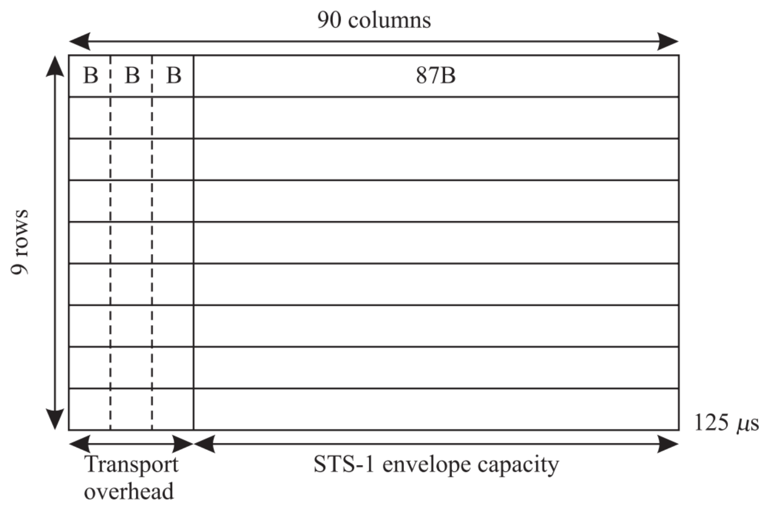


Figure 6.5 Structure of an STS-1 frame. B denotes an 8-bit byte.

The bytes are transmitted row by row, from left to right, with the most significant bit in each byte being transmitted first.

The first three columns are reserved for section and line overhead bytes. The remaining bytes carry the STS-1 SPE. The STS-1 SPE itself includes one column of overhead bytes for carrying the path overhead.

5.2 Optical Transport Network

The Optical Transport Network (OTN), sometimes referred to as G.709, was designed to transport data packet traffic such as IP and Ethernet over fiber optics, as well as legacy traffic and in particular SONET/SDH. It is called the *digital wrapper* technology because it wraps any client signal in overhead information for operations, administration, and management. Its line rates, OTU1, OTU2, and OTU3, are shown in Table 6.4. It builds on SONET/SDH concepts, and it features the following capabilities.

Table 6.4 OTN line rates compared with SONET/SDH line rates.

OTN (G.709)	Line Rates	SONET/SDH	Line Rates
OTU1:	2.666 Gb/s	STS-48/ STM-16:	2.488 Gb/s
OTU2:	10.709 Gb/s	STS-192/ STM-64:	9.953 Gb/s
OTU3:	43.018 Gb/s	STS-786/STM-128:	39.813 Gb/s

1. **Forward error correction (FEC):** OTN has been designed for high data transmission rates, as shown in Table 6.4. At very high data rates or over very long distances, noise is significant and becomes a problem when ensuring low bit error rates. Forward error correction (FEC) as we discussed in Section 4.5 is critical to achieving these low bit error rates. FEC had already been used in implementations of SDH. These are proprietary coding schemes that rely on making use of unused section overhead bytes to carry the redundant FEC bytes. However, the performance is limited since the number of bytes is limited, and interoperability with other vendor equipment cannot be assured. OTN has been designed to carry FEC overhead and employs stronger FEC using the (255,239) Reed-Solomon code (Section 4.5). Thus, for each 255 byte block, there are 16 redundant bytes. The FEC can correct errors in a block of up to 8 bytes of error and detect an error in a block with at most 16 bytes of error. The blocks are interleaved to increase the length of error bursts that can be corrected.
2. **Management:** As we have seen in the previous section, SONET/SDH supports monitoring and managing the signal at the section, line, and path levels. This overhead includes signal identification, BER measurement, and communicating alarm information. OTN provides structure for monitoring a connection end-to-end and over various segments. These segments may overlap with up to six such monitoring segments at any given point. An example application would be a connection of a network A that passes through another network B; that is, B is serving as a carrier for network A. Then the operators of both networks must monitor the connection as it passes through B, using their own set of monitoring and managing signals. These signals must be operating in tandem.

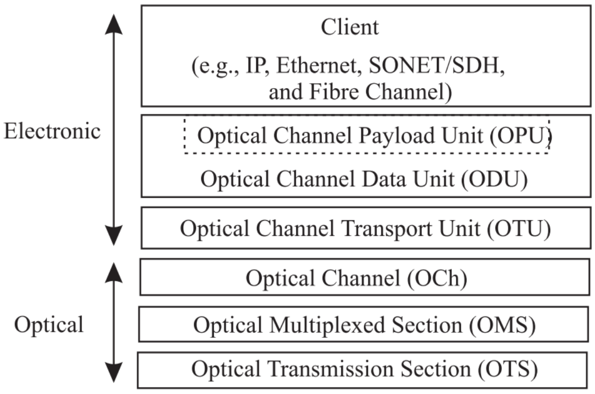

Client
(e.g., IP, Ethernet, SONET/SDH, and Fibre Channel)
Electronic
Optical Channel Payload Unit (OPU)
Optical Channel Data Unit (ODU)
Optical Channel Transport Unit (OTU)
Optical Channel (OCh)
Optical
Optical Multiplexed Section (OMS)
Optical Transmission Section (OTS)

Figure 6.10 OTN hierarchy.

3. **Protocol transparency:** OTN provides a constant bit rate service. It has operations, administration, and management of its connections that are transparent to its clients. It can carry all types of data packet traffic including IP and 10-Gigabit Ethernet, as well as SONET/SDH frames. OTN frames can carry entire SONET/SDH frames including overhead without modification. Table 6.4 shows that OTN line rates are 7% higher than SONET/SDH line rates, and this is due to its additional overhead and FEC information.
4. **Asynchronous timing:** OTN has an asynchronous mapping of client signals into OTN frames where the clock that generates the frames can be a simple free-running oscillator. To account for any mismatch between the clocks of the OTN frames and the client signal, the OTN payload floats within the frame. Using simple free-running oscillators can simplify implementation and reduce costs. OTN also has a synchronous mapping where the clock to generate the OTN frames is derived from the client signal.

5.3 Ethernet

Ethernet was created in the 1970s to be a packet-switched data link that connects computers and computer equipment over a single coaxial cable, that is, a bus. It is easy to understand, implement, manage, and maintain, and has led to low network costs. Ethernet has since evolved to include a variety of topologies including point-to-point, bus, star, and mesh as shown in Figure 6.14; and adapted to a variety of physical communication media, including coaxial cable, twisted pair copper cable, wireless media, and optical fiber. It has a wide range of rates. Typical rates today are 10 Mb/s, 100 Mb/s (Fast Ethernet), 1 Gb/s (Gigabit Ethernet or GbE), and 10 Gb/s (10-Gigabit Ethernet or 10 GbE). At the time of this writing 40 Gb/s and 100 Gb/s Ethernet are being developed. It was one of the first local-area network (LAN) technologies and has thrived to become the predominant LAN as well as a predominant link layer technology.

Media Access Control

In the original Ethernet, computers were attached to the network coaxial cable with a network interface card (NIC), and each NIC has a unique 6-byte Ethernet address that is assigned by the NIC manufacturer. A node can transmit a packet on the cable, and the transmission signal will be received by all the nodes. The coaxial cable was effectively a broadcast communication link. A problem with this configuration is that nodes transmitting at the same time can interfere with one another's transmissions, causing a transmission *collision*. Since such collisions mean transmissions are not received properly, they are a waste of link bandwidth.

Ethernet has a media access control (MAC) protocol to arbitrate transmissions between nodes. When a node has a packet to transmit, it listens to the link. When it detects that the link is idle (i.e., there are no transmissions), it transmits its packet and at the same time listens to the link. If it detects a collision, then it stops transmitting, avoiding further waste of bandwidth. Then it attempts to retransmit the packet after a randomly chosen delay. Since all nodes in a collision will retransmit after a randomly chosen delay, there is a high likelihood that exactly one of the nodes will retransmit before the others. Once this node retransmits, the other nodes will detect its transmission and wait until the link is idle again. The arbitration protocol is referred to as *carrier sense multiple access with collision detection* (CSMA/CD) because a node listens to the link before transmitting and stops transmitting upon detecting a collision.

To achieve high throughput, the time to detect a collision must be small relative to the time to transmit a packet. Then the fraction of time spent on collisions will be small compared to the fraction of time to successfully transmit packets. The collision detection delay is largely dependent on the propagation delay across the cable, which is dependent on the length of the cable. Therefore, collision detection delays were made small by limiting the length of the cable. For example, 10 Mb/s Ethernet networks have a maximum diameter of about 2500 m. Since packet transmission times are inversely proportional to transmission speeds, for 100 Mb/s Ethernet networks, the diameter limit was reduced by about a factor of 10 to 200 m. For Gigabit Ethernet, reducing the diameter by another factor of 10 leads to a diameter of about 20 m, which is too small to be practical for some important applications. Instead, the Ethernet packet lengths were increased by a factor of about 10. Another method to achieve high throughput is *frame bursting*, which allows a node to transmit frames consecutively without being interrupted. Then small frames can be put together and transmitted as a longer virtual frame.

Point-to-Point Link

An important application of Ethernet is as a point-to-point link connecting two end nodes. For twisted pair and fiber optic implementations, Ethernet has an option to operate as a full duplex link. Performance improves because both channels can be used simultaneously. As a consequence, CSMA/CD is unnecessary since the end nodes do not interfere with each other's transmissions. Without CSMA/CD, there are less constraints on frame lengths and link diameters. 10-Gigabit Ethernet only allows full duplex operation and uses ordinary Ethernet frames.

To realize flow control, a receiver can send a *pause frame* to the sender to make it stop sending. A pause frame indicates an amount of time the sender must wait before resuming transmissions.

Local-Area Network

Ethernet is the predominant LAN technology. Today the most popular Ethernet LAN configuration is the star topology, which has a hub at its center, connecting a number of Ethernet segments. The topology has management advantages since much of the network administration can be done at a single location. It fits naturally into the telecommunication infrastructure of office buildings. It also improves signal quality by repeating or regenerating signals.

One type of hub is a *repeater* that simply broadcasts the incoming signals to all Ethernet segments. Then the star topology behaves like a single Ethernet. This simple design does not scale well with the number of nodes because the network bandwidth is divided among all the nodes. This results in each node having access to network bandwidth that is inversely proportional to the number of nodes. Another type of hub that leads to better traffic throughput is a *switch* (or *multiport bridge*), which is discussed in Section 6.4.2. With switches, Ethernet can be extended from the star topology to operate on a mesh topology.

Another important Ethernet feature is the *virtual LAN* (VLAN). It allows the network bandwidth to be shared among groups of nodes, so that each group can communicate over its own VLAN. A VLAN has a distinct identifier called a *tag*. Ethernet VLAN packets have a field for their tag so that they can be distinguished and forwarded to the members of their VLAN group. VLAN technology can be used to implement *virtual private networks* (VPNs). In addition, Ethernet VLAN packets have a priority field to support quality of service. Note that Ethernet VLAN technology has similarities with MPLS technology (see Section 6.6 on MPLS), which is also used to forward packets, separate traffic, and support quality of service.

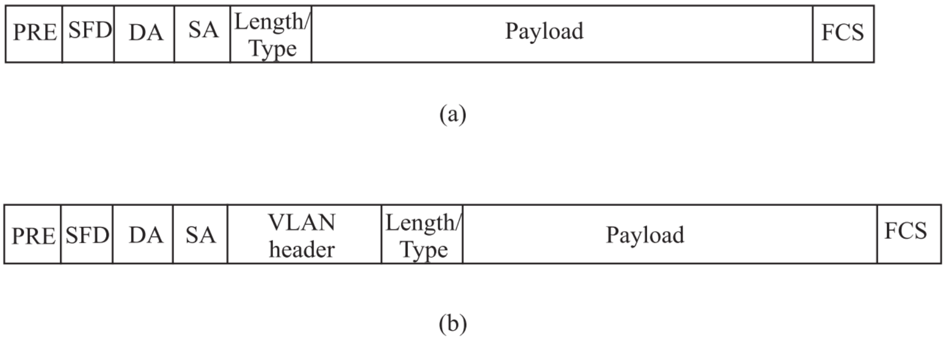


Figure 6.15 (a) Basic Ethernet frame and (b) VLAN Ethernet frame.

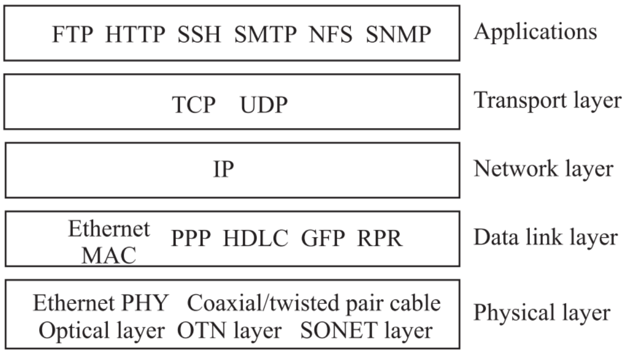
5.4 Internet Protocol (IP)

IP (Internet Protocol) is by far the most widely used wide-area networking technology today. IP is the underlying network protocol used in the all-pervasive Internet and is equally important in most private intranets to link up computers. IP is a networking technology, or protocol, that is designed to work above a wide variety of lower layers, which are termed *data link layers* in the classical layered view of networks (Section 1.4). This is one of the important reasons for its widespread success.

Figure 6.19 shows IP within the layered architecture framework. The traditional data link layers over which IP operates are Ethernet and the point-to-point protocol (PPP). IP operates over other low-speed serial lines as well as high-speed optical fiber lines using well-known data link layer protocols—for example, high-level data link control (HDLC).

Several layering structures are possible to map IP into the optical layer. The term *IP over WDM* is commonly used to refer to a variety of possible mappings shown in Figure 6.20. Figure 6.20(a) shows the packet-over-SONET (POS) implementation. Here, IP packets are mapped into PPP frames and then encoded into SONET frames for transmission over a wavelength. Figure 6.20(b) shows an implementation using Gigabit or 10-Gigabit Ethernet as the underlying link (media access control) layer and Gigabit/10-Gigabit Ethernet physical layer (PHY) for encoding the frames for transmission over a wavelength. We will study the implications of these different approaches in Chapter 13.

IP, being a network layer protocol, does not guarantee reliable, in-sequence delivery of data from source to destination. This job is performed by a transport protocol, typically the *transmission control protocol* (TCP). Another commonly used transport protocol for simple message transfers over IP is the *user datagram protocol* (UDP). Commonly used applications such the file transfer protocol (FTP), hypertext transfer



FTP HTTP SSH SMTP NFS SNMP	Applications
TCP UDP	Transport layer
IP	Network layer
Ethernet MAC PPP HDLC GFP RPR	Data link layer
Ethernet PHY Coaxial/twisted pair cable Optical layer OTN layer SONET layer	Physical layer

Figure 6.19 IP in the layered hierarchy, working along with a variety of data link layers and transport layers.

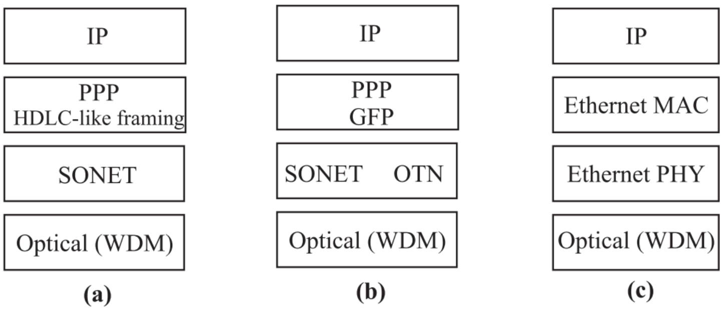
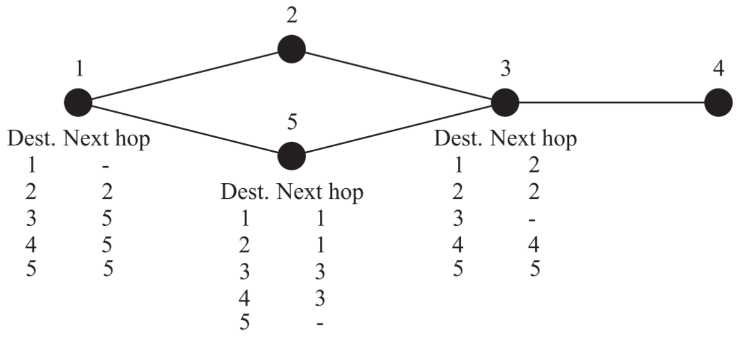


Figure 6.20 Various implementations of IP over WDM. (a) A packet-over-SONET (POS) variant, where IP packets are mapped into PPP frames then an HDLC-like framing and scrambling, and finally into SONET frames. (b) IP packets are mapped into PPP frames, and then framed using the Generic Framing Procedure (GFP) before mapped into a SONET or OTN path. (c) Using Gigabit or 10-Gigabit Ethernet media access control (MAC) as the link layer and Gigabit or 10-Gigabit Ethernet physical layer (PHY) for encoding the frames for transmission over a wavelength.

protocol (HTTP), secure shell (SSH), and simple mail transfer protocol (SMTP) use TCP as their transport protocol. Other applications use UDP for transport such as the network file system (NFS), which is used to share files across a network, and the simple network management protocol (SNMP), which is used for management. (We will discuss SNMP in Chapter 8.) UDP is also the transport protocol of choice for streaming media.

5.4.1 Routing and Forwarding

IP was one of the earliest packet-switching protocols. IP transports information in the form of packets, which are of variable length. An IP router is the key network element in an IP network. A router forwards packets from an incoming link onto an outgoing link. Figure 6.21 illustrates how packets are forwarded in an IP network. The nature of this routing is fundamental to IP. Here we describe the classical routing mechanism used by IP. Each router maintains a routing table. The routing table has one or more entries for each destination router in the network. The entry indicates the next node adjacent to this router to which packets need to be forwarded. The forwarding process works as follows. The router looks at the header in a packet arriving on an incoming link. The header contains the identity of the destination router for that packet. The router then does a lookup of its routing table to determine the next adjacent node for that packet and forwards the packet on the link leading to that node. In the example shown in Figure 6.21, consider a packet from node 1 destined for node 4. Node 1 looks at its table and forwards this packet to node 5. Node 5 forwards the packet to node 3, which in turn forwards the packet to node 4, its ultimate destination.



Dest.	Next hop
1	-
2	2
3	5
4	5
5	5

Dest.	Next hop
1	1
2	1
3	3
4	3
5	-

Dest.	Next hop
1	2
2	2
3	-
4	4
5	5

Figure 6.21 Routing in an IP network. The routing tables at some of the nodes are also shown. The tables contain the identity of the next hop node for each destination.

Clearly, maintaining these routing tables at the routers is central to the operation of the network. It is likely that links and nodes in the network may fail, or reappear, and new links and nodes may be added over the course of time. The routers detect these changes automatically and update their routing tables using a distributed *routing protocol*. The protocol works as follows. Each router is assumed to be capable of determining whether its links to its neighbors are up or down. Whenever a router detects a change in the status of these links, it generates a *link state packet* and *floods* it to all the routers in the network. Flooding is a technique used to disseminate information across the network. Each node, upon receiving a flood packet, forwards the packet on all its adjacent links except the link from which it came. Thus these packets eventually reach all the nodes in the network. A node receiving a link state packet updates its routing table based on the new information. Over time, all nodes in the network have updated routing tables that reflect the current network topology.

A number of subtle enhancements are needed to make the flooding process work reliably. For example, link state packets could take different paths through the network and undergo different delays. As a result, an older link state packet might arrive after a more recent up-to-date version. If left unchecked, this could cause damage. Consider what happens when a link goes down and comes back up. The first link state packet (packet X) says that the link is down, and the subsequent one (packet Y) indicates that the link is up. A node receiving packet X after packet Y will think that the link is down, even after it has come up! To prevent this phenomenon, the link state packets have a sequence number. If a router receives a link state packet whose sequence number is lower than a previously received link state packet, it simply discards the packet. Packets could also be lost in the network, so link state updates are generated periodically and not just after a link up/down event occurs.

Using these link state packets, each router can construct its view of the entire network topology. On this topology, each router then computes the shortest path from itself to all the other routers and stores the identity of the next router in the path for each destination node in its routing table. A typical shortest-path algorithm used for this purpose is the Dijkstra algorithm [Dij59].

5.5 Storage-Area Networks

Storage-area networks (SANs) are networks used to interconnect computer systems with other computer systems and peripheral equipment, such as disk drives, printers, and tape drives. These networks are built by enterprises that have medium to large data centers. Figure 6.27 shows a typical SAN interconnecting multiple CPUs and various types of peripheral devices. A key part of a SAN is a switch, which provides reconfigurable connectivity between the various attached devices. The SANs that we consider below all use a circuit-switched approach, where connections are rapidly established and taken down between the attached devices as needed.

In early installations, the entire SAN was located within a building or campus, but today the network is distributed over a wider metropolitan area, with some links extending into the long-haul network. One reason to do so is to be able to provide resilience against disasters. A common technique is to maintain two data centers, with data from one center backed up onto the other. Another reason to distribute the network is to locate peripherals and other equipment away from major downtown areas into cheaper suburban areas where real estate is less expensive.

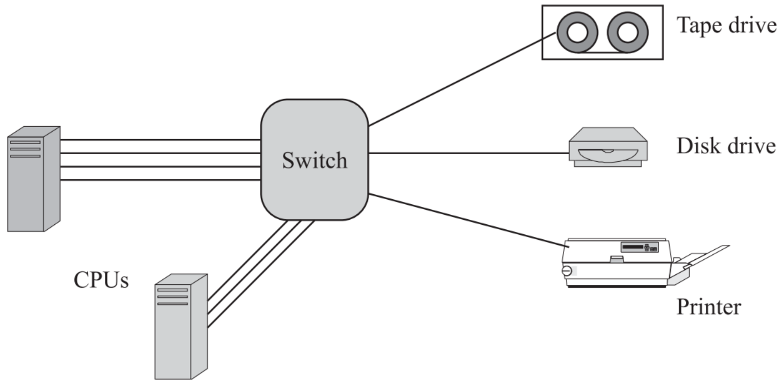
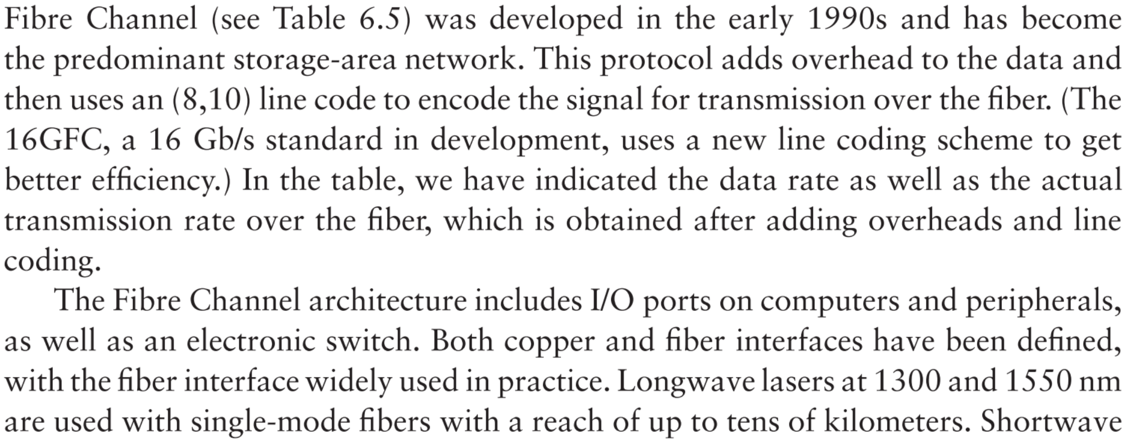


Figure 6.27 Architecture of a storage-area network.

SANs typically operate at bit rates ranging from 200 Mb/s to about 10 Gb/s and operate over fiber optic links in most cases. What makes them important from the perspective of the optical layer is that there can be a huge number of such connections between two data centers. Large mainframes have hundreds of I/O channels to connect them to other devices. It is not uncommon to see networks with hundreds to thousands of these links between two data centers.

5.5.1 Fiber Channel



Fibre Channel (see Table 6.5) was developed in the early 1990s and has become the predominant storage-area network. This protocol adds overhead to the data and then uses an (8,10) line code to encode the signal for transmission over the fiber. (The 16GFC, a 16 Gb/s standard in development, uses a new line coding scheme to get better efficiency.) In the table, we have indicated the data rate as well as the actual transmission rate over the fiber, which is obtained after adding overheads and line coding.

The Fibre Channel architecture includes I/O ports on computers and peripherals, as well as an electronic switch. Both copper and fiber interfaces have been defined, with the fiber interface widely used in practice. Longwave lasers at 1300 and 1550 nm are used with single-mode fibers with a reach of up to tens of kilometers. Shortwave lasers at 850 nm are used with multimode fibers with a reach of up to a few hundred meters.

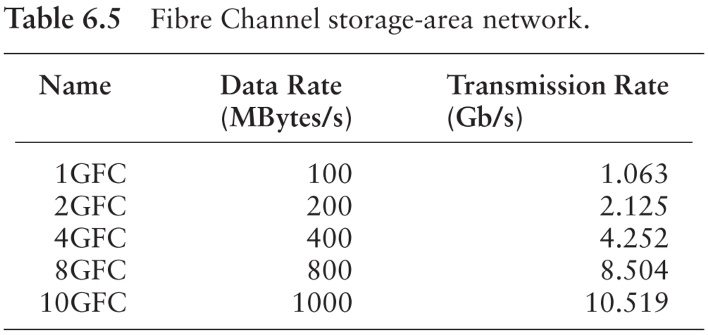


Table 6.5 Fibre Channel storage-area network.

Name	Data Rate (MBytes/s)	Transmission Rate (Gb/s)
1GFC	100	1.063
2GFC	200	2.125
4GFC	400	4.252
8GFC	800	8.504
10GFC	1000	10.519