# Chapter 3. Overview of Computer Networks

## 3.1. OSI Reference Model

The International Standards Organization (ISO) was among the first organizations to attempt to define a standard for global communications between various computer systems. Its seven-layer reference model (Figure 3.1) was the basis for the Open System Interconnection (OSI) architecture that separates each functional area within any computer into discrete layers. This effectively simplifies the complex processes involved in communication between or among various cooperating computer systems [10]. Even though this OSI method provides a sensible conceptual framework for independent computer networks, the ISO protocols themselves have been a commercial failure [6].

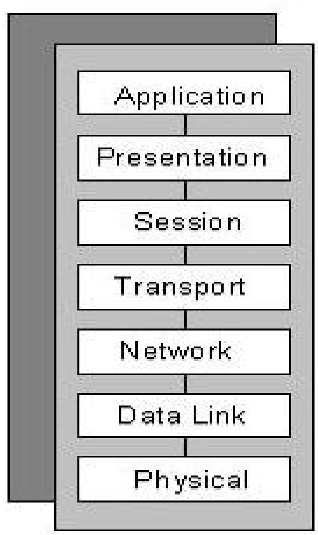


*Figure 3.1 ISO's OSI Reference Model*

## 3.2. SONET/SDH

Synchronous Optical NETwork (SONET) and Synchronous Digital Hierarchy (SDH) represent closely related sets of standards that govern interface parameters such as rates, formats, multiplexing methods, and Operations Administrations, Maintenance and Provisioning (OAM&P) for high speed optical transmission [11]. SONET is becoming the primary set of standards used in North America, while SDH is primarily used in Europe and Asia.

From a transmission perspective they both provide an adequate international basis for existing time division multiplexing (TDM) as well as the new (cell-multiplexed) services. Initially, the goal of these standards was to facilitate the interworking of multivendor equipment across a single fiber span. This would allow attachment of different vendors' equipment without loss of functionality of the overall system. Furthermore, the goal was to provide for future increases in data rates by defining the base signal rate, which could then be synchronously multiplexed to attain higher rates.

SONET represents a basic physical layer technology that can carry any type of payload, both isochronous (delay-sensitive) voice and video, or packet switched data. It is a transmission technology, which lies at the bottom of the Open System Interconnection (OSI) stack. Above SONET, it is possible to implement any Data Link technology, including Asynchronous Transfer Mode (ATM), Frame Relay, Switched Multi-megabit Data Service (SMDS), and even FDDI [9].

Both SONET and SDH make use of the basic building block signals. For SONET the first level, or the smallest building block, is a Synchronous Transport Signal Level-1 (STS-1) with a signal rate of 51.84 Mbps. STS-1 corresponds to

Optical Carrier Level –1 (OC-1) signal in SONET systems. SDH starts at the base rate of 155.52 Mbps, which is the rate of the fundamental building block, referred to as Synchronous Transport Module Level-1 (STM-1). Thus an STM-1 level in SDH corresponds to an OC-3 level in SONET. Since the rate of STM-1 is just 3 times the STS-1, once at the STM-1/STS-3 level and higher, SONET and SDH are completely interoperable. (Table 3.1)

*Table 3.1 SONET Transmission Rates*

| **SONET Signal** | **SDH Signal** | **Transmission Rate [Mb/s]** |
|---|---|---|
| OC-1 | | 51.84 |
| OC-3 | STM-1 | 155.52 |
| OC-12 | STM-4 | 622.08 |
| OC-24 | | 1244.16 |
| OC-48 | STM-16 | 2488.32 |

The transmission of data over SONET/SDH lines occurs in frames, which are depicted as rectangular octet-based units. Due to the compatibility issues, both systems send frames 8000 times per second, or one every 125 µs. Both frames have very similar structure except for their dimensions, which reflect the previously mentioned differences in basic building block rates. SONET STS-1 frame format is 9 rows by 90 byte columns. Keeping in mind that 1 byte equals 8 bits, then the frame consists of a total of 6480 bits. (9 x 90bytes x 8bits/byte). Since 8000 such frames are sent per second, then the aggregate rate is 51.84 Mbps. (6480 bits/frame x 8000frames/s).

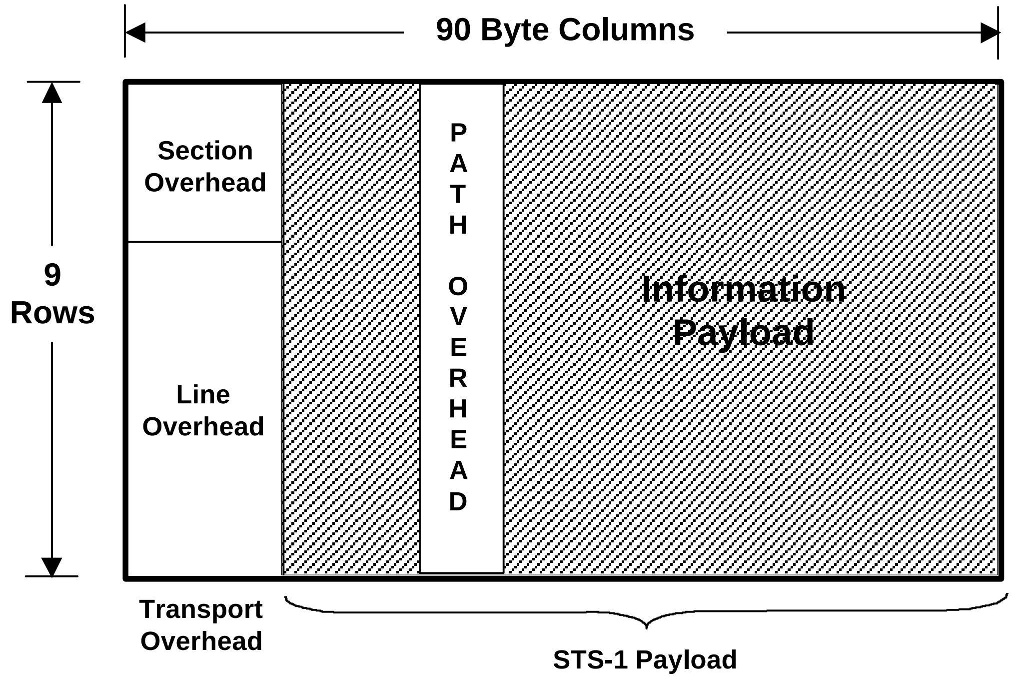


*Figure 3.2 SONET STS-1 Frame Structure*

The transmission of the actual frames occurs from the upper left to the lower right, just as if reading text on a page. It can be seen from Figure 3.2 that the first 3 byte columns of the SONET frame are dedicated to the Transport Overhead (TOH), which closely corresponds to the Section Overhead (SOH) – the first 9 byte columns in SDH STM-1 frame. Both TOH and SOH have a very important role in supporting transport capabilities such as framing, error monitoring, and management operations.

In SONET, the overhead and transport functions are grouped into layers. SONET layers, section, line, and path, have a hierarchical relationship, and can be viewed either from the top down or from the bottom up. (Figure 3.3) Each layer has its own management communication through the specific octets in the overhead.

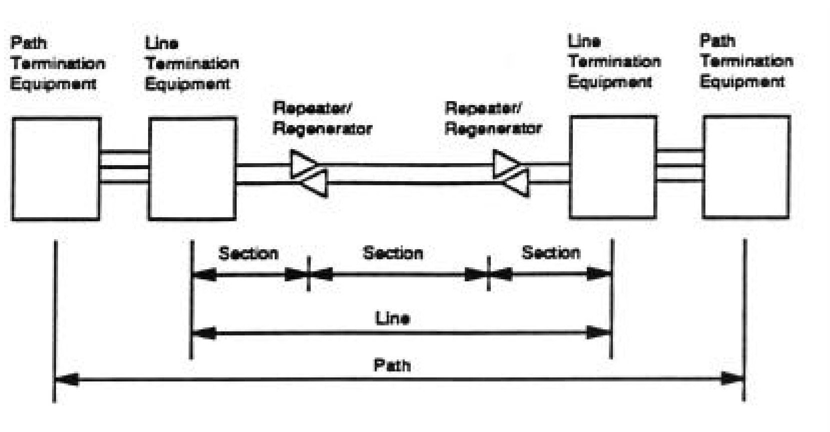


*Figure 3.3 Paths, Lines and Sections*

The TOH, which is composed of Line overhead and Section overhead, is added to the Synchronous Payload Envelope (SPE) to create the STS-1 frame. The SPE is further composed of the POH and the information payload. The functions for TOH and POH are given in Table 3.2.

*Table 3.2 TOH and POH Functions*

| **Transport Overhead Functions** | **Path Overhead Functions** |
|---|---|
| Framing | Path trace |
| Error detection | Error detection |
| Orderwire | Payload composition |
| User channel for network provider data communication (for operations) | Maintenance signaling |
| Pointer (for locating payload) | Far-end error information |
| Automatic protection switching signaling | Path user |
| Maintenance signaling | Payload-specific "signaling" |

### 3.2.1. SONET Multiplexing and Mapping

Although the STS-1 signal is the smallest building block in SONET, most of the end user devices do not operate at the rate of 51.84 Mbps. If the path terminating equipment does not need the entire SPE, the multiplexer can define smaller/slower paths within the SPE called Virtual Tributaries (VT) in SONET or Virtual Containers (VC) in SDH. The multiplexing of slower rates into SONET STS-1 is different and comparatively simpler than the multiplexing of slower rates into SDH STM-1 signals. It should be noted that SDH uses different standards for the slower signals, and is somewhat more complex than SONET in that it allows different mappings of the same payload, but also requires four hierarchical levels in forming the payload of an STM-1.

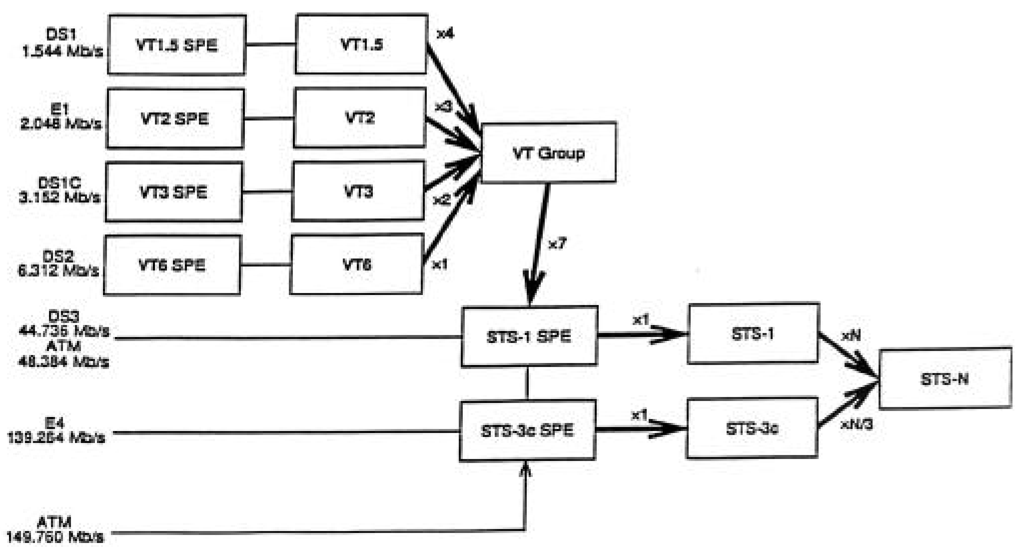


*Figure 3.4 SONET Multiplexing Structure*

The first step in SONET multiplexing is to map the slower signals into a Virtual Tributary Synchronous Payload Envelope (VTx-SPE) which corresponds to the appropriate size VT. Achieving greater SONET/SDH rates than the basic building blocks is more straightforward than mapping slower signals into the STS-1/STM-1 signals (See Figure 3.4)[6]. For SONET, higher rates are accomplished by byte-interleaving N frame-aligned STS-1 signals to create an STS-N signal. Since there is no overhead added during this multiplexing process, the rate of a newly created STS-N signal is exactly Nx51.84 Mbps, where values of N in current use are 1, 3, 12, 24 and 48.

However, when mapping high rate signals directly onto an STS higher than STS-1, such as ATM, then the frame aligning must be done with pointer concatenation which results in an STS-Nc concatenated frame signal with a locked (concatenated) STS-N payload (See Figure 3.5) [3]. After mapping these signals into a big enough STS-Nc SPE, they could be further multiplexed into an STS-N signal.

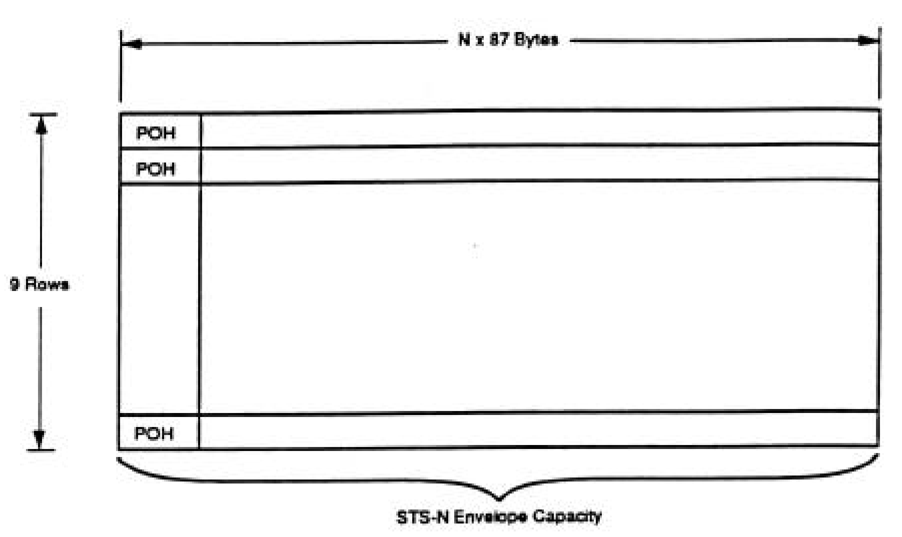


*Figure 3.5 Concatenated STS-Nc Payload*

### 3.3. Asynchronous Transfer Mode

The ATM concept can be approached from many different aspects but in its most essential sense ATM is a technology which is defined by standardized protocols. ATM is a connection oriented, cell-based switching and multiplexing technology designed to offer flexible transmission support for a wide range of services, including voice, video and data. ATM is connection oriented because it communicates with its peers by establishing virtual connections in a similar fashion to typical circuit-switched networks. It is a cell-based technology because ATM partitions all of its information into fixed-size packets or cells, which in turn simplify the switching and multiplexing functions. In addition to supporting a complete range of user traffic, simplified switching and multiplexing capabilities, ATM is well suited for high-speed networking too. This makes it the underlying technology of choice for B-ISDN.

#### 3.3.1. B-ISDN PRM

The ITU-T's B-ISDN protocol reference model, including ATM as its foundation, is shown in Figure 3.6 [9]. The B-ISDN protocol reference model (PRM) may be a useful model for conceptual visualization of certain functions which define ATM, but it certainly isn't the only one, nor do the manufacturers stick strictly to it when implementing ATM switches and interfaces. One of the reasons for this is that there are other standards setting bodies that define and further shape ATM. In addition to ITU-T, ANSI, and European Telecommunications Standards Institute (ETSI), one body that emerged as the dominant player in the ATM arena is the ATM Forum, founded by Northern Telecom, SPRINT, SUN Microsystems, and Digital Equipment Corporation (DEC) in 1991 [5].

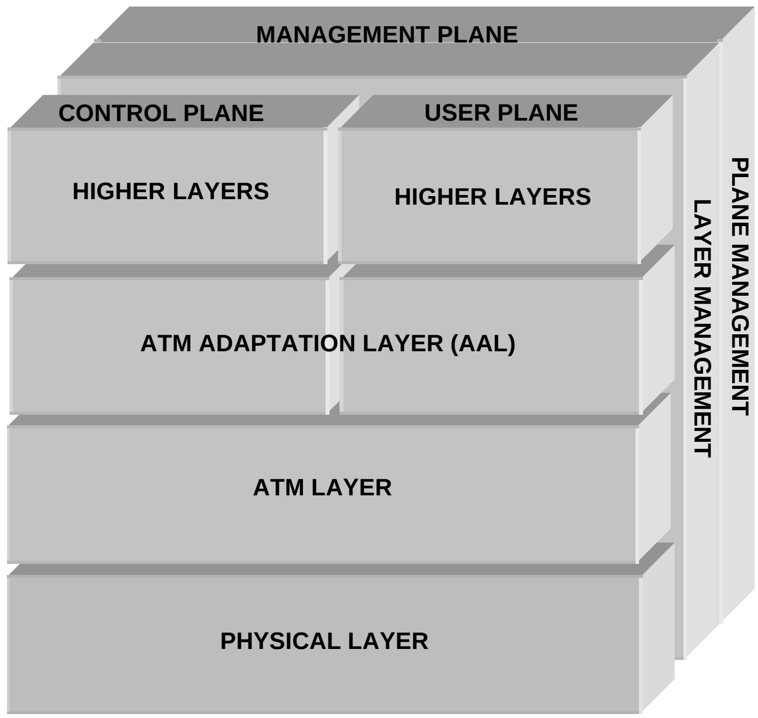


*Figure 3.6 B-ISDN Protocol Reference Model*

### 3.3.2. ATM Planes

The B-ISDN model consists of the three planes -- user, control and management planes -- which are labeled on top and stretch over the front and sides of the cube pictured in Figure 3.6. The user plane is responsible for transferring user information from applications, and the control plane is responsible for call and connection functions necessary for establishing the switched services. They both accomplish their functions by making use of the common underlying ATM and physical layers, as well as utilizing some of the same AAL layer protocols. The management plane is responsible for providing capabilities for exchanging information between the control and user planes, and is further divided into the layer management and plane management. The layer management is specifically responsible for management of functions specific to each layer, whereas plane management manages the entire system as a whole.

### 3.3.3. ATM Layers

Following its layered modeling practice, ITU-T offers further decomposition of the B-ISDN PRM into layers, and the detailed descriptions are contained in ITU-T recommendations I.321 and I.413 [9].

#### 3.3.3.1. Physical Layer

The physical layer provides for transmission of ATM cells over a physical medium connecting ATM devices; it is divided into two sublayers, the Physical Medium Dependent (PMD) and Transmission Convergence (TC) sublayer, as shown in Figure 3.7. One of the reasons for sublayers is to decouple the transmission from the physical medium to allow for many different physical media [13].

The PMD sublayer includes bit generation, transmission and reception capabilities and bit alignment. Line coding and possible electro-optical conversion are also functions of the PMD, as well as all other medium specific functions required for transmission and reception of data such as voltage levels, frequency, light intensity, wavelength, etc. For this project, the physical medium is multimode fiber. All timing, synchronization and framing methods are done according to SONET specifications.

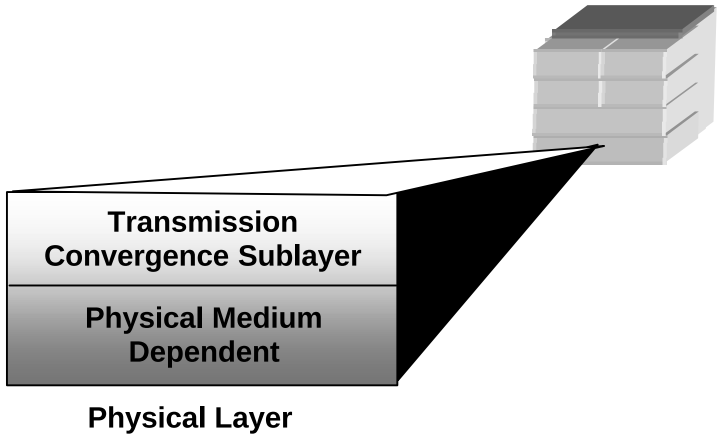


*Figure 3.7 Physical Layer of the B-ISDN PRM*

The TC sublayer maps the ATM cells to and from the PMD sublayer. In this case TC is responsible for mapping the ATM cells into SONET's OC-3c SPE. Since the integer number of ATM cells does not fit into the SPE, the TC sublayer continuously maps the cells across the SPE's boundaries. However, on the receiving end the TC needs to perform cell delineation, which is a process of recovering the ATM cells boundaries out of the continuous bit stream now arriving from the PMD.

Two functions helpful for cell delineation are scrambling / descrambling of the cell's information field and generation / verification of the Header Error Check (HEC). Scrambling / descrambling prevents the cell delineation mechanism from malicious attacks of bundled errored blocks of bits. Generation / verification of the HEC is a one-byte code applied to the header, capable of correcting any single-bit error in the header as well as detecting many patterns of multiple bit errors. In addition to these, the TC sublayer is also responsible for cell-rate decoupling, which is insertion at the transmission side and then suppression on the receiving side of the idle cells, which helps maintain a continuous flow of cells for the PMD-specific rate [5, 9, 13].

### 3.3.3.2. ATM Layer

The ATM layer can be further divided into the Virtual Channel (VC) level and Virtual Path (VP) level as shown in Figure 3.8.

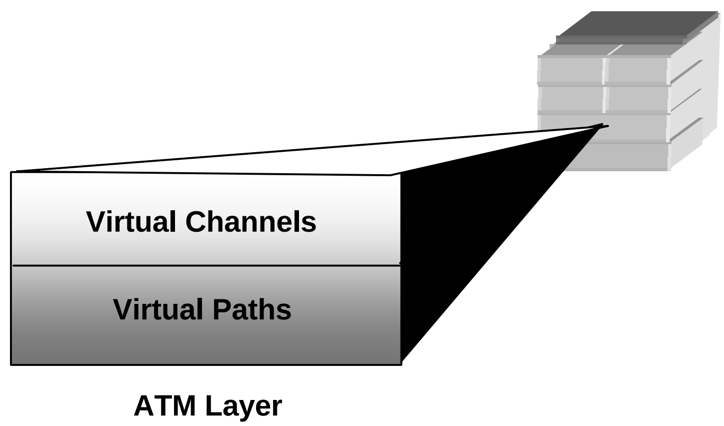


*Figure 3.8 ATM Layer of the B-ISDN PRM*

The VPs and VCs are instrumental to the ATM operations, and are defined in ITU-T Recommendation I.113 as [9]:

VC: "A concept used to describe unidirectional transport of ATM cells associated by a common unique identifier value."

VP: "A concept used to describe unidirectional transport of cells belonging to virtual channels that are associated by a common identifier value."

The relationship between VPs and VCs is shown in Figure 3.9.

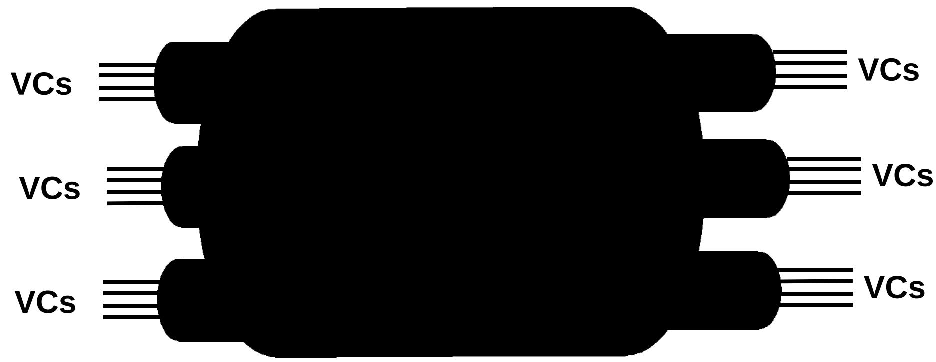


Figure 3.9 ATM Physical Circuit, VPs and VCs

It is the ATM layer that contains the functions that are most unique to ATM technology. First, at the transmitter, cells from individual VPs and VCs are multiplexed into one composite stream. At the receiving side, cells arriving from a composite stream are demultiplexed into individual cell flows designated for individual VP and VC cell streams.

Second, the ATM layer at the switching/routing nodes performs Virtual Channel Identifier (VCI) and Virtual Path Identifier (VPI) translation. VPIs and VCIs are the unique identifier values mentioned in the above ITU-T's definitions for VPs and VCs. VPI identifies a bundle of one or more VCs, and VCI identifies one unique VC in a particular VP. VCIs and VPIs are included in the ATM cell headers and are used for establishing end-to-end connections between ATM end devices as well as for necessary routing and switching during the connection setup phase.

VPIs and VCIs have only a local significance and therefore can be reused [5]. The ATM layer of each ATM device in a network between the end users assigns VPIs and VCIs necessary for creating an end-to-end VP or VC. Since it is possible that each switching node already uses a certain VPI or VCI, it is this case where the VPI/VCI translation is needed. In other words, switches are allowed to translate the values of individual VPIs and VCIs to any available value as long as that value describes the unique VP or VC and preserves the integrity of the end-to-end connection.

The third function of the ATM layer is generation/extraction of the header. After receiving cell information field from the higher layers, the ATM layer generates and adds an appropriate header, except for the HEC value, which is done by the physical layer. On the receiving side, it removes the header and passes the cell information field to the appropriate higher layer.

Finally, the ATM end device connecting to the network through the User to Network Interface (UNI) utilizes a Generic Flow Control (GFC) parameter which helps with the control of the ATM traffic flow in a customer network [9].

#### 3.3.3.3. ATM Adaptation Layer

The AAL layer offers the services provided by the ATM layer to the higher layers. Most of the higher layers hand down their information in variable size packets, or have some other unique bit rates or information transfer requirements. However, no other service communicates in 53 byte cells so it is the AAL's function to adapt all these different services to fit the ATM's 48 Byte information field or payload. In order to accomplish its task, the AAL's functions are divided into the Segmentation and Reassembly (SAR) Sublayer and the Convergence Sublayer (CS), as shown in Figure 3.10.

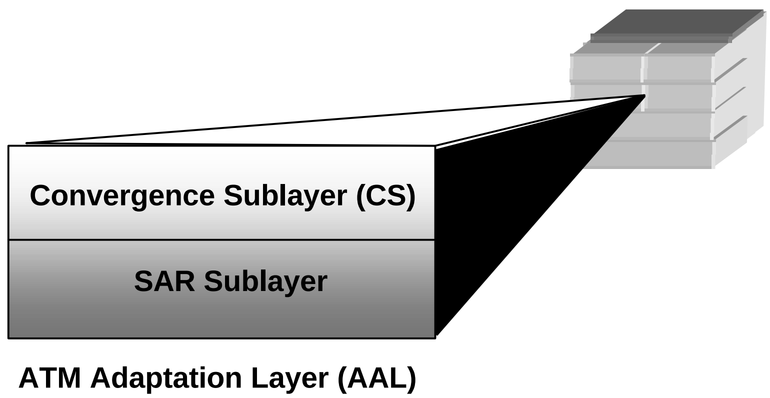


*Figure 3.10 ATM Adaptation Layers of the B-ISDN PRM.*

It is the SAR sublayer that is responsible for segmentation of higher layer information into the format suitable for consecutive ATM cells for a certain VC. The SAR layer is also responsible for extracting the higher layer information from the incoming ATM cells from a particular VC and reassembling them into a format usable by the higher layers.

The CS layer is responsible for identifying individual messages, recovery of timing, flow control, etc. It is further subdivided into the Service-Specific (SS) and Common Part (CP) sublayers. The SS sublayer performs certain functions particular to higher layer AAL user or, in absence of need for such functions, it may be null [5]. The CP layer, however, must always be complemented with the lower SAR layer, and is responsible for passing off SAR Protocol Data Units (PDU), which are essentially the ATM cell payload, to and from the ATM layer.

In general it is the AAL which needs to interface with an ever-expanding set of higher layer functions. The AALs were geared toward different classes of service, as shown in Table 3.3, and were classified based on the time relationship between source and destination, bit rate and connection mode [14].

*Table 3.3 AAL Service Classification [15]*

| | **CLASS A** | **CLASS B** | **CLASS C** | **CLASS D** |
|---|---|---|---|---|
| Timing Relation between Source and Destination | Required | | Not Required | |
| Bit Rate | Constant | Variable | | |
| Connection Mode | Connection Oriented | | | Connectionless |
| AAL Protocol | Type 1 | Type 2 | Type 3/4 Type 5 | Type 3/4 |

So far there are five AAL protocols than can provide different classes of service to the higher layers [9].

1. **AAL-0** is an AAL with empty SAR and CS, therefore adds no functionality between the ATM layer and higher layers, allowing information to be transparently transferred through the AAL.

2. **AAL-1** is generally recommended for constant bit rate applications sensitive to cell loss, delay and jitter, such as voice and constant rate video. It is also very appropriate for emulating constant-rate leased lines.

3. **AAL-2** provides transfer of variable bit rate applications with preservation of timing relationship between source and destination. It is commonly intended for the transport of variable bit rate compressed voice and video.

4. **AAL-3/4** was created by merging what were two separate AALs intended for two separate classes of service. However it was possible to merge the two AALs into one, which is able to satisfy demands from both classes of service. AAL-3/4 is mainly intended for applications that are sensitive to loss but not to delay. It can be used for both connection oriented as well as connectionless services.

5. **AAL-5** is mainly intended for data traffic although it has recently been used for all classes of service. AAL 5 offers same services as the AAL-3/4, but it does so with less overhead, no support for multiplexing functions and less complexity.

### 3.3.4. ATM Resources

The combination of fixed cell sizes, concept of virtual paths and virtual channels and five different categories of ATM adaptation layer protocols makes ATM one of the most flexible high speed networking technologies implemented today. In addition to it's flexibility for supporting voice, video and data and ability to scale to large networks, ATM is the first and so far the only widely implemented technology that can offer Quality of Service (QoS). Some of the ATM resources responsible for these achievements are discussed below.

### 3.3.4.1. ATM Cell

The transmission, switching and multiplexing unit of ATM technology is the 53-byte, fixed length packet, or cell. There are two formats for ATM cells, and which format is used depends on where they are used in the network. The traffic between the end user and the ATM network is carried on UNI cell format while the internetworking is done over the Network-to-Network Interface (NNI) cell formats. Both cells however are identical size of 53-bytes and both consist of 48-bytes payload and 5-bytes header, as presented in Figure 3.11 [5].

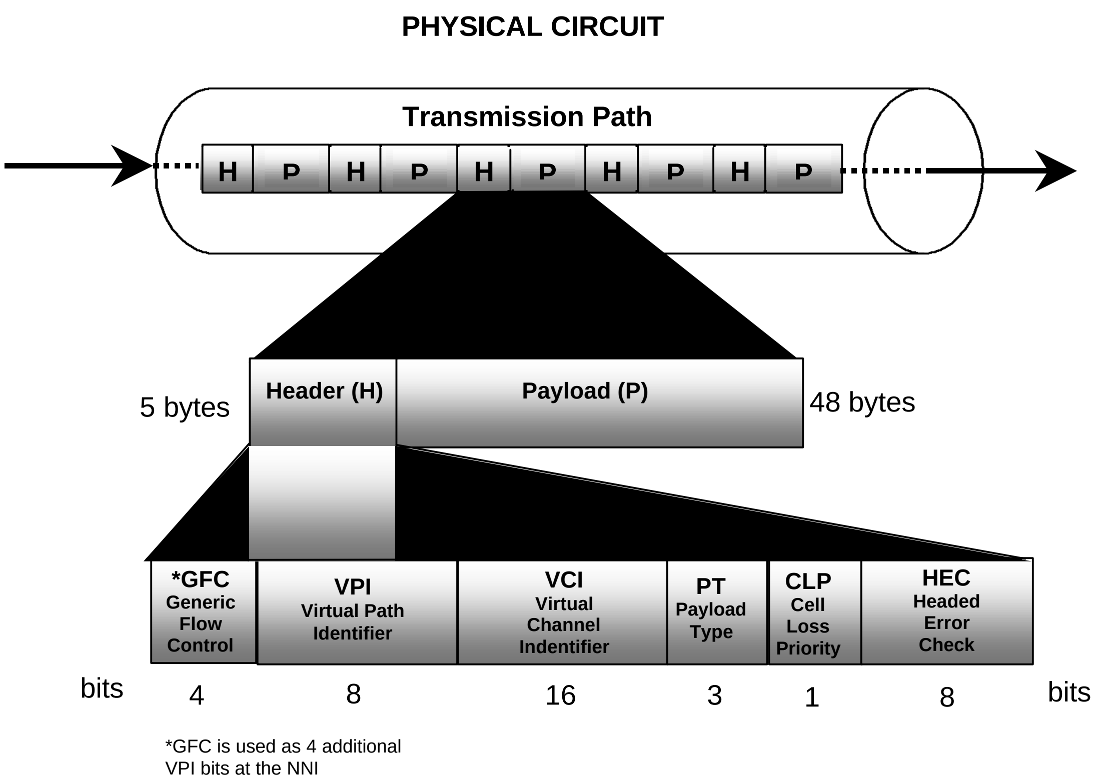


*Figure 3.11 ATM Cell Format*

The only difference between the two formats is the four-bit GFC parameter contained in UNI cell format, which are replaced by additional VPI field in a NNI cell format.

There are somewhat different stories behind the reasons for the 48-bytes payload size but all of them agree that it was an eventual compromise between the 32-Bytes and 64-Bytes sizes. The main reason for debate, however, was focused on the basic tradeoff between the smaller 32-bytes cells and larger 64-bytes cells. The smaller 32-bytes cells were favored in Europe where the small country sizes would enable the European telephone companies to achieve low enough latency in packing and transporting the voice via cell without the need to install echo cancellation devices. On the other hand the size of the US forces the phone companies to install echo cancellers anyway, so the large 64-bytes packets would be preferable since they would offer higher overall transmission efficiency by improving the header-to-payload ratio [6]. There is a mention that the size of the header was a separate tradeoff on its own between the increased function through an 8-bytes header and increased efficiency through a 3-bytes header [5].

As mentioned earlier, the switching and multiplexing of a cell is done by using VPIs and VCIs ,which are contained at the beginning of the cell header and allow for finding routes from input ports to output ports through individual switches by VPI/VCI translation, as well as establishing the end-to-end connection.

The next header field is the three-bit Payload Type (PT) Indicator which basically differentiates between the control data or the user data being carried in the payload of the cell. Aside from user data, the PT indicates VC Channel (VCC) -level Operation and Maintenance (OAM) information, Explicit Forward

Congestion Indication (EFCI), AAL information and the Resource Management (RM) information [16].

The Cell Loss Priority (CLP) bit differentiates between the high and low priority traffic and it is a topmost layer of priority setting for the traffic that has already passed through other congestion management systems [11]. For example, this bit can be set for low priority either by an application which could afford to lose certain cells or by a network element which encounters cells transmitted in excess of what was initially negotiated [6].

Finally the HEC, which uses an eight -bit Cyclic Redundancy Checksum (CRC) polynomial performs an error detection and single error correction on the header.

#### 3.3.4.2. Quality of Service

As is described in section 3.3.3.3, ATM provides 5 classes of service, which in turn can offer support to most existing data services and their various traffic patterns. Certain services can tolerate some cell loss while others cannot. Other services have timing constraints while some do not. And some may have multiple constraints while some have none. Thus, in order for ATM to be able to support all the different services, while using available network resources efficiently and providing specified and guaranteed levels of Quality of Service (QoS), it needs to bound and control the different traffic parameters associated with each service.

### 3.3.4.3. Contract Parameters

During the connection set-up a source can specify the following traffic parameters [17]:

***Peak Cell Rate (PCR)*** represents the maximum instantaneous rate at which the user will transmit. This parameter is the inverse of the minimum cell inter-arrival time.

***Sustained Cell Rate (SCR)*** is the average rate for a given connection over a long time period.

***Cell Loss Ratio (CLR)*** is the ratio of the number of lost cells sent by a source to the total number of cells sent by the source, as presented in Equation 3.1.

$$\text{CLR} = \frac{\#\ \text{Lost Cells}}{\#\ \text{Transmitted Cells}}$$

Equation 3.1

The cells are usually lost in the network due to error or congestion and buffer overflows. If congestion is the reason then the network will first discard the low priority cell or cells with CLP=1. Nevertheless, CLR can be set separately for both the low priority cells and high-priority cells.

***Cell Transfer Delay (CTD)*** is the total delay experienced by the cell from the time it enters the network until it leaves the network. This parameter includes propagation delay, queuing delays at intermediate ATM devices and service times at queuing points.

***Cell Delay Variation (CDV)*** is a measure of variance of CTD. It is a variation of delay or jitter experienced by a cell propagating through the network.

***Burst Tolerance (BT)*** is the maximum burst size that can be sent at the PCR and it is used to control the traffic entering the network via the leaky bucket algorithm which puts all arriving cells in a buffer (bucket) and sends them at the SCR. The Maximum Burst Size (MBS) is the maximum number of back-to-back cells that can be sent out of the bucket at the PCR. The relationship between the BT and MBS is shown in Equation 3.2.

***Minimal Cell Rate (MCR)*** is the minimum rate that the user desires.

$$BT = (MBS - 1)\left(\frac{1}{SCR} - \frac{1}{PCR}\right) \qquad \text{Equation 3.2}$$

### 3.3.4.4. Service Categories

Once the traffic parameters of a service are made known to the ATM network, but before the connection through the network is established, a contract for desired QoS based on the traffic parameters presented is being negotiated. If all intermediate ATM devices en route from source to destination can support the requested QoS, the connection gets established and all the traffic requirements for that particular service are guaranteed for the duration of the connection. In a broad sense all traffic contracts specify one of the five service categories defined by the ATM Forum so far. The five service categories are [5]:

***Constant Bit Rate (CBR)*** service category supports real-time applications, which require a fixed amount of bandwidth defined by the PCR. CBR is best suited for applications with stringent requirements on CTD and CDV such as voice and the constant-bit-rate video. When the CBR request is made to the network, all switches and intermediate devices must be able to provide the specified PCR or the connection will be refused.

***Real-Time Variable Bit Rate (rt-VBR)*** service category allows for rate variation while preserving high demands on CTD and CDV. The rate variation is bounded by the PCR and the "average" rate defined by the SCR and MBS. The three parameters, PCR, SCR and MBS, define a worst case scenario for which the given QoS traffic contract will hold. The applications usually requesting rt-VBR are delay-variation-sensitive and possibly bursty in nature. Variable-bit-rate video is an example of such traffic.

***Non-Real-Time VBR (nrt-VBR)*** is a service category geared toward applications bursty in nature but with no constraints on delay and jitter. The traffic contract is the same as the one for the rt-VBR and is specified by PCR, SCR and MBS. Applications such as airline reservations, banking transactions and other terminal sessions would use nrt-VBR.

***Unspecified Bit Rate (UBR)*** is also known as the "best effort" service, since it does not specify any bandwidth, has no guarantees for throughput, places no constraint on delay nor jitter and therefore offers no QoS. Since most LANs and IP networks also offer only best effort service, LAN emulation, IP over ATM and other non-mission-critical applications use UBR. In general UBR is a good service category for handling applications which have built-in retransmission schemes [18].

***Available Bit Rate (ABR)*** service category uses a closed-loop flow control, which together with the sources that vary their transmission rates based on the network feedback, tries to dynamically vary the traffic flow in order to use all the available bandwidth in the network not used by the other service categories, as shown in Figure 3.12 [18].

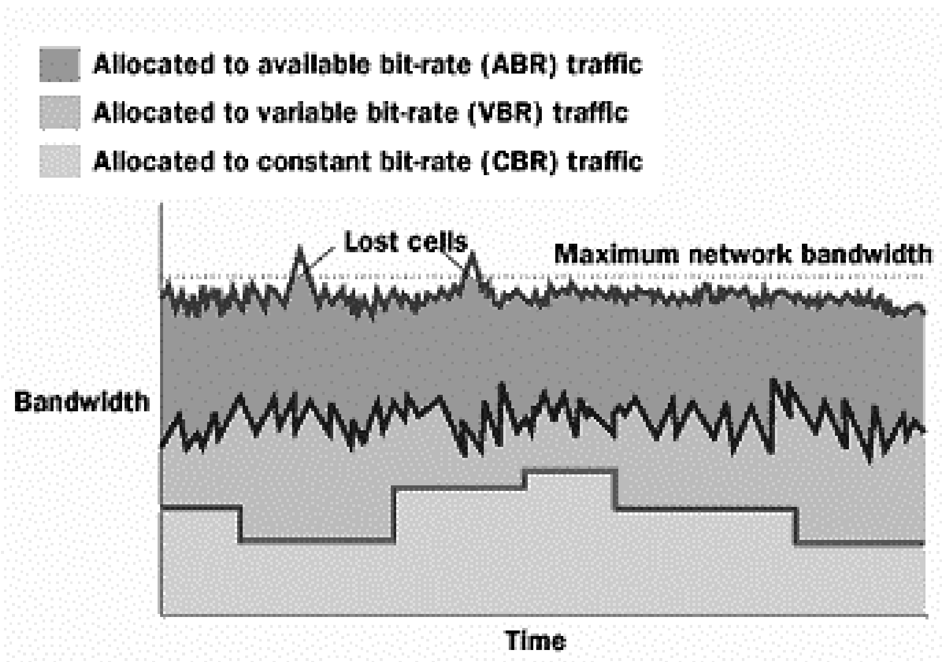


*Figure 3.12 Bandwidth Allocation for Different Service Categories* [18]

Thus, for users that can change their transmission rates, ABR can offer a relatively low loss service. The ABR service contract is based on the PCR and MCR parameters but offers no delay and jitter constraints. Therefore non-real-time applications such as LAN interconnections, web browsing, database archival and file transfers are good candidates for ABR.

### 3.3.4.5. Congestion Avoidance and Flow Control

As indicated previously, ABR and UBR support similar services but due to rate-based flow control protocols for avoiding congestion specified for ABR, ABR is becoming a preferred way for sending data traffic over ATM.

So far there are four options for rate-based flow control and they all rely on specialized resource management (RM) cells which provide feedback and some information about the congestion state in the network [19].

***Explicit Forward Congestion Indication (EFCI)*** is the simplest scheme for congestion avoidance. Any ATM node that is experiencing congestion can set the EFCI to indicate to the upstream nodes, and ultimately to the end-user, that once the EFCI is set no other network element can alter its value. The node sets EFCI to indicate congestion by setting first two bits of the PT field to 01, and once the end-user at the destination receives this indication it is up to higher layers to act, for instance by lowering the transmission rate at the source [15].

***Relative Rate (RR)*** marking is a rate-based method for flow control that uses Resource Management (RM) cells. Any node can set the Congestion Indication (CI) bit or the No-Increase (NI) bit in either forward or reverse direction of the traffic flow. As in the EFCI case, once any switch or node sets either one of the CI or NI bits, no other network element can alter them. When the source receives the CI bit, it decreases its rate, whereas the NI bit prevents the source from increasing the rate.

***Explicit Rate (ER)*** marking is another end-to-end flow control scheme and it uses the ER field in the RM cells for conveying information from the network nodes back to the source. The source sets the ER field to reflect the rate agreed on with the network at the connection set-up negotiation process. Usually the end-user requests the PCR; but since this is ABR, and there are no quality guarantees, the switches can change the ER in RM cells flowing in either direction to better reflect the resources available to that particular connection. When the source receives the RM cell with altered ER field, it has to adapt its transmission rate to the rate specified by the ER.

***Virtual Source/Virtual Destination*** control uses the same principles as listed above, except that instead of being an end-to-end flow control scheme, the control loop is broken into two or more segments, and the intermediate ATM nodes act as virtual sources and virtual destinations. In this way each segment can be treated as a smaller end-to-end problem and the congestion problem is, or can be, controlled on every hop.

If flow control fails, however, congestion can occur, and ATM has few congestion recovery techniques that may aid in restoration of normal operating conditions [5].

***Selective Cell Discard*** is a mechanism that allows the network to discard cells with low priority (CLP=1). This method can be a powerful tool for aiding congestion recovery, especially when used with Usage Parameter Control (UPC) tagging. UCP would change the CLP bit to indicate low priority for all cells that do not conform to the traffic parameters negotiated at the connection set-up.

***Early/Partial Packet Discard (EPD/PPD)*** is a scheme that maximizes the transfer of good packets by discarding on the packet level instead of the cell level. The main reason is that cells are generally much smaller than the higher layer variable-size packets and it takes many cells for transferring an entire packet. However, if one cell gets discarded then the SAR cannot complete the reassembly of the packet, and the request is usually sent for a retransmission of the entire packet -- not just the cell missing. Thus it is much better to free the buffer of the whole group of cells comprising the packet.

EPD prevents all the cells from a particular packet that has not entered the buffer yet from entering the buffer, and therefore reserving the buffer space for cells belonging to packets already admitted to the buffer. If a cell belonging to a certain packet already in the buffer gets dropped due to congestion, then PPD discards all remaining cells belonging to that packet, except the last one.

***Disconnection*** is another, and rather drastic, approach which can be used in times of prolonged severe congestion. In this scheme certain connections would be dropped regardless of the contracts. An example is the US government, which usually places requirements upon carriers to drop all other traffic if necessary in order to provide resources for priority national defense traffic or local emergency services.

### 3.4. LAN Emulation

There are obviously numerous advantages offered by ATM. However, most data traffic in existing Customer Premises Networks (CPN) is carried over LANs, Ethernet (IEEE 802.3) and Token Ring (IEEE 802.5) in [12]. In order for all disparate internet devices to operate over ATM, LANE is used to make ATM appear as if it were a traditional Ethernet or Token Ring. Basically, LANE disguises an ATM network so that it appears as a traditional network to the users and applications.

Disguising an ATM network to resemble a Legacy LAN is not an easy task considering that traditional LANs operate quite differently than ATM. Ethernet and Token Ring are both connectionless unlike the ATM which is connection oriented. The shared medium of LANs, in turn, make multicast and broadcast easily accomplished which isn't the case with ATM. Furthermore, Media Access Control (MAC) addresses used in LANs are burned in by the manufacturer and are completely independent from the network topology [12].

In order to overcome these major differences LANE is comprised of three major components [17]:

***LAN Emulation Client (LEC)*** maps the MAC address into an ATM address and then sets up an ATM channel between that LAN device's ATM address and an ATM address that corresponds to the MAC address of the target LAN device. The ATM channel is established by creating a direct VC between the corresponding ATM addresses followed by the necessary assignment of the VCIs.

***LAN Emulation Server (LES)*** is a LANE software entity that is mainly responsible for support of the LAN Emulation Address Resolution Protocol (LE-ARP). If a source LEC does not have the destination address in its address resolution table, it sends the LE-ARP request to the LES for the ATM address of the target LEC responsible for a certain destination MAC address.

***Broadcast / Unknown Server (BUS)*** has a main task of forwarding messages to certain LECs or broadcasting messages to all LECs attached to the given network.

## 3.5. Ethernet

In this project, Ethernet (IEEE 802.3) was emulated to run over the experimental laboratory ATM network. Ethernet is a very common LAN technology, which originated from early packet-radio networks, that uses Carrier Sense Multiple Access/ Collision Detection (CSMA/CD) over a 10 Mbps shared link. Basically Ethernet communicates across a medium shared by many users and thus it is a multiple access technology. "Carrier Sense" means that all users are capable of distinguishing between a busy or an idle link, and they can "listen" to determine whether they can transmit or not. "Collision Detection" means that the nodes can also "hear" whether the frame they just sent collided with some other user's frame. If so, the node will try to retransmit the frame after some random waiting period [6].

Generally, host adapters pack the data into a variable size payload of the Ethernet frame, which includes a CRC at the end of the frame and attaches the header at the beginning of the frame. The header is used to synchronize with the signal on the shared link, to include the destination and source address and to let the destination know which higher protocol the data is intended for and what type of data is contained in the payload, the maximum size of which is 1500 bytes.

## 3.6. Internet Protocol

Internet Protocol (IP) is a protocol that provides connectionless, best effort service of IP packets or datagrams across the Internet. Its wide acceptance is a result of its simplicity and ability to run over any kind of network.

However, this simplicity has a drawback: IP offers an unreliable service since IP just sends the packet with an included source and destination address, and then lets the network attempt to deliver it. Using the destination address from the header, the intermediate switching and routing devices know where to forward the packet next. The IP packets are also variable size with a maximum of 65535 bytes. In order to preserve flexibility and run over any network, IP supports fragmentation and reassembly so that its datagram size can always fit within each underlying network's Maximum Transmission Unit (MTU).

If the packet gets lost, however, IP has no way of telling, nor does it know, if there is congestion or even if the destination exists. If the packets don't get lost, they can still be delivered out of order, or sometimes any one packet can be delivered more than once. All in all, IP leaves it up to the higher layer protocols to handle these failure modes [6].

### 3.7. User Datagram Protocol

User Datagram Protocol (UDP) is an end-to-end transport protocol of the Internet that provides connectionless datagram services for processes-to-process communications between two or more hosts.

The main function of such a protocol is to be able to differentiate between multiple application processes running on a particular host, and be able to connect the same processes running on different hosts separated by some form of network. To successfully perform this function, UDP includes an identifier port on which the particular process is running for both the sender and the receiver of the message. Aside from this demultiplexing function, UDP offers only a "best effort" service, and implements neither flow control nor reliable/ordered delivery [6].

## 3.8. Transmission Control Protocol

Transmission Control Protocol (TCP) is, like UDP, a transport protocol. TCP also must allow multiple application programs on any given host to communicate simultaneously with its peers by supporting a demultiplexing mechanism. However, unlike UDP, TCP is a connection oriented, reliable/ordered delivery service, which implements a very efficient congestion control mechanism.

This mechanism in particular puts a limit on how fast the applications running over TCP can send data. Since TCP is a full-duplex protocol, supporting a connection in each direction across the network, it has feedback capabilities which can determine if there is a congestion problem in a network, and relay this data to the sender in order to slow down the transmission rate if needed [6]. In addition these functions can be performed and parameters can be changed dynamically for the duration of a particular connection, making the TCP very immune to constantly changing network conditions. This immunity was also manifested in this project.

## 3.9 The Big Picture

The setup assembled for this particular research project consists of most of these protocols stacked on top of each other as shown in Figure 3.13.

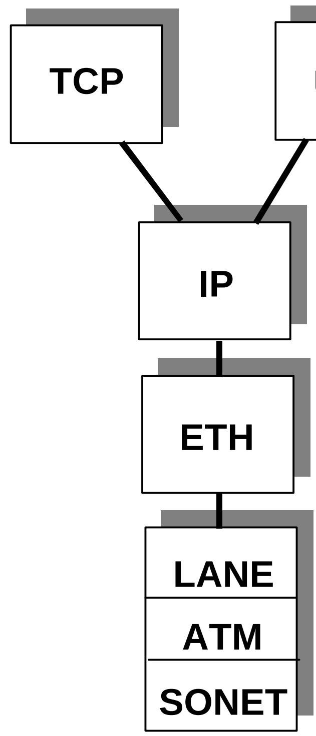


*Figure 3.13 Protocol Stack*

The application scripts described in Chapter 5 deliver data which represents a protocol data unit (PDU) to either UDP or TCP, which performs a variety of operations and adds its own header to create a segment. This segment is then passed down to IP which also performs certain operations and adds its own header creating a datagram in Internet terms [20]. This datagram is next passed down to Ethernet which appends its own header and trailer thereby forming a packet. LANE then takes the Ethernet packets, converts its MAC address into an ATM address, performs segmentation of the payload contained in the Ethernet packet, and adds an ATM header and trailer, thus forming ATM cells. The ATM cells are then placed within SONET's SPE which adds the POH, and finally forms a SONET frame. SONET frames are then carried across the network to the other host where processes are implemented in reverse.

It should be noted that partial decapsulation occurs in intermediary nodes, the two ATM switches in this particular case, which are needed for reading the addresses and further routing through the network toward the final destination.

This chapter points out that some of these protocols are connection-oriented and some are connectionless. Some have error checks over their headers while others have them over headers and payloads. Some are designed for simplicity and multiple access while others are complex and have a multitude of functions to provide reliable service.

All in all, it is a complex mix, which may have only one thing in common: the fiber. It is the fiber that ends up ultimately carrying all the information and is assumed to always perform well; if there is a problem, then it is assumed to be in higher network layers. In part because of this and the complexity of today's computer networks, it was decided in this project to investigate that assumption.

The goal is to gain some insight into whether or not user-level performance varies under slight degradations in the physical medium, and further how the degradation affects performance. The results could offer a better understanding of the interactions between network protocols and network elements. Perhaps the results will also yield that certain methods of configuring network parameters are better than others, and finally, what, if anything, should be done to avoid certain conditions or improve certain situations.