Chapter 10

Networking Systems

A *multimedia networking system* allows for the data exchange of discrete and continuous media among computers. This communication requires proper services and protocols for data transmission.

10.1 Layers, Protocols and Services

- A *service* provides a set of operations to the requesting application. Logically related services are grouped into *layers* according to the OSI reference model. Therefore, each layer is a *service provider* to the layer lying above. The services describe the behavior of the *layer* and its *service elements* (Service Data Units = SDUs). A proper service specification contains no information concerning any aspects of the implementation.

- A *protocol* consists of a set of rules which must be followed by peer layer instances during any communication between these two peers. It is comprised of the *format* (syntax) and the *meaning* (semantics) of the *exchanged data units* (Protocol Data Units = PDUs). The peer instances of different computers cooperate together to provide a service.

Multimedia communication puts several *requirements* on services and protocols, which are independent from the layer in the network architecture. In general, this set of requirements depends to a large extent on the respective application. However, without defining a precise value for individual parameters, the following requirements must be taken into account:

- Audio and video data processing need to be bounded by deadlines or even defined by a time interval. The data transmission – both between applications and transport layer interfaces of the involved components – must follow within the demands concerning the time domains.

- End-to-end jitter must be bounded. This is especially important for interactive applications such as the telephone. Large jitter values would mean large buffers and higher end-to-end delays.

- All guarantees necessary for achieving the data transfer within the required time span must be met. This includes the required processor performance, as well as the data transfer over a bus and the available storage for protocol processing.

- Cooperative work scenarios using multimedia conference systems are the main application areas of multimedia communication systems. These systems should support multicast connections to save resources. The sender instance may often change during a single session. Further, a user should be able to join or leave a multicast group without having to request a new connection setup, which needs to be handled by all other members of this group.

- The services should provide mechanisms for synchronizing different data streams, or alternatively perform the synchronization using available primitives implemented in another system component.

- The multimedia communication must be compatible with the most widely used communication protocols and must make use of existing, as well as future networks. *Communication compatibility* means that different protocols at least coexist and run on the same machine simultaneously. The relevance of envisaged protocols can only be achieved if the same protocols are widely used.

Many of the current multimedia communication systems are, unfortunately, proprietary experimental systems.

- The communication of discrete data should not starve because of preferred or guaranteed video/audio transmission. Discrete data must be transmitted without any penalty.

- The fairness principle among different applications, users and workstations must be enforced.

- The actual audio/video data rate varies strongly. This leads to fluctuations of the data rate, which needs to be handled by the services. Figure 10.1 shows data rates of three different situations: we distinguish among uncorrelated pictures, persons in a room and a news speaker. In spite of the high fluctuation

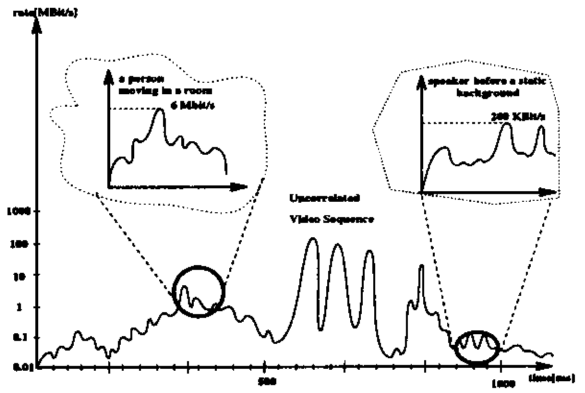

Figure 10.1: *Examples of rates in three situations.*

of the actual values (the peaks), the average values are relatively low. The information rate, shown in Figure 10.1, is based on a coding according to CCIR 601. Note that many compression schemes enforce a constant data rate, leading to a quality of video that depends on the content.

The above described requirements apply to different components of a communication system. By way of describing multimedia communication systems, let us consider the individual layers of the *ISO-OSI Reference Model* [Tan88], which provides at least the conceptual basis for any communication system:

1. *Physical Layer*

   The physical layer defines the transmission method of individual bits over the physical medium, such as fiber optics.

   For example, the type of modulation and bit-synchronization are important issues. With respect to the particular modulation, delays during the data transmission arise due to the propagation speed of the transmission medium and the electrical circuits used. They determine the maximal possible bandwidth of this communication channel. For audio/video data in general, the delays must be minimized and a relatively high bandwidth should be achieved.

2. *Data Link Layer*

   The data link layer provides the transmission of information blocks known as *data frames*. Further, this layer is responsible for access protocols to the physical medium, error recognition and correction, flow control and block synchronization.

   Access protocols are very much dependent on the network. Networks can be divided into two categories: those using point-to-point connections and those using broadcast channels, sometimes called *multi-access channels* or *random access channels*. In a broadcast network, the key issue is how to determine, in the case of competition, who gets access to the channel. To solve this problem, the *Medium Access Control (MAC)* sublayer was introduced and MAC protocols, such as the *Timed Token Rotation Protocol* and *Carrier Sense Multiple Access with Collision Detection (CSMA/CD)*, were developed. The MAC sublayer is especially important in Local Area Networks (LANs), nearly all of which use multi-access channels as the basis for their communication.

   Continuous data streams require *reservation* and *throughput guarantees* over a line. To avoid larger delays, the error control for multimedia transmission needs a different mechanism than retransmission because a late frame is a lost frame. However, because of the new high-speed networks based on fiber optics, there may not be a need for any error control at this layer. These networks favor multimedia transmission because of their very low transmission error rate. Further, a fixed-size information block (cell), such as in Asynchronous Transfer Mode (ATM) networks, allows for an efficient protocol implementation providing, reservations and guaranteed throughput.

3. *Network Layer*

   The network layer transports information blocks, called *packets*, from one station to another. The transport may involve several networks. Therefore, this layer provides services such as addressing, internetworking, error handling, network management with congestion control and sequencing of packets.

   Again, continuous media require *resource reservation* and *guarantees* for transmission at this layer. A request for reservation for later resource guarantees is defined through *Quality of Service (QoS) parameters*, which correspond to the requirements for continuous data stream transmission. The reservation must be done along the path between the communicating stations. For this purpose, connection-oriented behavior is needed where the reservation is made during the connection setup. Through this approach, end-to-end delay with small jitter and correct packet ordering can be enforced. If internetworking is included, for different communication structures in multicasting or broadcasting connections, duplication of packets can follow, which may introduce further complexity into the reservation process. The network QoS for a connection should be negotiated at this layer.

4. *Transport Layer*

   The transport layer provides a process-to-process connection. At this layer, the QoS, which is provided by the network layer, is enhanced, meaning that if the network service is poor, the transport layer has to bridge the gap between what the transport users want and what the network layer provides. Large packets are segmented at this layer and reassembled into their original size at the receiver. Error handling is based on process-to-process communication.

   With respect to continuous data streams, the QoS parameters must match their requirements. Again, the error handling does not include retransmission because this mechanism would introduce high end-to-end delays and strong jitter. Also, synchronization, which allows time relations between two LDUs (Logical Data Unit) of one connection, and between SDUs (Session Data Unit) of different connections, is often said to be a function of this layer.

5. *Session Layer*

The session layer guarantees the existence of a connection during a session. Types of sessions include point-to-point sessions, multicast sessions (a connection to many destinations) and multidrop sessions (a connection from many sources).

In the case of continuous media, *multimedia sessions* which reside over one or more transport connections, must be established. This introduces a more complex view on connection reconstruction in the case of transport problems. Another aspect is data coding. It is important for the presentation layer to know when an LDU is ready for presentation, especially if, for example, an intraframe was compressed.

6. *Presentation Layer*

   The presentation layer abstracts from different formats (the local syntax) and provides common formats (transfer syntax). Therefore, this layer must provide services for transformation between the application-specific formats and the agreed-upon format. An example is the different representation of a number for Intel or Motorola processors.

   The multitude of audio and video formats also require conversion between formats. This problem also comes up outside of the communication components during exchange between data carriers, such as CD-ROMs, which store continuous data. Thus, format conversion is often discussed in other contexts.

7. *Application Layer*

   The application layer considers all application-specific services, such as *file transfer service* embedded in the file transfer protocol (ftp) and the *electronic mail service*.

   With respect to audio and video, special services for support of real-time access and transmission must be provided. For example, in the case of an application such as video-on-demand, special services on the video server side for support of real-time database access and transmission must be developed [RVG+93].

The following section describes the principles of different networks according to the ISO-OSI Reference Model (i.e., physical layer, data link layer and their sublayers).

The other layers (i.e. network, transport, session, presentation and application layers) are discussed in Chapter 11. The presentation mainly includes issues regarding audio and video data transmission and the analysis of networks with respect to their multimedia capabilities. We will assume a fundamental understanding of networks, outlined in network textbooks such as [Tan88, Sta92, BG87, Pry93, Par94], ISO standards and many technical product descriptions.

10.2 Networks

The main goal of distributed multimedia communication systems is to transmit all their media over the same network. We will examine in the remainder of this chapter different networks with respect to their multimedia transmission capabilities.

Depending mainly on the distance between end-points (stations/computers), networks are divided into three categories: *Local Area Networks (LANs)*, *Metropolitan Area Networks (MANs)*, and *Wide Area Networks (WANs)*.

10.3 Local Area Networks (LANs)

A LAN is characterized by (1) its extension over a few kilometers at most, (2) a total data rate of at least several Mbps, and (3) its complete ownership by a single organization. Further, the number of stations connected to a LAN is typically limited to 100. However, the interconnection of several LANs allows the number of connected stations to be increased. The basis of LAN communication is *broadcasting* using *broadcast channel* (multi-access channel). Therefore, the MAC sublayer is of crucial importance in these networks.

10.3.1 High-speed Ethernet

Ethernet is the most widely used LAN. Currently available Ethernet offers bandwidth of at least 10 Mbps, but new fast LAN technologies for Ethernet with bandwidths in the range of 100 Mbps are starting to come on the market. This bus-based network uses the *CSMA/CD protocol* for resolution of multiple access to the broadcast channel in the MAC sublayer – before data transmission begins, the network state is checked by the sender station. Each station may try to send its data only if, at that moment, no other station transmits data. Therefore, each station can simultaneously *listen* and *send*. If multiple stations start to transmit data at the same time, the sending stations detect the collision by recognizing errors in their own data. Subsequently, stations try again to transmit the data after a randomly computed time, assuming that no other station has begun transmitting in the meantime.

The difficulty in building any CSMA/CD high-speed Ethernet on top of fiber optics is getting the collision detection to work. Several methods (e.g., power sensing, directional coupling) are possible using the passive star configuration [RJ85]. Fibernet II from Xerox [SRN+83] uses the active star configuration for collision detection.

The communication of continuous data requires guarantees of a maximal end-to-end delay and a minimal jitter. In principle, Ethernet cannot guarantee these requirements. However, there are several possibilities for using this type of LAN for audio and video transmission.

Handling of Continuous Data as any Other Data

In the first variant, continuous data are handled in the same way as other data. If the maximal utilization of the network is bounded, the number of errors due to delays is extremely small. This approach is currently the most used solution, but it is not satisfactory if the network load is high.

Dynamic Adaptation

To avoid the errors due to a high network load, continuous data transmission can be dynamically adapted to the network load. This means that if the network load is high, the data rate of continuous media can be reduced dynamically. This is possible with the help of scalable coding (see Section 6.7.4), but here again, transmission errors are possible.

Dedicated Ethernet

Another possibility for the transmission of audio/video data is to dedicate a separate Ethernet LAN to the transmission of continuous data. This solution requires compliance with a proper additional protocol. Further, end-users need at least two separate networks for their communications: one for continuous data and another for discrete data. This approach makes sense for experimental systems, but means additional expense in the end-systems and cabling.

A similar solution is comprised of one LAN as a digital network for transmission of control and discrete data and another LAN as an analog network for transmission of video and audio. This approach can be found in *Media Spaces*, created at Xerox PARC [BBI93] and in the *Cruiser Environment*, developed at Bellcore [FKRR93].

Hub

A very pragmatic solution can be achieved by exploiting an installed network configuration. Most of the Ethernet cables are not installed in the form of a bus system. They make up a star (i.e., cables radiate from the central room to each station). In this central room, each cable is attached to its own Ethernet interface.

Instead of configuring a bus, each station is connected via its own Ethernet to a *hub*. Hence, each station has the full Ethernet bandwidth available, and a new network for multimedia transmission is not necessary. Additional cost is created by the usage of the hub. In this solution, it is assumed that each Ethernet provides enough bandwidth to each station, but a file transfer may interfere with the audio and video transmission.

Fast Ethernet

Fast Ethernet, known as *100Base-T* offers throughput speed of up to 100 Mbits/s, and it permits users to move gradually into the world of high-speed LANs.

The Fast Ethernet Alliance, an industry group with more than 60 member companies, began work on the *100-Mbits/s 100 Base-TX* specification in the early 1990s. The alliance submitted the proposed standard to the IEEE and it was approved. During the standardization process, the alliance and the IEEE also defined a *Media-Independent Interface (MII)* for fast Ethernet, which enables it to support various cabling types on the same Ethernet network. Therefore, fast Ethernet offers three media options: 100 Base-T4 for half-duplex operation on four pairs of UTP (Unshielded Twisted Pair cable), 100 Base-TX for half- or full-duplex operation on two pairs of UTP or STP (Shielded Twisted Pair cable), and 100 Base-FX for half- and full-duplex transmission over fiber optic cable.

Like 10 Mbits/s Ethernet, 100 Mbits/s fast Ethernet can be configured in switched or shared-media implementations. Full-duplex operation requires a switch, which allows individual nodes to transmit and receive data simultaneously. Full-duplex fast Ethernet switches will effectively increase network throughput speeds to 200 Mbits/s [Mel94].

Fast Ethernet requires new workstation adapter cards, as well as new hubs or switches equipped with 100 Mbits/s transceivers. Not required are changes to existing Ethernet applications. Another of Fast Ethernet's strengths is that it can be added easily to shared-media and switched 10 Base-T (10 Mbits/s Ethernet, connected with TPC) networks.

10.3.2 Token Ring

The Token Ring is a LAN with 4 or 16 Mbits/s throughput. All stations are connected to a logical ring. In a Token Ring, a special bit pattern (3-byte), called a *token*, circulates around the ring whenever all stations are idle. When a station wants to transmit a frame, it must get the token and remove it from the ring before transmitting. Ring interfaces have two operating modes: *listen* and *transmit*. In the *listen* mode, input bits are simply copied to the output. In the *transmit* mode, which is entered only after the token has been seized, the interface breaks the connection between the input and the output, entering its own data onto the ring. As the bits that were inserted and subsequently propagated around the ring come back, they are removed from the ring by the sender. After a station has finished transmitting the last bit of its last frame, it must regenerate the token. When the last bit of the frame has gone around and returned, it must be removed, and the interface must immediately switch back into the *listen* mode to avoid a duplicate transmission of the data.

Each station receives, reads and sends frames circulating in the ring according to the *Token Ring MAC Sublayer Protocol* (IEEE standard 802.5). Each frame includes a *Sender Address* (SA) and a *Destination Address* (DA). When the sending station drains the frame from the ring, a *Frame Status* field is updated, i.e., the A and C bits of the field are examined. Three combinations are allowed:

- A=0, C=0 : destination not present or not powered up.
- A=1, C=0 : destination present but frame not accepted.
- A=1, C=1 : destination present and frame copied.

Most of the time, after reading a frame, the station sends the frame to a neighboring station because the DA of the received packet is not the same as the station's own address. If the DA is identical to the owner's address, this packet is stored in the local storage (as long as there is enough buffer space). If a frame is rejected at the receiver station due to a lack of a buffer space or some other reason, the sender has the option to try again (in a while).

The IEEE Standard 802.5 protocol includes an elaborate scheme for handling *multiple priority schemes.* The token contains a field which reflects the priority of the token. In a frame, a priority is indicated in the *Access Control (AC)* field of the frame header. When a frame is transmitted, it inherits the priority from the token that was captured, and the priority is stored in the AC field.

The *priority operation* works as follows: a station can transmit a frame at a given priority using any available token with a priority less than or equal to that of the frame. If an appropriate token is not available, the station may reserve a token of the required priority in a passing token or frames as follows:

- If another station has reserved an equal or higher priority in a passing token or frame, the station cannot make any reservation in this frame or token.

Priority	Application
0	Free usable; used by most applications
1-3	Free usable
4	Used by bridges
5-6	Reserved but not used
7	Used by ring management

Table 10.1: *Priorities used in Token Ring.*

- If the reservation bits have not been set, or if they have been set to a lower priority than that required by the station, it sets the reservation bits to its required priority.

When a station removes one of its frames from the Token Ring and finds nonzero values in the reservation bits, it must originate a priority token equal to the reservation bits. To prevent a station from continuously transmitting priority frames, the Token Ring provides a *fairness* scheme. Although a priority can be preempted at any time by the request of a higher priority, once the highest priority has been satisfied, the priority reverts to a lower priority, and eventually to a normal priority. Priorities are ordered according to their importance, as shown in the Table 10.1.

The *priority scheme*, together with the *fixed maximal propagation delay* of a frame, makes it possible to support a guaranteed data transmission of continuous media. The principal strength of this purpose is the *predictable* nature of a station's opportunity to transmit. The worst case, or the longest period of time that a station must wait until it gets a token, is deterministic and can be calculated. For example [BPSWL93], given a ring of N stations, the *token rotation time* t_{trt} can be characterized as

$$t_{trt} = \tau_l + \sum_{i=0}^{N-1} \tau_i$$

where τ_l is the *ring latency* (the fixed delay due to attachment delays and the physical ring length) and τ_i is the *token holding time* of station i for a given rotation of the token. The largest allowable frame size is 16 K bytes and the maximal token holding time τ_{max} is 10 ms. (In today's attachments, we usually encounter

$\tau_{max} = 2ms.$) Hence, in the worst case, a station may have to wait $(N-1) \times \tau_{max} + \tau_l$ seconds between consecutive opportunities to transmit data. Note that typically, τ_l is negligible compared to $(N-1) \times \tau_{max}$.

A reservation mechanism is useful for bounding the access delay of high-priority traffic. For example [BPSWL93], if at a given time there is only one station wishing to transmit a high-priority packet, the worst access delay is bounded by

$$t_{access} \leq 2 \star \tau_{max}$$

The upper bound is higher if there are multiple high-priority frames to transmit from different stations. Given M stations transmitting high-priority data, the worst case delay is bounded by

$$t_{access} \leq (M-1)\tau_{mm} + 2\tau_{max}$$

where τ_{mm} is the transmission time of high-priority multimedia data. The component $2\tau_{max}$ in the above expression means that the high priority frame may wait one round (τ_{max}) to reserve the token if the token's priority field is free, and then wait another round (τ_{max}) to grasp the token. The component $(M-1)\tau_{mm}$ means that, in the worst case, the *M-th* high-priority station must wait (*M-1*) rounds to reserve the priority token.

In this LAN, there are – similar to Ethernet – different possibilities for supporting a *continuous media transmission.* Some variants, described in the Ethernet section, can be adopted to the Token Ring. For example, scaling of media can be used in exactly the same manner. The following approaches can also be considered for continuous media transmission:

Using the Existing Priority Mechanism

The existing priority mechanism, without modification of any communication components, allows for the use of these LANs for continuous media transmission. The continuous data streams all get the same higher priority than discrete media streams.

However, with a high share of continuous media in the Token Ring traffic, undesirable delays of data frames occur. To achieve good performance, a simple form of bandwidth management is to limit the number of active multimedia sessions on a Token Ring segment. For example, when DVI-PLV multimedia data are transmitted with a high priority in a client/server application, it is possible to run at most 13 multimedia sessions independently of the data traffic on one Token Ring [ACT93].

Other results show that using the existing priority mechanism with allocation of separate queues at the end-point stations for different priority frames reduces the delay of multimedia traffic. Under heavy load, when the delay is most critical, the provision of multiple queues is essential to reduce the delay. Additionally, priority access mechanisms help to reduce the delay, particularly under light load conditions [BPSWL93].

Using Various Priorities for Continuous Media Streams

Within continuous media streams, priorities can be assigned according to the *importance* of the particular stream. For example, all audio streams are assigned a higher priority level than any other continuous media streams. For this approach, two higher priority levels are needed, for example, priorities 5 and 6 can be used. This approach obtains good results, although by many audio streams, failures can also occur.

Resource Reservation

It is possible to admit audio and video traffic of multiple streams at the same higher priority levels, but a resource reservation for these connections will also have to be performed. This can be done in at least three ways:

- *Static Resource Reservation*

  Resource reservation can be done *statically* by a predefined distribution of resources according a rule like "the 8 multimedia-capable stations can each send up to maximum of 1.2 Mbits/s of continuous media data". Only a small environment and applications with well-known throughput requirements can use this approach.

- *Dynamic Central Resource Reservation*

  As an alternative, each station can "talk" during connection establishment to central management to get an agreement and an assignment of tight resources. Finally, the data are transmitted according to this agreement.

  This central management can be, for example, a *bandwidth manager* which has all the necessary information to distribute the requested resource capacity (i.e., the required bandwidth). There are already suitable protocols for such a central bandwidth management. They might also allow for network resource reservation across LAN boundaries in the future, although the connection establishment phase will take a much longer time than it does in a LAN setup.

  The problem of resource reservation also touches on operating system resources. This area is discussed in Chapter 8 on multimedia operating systems. The reservation tests for non-interruptive planning methods with respect to the Token Ring are presented in [NV92]. They are applicable toward reservation of the operating system resources, as well as toward reservation of the network bandwidth.

- *Dynamic Distributed Resource Reservation*

  Distributed bandwidth management can allocate a required resource to the connection faster than a central solution. In the following paragraphs, a *time-optimized* mechanism is briefly described [Ste93a].

  Each multimedia station contains an internal table *Available Resource Table "art"*. This table includes records about the available network bandwidth, together with the bandwidth, already in use, for the ongoing audio and video connections. Hence, during the initialization of the Token Ring we can reduce the common available bandwidth to 80% of the real bandwidth (i.e., 3.2 Mbits/s for the 4 Mbits/s, respectively 12.8 Mbits/s for the 16 Mbits/s, early Token Ring release version). The remaining 20% includes the control traffic for the ring management (typically, this is around 3%), which has the highest priority (7) in the priority scale of traffic in the ring. This principle leads to sufficient capacity for other media. The 80% bound is an experience value, which can be adjusted according to the application and configuration scenario.

The *resource managers* of all multimedia stations belong to a group with a common address in the Token Ring. Therefore, each resource manager can send messages to all other managers using this address. Hereby, the *functional addresses* in Token Ring can be used.

In the first phase of a connection setup, the local resource manager compares the required capacity with the currently available capacity. If the bandwidth, recorded in *art* and actually available, is less than the required bandwidth, the connection request is immediately rejected. In the case of having sufficient bandwidth, which is recorded in the local *art*, at the own station, the resource request is delegated to the group of all other resource managers. For example, consider a bandwidth requirement (wish) of 1.41 Mbits/s. This requirement is sent to all members of the group. Let the available bandwidth at that moment be 10 Mbits/s. In the first step, the requesting station compares the required bandwidth with its own available bandwidth value (recorded in the local *art*). The required bandwidth can be provided and the value is modified in *art* from 10 Mbits/s to 8.59 Mbits/s. In the next step, the requesting station sends the resource reservation request with bandwidth value -1,410,000 (bits) to the group. This information is used by all other stations for comparison and modification of the locally stored bandwidth values in their *arts*.

After all other resource managers check their capacity, the frame with the resource request comes back to the requesting station, also including a response. If the response is positive, the required station knows that all members of the group are informed and their local *arts* are adjusted. In the case of a negative response of at least one station in the group, the resource request is rejected and all members of the group are informed. The possibility of rejection can occur because of a collision of simultaneous connection requests, issued by various stations attached to the early token release Token Ring. To avoid such an access collision concerning the available bandwidth, Ethernet-similar access methods should be used. For example, only three access trials are allowed, and they are distributed according to a particular statistic. Hence, the probability of collisions decreases dramatically. The protocol was optimized for establishment of a high-speed connection with the assumption that few collisions occur.

During the *connection close up*, all stations must be informed, and the values in their *arts* must be reset. In the example mentioned above, the value 1,410,000 is sent to the whole group.

The initialization of a new station in a ring happens through a query sent to all available *arts* of all other participants. If all other stations respond and send their *art* bandwidth values, and the values are all the same, the new station initializes its own *art* with this value. Inconsistences will be handled through an additional management protocol.

This method can be used for streams with a variable bit rate also. For the reservation, the maximal values are given. Yet, the resources will not be wasted because the gaps can be used by discrete data traffic.

Further, the distributed resource reservation method can be used not only for bandwidth reservation, but also for reservation of other resources within a ring. For example, the limited number of available *Functional Addresses* in Token Ring can be handled in a similar way.

The distributed resource reservation method assumes that each station is handled according to the bandwidth agreement. For violation cases, a separate control and monitoring component needs to be implemented.

The most important network resource to be managed is the *bandwidth*, which can be distributed using the above described methods. An increase in bandwidth allows a higher number of concurrent continuous media streams. Yet, a main criterion for the effectivity of the access protocol is the *bit length* compared to the size of the physical network. The following example shows: with bandwidth of 100 Mbits/s, a *bit* in fiber optics has the following "size":

$$Length = \frac{propagation\ speed\ in\ fiber\ optics}{data\ rate}$$

$$Length \approx \frac{2 \times 10^8 m/s}{100 \times 10^6 bits/s} = 2m/bit$$

On the other hand, a bit in a copper cable with 64 kbits/s bandwidth has the following "size":

$$Length = \frac{propagation\ speed\ in\ copper}{data\ rate}$$

$$Length \approx \frac{2.5 \times 10^8 m/s}{64 \times 10^3 bits/s} = 3.9 km/bit$$

In the Token Ring, the data are always read, written and sometimes copied and changed from and to the cable. Therefore, some bits are not "on" the cable, but "in" the attachment. Here, each station buffers at least 64 bits of the forwarded data. Therefore, with a higher data rate, several frames could be simultaneously on the Token Ring. The transition toward higher rates led to the introduction of the *Early-Token-Release-Principle.* This method allows stations to release the token before receiving their frame back. It means that after sending data, the sending station immediately releases the token again. Here, several frames may be on the ring. Due to the *Early Release* policy, the priority reservation mechanism may be less effective [BPSWL93]. In most cases, however, at a physical ring speed of 5 μsec/km, the ring must be very large to accommodate an entire frame. For example, for a 4 km ring of 20 stations, the ring latency τ_l is approximately 20 μsec, and can hold to 320 bits (at the 16 Mbps rate). Hence, in the worst case, a station wishing to transmit a priority frame may have to wait $N \times \tau_l$ seconds to gain access to the ring. This may only occur if all stations are sending frames which are smaller than the ring size and the reservation requests are essentially chasing the token around the ring. This, however, may be only a problem in very large rings with many stations.

10.3.3 FDDI

The *Fiber Distributed Data Interface (FDDI)* is a high-performance fiber optic LAN, which is configured as a ring. It is often seen as the successor of the Token Ring IEEE 802.5 protocol. The standardization began in the *American Standards Institute (ANSI)* in the group X3T9.5 in 1982. Early implementations appeared in 1988.

Compared to the Token Ring, FDDI is more a backbone than a LAN only because it runs at 100 Mbps over distances up to 100 km with up to 500 stations. The Token Ring supports typically between 50-250 stations. The distance of neighboring stations is less than 2 km in FDDI.

The FDDI design specification calls for no more than one error in 2.5×10^{10} bits. Many implementations will do much better. The FDDI cabling consists of two fiber rings, one transmitting clockwise and the other transmitting counter-clockwise. If either one breaks, the other can be used as a backup.

FDDI supports different transmission modes which are important for the communication of multimedia data (see Figure 10.2). The *synchronous mode* allows a

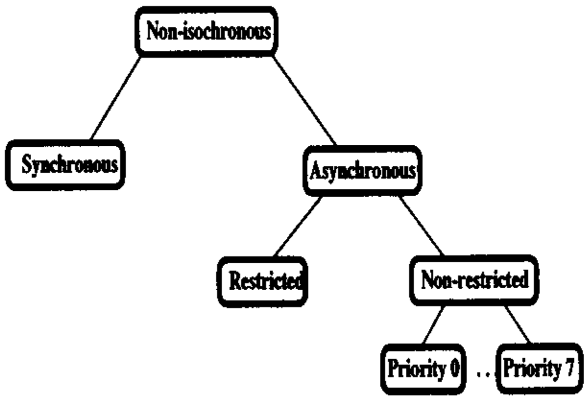

Figure 10.2: *Communication possibilities in FDDI: an overview of data transmission modes.*

bandwidth reservation; the *asynchronous mode* behaves similar to the Token Ring protocol. Many current implementations support only the asynchronous mode. Before diving into a discussion of the different modes, we will briefly describe the topologies and FDDI system components. The details are described in standard documents, as well as, for example, in [MK93].

Topology

The main *topology* features of FDDI are the two fiber rings, which operate in opposite directions (*dual ring topology*). The *primary ring* provides the data transmission, the *secondary ring* improves the fault tolerance. Individual stations can be - but do not have to be - connected to both rings. FDDI defines two classes of stations, A and B:

- Any class A station (*Dual Attachment Station*) connects to both rings. It is connected either directly to a primary ring and secondary ring or via a concentrator to a primary and secondary ring (see Figure 10.3).
- The class B station (*Single Attachment Station*) only connects to one of the rings. It is connected via a concentrator to the primary ring.

The concentrator-station can be connected to more than two stations. Further, it is always connected to the primary and secondary rings. Figure 10.3 shows a complete configuration with different stations. This is a dual ring with trees. In case of a

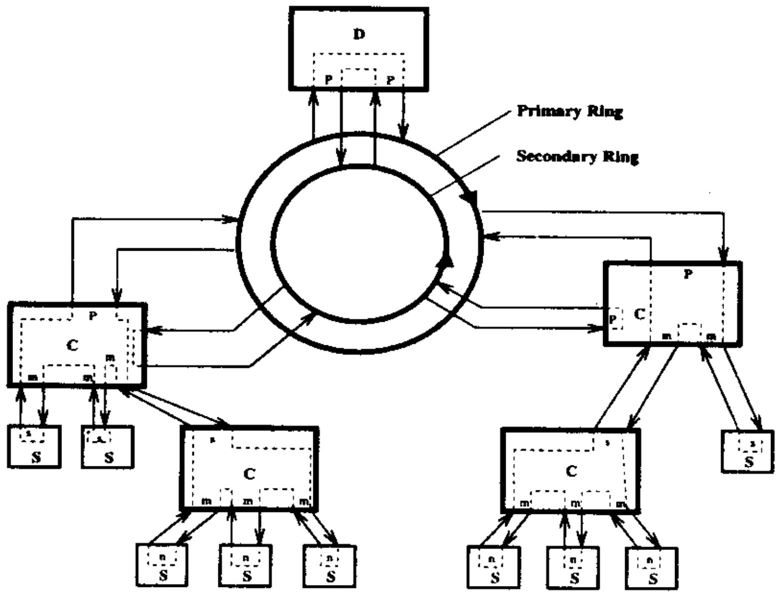

Figure 10.3: *A possible connection of different FDDI stations: a dual ring with several trees.*

failure, the FDDI ring will be newly reconfigured.

Reliability

In addition to the fault tolerance level provided by the dual rings (secondary ring), the *reliability* of an FDDI network is enhanced by the use of *station bypass switches* or *concentrators (CON)*. A station equipped with a bypass switch is "extracted" from the ring when the station experiences a power failure. A CON facilitates the connection of stations to the ring and it extracts from the ring any faulty station connected to it. The use of reliable CONs to interconnect stations is the heart of the *dual homing configuration*. In most cases, dual homing is more reliable than dual ring; however, the dual homing technique is beneficial only if the following conditions are met [Ngu93]:

- The number of dual attachment stations must be at least four.
- The link reliability should be at least $\sqrt{2}/2$, where the link reliability is defined as the difference probability: (1 - link error probability).
- The primary ring comprising CONs must be reliable.
- The number of CON pairs on the primary ring should be as small as possible.

Another approach to make complex FDDI networks more reliable is to use *creative topologies*, which means to set up configurations that reduce link vulnerability and/or to use redundant resources. Such a configuration approach is *a concentrator tree*. A concentrator tree is a method of interconnecting concentrators such that the A and B ports of each concentrator in the tree connect to M ports of the same concentrator higher in the tree. The concentrator tree alone is no more reliable than approaches such as dual-homing because damage to any set of links still results in isolation. Mechanisms which can survive more than two link failures are of interest to some applications. One technique involves a concentrator tree with connections between the top and bottom of the tree. This configuration is called the *concentrator tree with loop-back* [HM93].

FDDI Architecture

FDDI includes the following components which are shown in Figure 10.4:

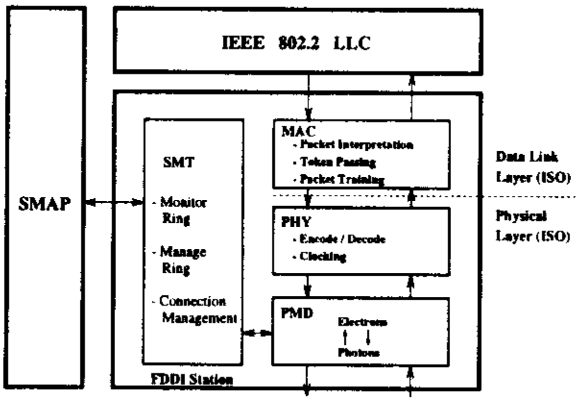

Figure 10.4: *FDDI reference model*

- *PHYsical Layer Protocol (PHY)*

  is defined in the standard *ISO 9314-1 Information Processing Systems: Fiber Distributed Data Interface - Part 1: Token Ring Physical Protocol.*

- *Physical Layer Medium-Dependent (PMD)*

  is defined in the standard *ISO 9314-3 Information Processing Systems: Fiber Distributed Data Interface - Part 1: Token Ring Physical Layer, Medium Dependent.*

- *Station ManagemenT (SMT)*

  defines the management functions of the ring according to *ANSI Preliminary Draft Proposal American National Standard X3T9.5/84-49 Revision 6.2, FDDI Station Management.*

- *Media Access Control (MAC)*

  defines the network access according to *ISO 9314-2 Information Processing Systems: Fiber Distributed Data Interface - Part 2: Token Ring Media Access Control.*

Hence, the physical layers *PMD* and *PHY* provide the transmission over fiber optics. FDDI uses multimode fibers with a diameter of 62.5μm because the additional expense of single mode fibers is not needed for networks running at only 100 Mbps.

However, it is also possible to take monomode fiber with a diameter of 125 μm. Further, it uses LEDs with wavelength of 1.320 nm, rather than lasers, not only due to lower cost, but also because FDDI may sometimes be used to connect directly to end systems. The physical layer uses for encoding a scheme known as *"4 out of 5"*: each group of four MAC symbols (0s, 1s and certain nondata symbols such as start-of-frame) are encoded as a group of five bits (four bits for encoding of the four MAC symbols, plus an additional *Non Return to Zero Inverted (NRZI)* bit). Therefore, the physical data rate increases to 125 Mbits/s.

The *SMT* functions are the *control, supervision* and *management* of the connected stations and the network itself. Here with respect to the stations, the initialization, activation and supervision of the performance and error handling activities are meant. With respect to the network, the main functions are addressing, reservation of the bandwidth and configuration.

The LAN access component, *MAC*, decides which station may access the ring. It provides address recognition functions and repeats, removes or inserts frames from/to the network. The length of the frame varies, but it is never longer than 4,500 bytes. The addressing allows, besides point-to-point communication, also *multicast* and *broadcast* communications.

FDDI Timed Token Rotation Protocol

The FDDI-specific access protocol is the *Timed Token Rotation Protocol.* This protocol provides support for two types of service: *synchronous* and *asynchronous.* Each station may be allocated a portion of the network bandwidth for its synchronous traffic. If a station receives the token, it can transmit messages in the synchronous mode for at least its preallocated time (called *synchronous capacity*) before releasing the token to its downstream neighbor. Messages in the asynchronous mode are transmitted only if certain time constraints are valid.

For the protocol, a *Target Token Rotation Time (TTRT)* is introduced. This value represents a typical and desired time for a ring round-trip of a packet. This value is set up during the initialization of the ring through a query to the SMT components of all stations. Each station stores this value. The TTRT must be, according to the

FDDI specification, more than 4ms and less than 165ms. A typical TTRT value is around 50ms. Such values are met, for example, in networks with high utilization and 75 connected stations using approximately 30 km of fiber optics.

Each station measures continuously the *real round-trip time* of a token and stores this time as the *Token Rotation Time (TRT)*. The TRT represents the last measured duration which, with respect to a particular station, a token needed for its last round trip. With respect to the asynchronous and synchronous traffics, the following considerations apply:

- *Asynchronous Traffic*

  The *asynchronous* traffic can be sent only if the network has free capacity. The criterion for this free capacity is the comparison between the last measured TRT and the preassigned TTRT. A station can always send asynchronous data if the following condition is true: $TRT < TTRT$. Hence, asynchronous traffic occurs in the network at most up to the time duration of TTRT.

  Some other approaches, describing how asynchronous traffic can coexist with synchronous traffic without too much delay penalty, are:

  - To transmit the non-real-time messages ahead of real-time messages, unless it is absolutely necessary to transmit real-time messages first to meet their deadlines [HR93]. The determination by a station, if it can defer the transmission of its real-time messages to a later time and still guarantee timely delivery of the real-time messages, is essential in this case.
  - To use a *Restricted Token* according to Figure 10.2. In this case, the total asynchronous bandwidth is reserved for dialogues between two stations as follows: the sending station informs the receiving station about a dialogue wish. This happens through normal (*non-restricted*) asynchronous traffic. The next step is the transmission of the corresponding packets from the sending station as additional packets together with the *restricted token*. At this moment, no other station is allowed to use the asynchronous bandwidth. The receiving station can continue the ongoing dialogue. This dialogue ends when the sending station removes the *restricted token* and places a *non-restricted token* on the network. The asynchronous

traffic with the *non-restricted token* distinguishes among eight priorities, which is similar to Token Ring.

- *Synchronous Traffic*

  To guarantee that the deadlines of synchronous messages are met, network parameters such as *synchronous bandwidth*, the *TTRT* and the *buffer size* must be chosen carefully [MZ93]:

  - The *synchronous bandwidth* is the most critical parameter in determining whether message deadlines are met. Through the SMT procedures, bandwidth for synchronous traffic is allocated to each station. The bandwidth allocation is called *Synchronous Allocation (SA)*.

    If the synchronous bandwidth is too small, the node may not have enough network access time to transmit messages to meet their deadlines. Conversely, large synchronous bandwidths can result in a long TRT, which can also cause message deadlines to be missed.

  - Proper selection of *TTRT* is also important because the time duration for all synchronous connections in their sum cannot cross the TTRT value. A smaller TTRT results in less available utilization and limits network capacity. On the other hand, if TTRT is too large, the token may not arrive often enough at a node to meet message deadlines.

  - Each node has a *buffer* for outgoing synchronous messages. The size of this buffer also affects the real-time performance of the network. The size of the buffer for incoming messages also affects the real-time network performance. The receiving node should be able to keep pace with incoming messages. A buffer that is too small can result in messages being lost due to buffer overflow. A buffer that is too large wastes memory.

The TRT value is maximally double the value of TTRT [SJ87]. For example, with a TTRT value of 50 ms, the maximal round-trip time is limited to 100 ms. The TRT is therefore a metric for the current utilization of the ring.

FDDI Audio and Video Transmission

For audio and video transmission in FDDI, the synchronous mode is very adequate. Yet the time spent on bandwidth reservation cannot be neglected. Dependent on the relation between TRT and TTRT during data transmission, we distinguish between the following cases:

1. *TRT* < *TTRT*

   Figure 10.5 shows this relation. The station can send data as long as the local

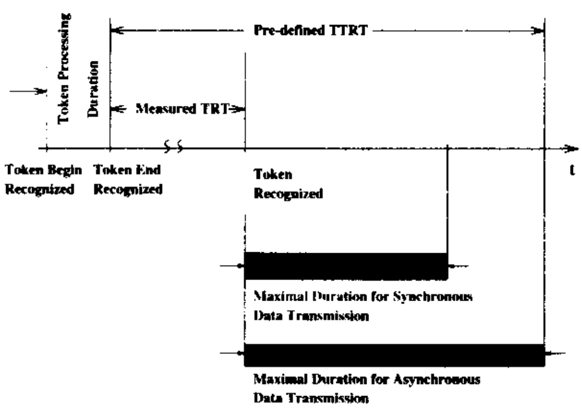

Figure 10.5: *Synchronous and asynchronous traffic on an FDDI with* $TRT < TTRT$.

   TRT counter does not cross the TTRT value.

2. *TRT* > *TTRT*

   Figure 10.6 shows this relation. The station can transmit yet only its syn-

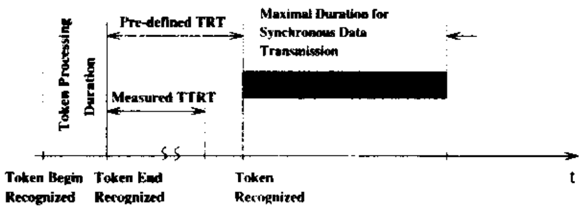

Figure 10.6: *Synchronous traffic on an FDDI with* $TRT > TTRT$.

   chronous traffic.

Next, an example of the *Timed Token Rotation Protocol* is presented. Let us assume that the TTRT takes eight time units and there are three stations connected to the FDDI ring. Each station has reserved one time unit (SA) for the synchronous traffic and sends according to the reservation. All stations would like to transmit as much non-restricted asynchronous traffic as the network admits. Table 10.2 shows the values of the corresponding TRT counters per station together with the transmitting units. After Station 1, Station 2 receives data, after that. Station 3 receives data and then again Station 1. Before the system is balanced, the measured TRT time units are of different values (e.g., 0, 9, 10). At each station, the decision concerning which data to send is determined as shown in Figures 10.5 and 10.6. It means that if $TRT < TTRT$, synchronous and asynchronous traffic can be sent. If $TRT \geq TTRT$, only synchronous data units are sent to the other station. This process repeats until the system is balanced. For example, at the beginning, where measured TRT is 0 time units and pre-defined TTRT is 8 time units, 1 synchronous data unit and 8 asynchronous data units can be sent. After the system is balanced, the behavior from Figures 10.5 and 10.6 is applied further, i.e., Station 1 sends 1 synchronous data unit to Station 2, where TRT is 8 time units ($TRT = TTRT$), Station 2 sends 1 synchronous data unit to Station 3, where TRT is 8 time units, and Station 3 sends 5 asynchronous data units and 1 synchronous data unit to Station 1, where TRT is 3 time units ($TRT < TTRT$), and so on. The joined relations, which also demonstrate the proper distribution of the asynchronous traffic, are the result (Table 10.2).

Further Properties of FDDI

- *Multicasting*

  The multicasting service became one of the important networking services to support multimedia applications, such as cooperative applications. It is very advantageous to have multicasting as part of the network lower layers, such as the data link layer or the MAC layer, because the higher layers do not have to spend any time and resources for multiplying (making copies) data to be sent. Therefore, overhead transmission delays are avoided.

  FDDI supports *group addressing*, hence multicasting is part of the network. We discuss the multicasting service in Chapter 11, where communication sys-

Station 1			Station 2			Station 3		
TRT	*syn*	*asyn*	*TRT*	*syn*	*asyn*	*TRT*	*syn*	*asyn*
0	1	8	9	1		10	1	
11	1		3	1	5	8	1	
System is balanced, first cycle starts:								
8	1		8	1		3	1	5
3	1	5	8	1		8	1	
8	1		3	1	5	8	1	
Cycle starts again:								
8	1		8	1		3	1	5

Table 10.2: *Example: TRT values per Station*

tems above the data link layer are described. In this case, the multicasting service is not part of the lower layers of the network, hence other solutions must be implemented.

- *Synchronization*

  The synchronization among different data streams is not part of the network, therefore it must be solved separately. Here, the relation between synchronous and asynchronous data is of interest. A tight time relationship, existing at the sender, does not have necessarily to exist at the receiver side because of the Timed Token Rotation Protocol.

- *Packet Size*

  The size of the used packets can directly influence the data delay in applications. For example, in the case of speech/audio transmission with an 8KHz sampling rate and a data rate of 64 kbits/s, the audio packets must be collected until an FDDI frame is complete. Here, the desirable FDDI frame size would be very small.

- *Implementations*

  Unfortunately, many FDDI implementations do not support the synchronous mode, which is very useful for the transmission of continuous media. In asynchronous mode additionally, the same methods can be used as described previously by Token Ring. In the case of an available synchronous mode, the continuity of the asynchronous traffic can be disrupted. Therefore, guarantees, as described by Token Ring, can be given only if no station uses the synchronous mode.

- *Restricted Token*

  If only two stations interact by transmitting continuous media data, then one can also use the asynchronous mode with *Restricted Token*. This leads to small end-to-end delays, but it inhibits any other asynchronous traffic on the LAN. Hence, this mode should be only used for the transmission of continuous media, if no other station requests such a service.

FDDI-II

The *synchronous mode*, in general, requires non-negligible times for the reservation of the synchronous bandwidth. Further, the synchronous mode guarantees a bandwidth with a maximal delay. This delay does not exceed double of TTRT value. Of course, data in the synchronous mode can arrive much earlier at the receiver also. The jitter depends heavily on the utilization of the ring. Therefore, the delay is bounded by a maximal value, which can reach 100 ms, and may play an important role in dialogue applications. It means, additionally, that the early arrived data must be stored. For this purpose, a buffer for the duration of at least one TTRT is needed. This argument led to the introduction of an additional *isochronous mode* in *FDDI-II* (see Figure 10.7).

The design of FDDI-II started in 1984 and was planned as an addition to the original FDDI. FDDI-II is a hybrid High-Speed Local Area Network (HSLAN). It integrates *circuit switched services* (isochronous services) for delay-sensitive applications, such as voice and video, and *packet services* (synchronous and asynchronous services), as available in basic FDDI on the same physical medium. Although the packet switched service remains connection-less, the circuit switched service is connection-oriented. Figure 10.8 illustrates the relationship between FDDI-II and the OSI reference model.

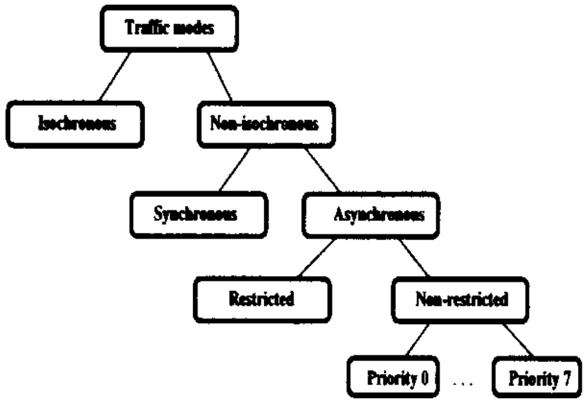

Figure 10.7: *Communication possibilities in FDDI-II: overview of the transmission mode.*

The *FDDI-II Hybrid Ring Control (HRC)* protocol integrates isochronous and asynchronous packet data by carrying both in special fixed-length, fixed-duration frames called *cycles*. The HRC consists of two additional data link layer entities, the *Hybrid MUltipleXer (HMUX)*, and the *Isochronous MAC*, in addition to the MAC entity of FDDI.

FDDI-II networks have two major modes of operation: *basic* and *hybrid*. The basic mode provides packet data services and is "compatible" with the FDDI protocol. Note, "compatible" does not mean that FDDI and FDDI-II stations can simultaneously operate on the same ring. Following initialization in the basic mode, FDDI-II networks may switch to hybrid mode operation, where packet and isochronous services are both provided within the FDDI-II cycle structure. In steady-state hybrid mode operation, one station is designed to be the *cycle master*, which is responsible for generating all cycles on the ring, and for maintaining cycle timing. All other stations, called *slaves*, repeat incoming cycles for a transmission onto the ring after inserting their own packet and isochronous data into the cycles. A station which is capable of becoming the cycle master and generating the cycles, is classified as the *monitor station*. Otherwise, the stations are classified as *non-monitor stations*. In hybrid mode, a new cycle is generated by the cycle master every 125μs with the cycle length of 125 μs[BK93]. Cycles circulate around the ring and return to the cycle master. Multiple cycles can be present on the ring at the same time.

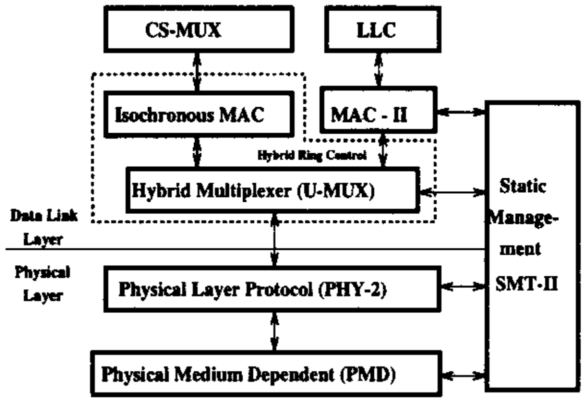

Figure 10.8: *FDDI-II reference model.*

A cycle is composed of four parts: a *preamble*, a *cycle header*, a *Dedicated Packet Group (DPG)*, and 96 *cyclic groups*. Collectively, the DPG and cyclic groups are referred to as the *cycle body*. The cycle header contains all of the peer-to-peer information. Sixteen time slots, known as *Wide Band Channels (WBCs)*, are available in each cycle. These slots are byte-interleaved across the 96 cyclic groups. A single WBC allows data transfer at 6.144 Mbps. This value (6.144 Mbps) is the smallest common multiple of the American (four times) and European (three times) primary rate of the Narrowband ISDN (American Narrowband ISDN rate is 1.536 Mbps, European Narrowband ISDN rate is 2.048 Mbps.)[TG90]. WBCs can be allocated *individually* or *combined*, and they can be used between two or more stations as the transmission channel for duplex connections.

Each WBC can be dynamically allocated for either the packet or isochronous data by the cycle master through updating the *Programming Template*, which is transmitted in each cycle according to the WBC request from SMT. An isochronous WBC can split the available bandwidth into multiples of 8 kbits/s as virtual connections. Therefore, different channels can be implemented, for example, 16 kbits/s, 64 kbits/s, 128 kbits/s, 1,536 kbits/s, 2,048 kbits/s. The maximum number of WBCs for isochronous traffic is set up during the initiatization phase. If all WBCs are used for isochronous data in the 100 Mbps bandwidth range, 98.304 Mbps bandwidth would be available for this class of traffic [KJ93].

For *packet transmission*, DPG guarantees a minimal channel bandwidth. It is, in turn, concatenated with any WBC, which is allocated to packet data to form a *single packet data channel*. The single packet data channel provides a virtual ring over which the FDDI timed token protocol operates.

For *isochronous data transmissions*, the channel bandwidth is centrally allocated on a demand basis, using a connection-oriented service. The station wishing to establish an isochronous connection must send a connection setup request to the cycle master prior to a conversation. Then it must issue a termination request upon finishing the conversation. The signaling, required for setting up isochronous connections, would normally be carried out in-band over the packet data channels. Consequently, using such a connection-oriented and centralized access control, FDDI-II is able to fairly and accurately provide bandwidth access and control. However, it reveals several performance weaknesses, such as bandwidth waste in HSLAN, degradation of packet data transmission performance, and finally, upon finishing conversations, the freed channels remain idle and unused for an average of half the propagation delay of the ring before being reassigned by the cycle master [YH93].

The alternative would be to provide distributed control and access of bandwidth using a connectionless service for both packet and isochronous data transmission. This access scheme is implemented in $FDDI-II^*$ [YH93].

FDDI-II is very good for the transmission of continuous data because of the isochronous service. The commercial relevance of FDDI-II as opposed to the original FDDI is still doubtful. One reason may involve the incompatibility of both systems. Existing FDDI systems cannot be connected to an FDDI-II, they must be replaced.

10.3.4 Local ATM Networks

In this section we first briefly discuss ATM characteristics and the ATM architecture, and then we present local ATM network issues. In Section 10.5 we discuss the switching properties of WAN ATM networks.

ATM Characteristics

The *Asynchronous Transfer Mode (ATM)* is a cell (packet) switching concept with minimal function in network. The ATM concept uses a *fixed length packet (cell)* of size 53 bytes - 48 bytes of user data and a 5 bytes header (see Figure 10.9). ATM

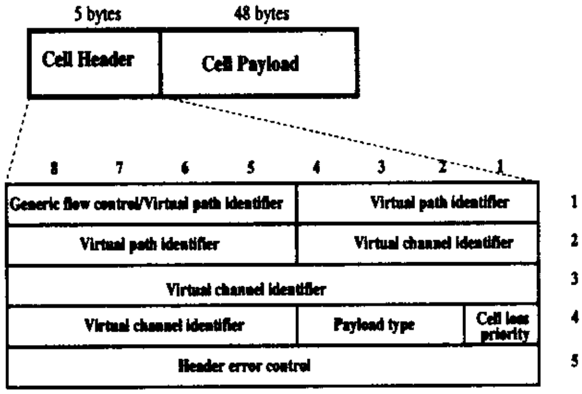

Figure 10.9: *ATM cell.*

allows the systems to operate at a much *higher rate* than usual packet switching systems. The high rate is achieved because of the following characteristics:

1. *No error protection or flow control on a link-to-link basis.*

   If a link of the connection, either the user-to-network link or a link between two ATM switches, introduces an error during transmission, or is temporarily overloaded thereby causing the loss of packets, no special action is taken on that link (e.g., no request for retransmission is made on that link, as is common in packet switching). It means nothing is done to correct errors inside the network. The ATM network relies on end-to-end protocols.

2. *ATM operates in a connection-oriented mode.*

   Before data are transferred from a multimedia terminal to the network, a logical/virtual connection setup phase must allow the network to perform a reservation of the necessary resources. If no sufficient resources are available, the connection is refused and the requesting terminal is notified. After the end of the data transfer phase, the allocated resources are released. This connection-oriented mode allows the network to provide (in all cases) a minimal packet loss ratio.

3. *The header functionality is reduced.*

   To guarantee fast processing in the network, the ATM header has a very limited function. Its main function is the identification of the *virtual connection* by an identifier, which is selected at a call setup time and is guaranteed a proper routing of each packet in the network. In addition, it allows an easy multiplexing of different virtual connections over a single link.

   In addition to the *Virtual Connection Identifier (VCI)*, a very limited number of functions are supported by the header, mainly related to maintenance. Due to the limited functionality of the header, the implementation of the header processing in the ATM nodes is simple and can be done at very high speeds (150 Mbits/s up to Gbits/s). This results in low processing and queuing delays.

4. *The information field length is relatively small.*

   To reduce the internal buffers in the switching nodes, and to limit the queuing delays in those buffers, the information field length is kept relatively small. This assures a small end-to-end delay caused by the packetization of audio and video data into ATM cells.

The ATM concept guarantees the possibility of transporting any data independent of its characteristics such as the bit rate, its quality requirements or its bursty nature. This means that the ATM networks are suitable, with their speed and bandwidth characteristics, to offer a transmission service for the multimedia traffic with different quality requirements. This capability was one of the main motivations for CCITT to decide that ATM will be the transfer mode of the future *Broadband-ISDN*, which represents the concept for high-speed WANs.

ATM Architecture

The same logical hierarchical architecture as used in OSI is used for ATM B-ISDN network in the CCITT Standard I.321 (see Figure 10.10). Only lower layers of ATM

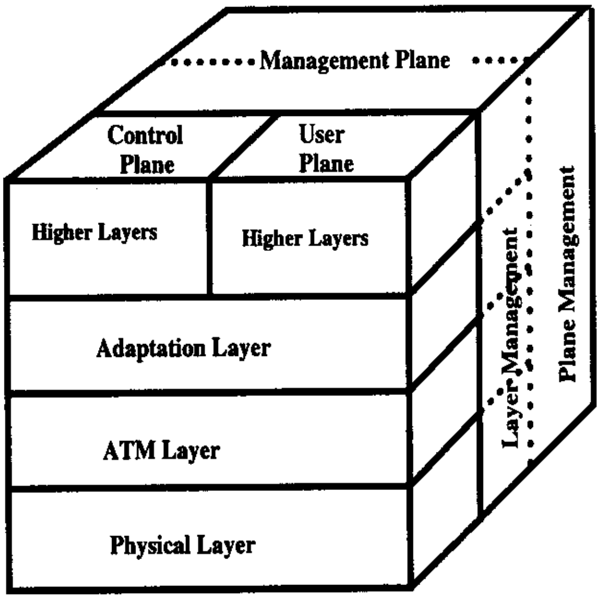

Figure 10.10: *B-ISDN ATM protocol reference model.*

B-ISDN are specified by CCITT. The B-ISDN protocol model for ATM contains three planes: *a user plane* to transport user information, *a control plane* mainly composed of signaling information and *a management plane*, used to maintain the network and to perform operational functions. For each plane, a layered approach is used with an independence between the layers. According to CCITT, there are three lower layers:

- *The PHYsical Layer (PHY)*

  carries data (bits/cells). It is composed of two sublayers: *the Physical Medium (PM)* sublayer which supports pure medium-dependent bit functions, and *the Transmission Convergence (TC)* sublayer, which converts the ATM cell stream into bits to be transported over the physical medium.

  The preferred medium for the full-duplex 155.52 Mbps service, as well as 622.08 Mbps service, is the *optical fiber*.

  The transmission structure (for 155.52 Mbps data rate) used to multiplex cells from various virtual channels can be either a *continuous stream of cells* with no multiplex frame structure imposed at the interface, or a *placement of cells in a synchronous time-division multiplex envelope*. In this case, the bit stream at the interface has an external frame based on the Synchronous Digital Hierarchy (SDH). In the U.S., this frame structure is referred to as SONET (Synchronous Optical NETwork). The SDH standard G.709 defines a hierarchy of data rates, all of which are multiples of 51.84 Mbps.

- *The ATM Layer*

  is fully independent of PHY. The main functions are:

  - *Multiplexing* and *demultiplexing* of cells of different connections into a single cell stream on a physical layer.
  - *Translation* of the cell identifier.
  - *Provision with one of the service classes.*
  - *Management functions.*
  - *Extraction (addition)* of the cell header.
  - *Implementation of a flow control mechanism* at the user-network interface.

- *The ATM Adaptation Layer (AAL)*

  enhances the service provided by the ATM layer. It performs functions for the user, control and management planes and supports the mapping between the ATM layer and the next higher layer. AAL consists of a *Segmentation And Reassembly (SAR)* sublayer, where the higher layer information is segmented into ATM cell size stream at the sender, and the cell stream is reassembled into the higher layer data units at the receiver. AAL further includes the *Convergence* Sublayer (CS) with functions like message identification, time/clock recovery, etc.

  To accommodate various services (e.g., voice, video, data, ...), several types of AAL have been defined. Up to now, five AALs have been defined, which include support for connection-oriented, connectionless, Variable Bit Rate (VBR) and Constant Bit Rate (CBR) services:

- *AAL1* supports CBR services after a virtual connection is established. This service class is advantageous for high-quality constant bit rate audio and video.
- *AAL2* offers a transfer of data with a VBR. In addition, timing information is transferred between source and destination. In the CS sublayer, the following functions are performed: *clock recovery* (e.g., time stamps), *handling of lost or misdelivered cells* and *Forward Error Correction (FEC)* for audio and video services.
- *AAL3/4* should be used for transfer of data which are sensitive to loss, but not to delay. This AAL may be used for connection-oriented, as well as for connectionless, data communication (e.g., multimedia file transfer, multimedia e-mail).
- The ATM Forum defined a different AAL for high-speed data transfer (e.g., transaction transmission), called *AAL5*. An AAL5 packet has far less overhead than AAL3/4. Further, this layer minimizes the computer's cost in handling cells and behaves like data communications interfaces for Ethernet and FDDI (LAN emulation), so that existing data communication software can easily be ported to support ATM.

ATM Cell Information

ATM services and capabilities are achieved using the information in the ATM cell header. Hence, we describe some header information which is important for the support of multimedia data transmission:

- *Virtual Connections*

  Functions such as the source, destination address and sequence number are not required to be part of an ATM cell. Every virtual connection is identified by a number, which has only a local significance (per link). The virtual connection is identified by two subfields of the header: *Virtual Channel Identifier (VCI)* and *Virtual Path Identifier (VPI)*.

- *Virtual Channels*

Since the ATM network is connection-oriented, a VCI is assigned to each connection during the connection setup procedure . A VCI has only a local significance at the link between two ATM nodes. In each ATM node along the path between the sender and the receiver, the VCI is translated. If the connection is released, the VCI values of the involved path are released and can be reused by other connections.

An interesting advantage of this VCI principle is the use of multiple VCI values for multicomponent services (e.g., video telephony, TV etc.). For example, a video telephone call may use three communication streams: voice, video and data, each of which will be transported over a separate VCI. It allows the network to add or remove streams during transmission. This means that the video telephony service can start with voice only (single VCI) and the video stream can be added (and removed) over a separate VCI later on. The signaling for managing the particular connections is transported over a separate VCI. This VCI principle has an implication to call establishment and management in the higher layers, where a video telephone call between sender and receiver (one logical call connection) can be mapped to three virtual connections, dependent on which media the sender or receiver wants to use for communication

- *Virtual Path*

  Future broadband networks will support semi-permanent connections between end-points. This concept is known as a *virtual path* or a *virtual network*. To perform this virtual network, another field is defined in the header *Virtual Path Identifier.*

- *Priorities*

  The ATM header supports a differentiation of logical connections through different priorities. Two types of priorities exist: *time priority* and *semantic priority*. In the time priority system, it is assumed that some cells may remain longer in the network than others. In a semantic priority system, some cells have a higher probability of being lost, giving rise to a higher cell loss ratio.

  Priorities can either be assigned on a per connection basis (per VPI or per VCI), or on a per cell basis. In the first option, all cells in the virtual channel/path have the same priority. In the second case, cells within a virtual channel/path may have different priorities. No priorities at all lead to the best sharing of resources, but do not allow a differentiation between services with different quality requirements. With priorities applied to continuous media streams, guarantees are provided on a statistical basis with a very high probability.

- *Maintenance*

  To maintain the overall network and to provide performance monitoring (which implies a monitoring of the media quality) of the ATM connections, some additional bits are useful. The *Payload Type Identification (PTI)* distinguishes, using the PTI information field, which type of data (e.g., video, data, control, ...) is carried in the ATM cell.

- *Multiple Access*

  In addition, to allow multiple terminals (users) to be connected to the same physical link, a point-to-multipoint protocol is specified. This protocol is useful, for example, for video-conferencing applications. To perform this point-to-multipoint function, the *Generic Flow Control (GFC)* field in the ATM header is used. GFC provides the possibility to negotiate shared access to the network at the *User-Network Interface (UNI)*, i.e., how to multiplex the shared network among the cells of the various ATM connections.

- *Header Error Protection*

  The header of the ATM cell is the most sensitive part to be protected against corruption. Therefore, it should be protected against single bit errors and, if possible, also against burst errors. To protect the header, a coding principle based on a generalization of Hamming codes is very suitable, namely the BCH codes (Bose-Chadhuri-Hocquenghem) [Pry93].

ATM Connectivity between Traditional LANs and ATM Networks

The ATM WANs will enforce the LANs to provide for an *ATM connectivity* if they want to internetwork over long distances.

There are several possibilities to have an ATM connectivity of a traditional LAN (e.g., Ethernet, FDDI, Token Ring) to the ATM networks:

- to introduce additional functionality at the gateways between LAN and ATM WAN and at the *ATM hubs* (bridges) between traditional LANs and ATM LAN, Figure 10.11 shows a possible topology. ATM intelligent hubs support

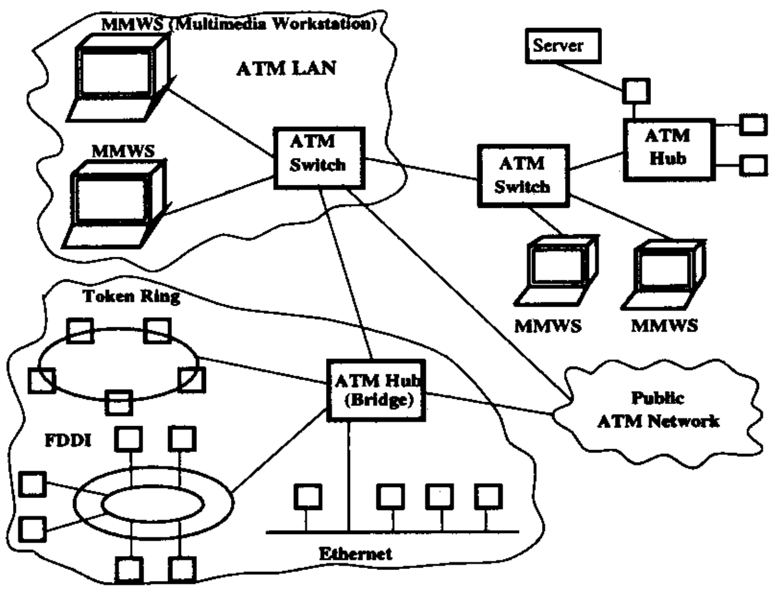

Figure 10.11: *ATM connectivity of traditional LANs and an ATM LAN.*

  multiple LAN segments of multiple LAN types (e.g., Ethernet, Token Ring and FDDI) and provide bridge and routing functionalities. Further, they give every terminal the full capacity by relying on the high switching capabilities of the hubs.

- To implement additional ATM software at the end-points above the traditional LAN protocol stack [AA93].

  For example, between two Ethernet interfaces, the ATM concept can be emulated. The ATM cells are encapsulated inside the Ethernet packets. A network architecture, consisting of at least one LAN/fiber switch node whose job is to switch cells between in-house physical LAN links and an external B-ISDN fiber, is proposed in [AA93]. A possible structure of an ATM/Ethernet interface at a multimedia end-user workstation is shown in Figure 10.12. There

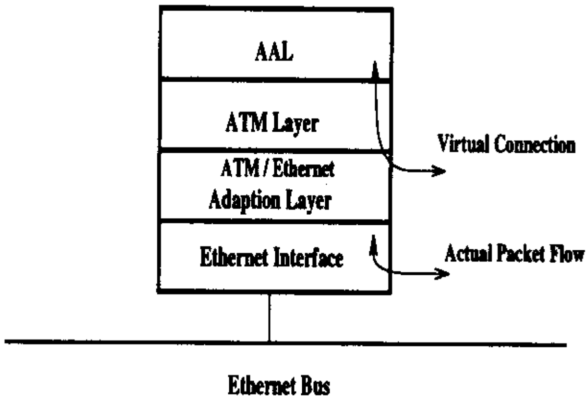

Figure 10.12: *Example: ATM connectivity in Ethernet.*

may be a decrease in throughput, but considering the connectivity to a wide range of multimedia services over the WAN, it is certainly worthy to take into account this "speed penalty".

Local ATM Network

An ATM LAN connects ATM multimedia terminals, i.e., workstations or servers with an ATM host interface, either via a private UNI (User Network Interface) with a private ATM switch, or via a public UNI with a public ATM network, as shown in Figure 10.13. A simple star configuration is used; hence, no MAC protocol must be defined, as is done in other HSLAN systems. Hence , for this star architecture, only a point-to-point interface is defined.

An ATM LAN can be used to cover a small geographical area as well as a large one. It can have a small number of terminals or a large number of terminals just by dimensioning the size of the ATM LAN switch, and by using interfaces with short or long distances. There are some differences in the service requirements between ATM LAN switches and ATM switches for B-ISDN:

- An ATM LAN switch may provide only a limited number of ports (typically

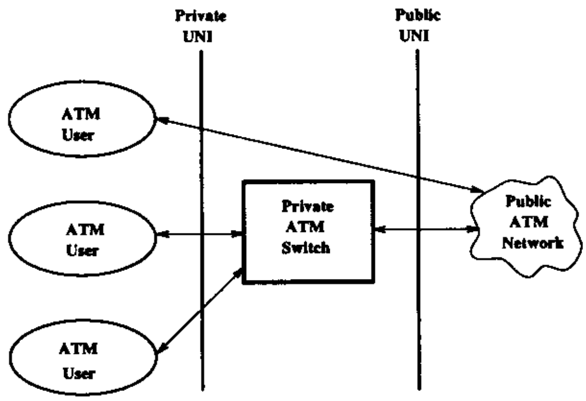

Figure 10.13: *Local ATM network.*

below 1024 ports) due to the fact that the amount of attached stations in a LAN environment is limited.

- Delays inside the switches themselves will dominate the link delays. To provide a total delay comparable to traditional LANs, such switches must minimize this value. Therefore, switching delay is much more important in ATM LANs than in comparable WANs.

- Several LAN applications rely on data to be exchanged between all or a set of stations. Therefore, an ATM LAN switch must provide capabilities for both multicasting and broadcasting.

- Reliability is crucial for WANs which carry voice data. In ATM LANs, we may not encounter the same stringent requirements.

- The connection establishment in a WAN is similar to the telephone paradigm where we are accustomed to wait for some time. This is difficult in LANs. There, we encounter an immediate data transmission. Hence, ATM LANs should have a very short connection establishment.

A high-performance, self-routing ATM LAN switch is proposed in [TYL93]. We discuss ATM switching, routing, and multicasting properties in Section 10.5. because these properties are common for ATM LANs and ATM WANs.

The end-points of ATM LANs require high-performance end-to-end protocols because the delays in the end-points will dominate link and ATM LAN switch delays. As mentioned above, it is proposed to use the AAL5 to carry signaling data. For user data, the AAL1/2 should be used for voice/video services, AAL3/4 for the other services.

10.4 Metropolitan Area Networks (MANs)

Between LAN and WAN we have the *Metropolitan Area Network (MAN)*, which covers an entire city, but uses LAN technology. MAN uses a shared medium with distributed switching and Medium Access Control (MAC). MANs have generally higher data rates than LANs, i.e., more than 100 Mbits/s. The administration can be either public or private. The number of stations connected in a MAN is mostly around thousands. A MAN's services include:

- *The interconnection of different LANs.*

  For *LAN interconnection*, a gateway/bridge performs functions such as protocol conversion, address mapping, access control, etc. depending on the compatibility of the interconnected LANs. Since most LANs operate in connectionless mode, it is appropriate that the MAN also interconnects these LANs in a connectionless way. This means that no resources are usually allocated in the MAN. However, continuous media transmission requires guarantees which can easily be provided by a resource allocation.

- *Host-to-host computer internetworking.*

  MAN host-to-host networking can be supported by [Pry93]:

  1. Providing a semi-permanent point-to-point connection with high throughput and high reliability. In this case, a connection is established at installation time, and ensures that enough resources are available on the MAN. Therefore, the mode of operation is connection-oriented, and resources are assigned semi-permanently.
  2. Offering a number of isochronous slots, requested on demand by signaling. This solution is more comparable to a circuit switched solution, where a

TDM (Time Division Multiplexing) approach is taken. This functionality may not be available in all MANs.

3. Offering a number of non-isochronous slots. In this case, resources of the MAN are only occupied when data must be transported (on-demand). This type of service is provided in the connectionless mode.

- *Voice and video communication.*

  For voice and video services, the three alternatives described in host-to-host internetworking can be used depending on the quality of requirements. If voice and video are offered via alternatives 1 or 3, the jitter must be removed at the receiving terminal.

In addition to these services, functions like broadcasting and multicasting, etc. can be offered.

There are two main proposals for MAN standards: (1) *FDDI* from ANSI, which was initially proposed as a high-speed LAN but achieved a span of up to 100 km, and (2) *Distributed Queue Dual Bus (DQDB)* from the group IEEE 802.6. In Europe, an additional MAN mechanism, called *Orwell*, has been considered and developed by researchers of British Telecom. However, it has not been retained by a standards body. We described the FDDI mechanisms already in Section 10.3.3, therefore in this section we concentrate on the DQDB and Orwell mechanisms, as well as on MAN connectivity to ATM B-ISDN.

10.4.1 Distributed Queue Dual Bus (DQDB)

The MAC mechanism using a token is not effective anymore in the data rate range above 100 Mbits/s and a geographic range of a network above 100 km [RB90]. The reason is the ring or bus latency, which requires the token to circulate to the next active station and along the entire ring or bus. Therefore, if the geographic range of a network is large, it takes a long time until the token comes back to the sending station, and a release of a token occurs. If the packet is large, the sending station does not release the token until it sends all the data which may occupy several frames, hence the latency is equal to multiple TRTs.

This argument supported a development of a further network which originally became known as *Queued Packet Synchronous Exchange (QPSX)* [NBH88]. This network was a cooperative effort of the company QPSX, the University of West Australia and Telecom Australia. Because of the naming collision between the standard name and the company name, this network was renamed and is now the *Distributed Queue Dual Bus (DQDB)*.

This IEEE 802.6 MAN standard is characterized by a bus with a data rate of 2×150 Mbits/s. It uses different kinds of cables. DQDB is based on two counter rotating symmetric bus-systems as shown in Figure 10.14. In contrast to FDDI,

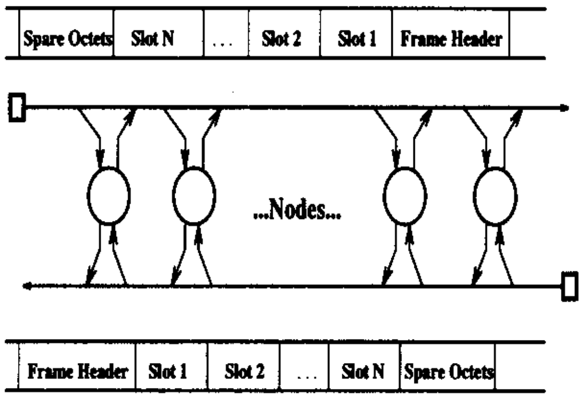

Figure 10.14: *Distributed Queue Dual Bus - principal topology with standard names.*

both busses carry data and therefore, are not only used for fault tolerance. The data are transmitted on the bus - analog to FDDI-II - in 125 μs frames. Each frame includes further time slots of fixed length, called *segments.* The time slot transports the data between the nodes. It means that the bus capacity is allocated in slots of 53 bytes to *Access Units (AUs)* via which user terminals gain access to the network. Slots are generated at the *head* of each bus.

The DQDB MAC mechanism is fundamentally different from most of the other LAN/MAN protocols with shared access. In these other protocols, there is no continuous record kept of the network state in the nodes. In those networks, state information must be derived from the medium first before accessing it. This feature makes these systems' performance very sensitive to the size of the network. For example, in FDDI, the performance is negatively influenced by the length of the ring. DQDB is based on the *distributed queuing algorithm*, where the current state information of the network is stored in each node; namely every node knows the exact number of segments still waiting to be put on the bus. If a node has a segment to transmit, it uses this local information (stored in a counter) to determine the position of the segment.

The DQDB standard includes one complete specification of the *connectionless packet service* for asynchronous data transmission. The *isochronous service* is nearly completed. Work on the so-called *Guaranteed Bandwidth (GBW) protocol* for *connection-oriented data service* has just begun.

Connectionless Packet Service

The *asynchronous data transmission* is implemented by the *Queued Arbitrated (QA) Access* function. QA accesses use the MAC mechanism, which is based on *distributed queues*. The system (i.e., one bus) has logically one queue for the storage of data to be transmitted in one direction, but this queue is physically distributed over the nodes. The service discipline of the queue is FIFO. Hence, in a DQDB system, there are two logical queues available, one for the traffic in each direction. One bus is used to coordinate the queuing discipline of the segments to be transferred over the other bus, and vice versa. This coordination is achieved by collecting requests from the nodes when these nodes have segments to transmit. The following four steps explain this process taking the upper bus, shown in Figure 10.14, as an example for data to be transmitted:

1. Each *listening station* (node in Figure 10.14) counts all, from the right, requests coming in on the lower bus. These requests are stored in the internal queue. Now, if from the left on the upper bus, a free time slot arrives, the oldest request in the queue is satisfied. Hence, the request entry can be deleted from the queue. Afterwards, the queue contains all, further right, open "send requests" which must be transmitted on the upper bus. The same principle is applicable to the lower bus with left-ordered stations and with a second queue.

2. A *station, willing to send, issues a "send request"*. To perform this action, the station waits for a free request field at the lower bus. If such a free request field arrives at the lower bus, the station makes its request by marking the *Request Field* on the lower bus and it queues its own "send request" in the internal queue for the upper bus.

3. The *station, willing to send, waits to send data.* It can always store, at most, one of its own requests, according to the above described principle, in the queue. Further, it must continuously trace the bus and take away the requests from the internal queue, corresponding to the free time slots arriving at the upper bus. Additionally, it registers further requests from the lower bus.

4. The *station sends the data.* The station's own "send request" has the highest preference in the queue. Now, at the first free time slot occurring at the upper bus, the data are transmitted and the request entry is removed from the queue. The free time slot is marked as *busy.* Further, the incoming "send requests" are queued and processed as described above.

In summary, the DQDB protocol is based on counters keeping track of requests by different stations and making sure that a station refrains from transmission until reservations made prior to its own request have been satisfied.

The standard states that *priorities* of transmission requests are distinguished by the use of different request bits to be queued in different priority queues. It also states that the connectionless service is restricted to low priority. It means that the connectionless traffic may only be transmitted if the high-priority queues are empty. Unfortunately, not all implementations follow this approach with multiple priority queues.

Isochronous Service

For continuous data, *isochronous data transmission* is essential. This kind of transmission is implemented by the *pre-arbitrated function.* Certain time slots of this particular type, known as SLT, are marked at the head stations. The slot type *SLT* gets the value 1. The SLT slots can only be used according to the reservation previously done. The SLT slots arrive every 125 μs and therefore the slot frequency corresponds to 8 kHz. Hence, the isochronous mode of the DQDB can be connected with the PCM hierarchy to the WANs. Further, the 8 kHz frequency allows integration of audio and video data, which are often sampled in multiples of 8 kHz.

Connection-oriented Data Transmission

The connection-oriented data transmission is provided by the *Guaranteed Bandwidth Protocol (GBW)*. GBW represents an enhancement of the basic DQDB protocol and it is tailored to the requirements of the Variable Bit Rate (VBR) traffic requiring bandwidth guarantees. Hence, GBW is, using the priority mechanism, compatible with the basic DQDB protocol at the lowest priority and is able to allocate a required bandwidth for connection-oriented data service at higher priorities 1 or 2.

According to the GBW protocol, the queues for each priority include all segments not yet transmitted but "accepted" for immediate transmission. In this context, accepted means that the segment has passed the *traffic shaper* described in the next paragraph. The GBW protocol also uses a distributed queuing algorithm as follows: Each queue can be described as a linked list of 0's and 1's where a "1" represents a segment enqueued by the station itself and a "0" represents a request by a downstream station. The queue is updated each time a segment is transmitted, a segment is admitted by the traffic shaper or a request is received. By this kind of distributed queue, GBW allows multiple outstanding requests.

GBW *shapes the traffic* by limiting the rate of enqueuing segments for a specific connection according to the bandwidth accepted at call setup. Hence, to shape the traffic, a variable *credit* and the system parameters *income*, *slotcost* and *creditmax* are used. For each slot passing on the bus to be used for transmission, the parameter *income* is added to the current value of *credit*. A connection is allowed to enqueue a segment if the value of credit exceeds the parameter *slotcost* indicating the amount of "money" to be spent for the transmission of one segment. At the time of enqueuing the segment, the current value of credit is immediately reduced by *slotcost*. To reduce burstiness, the *credit* value is upper bounded by *creditmax*. The parameters of GBW are determined by the network management. During the enabling of high priorities, the fairness scheme "Bandwidth Balancing" must be disabled. Otherwise, high-priority transmission does not work [As90].

Low-priority traffic is controlled in a different way. GBW allows connectionless traffic to be transmitted at priority 0 with a mechanism fully compatible to the DQDB protocol, i.e., with dynamic bandwidth sharing without bandwidth guarantees.

Reachability and Fairness

Several critical issues of the basic DQDB protocol are subjects for improvement. One of these is the *reachability problem.* The solution to this problem is: each station must know which bus is reachable to other stations. Besides the reachability problem, the location of the station with respect to the *head stations* has an impact on the *fairness.* The stations, which are located closer to the *head end*, can transmit earlier data in the direction of this *head end* than other stations. For the proper operation of the protocol, it is important to transport the data until the *head ends* are reached, although utilization is lowered and it is not necessary for the actual data transmission.

Several solutions have been devised to counteract the potential unfairness of DQDB. One prominent example is *Bandwidth Balancing (BWB)*, which was incorporated into the IEEE 802.6 draft standard for MANs. This scheme attempts to prevent unfair allocation of bus capacity by artificially forcing an Access Unit (AU) to forego a certain fraction of its bus access entitlement. The scheme is such that the greater the fraction, the smaller the potential for unfairness. In general, this technique achieves fairness but at the expense of a wasted bandwidth.

Another approach for fairness improvement is *Reactive DQDB protocol (R-DQDB)*, where more status information about the bus activity is collected via special control slots dedicated to this purpose. This additional information allows an AU to discover if it is experiencing unfairness relative to other users. If this is the case, the reactive protocol allows the disadvantaged AU to issue additional requests for capacity, which has the effect of immediately establishing fairness. The overhead of R-DQDB consists of periodically collecting bus status information to determine the number of active AUs. A bus sampling rate of one per 100 slots was found to be very adequate [OF93].

The isochronous mode of DQDB fits the requirements for the transmission of continuous data very well. However, most implementations do not support this mode yet. The simultaneous access of several stations to the network data also allows an effective and high data rate for the networks with larger spatial expansion.

All networks considered so far offer their services to the higher·component of the data link layer, the *logical link control*, according to the IEEE Standard 802.2. This layer must not have specific properties to support multimedia transport. It must provide only an effective access to all, with the required parameters corresponding MAC services.

10.4.2 Orwell

The *Orwell mechanism* for MANs is based on the *slotted ring* approach. The use of fixed slots allows an easy synchronization mechanism, whereas high performance is achievable at very high rates. This scheme allows the integration of asynchronous (packet-like) and isochronous traffic (continuous-like).

The basic principle of the slotted ring is as follows: the ring is partitioned into slots of equal length; every slot has bits on the transmission line and in the node(s); to ensure that an integer number of slots is present on the ring, a latency register, located in a monitor node, is introduced to virtually lengthen the ring to a multiple number of slots; these slots circulate around the ring and are either empty or full, indicated by a single bit; if the content of a slot arrives at the destination node, it will be read; in the case of a node wanting to send data, writing of data can be performed if an empty slot is passing by.

In general, the deletion of a filled slot can be done either in the source or destination node, which provides two different access methods (source releasing approach/destination releasing approach). Orwell defines the *destination releasing approach*. This results in better performance since the released slot can already be used by another node between the destination and source node. To ensure fairness on the ring and to prevent a node from monopolizing the ring, the access is organized in *cycles*, also called *reset intervals*. Each node on the ring has a counter indicating how many slots may be occupied by that node during a cycle.

Two classes of services are defined: *class 1 for isochronous traffic* and *class 2 for asynchronous traffic*. Four priority levels are provided, implemented by four queues in each node. This is also reflected in the Orwell slot header. To increase the bandwidth of the Orwell ring, several rings can be assembled in parallel. This configuration allows a flexible increase of the network capacity, as well as providing higher network reliability. The slot structure was lined up close to the ATM structure. Therefore, ATM cells can be completely transported by the Orwell ring.

The performance of the Orwell ring is very good because of the destination releasing approach. The ring can be loaded with up to 240 Mbits/s; however, at these high loads, the delay on the ring becomes rather large (more than 1ms) [Pry93].

10.4.3 MAN Connectivity to ATM Networks

The physical limitations of a MAN prevent it from being used as a wide-area solution, so ATM B-ISDN (ATM WAN) systems must provide internetworking among different MANs. The B-ISDN can function as a higher hierarchical level of switching, interconnecting different MANs.

In the case of a FDDI MAN, there is a large incompatibility between ATM and FDDI, requiring additional adaptation functions in the gateways or bridges. For example, the cell size of ATM and the frame size of FDDI are incompatible. This requires a permanent segmentation and reassembly at the internetworking unit, which influences the end-to-end delays. It also requires additional substantial processing capabilities to maintain the throughput. The segmentation and reassembly can be achieved, for example, by the AAL4 layer.

The specification of DQDB and Orwell was performed in parallel with the ATM specification, therefore it will become easier to interconnect them. The DQDB protocol uses slot-based transmission where the slot size (53 bytes) was adopted from ATM. Sometimes, DQDB is called the "Shared Medium ATM", but the similarity between the ATM and DQDB is limited to the payload field length.

The future and relevance of DQDB depends highly on availability of gateways between DQDB and ATM B-ISDN. Internetworking means mapping DQDB slots to ATM cells and vice versa, as well as service internetworking. Some internetworking solutions, such as *cell-to-slot internetworking* and *frame internetworking* were tested for DQDB connection-oriented mode through simulations. The cell-to-slot internetworking modifies ATM cells to DQDB slots and vice versa. The frame internetworking reassembles cells/slots and modifies AAL frames to DQDB initial MAC PDU and vice versa. Both modes were compared, receiving aggregated traffic from B-ISDN and using high-priority transmission controlled by GBW protocol for access to DQDB. The cell-to-slot mode performed better and showed to be more advantageous [MR93a].

More generally, ATM connectivity of MANs may be provided as follows:

- *Semi-permanent Basis*

  As described in Section 10.4.1, the DQDB standard up to now only covers the *connectionless packet service.* The final specification of the *isochronous service* and the *connection-oriented data service* are still in progress.

  For the connectionless mode, some precautions must be taken due to the connection-oriented nature of the ATM concept. Solutions may be:

  - The installation of semi-permanent connections, using *virtual paths* and transporting different *virtual channels* between all MANs to be interconnected. This may result in a *virtually meshed network* [TP90].
  - The provision of special Message IDentification (MID) values in the AAL.
  - The provision of very fast call setup [Pry89], also called *fast reservation protocol* [TBR92].

- *Connectionless Servers*

  Because of the connectionless transport in MANs, instead of virtual linking all the MANs, the linking one or more servers that are capable of serving connectionless data is possible. A connectionless server based on the *Switched Multimegabit Data Service (SMDS)* or the *Connectionless Broadband Data Services (CBDS)* might be appropriate [HL90].

- *Direct ATM Connection via an ATM LAN*

  Traffic increase may continue until the final solution is that each device within the user's premises will have its own direct connection to the B-ISDN via an

ATM LAN. This would result in a full star-like topology with every subscriber directly connected to the ATM network, with the maximal functions expressed in terms of B-ISDN traffic.

In principle, DQDB as a whole turns out to be well-suited for continuous media traffic. However, the required modes are not fully specified and only connectionless packet service is available, which is not suitable to transport audio and video data. Orwell is an experimental MAN which so far is well-designed for multimedia data transport. However, it is questionable if Orwell will receive wide attention in the real world. Whereas a few years ago, the necessity to have MANs was well-accepted, today's high-speed LANs on the one side and the ATM LANs on the other side show that we might encounter LANs and WANs only in the future.

10.5 Wide Area Networks (WANs)

WANs typically span entire countries. The data rates of current WANs (e.g., Internet's T3 NFSNET backbone has a data rate of 45 Mbits/s) are lower than those of LANs and MANs, but this changes with the up-coming Broadband ISDN and its new technologies for high-speed networking. The ownership is split among multiple organizations (the carrier owns the communications subnet and numerous clients own the hosts). The basis of a WAN's communication is point-to-point connection, except for satellite networks.

Currently we see two main wide area networks which are/will be used for multimedia transmission: *traditional networks*, such as Internet, and *ATM B-ISDN* We describe in this section the basic mechanisms and principles of these WAN systems below the network layer which are relevant to multimedia transmission. With respect to traditional WAN systems, we present the connectivity issues and a brief overview of the Internet network. Internet functions, such as *multicasting* and *routing* at the network layer, are crucial for multimedia applications, such as collaborative computing. These issues will be discussed in Chapter 11. In the case of ATM B-ISDN, we discuss *switching* functions, *routing* and *host interfacing* issues important for multimedia transmission.

10.5.1 Traditional WAN's

Connectivity

Connectivity of different local networks over a wide area is an important and crucial feature for WANs. This function allows users to execute multimedia networked applications which provide for remote project collaborations. The connectivity of the systems can be described in terms of three levels:

- *Basic Level*

  At the first level, basic *link connectivity* is provided. This is commonly provided either through the use of leased lines or by using the available public packet data network infrastructure, which is commonly X.25-based.

  In 1990, the principal backbone networks for the Internet in the USA became NSFNET (National Science Foundation Network) and other agency networks such as NASA's NSI and Department of Energy networks. NSFNET is organized to interconnect regional and local networks that serve the academic community. NSI connectivity reflects the networking needs of NASA science and research projects.

  The link interconnectivity between USA and Europe is based on:

  - *Specific lines*, which mostly operate at approximately 64 kbits/s or below and tend to serve a specific need. An example is the line between Oslo and College Park, which connects NORDUNET and the US research network infrastructure.
  - *Fat pipes* in which larger bandwidth (typically a fraction of T1) is purchased and shared among several networks. An example is a T1 link (1.5 Mbps) for the *Multicast Backbone (MBone)* service connectivity [MB94]. MBone is a virtual network that shares the same physical media as the Internet. For multimedia traffic in this basic level, Internet (MBone) can make use of switched lines which have a guaranteed bandwidth. Also, experiments showed [MB94] that bandwidth capacities lower than T1 are unsuitable for MBone video, although some users, even sometimes entire

countries, have managed on special configured networks to use MBone applications at 54 and 64 kbps.

- *Network Level*

  The next level of connectivity provides *network connection* so that an end-to-end connectivity may be established. This level of connection is typically provided through the use of a single protocol suite, or by sharing link or physical connectivity between networks using multiple protocol suites.

  The focus of international internetworking in addressing multiple protocol suites is to find a method for coexistence and interoperability. In our context, we see the need for this to be provided for continuous streams. An example is Internet's research to develop mechanisms for an *integration of multiple protocol suites* within the Internet that maximizes the functionality and performance of the end-to-end systems for different quality of media.

  Another challenge is to provide *multicasting* functions in the Internet routers when continuous media are transmitted. Currently, MBone can support multicast using a network of routers (mrouters) which can support multicasting. The reason for the need of the multicasting functions, when continuous media are transmitted, is the bandwidth. A multicast stream is *bandwidth-efficient* because one packet can touch all workstations on a network. Thus, a 128 kbits/s video stream (typically 1-4 frames per second) uses the same bandwidth, whether it is received by one workstation or 20.

- *Application Level*

  In the upper level, user network services are provided, such as electronic mail, conferencing and application sharing, by providing the connection of similar networking services between different protocol suites. This often takes the form, for example, of application gateways (packet format translations) between OSI and Transmission Control Protocol/Internet Protocol (TCP/IP) services and translations of media qualities from one end-point to another.

  An example where application-to-application connectivity is visible is the set of MBone applications which connect users all over the world. The Mbone applications, such as *nv* (net video), *vat* (visual audio tool) and *wb* (white-board)are based on IP multicast and some are using the *Real-Time Protocol*

(RTP) on top of the *User Datagram Protocol (UDP)*/IP. These applications provide connectivity for users who are running MBone tools on workstation architectures, such as Sun, Silicon Graphics, DEC, Hewlett-Packard and PC architecture Macintosh for multimedia conferencing over the WAN.

Internet

The internetting experiment by DARPA in 1973 led to the evolution of a system of networks, called *Internet*, which is global in scope. There are over 7500 networks interlinked, supporting over 1,000,000 computers, and millions of users in over three dozen countries [LR93].

Internet is layered similar to the OSI reference model shown in Figure 10.15. The

Layer	Example
Application	FTP, Telnet, SMTP, X-Windows
Transport	UDP, TCP, TP4, Routing
Internet	ICMP, IP, CLNP
Subnetwork	Ethernet, X.25, FDDI, Token Ring
Link	HDLC, PPP, SLIP
Physical	RS 232, V.35, 10 Base T, Fiber optic, etc.

Figure 10.15: *Internet layering.*

Internet's major backbones currently support transfer rates from T1 (1.5 Mbits/s) to T3 (45 Mbits/s). Internet's transmission mode is based on packet-switching technology. This is implemented by the *Internet Protocol (IP)* environment which provides the *network connectivity*. Its multicasting capability provides a strong support for multimedia collaborative applications.

Application-to-application connectivity is provided by the *Transmission Control Protocol (TCP)*, as well as other transport (e.g., UDP) and higher layer protocols (e.g., RTP).

The Internet is a cooperative effort among all the diverse networks that make it up. The Internet essentially provides electronic mail, file transfer capabilities and remote login. Special electronic mail procedures are also used to support news distribution applications and bulletin boards. New Internet applications support audio (e.g., vat) and video (e.g., nv), conferencing capabilities, as well as application sharing tools (e.g., wb).

Interconnection Devices

In a WAN, the source and destination are connected by a sequence of *interconnection devices* (packet switches). The packet switches must cooperate to calculate a path through the network. They are referred to:

- At the network layer as *routers*.

  Routers are used if LANs or MANs must be connected over longer distance and service like "isolation required to ensure that local problems do not affect other areas" is needed. Routers verify and modify the packets that they forward, and recalculate new checksums. Routers are not end-node transparent. If an end-node needs to send a packet to a node on some other LAN through a router, the packet must be addressed to the router's hardware address. The routers differ from "routing bridges" because they use software that implements a routing protocol. This protocol determines the next node along the path that a particular packet is to take. For multimedia transmission, the routers need to be extended to support *multicasting functions*. Current state-of-the-art in MBone is that *mrouters* support multicasting. *mrouters* are either upgraded commercial routers, or dedicated workstations with modified kernels in parallel with standard routers [MB94].

  The most popular types of routing protocols are

  - *Distance vector* (e.g., ARPANET routing algorithm, DECnet Phases III and IV, AppleTalk's routing algorithm)

    Each router is responsible for keeping track and informing its neighbors of its distance to each destination. The router computes its distance to a destination based on its neighbor's distance to the destination. The only information a router must know a priori is its own ID and the cost of its links to each neighbor.

For multicasting in mrouters, the *Distance Vector Multicast Routing Protocol* is used [WPD88]. Some researchers consider this protocol inadequate for rapidly changing network topologies because the routing information propagates too slowly [Per93].

- *Link state* (e.g., DECnet Phase V and OSPF)

  Each router is responsible for determining the identities of its neighbors and constructing a *Link State Packet (LSP)*, which lists its neighbors and the cost of the link to each. The LSP is transmitted to all of the other routers, which are responsible for storing the most recently generated LSP from each other router. Given this information (the LSP database), it is possible to calculate routes.

  An example of an LSP protocol, which will also be used in mrouters, is the *open shortest path link state protocol* proposed by the *Open Shortest Path Working Group* [MB94]. This link state protocol is based on an algorithm developed by Deering [Moy93]. Note that using this protocol and also the distance vector multicast routing protocol, mrouters dynamically compute a source tree for each participant in a multicast group.

- At the data link layer as *bridges.*

  Bridges are used to connect LANs. They conditionally forward packets from one network port to another. Since bridges operate in store-and-forward mode with packets, they present a much longer processing delay than repeaters. There are three types of bridges:

  - *Transparent bridges* are used in networks like Ethernet and 802.3, and they are completely transparent to the end-nodes on the network and are protocol-independent. The basic mechanism behind a transparent bridge is that it acts as a station on two or more LANs, listens promiscuously to each packet, stores the packet for forwarding and forwards it onto every other LAN to which it is connected when the LAN arbitration protocol for the destination LAN indicates the medium is available. This idea works as long as there are no loops in the topology. To enable people to plug transparent bridges into arbitrary topologies, bridges run a *spanning* *tree protocol.* The protocol matches the current topology into a spanning tree, i.e., a loop-free subset of the topology that has maximal coverage.

- *Source route bridges* are used in Token Ring networks, and they are protocol-independent and not end-node transparent.

  The mechanism behind source route bridging is that the source end-node puts a route into each data packet's header. The way that the source end-node discovers a route is by issuing an "explorer" packet that makes copies of itself for every possible path, with each copy keeping a diary of where it has been. Then, the destination can choose a route, based on the received explorer packets, or send them back to the source and let the source choose a route.

- *Combination bridges* are used to connect, for example, Token Ring to FDDI.

Bridges for multimedia are easier to implement than routers. One requirement is that they must provide guarantees to the payer if possible.

- At the physical layer as *repeaters.*

  Repeaters unconditionally copy bits of data from one port to another. They do not check or regenerate checksums in the data stream. They are transparent to protocols and are end-node transparent.

  There is no special processing required for multimedia.

10.5.2 B-ISDN: ATM

Broadband Integrated Services Digital Network (B-ISDN) is a network concept which represents the extension of the *Narrowband ISDN (N-ISDN).* The goal of B-ISDN is to define an application interface and a corresponding WAN with conversational, distributed, messaging and query services of different bandwidth requirements. Further, the goal is to provide connectionless and connection-oriented services for transmission of different media. CCITT Study Group XVIII works on B-ISDN standards.

Channel	**Bandwidth**
B	64 kbits/s
H_0	385 kbits/s
H_1	1,920 kbits/s (Europa), 2,048 kbits/s (USA)
H_2	32,768 kbits/s (Europa)
H_4	132,032 - 139,264 kbits/s

Table 10.3: *Channels and the corresponding data rates.*

Transfer Modes

Until 1987, the B-ISDN was based on *Synchronous Transfer Mode (SMT)*. Since 1988, the *Asynchronous Transfer Mode* is the basis of B-ISDN [HH91, Sta92].

- *Synchronous Transfer Mode*

  STM in the N-ISDN works in a connection-oriented mode with fixed assigned bandwidth according to the time multiplexing method. Therefore, already at the lowest layer, a guaranteed transmission with a determined bandwidth (as in N-ISDN) can be supported. Only small end-to-end delays occur. A time window is called *a slot*. These slots are reserved for the duration of a connection. They lay within periodically repeated structures called *frames*. The assignment of time slots and frames is shown in Figure 10.16. Each slot

Synchronous Transmission Mode

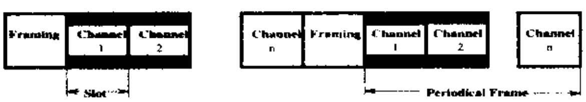

Figure 10.16: *Assignment of time slots in STM (time multiplexing method).*

  has a fixed duration. STM adapts well to the PCM-transmission hierarchies and therefore to the N-ISDN. The channels with correspondent data rates are shown in Table 10.3. This approach is not flexible because of the fixed data rates and the fixed assignment of bandwidth to connections. There are several solutions to improve the flexibility:

- *Compression methods*

  For example, the video compression methods described in Chapter 6, may be used. Many multimedia systems already use compression methods such as JPEG or MPEG.

- *Additional slots*

  One solution is to introduce many slots, for instance, 2,048 8 bit-slots for 15 Mbits/s. The disadvantage of this solution is the increased management overhead by many possible combinations.

- *Container solution*

  As a compromise, a *container* solution was discussed, which allows a limited number of partitions. For example, these can be $H_4+4\times H_1+n\times B+D$ with containers in H_1 and H_4 channels. The static STM mechanism remains and the left bandwidth is not continuously used through this partial reservation.

If data with fixed data rates are used, the STM approach is appropriate for the transmission. It means STM is suitable for transmission of continuous media because it utilizes the bandwidth and end-to-end delays are low.

- *Asynchronous Transfer Mode*

  The *Asynchronous Transfer Mode (ATM)* concept was introduced in Section 10.3.4. The ATM approach is more efficient and flexible for a transmission of data streams with variable data rates than the STM approach because the total bandwidth is better utilized.

We will concentrate in the next paragraphs on ATM WAN issues such as *switching* and ATM *host interfacing* to multimedia workstations, which provide the connectivity functionality for multimedia communication systems.

Switching

In the past, various *switching* architectures were developed for different applications such as voice and data, based on transfer modes like STM and packet switching. The switching architectures developed for STM and packets switches are not applicable to ATM because of (1) the high speed at which the switch must operate (from 150 up to at least 600 Mbits/s today), and (2) the statistical behavior of the ATM streams passing through the ATM switching systems.

Today, ATM switches as products are installed by public operators to offer a public wide broadband service (these systems are called an *ATM Central Office*), and by private users to fulfill internal high-speed communication needs (ATM LAN).

The switching requirements for ATM correspond to the requirements of B-ISDN, which must be capable of transporting all kinds of information, ranging from telecontrol over voice to high-quality video. As we described before, these services have different requirements in terms of *bit rate* (from a few kbits/s up to hundreds of Mbits/s), *behavior in time* (constant bit rate or variable bit rate), *semantic transparency* (cell loss rate, bit error rate) and *time transparency* (delay, delay jitter). We describe some of the major service requirements the ATM switches have to cope with:

- *Bandwidth*

  The usual bit rate, which ATM switches are able to switch, is around 150 Mbits/s. This does not mean that ATM switches have to operate internally at 150 Mbits/s. Switching of a single path can be realized in parallel so that a lower speed can be used internally, or several 150 Mbits/s paths can be multiplexed on a single link so that internally higher speeds than 150 Mbits/s are implemented (potentially in the Gbits/s domain).

- *Broadcast/Multicast*

  In classical STM and packet switches, only point-to-point connections are available because data are switched from one logical inlet to another logical outlet. However, the future broadband network requires multicast and broadcast functionality. Therefore, there is a requirement on future ATM switches to provide these functionalities. These functions are required for services, such as conferencing, concurrent video library access and TV distribution.

- *Performance*

In ATM environments, the following performance parameters are important for multimedia applications:

- *Throughput and Bit Error Rate*

  In ATM switches, the *throughput* and *bit error rate* are mainly determined by the high speed hardware technology (CMOS, BICMOS or ECL), where bit rates of hundreds of Mbits/s can easily be achieved with an acceptable bit error rate.
- *Connection Blocking*

  Connection blocking is determined as the probability that not enough resources can be found between input and output to guarantee the quality of all existing connections and the new connection. If the switch has enough resources (i.e., bandwidth and header values) on the input and the output of the switch, no connection blocking occurs internally. If the switch does not have enough resources, an explicit check of internal switch resources has to occur. So, a new connection will be accepted, and resources will be allocated only if enough resources are also available on the external links.
- *Cell Loss Probability*

  It is possible that more cells arrive simultaneously and compete for a buffer than a buffer in the switch can store. Hence, some cells will be dropped and therefore lost. Typical values for *cell loss probability* mentioned for ATM switches range between 10^{-8} to 10^{-11}. This is considered to be harmless as the header information of the cell is typically protected by some kind of FEC.
- *Cell Insertion into other Connections*

  It is also possible that ATM cells are internally misrouted in the switch, so they arrive erroneously on another logical connection. The probability of this *cell insertion* must be kept within limits, and values of 1000 times or better than the cell loss rate are typically mentioned in the literature (10^{-14}) [Pry93].
- *Switching Delay*

  The time to switch an ATM cell through a switch is of importance for the end-to-end delay. Typical values for the delay of ATM switches are between 10 and 1000 μs with a jitter of a few 100 μs or less [Pry93]. For any continuous media applications, an end-to-end delay in the range of 100 ms is acceptable. Hence, the switching delay coming from the switch is a minor portion, even if we encounter, for example, several ATM switches in the data path.

An ATM *switching fabric* is composed of basic ATM switching building blocks, also called *switching elements*. An ATM switching element size can be from two inputs and two outputs at 150 Mbits/s, up to 16 inputs and 16 outputs at 2.4 Gbits/s. A switching fabric, composed of a large number of identical basic switching building blocks, is called *Multi-stage Interconnection Network (MIN)*.

The ATM *switching architecture* consists of the *control part* and the *transport part*. The *control part* of the switch controls the cell transport. It decides, for example, which input to connect to which output. The control part uses QoS parameters to control the performance of its own services, such as call setup, as well as the transport ATM services, such as cell transmission. Quality of service parameters for the control network are related to the signaling protocols, e.g., the call setup time, call release time, etc. The *transport part* is responsible for the correct transportation of cells from an ATM input to an ATM output, within the QoS specifications of ATM. Typical QoS parameters for the transport part are the cell loss rate, bit error rate, cell delay, cell delay jitter, etc.

The transportation of data from an incoming logical ATM channel to an outgoing logical ATM channel means selection between a number of outgoing logical channels. A logical ATM channel is characterized by a physical input/output (physical port number) and by a logical channel on the physical port (VCI and/or VPI). To provide the *switching function*, two other functions must be implemented:

- *Routing* (space switching).

  Routing in an ATM switch means to route the cell internally from the input to the output.

- *Queuing* (time switching).

  Queuing cells means to transport a cell from a slot k to a time slot l. Since the pre-assigned time slot concept disappears in an ATM system, contention problems arise if two or more logical channels compete for the same time slot. This can be solved by temporarily queuing the ATM cells before sending them out.

We discuss briefly the *routing* of cells in MINs because this function is closely related to the multicasting function which is crucial for multimedia applications.

ATM is based on a connection-oriented approach, therefore, the routes (logical ATM channels) from the source to the destination are determined during the connection setup. The header values (VPI/VCI) are assigned to each section of a connection, and translated when switched from one section to another to get the cell routed. When switching/multiplexing in cells is to be performed, the header/link translation (routing) from the input data (incoming header) to the output data (outgoing header) occurs.

The routing function can be categorized with respect to the *routing decision time* and the *routing information place*: The time parameter decides when the translation decision is made namely, if the routing decision is made once for a connection or every time a cell arrives at a switch. The place parameter specifies where the routing information is stored.

- *Routing Decision Time*

  Routing translation can be performed either once for the complete duration of the connection (*connection-based routing*), or it can be performed for every cell separately (*cell-based routing*). For the first case, it means that the MIN is internally connection-oriented or uses pre-set path routing. This case is more suitable for continuous media transmission because the data arrive at the receiver in order. A problem may occur if a contention occurs along the pre-assigned path, which causes an increase of the end-to-end delay for the media PDUs. This problem can be solved using resource reservation mechanisms at the switches.

  In the second case, the MIN operates internally connectionless. In the first case, all cells of the virtual connection follow the same route/path through the MIN; in the second case, they do not. The cell-based mode of operation should be chosen when data can arrive at the receiver out of order, or the urgency of the transmitted data is high (e.g., transactions). This mode allows each cell to choose a path which has enough resources to handle congestion. This implies functions which report resource availability in the switches.

- *Routing Information Place*

  Routing information can either be transported by each cell itself via a so-called *routing tag*, or it can be stored in *routing tables* in the switching blocks. In the case where a routing table is used, this table must be accessed via an entry. For multimedia transmission, the routing tables are preferable because of multicasting. The routing tables become multicast routing tables. An example of an ATM switch which supports multicasting in the cell-based mode is Roxane [Pry93]. Most of the currently available ATM switches do not support multicasting .

Host-interfacing to Multimedia Workstations

Another important aspect of ATM B-ISDN is the *impact of ATM at the end-points* (hosts) which are represented through multimedia workstations. We discuss some opportunities of how the end-points can take advantage of the ATM concept for transmission of multimedia:

- *Support of VBR Coding*

  When video signals are encoded in digital format using a simple Pulse Code Modulation (PCM) method, the resulting bit rate is fixed and simply the product of the sampling rate and the number of bits per sample. However, as soon as a compression algorithm is used, the resulting bit rate varies in time. Hence, the bit rate generated by the future video codec will fluctuate in time or will be different for different qualities of video images like in MPEG 1, 2, 4.

  In classical STM networks, this fluctuating rate must be converted into a Constant Bit Rate (CBR), namely that at which this STM network is operating. For example, 64 kbits/s are required for N-ISDN and 139,264 kbits/s are needed for videophone or TV distribution service (Table 10.3). This bit rate equalization can be realized by an output buffer between the encoder and the network and a feedback signal between encoder and buffer.

In ATM networks, the limitation of working at a constant bit rate disappears because of the AAL2, so the output buffer is basically no longer required at the output of the encoder. The output of the encoder can be directly fed into the ATM network, resulting in a *VBR* video encoder.

- *Support of Medium Layered Coding*

In future B-ISDN, we see the new multimedia services to consist of, for example, one or more of the following service components: *audio*, *standard video*, *high-definition video overhead*, *teletext*, and *data*. For example, HDTV is composed of all five different service components. All these individual components can be transported individually over separate virtual channels. However, some restrictions must be taken into account between those different virtual channels, mainly with respect to the relative delay through the network. For example, the lip synchronization between voice sound and video image requires delay (skew) difference $\leq$ 80 ms [SE93]. Therefore, the individual service components can be further divided into *layers*. Every higher layer uses the information of the lower layer to construct the image of that layer with the required quality. In terms of resolution, this is known as *hierarchical encoding*. An example of such a layered model (hierarchical encoding) for a video medium is shown in Figure 10.17. Sound (HIFI) and data (subtitling) represent *non-hierarchical encoding*. This implies that each layer of the medium quality division will be assigned a separate VCI/VPI and different service, which may help in the adaptive behavior of each medium during the transmission.

An important advantage of this layered approach is the compatibility between different services and terminals. An additional advantage of this layered coding principle is its possibility to cope efficiently with the cell loss caused by the ATM network [Pry93]. However, a wide acceptance of a specific layering tends to be difficult to achieve.

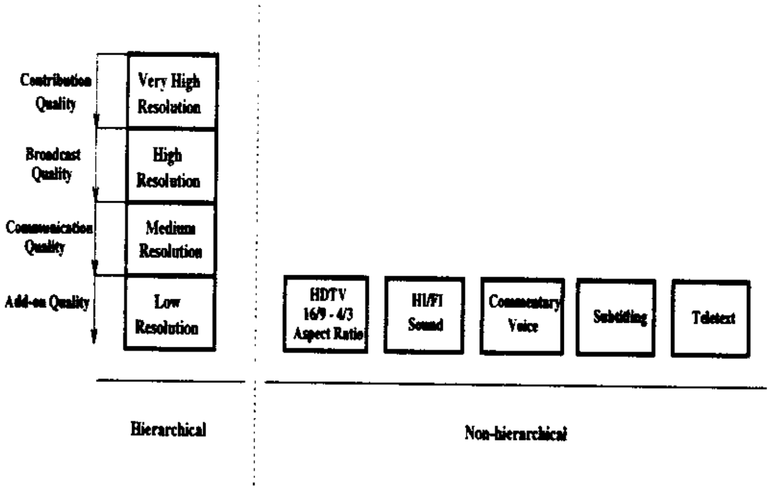

Figure 10.17: *Layered video encoding.*

10.6 Conclusion

Most current multimedia network systems are available as commercial products only for single user computers (PCs). Most experimental and research software provides the capabilities of video/audio data transfer on multiuser UNIX machines (e.g., MBone tools). The most popular networked multimedia applications are interactive video retrieval, TV and audio broadcasting, conferencing and electronic mail.

Most cooperative multimedia applications use the underlying networks, for example, Ethernet or Internet, without any specific changes. Video and voice teleconferencing have been demonstrated on the Internet with datagram-based protocol in [TP91, NS92, CCH+93b] and other implementations. Hence, it is not true that multimedia communication systems need at least the bandwidth of 100 Mbits/s to provide acceptable communication. Current usage of MBone (1.5Mbits/s) with the video and audio conferencing software nv/vat supports remote presence at conferences and other technical meetings [Moy93].

Current existing network solutions may work for multimedia traffic if the utilization of the network is low. However, this situation will change as soon as these applications become more widely used. The number of end-users increases, and with it the network traffic. Large quantities of such traffic cannot be supported anymore without handling different types of service at gateways and routers, and in the end-to-end protocols. For example, Internet already faces challenges in routing today:

- Routing tables expanded beyond router capabilities.
- Premature exhaustion of assigned IP network numbers due to inefficient allocation of the IP address space.
- IP's inability to address more than four billion hosts connected to a single Internet.

In the case of end-to-end protocols, TCP's window mechanism, because of its limited window size, does not effectively utilize the potential transmitting capacity of high-speed networks. The already installed base of networks will not be replaced just for providing multimedia capabilities everywhere. These networks, especially in the LAN domain, will remain and here we can exploit their properties for carrying continuous media traffic as was outlined in this chapter.

New commercial services, such as SONET fiber links and gigabit ATM LANs and WANs, together with new workstation capabilities including sound and video, produce demanding requirements in terms of multiple types of service. Hence, existing networks will have to change/expand towards supporting multiple types of service and/or new high-speed networks, such as ATM LANs and ATM WANs, will have to be installed. For example, the new challenges in Internet already lead to modifications of its current protocol suite. Several methods and modifications are described in [JBZ90, JB88, BCS93, ZDE+93].

Several new protocols at the network/transport layers in Internet (e.g., ST-II) and higher layers in B-ISDN are currently centers of research to support more efficient transmission of multimedia and multiple types of service. We discuss the changes in the transport system toward accommodation of integrated services, as well as new protocols in higher layers, in Chapter 11.