NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Computer Engineering

ATM Networking

Graduation Project
Com 400

Student: Amjad Hussain Malik (980848)

Supervisor: Prof. Dr. Fakhreddin Mamedov

NICOSIA 2001

TABLE OF CONTENTS

ACKNOWLEDGEMENT

First of all I am happy to complete the task which I had given with blessing of God, I would like to thanks my dearest parents for their moral and financial support to complete my ambition and without their endless support I would never achieve my ambition.

I wish to thank my supervisor Prof. Dr. Fakhreddin Mamedov for intellectual support, encouragement, and enthusiasm, which made this project possible, and his patience for correcting both my stylistic and scientific errors.

My sincerest thanks go to my advisor Mr. Tayseer Al-Shanableh and my teacher Miss. Besime Erin, who shared their suggestions and evaluations throughout the completion of my project.

ABSTRACT

Asynchronous Transfer Mode (ATM) is the optimal solution developed and accepted as the standard for public and private networks to implement and begin there evolution towards the Integrated Broadband Communication Network (IBCN). It is a flexible service available today.

Due to the direction of international standards it is important for everyone to become familiar with the language and principles of ATM and begin to incorporate this information into their own planning process. This project is in depth introduction what ATM is exactly, as well as where and why this technology originated and emerged as a revolutionary technological breakthrough. ATM is presented both in developmental and deliverable perspectives, which are then contrasted with other services that are utilized today. This contrast clearly delineates the differences that exist between these current services and ATM, and highlights the real values that can be achieved with ATM as a major part of the user network.

ATM, with its powerful and efficient aspects and attitudes is considered as the living fact of the future world.

INTRODUCTION

In the emerging field of high-speed virtual networking, Asynchronous Transfer Mode (ATM) is a key component. ATM is a telecommunications concept defined by ANSI and ITU (formally CCITT) standards for carriage of a complete range of user traffic, including voice, data, and made signals, on any User-to-Network Interface (ITN). As such, ATM is extremely well suited to high speed networking in the 1990s. ATM technology can be used to aggregate user traffic from existing applications onto a single UNI (e.g. PBX tie trunks, host-to-host private lines, video conference circuits), and to facilitate multi-media networking between high speed devices (e.g. workstations, supercomputers, routers or bridges) at multi-megabit Speeds (e.g. 150-M bit/s).

On the basis of its numerous strengths, ATM has been chosen by standards committees (e.g. ANSI T1, ITU SG XIII) as an underlying transport technology within much Broadband Integrated Services Digital Network (B-ISDN) protocol stacks. In this context, "transport" refers to the use of ATM switching and multiplexing techniques at the data link layer (i.e., OSI Layer 2) to convey end-user traffic from source to 35destination within a network.

While B-ISDN is a definition for public networks, ATM can also be used within private networking products. In recognition of this fact, and for clarity, this document defines two distinct forms of ATM UNI:

Public UNI - which will typically be used to interconnect an ATM user with an ATM switch deployed in a public service provider's network,

Private UNI - which will typically be used to interconnect an ATM user with an ATM switch that is managed as part of the same corporate network (e.g., MIS department responsible for the user device is also responsible for the private ATM switch).

The primary distinction between these two classes of UNI is physical reach. There is also some functionality differences between the public and private UNI due to the applicable requirements associated with each of these interfaces. Both UNIs share an ATM layers specification, but may utilize different physical media. Facilities that connect users to switches in public central offices must be capable of spanning long distances. In contrast, private switching equipment can often be located in the same room as the user device (e.g. computer, PBX), and hence can use limited distance technologies.

CHAPTER 1
ATM SWITCHING

1.1 Definition

Asynchronous transfer mode (ATM) is a high-performance, cell-oriented switching and multiplexing technology that utilizes fixed-length packets to carry different types of traffic. ATM is a technology that will enable carriers to capitalize on a number of revenue opportunities through multiple ATM classes of services; high-speed local-area network (LAN) interconnection; voice, video, and future multimedia applications in business markets in the short term; and in community and residential markets in the longer term.

1.2 Overview

Changes in the structure of the telecommunications industry and market conditions have brought new opportunities and challenges for network operators and public service providers. Networks that have been primarily focused on providing better voice services are evolving to meet new multimedia communications challenges and competitive pressures. Services based on asynchronous transfer mode (ATM) and synchronous digital hierarchy(SDH)/synchronous optical network (SONET) architectures provide the flexible infrastructure essential for success in this evolving market shown in figure 1.1.

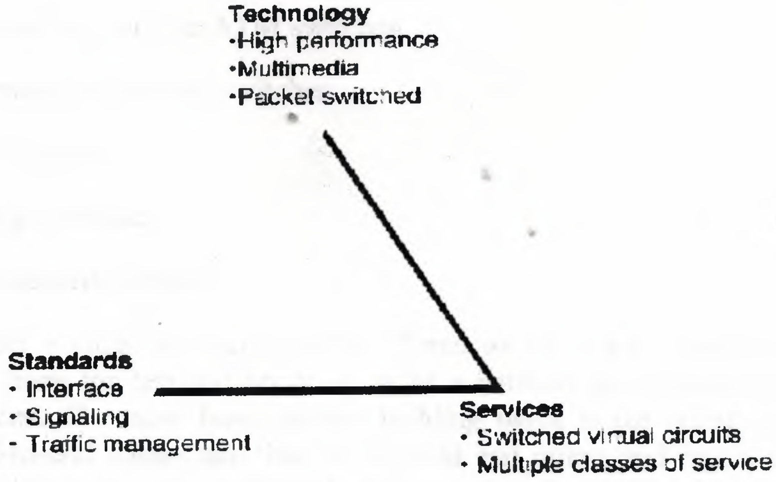


Figure 1.1 An ATM Model

ATM, which was once envisioned as the technology of future public networks, is now a reality, with service providers around the world introducing and rolling out ATM and ATM–based services. The ability to exploit the benefits of ATM technology within the public network successfully will provide strategic competitive advantage to carriers and enterprises alike.

In addition to revenue opportunities, ATM reduces infrastructure costs through efficient bandwidth management, operational simplicity, and the consolidation of overlay networks. Carriers can no longer afford to go through the financial burden and time required to deploy a separate network for each new service requirement (e.g., dedicating a network for a single service such as transparent LAN or frame relay). ATM technology will allow core network stability while allowing service interfaces and other equipment to evolve rapidly

1.3 Definition of ATM

Asynchronous transfer mode (ATM) is a technology that has its history in the development of broadband ISDN in the 1970s and 1980s. Technically, it can be viewed as an evolution of packet switching. Like packet switching for data (e.g., X.25, frame relay, transmission control protocol [TCP]/Internet protocol [IP]), ATM integrates the multiplexing and switching functions, is well suited for bursty traffic (in contrast to circuit switching), and allows communications between devices that operate at different speeds. Unlike packet switching, ATM is designed for high-performance multimedia networking. ATM technology has been implemented in a very broad range of networking devices:

- PC, workstation, and server network interface cards
- Switched-Ethernet and token-ring workgroup hubs
- Workgroup and campus ATM switches
- ATM enterprise network switches
- ATM multiplexers
- ATM–edge switches
- ATM–backbone switches

ATM is also a capability that can be offered as an end-user service by service providers (as a basis for terrified services) or as a networking infrastructure for these and other services. The most basic service building block is the ATM virtual circuit, which is an end-to-end connection that has defined end points and routes but does not have bandwidth dedicated to it. Bandwidth is allocated on demand by the network as users have traffic to transmit. ATM also defines various classes of service to meet a broad range of application needs.

ATM is also a set of international interface and signaling standards defined by the International Telecommunications Union–Telecommunications (ITU–T) Standards Sector (formerly the CCITT). The ATM Forum has played a pivotal role in the ATM market since its formulation in 1991. The ATM Forum is an international voluntary organization composed of vendors, service providers, research organizations, and users. Its purpose is to accelerate the use of ATM products and services through the rapid convergence of interoperability specifications, promotion of industry cooperation, and other activities. Developing multivendor implementation agreements also furthers this goal.

1.4 Benefits of ATM

The benefits of ATM are the following:

- High performance via hardware switching
- Dynamic bandwidth for bursty traffic
- Class-of-service support for multimedia
- Scalability in speed and network size
- Common LAN/WAN architecture
- Opportunities for simplification via VC architecture
- International standards compliance

The high-level benefits delivered through ATM services deployed on ATM technology using international ATM standards can be summarized as follows:

- High performance via hardware switching with terabit switches on the horizon
- Dynamic bandwidth for bursty traffic meeting application needs and delivering high utilization of networking resources; most applications are or can be viewed as inherently bursty; data applications are LAN–based and are very bursty, voice is bursty, as both parties are neither speaking at once nor all the time; video is bursty, as the amount of motion and required resolution varies over time
- Class-of-service support for multimedia traffic allowing applications with varying throughput and latency requirements to be met on a single network
- Scalability in speed and network size supporting link speeds of T1/E1 to OC–12 (622 Mbps) today and into the multi–Gbps range before the end of the decade; networks that scale to the size of the telephone network (i.e., as required for residential applications) are envisaged
- Common LAN/WAN architecture allowing ATM to be used consistently from one desktop to another; traditionally, LAN and WAN technologies have been very different, with implications for performance and interoperability

- Opportunities for simplification via switched VC architecture; this is particularly for LAN–based traffic that today is connectionless in nature; the simplification possible through ATM VCs could be in areas such as billing, traffic management, security, and configuration management

- International standards compliance in central-office and customer-premises environments allowing for multivendor operation

1.5 ATM Technology

In ATM networks, all information is formatted into fixed-length cells consisting of 48 bytes (8 bits per byte) of payload and 5 bytes of cell header (see *Figure 1.2*). The fixed cell size ensures that time-critical information such as voice or video is not adversely affected by long data frames or packets. The header is organized for efficient switching in high-speed hardware implementations and carries payload-type information, virtual-circuit identifiers, and header error check.

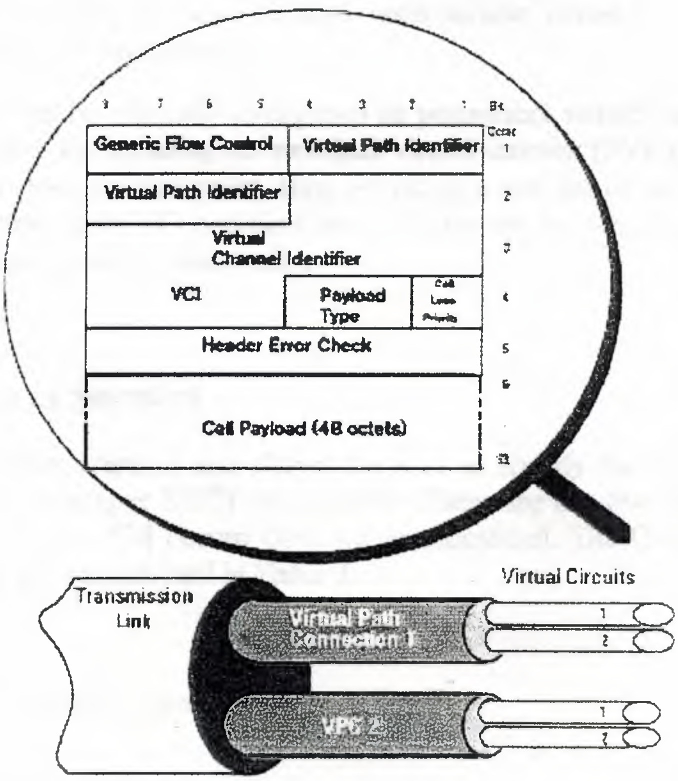


Figure 1.2 Fixed-Length Cells

ATM is connection oriented. Organizing different streams of traffic in separate calls allows the user to specify the resources required and allows the network to allocate resources based on these needs. Multiplexing multiple streams of traffic on each physical facility (between the end user and the network or between network switches)combined with the ability to send the streams to many different destinations enables cost savings through a reduction in the number of interfaces and facilities required to construct a network.

ATM standards defined two types of ATM connections: virtual path connections (VPCs), which contain virtual channel connections (VCCs). A virtual channel connection (or virtual circuit) is the basic unit, which carries a single stream of cells, in order, from user to user. A collection of virtual circuits can be bundled together into a virtual path connection. A virtual path connection can be created from end-to-end across an ATM network. In this case, the ATM network does not route cells belonging to a particular virtual circuit. All cells belonging to a particular virtual path are routed the same way through the ATM network, thus resulting in faster recovery in case of major failures.

An ATM network also uses virtual paths internally for the purpose of bundling virtual circuits together between switches. Two ATM switches may have many different virtual channel connections between them, belonging to different users. These can be bundled by the two ATM switches into a virtual path connection. This can serve the purpose of a virtual trunk between the two switches. This virtual trunk can then be handled as a single entity by, perhaps, multiple intermediate virtual path cross connects between the two virtual circuit switches.

Virtual circuits can be statically configured as permanent virtual circuits (PVCs) or dynamically controlled via signaling as switched virtual circuits (SVCs). They can also be point-to-point or point-to-multipoint, thus providing a rich set of service capabilities. SVCs are the preferred mode of operation because they can be dynamically established, thus minimizing reconfiguration complexity.

1.6 ATM Classes of Services

ATM is connection oriented and allows the user to specify the resources required on a per-connection basis (per SVC) dynamically. There are the five classes of service defined for ATM (as per ATM Forum UNI 4.0 specification). The QoS parameters for these service classes are summarized in *Table 1*.

Table 1.1 ATM Service Classes

Service Class	Quality of Service Parameter
Constant bit rate (CBR)	This class is used for emulating circuit switching. The cell rate is constant with time. CBR applications are quite sensitive to cell-delay variation. Examples of applications that can use CBR are telephone traffic (i.e., nx64 kbps), videoconferencing, and television.
Variable bit rate–non-real time (VBR–NRT)	This class allows users to send traffic at a rate that varies with time depending on the availability of user information. Statistical multiplexing is provided to make optimum use of network resources. Multimedia e-mail is an example of VBR–NRT.
Variable bit rate–	This class is similar to VBR–NRT but is designed for applications that are

real time (VBR–RT)	sensitive to cell-delay variation. Examples for real-time VBR are voice with speech activity detection (SAD) and interactive compressed video.
Available bit rate (ABR)	This class of ATM services provides rate-based flow control and is aimed at data traffic such as file transfer and e-mail. Although the standard does not require the cell transfer delay and cell-loss ratio to be guaranteed or minimized, it is desirable for switches to minimize delay and loss as much as possible. Depending upon the state of congestion in the network, the source is required to control its rate. The users are allowed to declare a minimum cell rate, which is guaranteed to the connection by the network.
Unspecified bit rate (UBR)	This class is the catch-all, other class and is widely used today for TCP/IP.

The ATM Forum has identified the following technical parameters to be associated with a connection. These terms are outlined in Table 1.2.

Table 1.2 ATM Technical Parameters

Technical Parameter	**Definition**
Cell loss ratio (CLR)	CLR is the percentage of cells not delivered at their destination because they were lost in the network due to congestion and buffer overflow.
Cell transfer delay (CTD)	The delay experienced by a cell between network entry and exit points is called the CTD. It includes propagation delays, queuing delays at various intermediate switches, and service times at queuing points.
Cell delay variation (CDV)	CDV is a measure of the variance of the cell transfer delay. High variation implies larger buffering for delay-sensitive traffic such as voice and video.
Peak cell rate (PCR)	The maximum cell rate at which the user will transmit. PCR is the inverse of the minimum cell inter-arrival time.
Sustained cell rate (SCR)	This is the average rate, as measured over a long interval, in the order of the connection lifetime.
Burst tolerance (BT)	This parameter determines the maximum burst that can be sent at the peak rate. This is the bucket-size parameter for the enforcement algorithm that is used to control the traffic entering the network.

Finally, there are a number of ATM classes of service. These classes are all outlined in Table 1.3.

Table 1.3 ATM Classes of Services

Class of Service	CBR	VBR–NRT	VBR–RT	ABR	UBR
CLR	yes	yes	yes	yes	no
CTD	yes	no	yes	no	no
CDV	yes	yes	yes	no	no
PCR	yes	yes	yes	no	yes
SCR					
no	yes	yes	no	no	
BT @ PCR	no	yes	yes	no	no
flow control	no	no	no	yes	no

Its extensive class-of-service capabilities make ATM the technology of choice for multimedia communications.

1.7 ATM Standards

The ATM Forum has identified a cohesive set of specifications that provide a stable ATM framework. The first and most basic ATM standards are those that provide the end-to-end service definitions as described in *Topic 4*. An important ATM standard and service concept is that of service interworking between ATM and frame relay (a fast-growing pervasive service), whereby ATM services can be seamlessly extended to lower-speed frame-relay users. Frame relay is a network technology that is also based on virtual circuits using variable-length frame transmission between users.

ATM user network interface (ATM UNI) standards specify how a user connects to the ATM network to access these services. A number of standards have been defined for T1/E1, 25 Mbps, T3/E3, OC–3 (155 Mbps) and OC–12 with OC–48 (2.4 Gbps) in the works. OC–3 interfaces have been specified for use over single-mode fiber (for wide-area applications) and over unshielded twisted pair or multimode fiber for lower-cost, in-building applications.

The following two ATM networking standards have been defined that provide connectivity between network switches and between networks:

- Broadband intercarrier interface (B–ICI)
- Public network-to-network interface (P–NNI)

P–NNI is the more feature-rich of the two and supports class of service-sensitive routing and bandwidth reservation. It provides topology-distribution mechanisms based on advertisement of link metrics and attributes, including bandwidth metrics. It uses a multilevel hierarchical routing model providing scalability to large networks. Parameters used as part of the path-computation process include the destination ATM address, traffic class, traffic contract, QoS requirements and link constraints. Metrics that are part of the ATM routing system are specific to the traffic class and include quality of service-related metrics (e.g., CTD, CLR) and bandwidth-related metrics (e.g., PCR). The path computation process includes overall network-impact assessment, avoidance of loops, minimization of rerouting attempts, and use of policy (inclusion/exclusion in rerouting, diverse routing, and carrier selection). Connection admission controls (CACs) define procedures used at the edge of the network, whereby the call is accepted or rejected based the ability of the network to support the requested QoS. Once a VC has been established across the network, network resources have to be held and quality service guaranteed for the duration of the connection.

All ATM traffic is carried in cells, yet no applications use cells. So, specific ways of putting the data into cells are defined to enable the receiver to reconstruct the original traffic. Three important schemes are highlighted in *Figure 1.3* and discussed in detail later in the tutorial.

- RFC1483, which specifies how interrouter traffic is encapsulated into ATM using ATM adaptation Layer 5 (AAL–5); AAL–5 is optimized for handling framed traffic and has similar functionality to that provided by HDLC framing in frame relay, SDLC, and X.25
- ATM LAN emulation (LANE) and multiprotocol over ATM (MPOA), which are designed to support dynamic use of ATM SVCs primarily for TCP/IP; LANE, which is a current standard that is widely deployed and will be a subset of the MPOA standard (which is targeted for standardization only in mid-1997), will be discussed later in the tutorial
- Voice and video adaptation schemes that can use AAL–1, which is defined for high efficiency—for traffic that itself has no natural breaks, such as a circuit carrying bits at a fixed rate

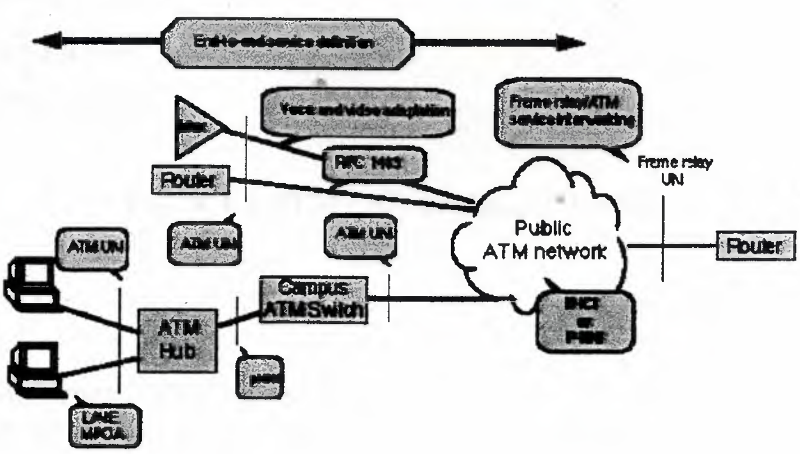


Figure 1.3 Data Insertion in Cells

1.8 ATM LAN Emulation

ATM–based Ethernet switches and ATM workgroup switches are being deployed by end users at various corporate sites. The most widely used set of standards in local ATM environments is ATM LAN emulation (LANE) (see *Figure 1.4*). ATM LAN emulation is used to make the ATM SVC network appear to be a collection of virtual-Ethernet/IEEE 802.3 and token-ring/IEEE802.5 LANs. The replication of most of the characteristics of existing LANs means that LAN emulation enables existing LAN applications to run over ATM transparently, this latter characteristic leading to its wide deployment. In ATM LAN emulation, most unicast LAN traffic moves directly between clients over direct ATM SVCs, while multicast traffic is handled via a server functionality. Bridging is used to interconnect real LANs and emulated LANs running on ATM, while routing is used to interconnect ATM–emulated LANs and other WAN or LAN media for purposes of routing scalability, protocol spoofing, or security firewalls.

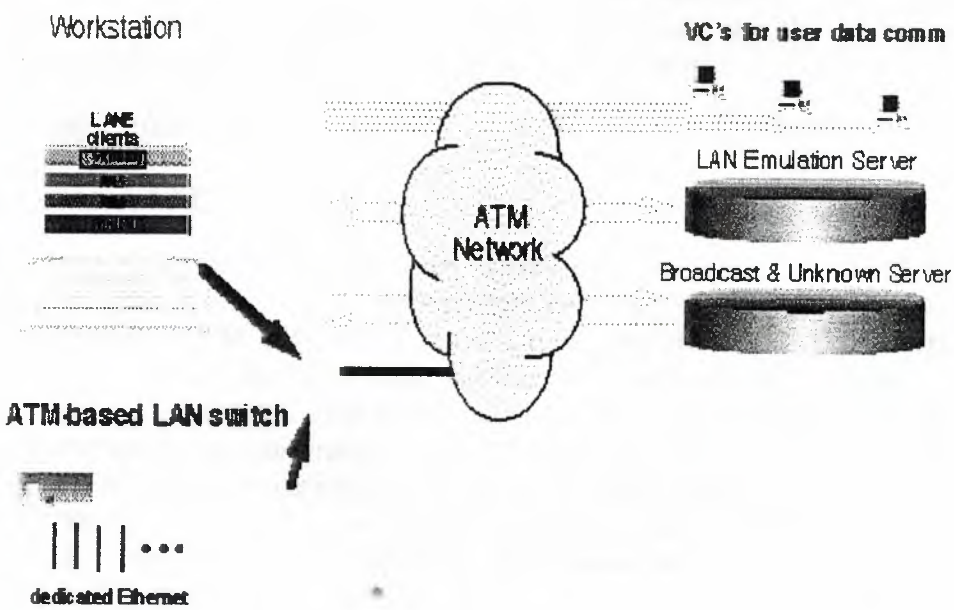


Figure 1.4 ATM LAN Emulation (LANE)

The ATM Forum LANE implementation agreement specifies two types of LANE network components connected to an ATM network.

- LANE clients which function as end systems, such as computers with ATM interfaces that operate as file servers; end-user workstations or personal computers; Ethernet or token-ring switches that support ATM networking; and routers, bridges, and ATM ENS with membership in an emulated ATM LAN

- LANE servers that support ATM LANE service for configuration management, multicast support, and address resolution

The LAN–emulation service may be implemented in the same devices as clients or involve other ATM network devices. The communications interface, LAN emulation user-network interface (LUNI), is the sequence and contents of the messages that the clients ultimately use to transfer traffic of the type expected on IEEE 802.3/5 LANs. The component of the LAN–emulation service that deals with initialization (i.e., emulates plugging the terminal into a LAN hub), is the LAN emulation configuration server (LECS). It directs a client to connect to a particular LAN emulation server (LES). The LES is the component of the LAN–emulation service that performs the address registration and resolution. The LES is responsible for mapping IEEE 48-bit MAC addresses and token-ring route descriptors to ATM addresses. One very important MAC address for clients is the MAC–layer broadcast address that is used to send traffic to all locations on a LAN. In LAN emulation, this function is performed by the broadcast and unknown server (BUS). ATM LANE is a comprehensive set of capabilities which has been widely deployed in ATM networks.

ATM LANE is an element of the multiple protocol over ATM (MPOA) architecture that is being defined by the ATM Forum. This work is addressing encapsulation of multiple protocols over ATM, automatic address resolution, and the routing issues associated with minimizing multiple router hops in ATM networks.

1.9 Voice over ATM

As real-time voice services have been traditionally supported in the WAN via circuit-based techniques (e.g., via T1 multiplexers or circuit switching), it is natural to map these circuits to ATM CBR PVCs using circuit emulation and ATM adaptation Layer 1 (AAL1). However, there are significant disadvantages in using circuit emulation in that the bandwidth must be dedicated for this type of traffic (whether there is useful information being transmitted or not), providing a disincentive for corporate users to implement circuit emulation as a long-term strategy. For example, a T1 1.544–Mbps circuit requires 1.74 Mbps of ATM bandwidth when transmitted in circuit-emulation mode. This does not downplay its importance as a transitional strategy to address the installed base.

As technology has evolved, the inherent burstiness of voice and many real-time applications can be exploited (along with sophisticated compression schemes) to decrease the cost of transmission significantly through the use of VBR–RT connections over ATM.

VBR techniques for voice exploit the inherently bursty nature of voice communication, as there are silence periods that can result in increased efficiency. The following silence periods (in decreasing levels of importance) arise:

- When no call is up on a particular trunk; that is, the trunk is idle during off-peak hours (trunks are typically engineered for a certain call-blocking probability: at night, all the trunks could be idle)
- When the call is up, but only one person is talking at a given time

- When the call is up, and no one is talking

Work is just starting in the ATM Forum on ATM adaptation for VBR voice.

The addition of more bandwidth-effective voice coding (e.g., standard voice is coded using 64–kbps PCM) is economically attractive, particularly over long-haul circuits and T1 ATM interfaces. Various compression schemes have been standardized in the industry (e.g., G720 series of standards). Making these coding schemes dynamic provides the network operator the opportunity to free up bandwidth under network-congestion conditions. For example, with the onset of congestion, increased levels of voice compression could be dynamically invoked, thus freeing up bandwidth and potentially alleviating the congestion while diminishing the quality of the voice during these periods.

A further enhancement to the support of voice over ATM is to support voice switching over SVCs. This entails interpreting PBX signaling and routing voice calls to the appropriate destination PBX (see *Figure 1.5*). The advantage from a traffic management perspective is that connection admission controls can be applied to new voice calls; under network congestion conditions, these calls could be rerouted over the public network and therefore not cause additional levels of congestion.

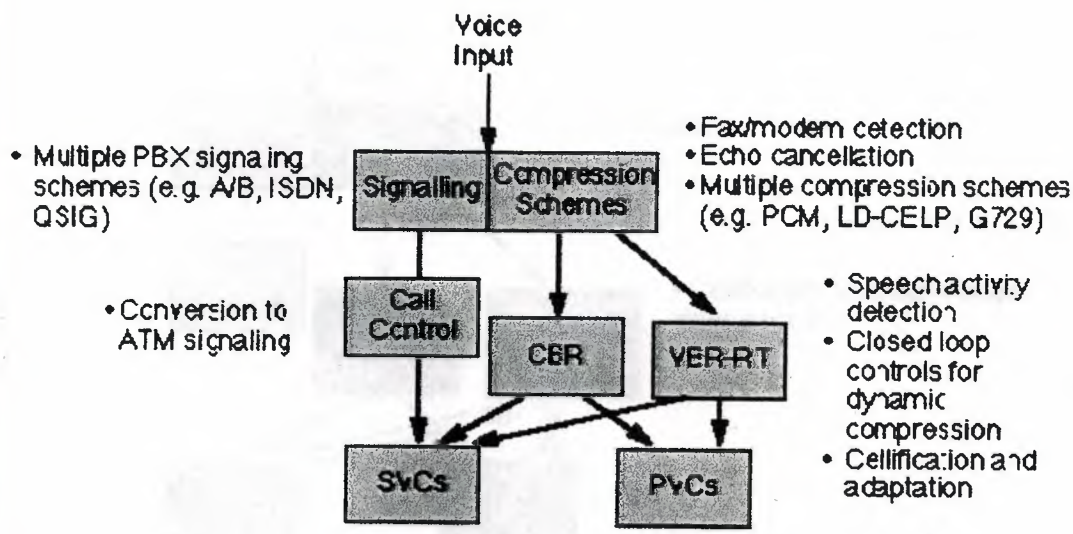


Figure 1.5 Voice Switching over SVCs

The ATM Forum is currently focusing its efforts on voice handled on CBR SVCs. VBR–RT voice is a future standards activity .

1.10 Video over ATM

While circuit-based videoconferencing streams (including motion JPEG running at rates around 10 Mbps) can be handled by standard circuit emulation using AAL–1, the ATM Forum has specified the use of VBR–RT VCs using AAL–5 for MPEG2 on ATM for video-on-demand applications, as this approach makes better use of networking resources.

MPEG is a set of standards addressing coding of video and surround-sound audio signals and synchronization of video and audio signals during the playback of MPEG data. It runs in the 2 Mbps to 15 Mbps range (with bursts above these rates) corresponding to VCR and broadcast quality respectively. The initial MPEG standard (MPEG1) was targeted at VHS–quality video and audio. MPEG2 targets applications requiring broadcast-quality video and audio and HDTV. MPEG2 coding can result in one of the following two modes:

- **Program streams**—variable-length packets that carry a single program or multiple programs with a common time base
- **Transport streams**—188-byte packets that contain multiple programs (for examples, see Figure 1.6).

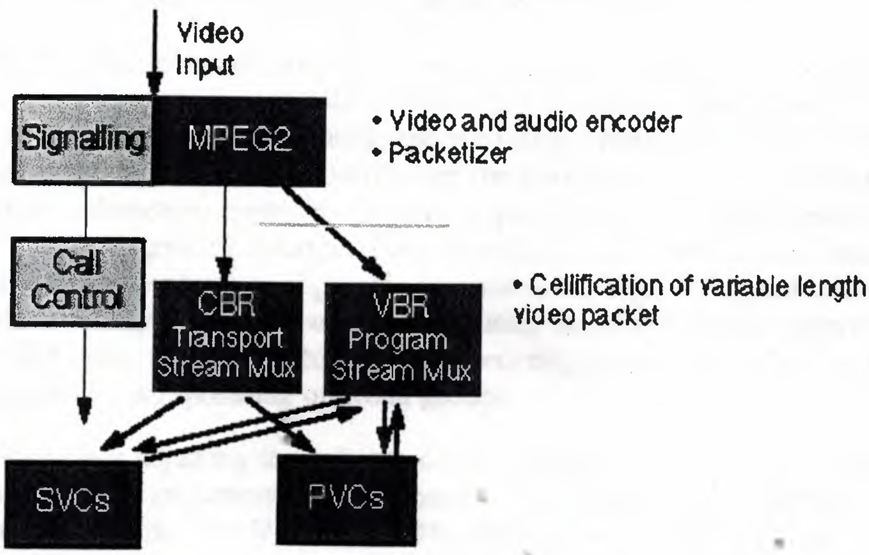


Figure 1.6 Transport Streams

In both cases, time stamps are inserted into MPEG2 packets during the encoding and multiplexing process. MPEG2 assumes a constant-delay model across the network, thus allowing the decoder to exactly follow the original encoder source clock. Due to the cost of coding, MPEG2 is primarily used in a non-interactive broadcast mode as would be the case for a point-to-multipoint broadcast in residential video on demand applications and in a business TV application for training or employee communications.

1.11 ATM Traffic Management

Broadly speaking, the objectives of ATM traffic management are to deliver quality-of-service (QoS) guarantees for the multimedia applications and provide overall optimization of network resources. Meeting these objectives enables enhanced classes of service and offers the potential for service differentiation and increased revenues, while simplifying network operations and reducing network cost.

ATM traffic management and its various functions can be categorized into three distinct elements based on timing requirements.First, are nodal-level controls that operate in real time. These are implemented in hardware and include queues supporting different loss and delay priorities, fairly weighted queue-servicing algorithms, and rate controls that provide policing and traffic shaping. Well-designed switch-buffer architectures and capacity are critical to effective network operation. Actual network experience and simulation has indicated that large, dynamically allocated output buffers provide the flexibility to offer the best price performance for supporting various traffic types with guaranteed QoS. Dynamically managing buffer space means that all shared buffer space is flexibly allocated to VCs on an as-needed basis. Additionally, per virtual connection (VC) queuing enables traffic shaping, and early and partial packet-level discard have been shown to improve network performance significantly.

Second, network-level controls operate in near real time. These are typically, but not exclusively, implemented in software including connection admission control (CAC) for new connections, network routing and rerouting systems, and flow-control-rate adaptation schemes. Network-level controls are the heart of any traffic-management system. Connection admission controls support sophisticated equivalent-bandwidth algorithms with a high degree of configuration flexibility, based on the cell rate for CBR VCs, average cell rate plus a configurable increment for VBR VCs, and minimum cell rate for ABR VCs. Dynamic class-of-service routing standards define support for fully distributed link-state routing protocols, auto-reconfiguration on failure and on congestion, and dynamic load spreading on trunk groups.

Flow control involves adjusting the cell rate of the source in response to congestion conditions and requires the implementation of closed loop congestion mechanisms. This does not apply to CBR traffic. For VBR and UBR traffic, flow control is left as a CPE function. With ABR, resource management (RM) cells are defined, which allow signaling of the explicit rate to be used by traffic sources. This is termed rate-based flow control. ABR is targeted at those applications that do not have fixed or predictable bandwidth requirements and require access to any spare bandwidth as quickly as possible while experiencing very low cell loss. This allows network operators to maximize the bandwidth utilization of their network and sell spare capacity to users at a substantial discount while still providing QoS guarantees. To enhance the effectiveness of network-resource utilization, the ABR standard provides for end-to-end, segment-by-segment, and hop-by-hop service adaptation.

Third, network engineering capabilities operating in nonreal time support data collection, configuration management, and planning tools (see *Figure 1.7*).

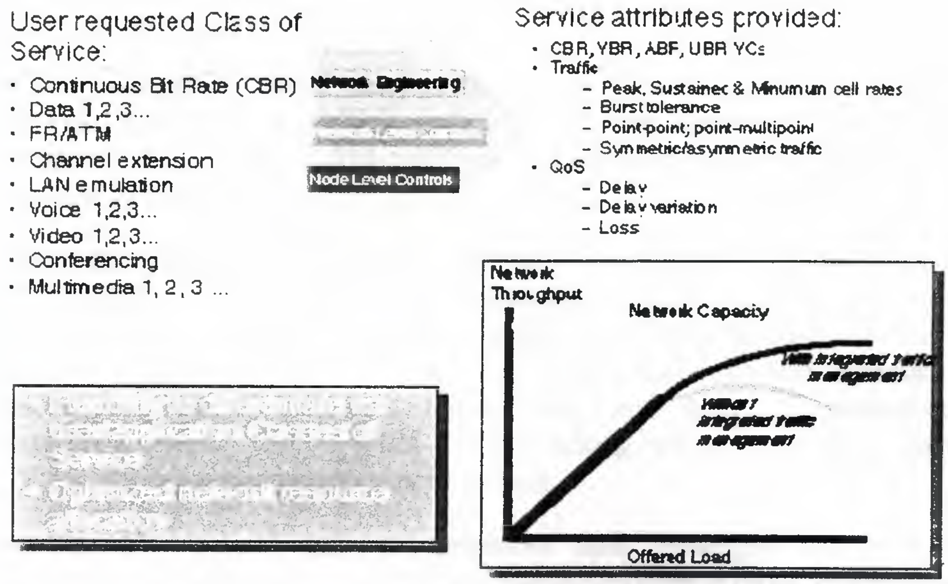


Figure 1.7 Network Engineering Capabilities

1.12 Applications of ATM

ATM technologies, standards, and services are being applied in a wide range of networking environments, as described briefly below (see Figure 1.8):

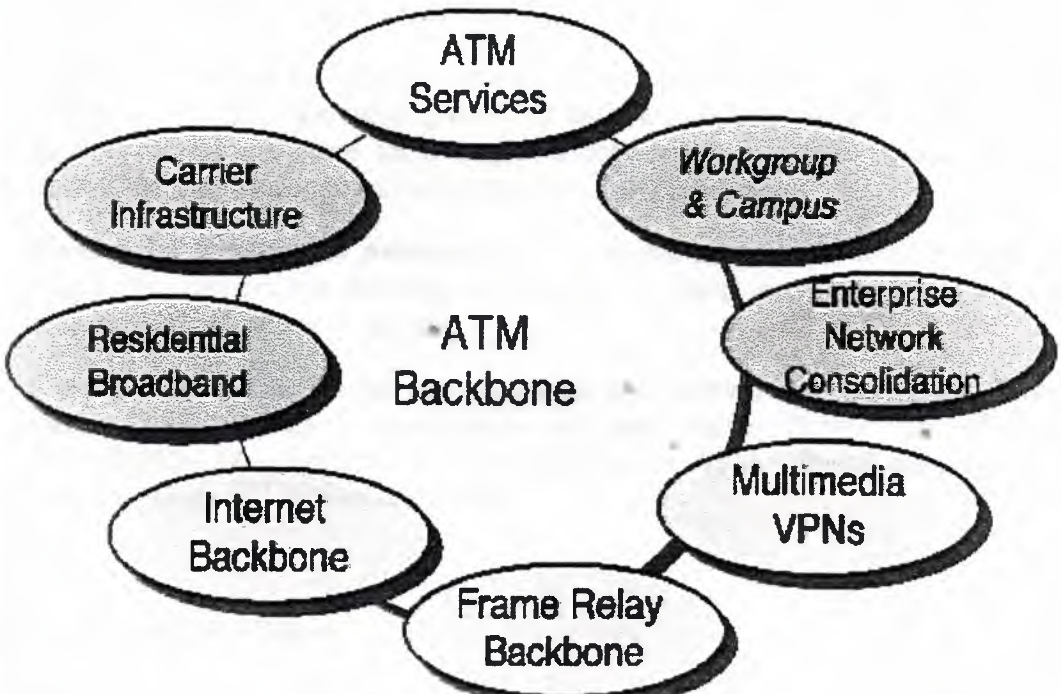


Figure 1.8 ATM Technologies Standards, and Services

- **ATM services**—Service providers globally are introducing or already offering ATM services to their business users.

- **ATM workgroup and campus networks**—Enterprise users are deploying ATM campus networks based on the ATM LANE standards. Workgroup ATM is more of a niche market with the wide acceptance of switched-Ethernet desktop technologies.

- **ATM enterprise network consolidation**—A new class of product has evolved as an ATM multimedia network-consolidation vehicle. It is called an ATM enterprise network switch. A full-featured ATM ENS offers a broad range of in-building (e.g., voice, video, LAN, and ATM) and wide-area interfaces (e.g., leased line, circuit switched, frame relay, and ATM at narrowband and broadband speeds) and supports ATM switching, voice networking, frame-relay SVCs, and integrated multiprotocol routing.

- **Multimedia virtual private networks and managed services**—Service providers are building on their ATM networks to offer a broad range of services. Examples include managed ATM, LAN, voice and video services (these being provided on a per-application basis, typically including customer-located equipment and offered on an end-to-end basis), and full-service virtual private-networking capabilities (these including integrated multimedia access and network management).

- **Frame-relay backbones**—Frame-relay service providers are deploying ATM backbones to meet the rapid growth of their frame-relay services to use as a networking infrastructure for a range of data services and to enable frame relay to ATM service interworking services.

- **Internet backbones**—Internet service providers are likewise deploying ATM backbones to meet the rapid growth of their frame-relay services, to use as a networking infrastructure for a range of data services, and to enable Internet class-of-service offerings and virtual private intranet services.

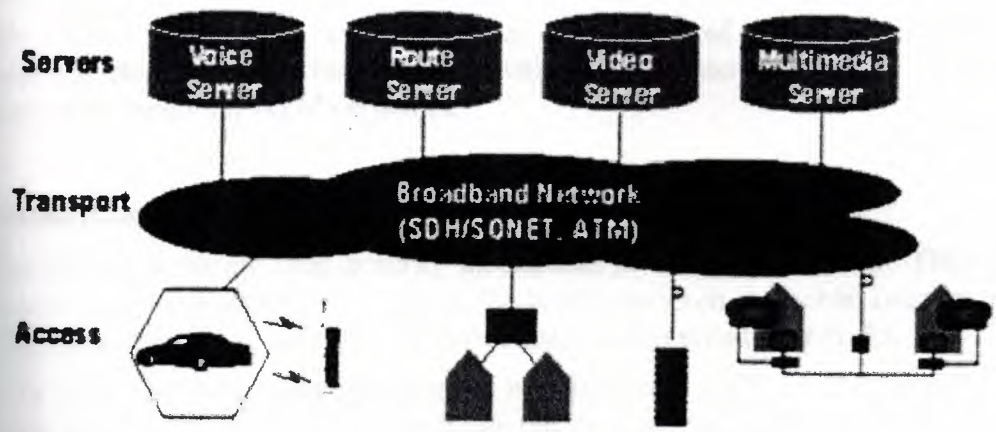
- **Residential broadband networks**—ATM is the networking infrastructure of choice for carriers establishing residential broadband services, driven by the need for highly scalable solutions.

- **Carrier infrastructures for the telephone and private-line networks**—Some carriers have identified opportunities to make more-effective use of their SONET/SDH fiber infrastructures by building an ATM infrastructure to carry their telephony and private-line traffic.

1.13 Nortel's ATM Vision

Nortel believes that ATM is the only viable backbone networking technology that can meet the objective of making multimedia calls as easy, reliable, and secure as voice calls are today.

ATM coupled with SONET/SDH for fiber transport, sits at the core of Nortel's long-term architectural vision. That vision embraces various residential, business, and mobile access arrangements with a set of voice/data/video and, ultimately, multimedia servers. There will be many ways of accessing ATM networks including desktop ATM, switched Ethernet, wireless, and xDSL, to name a few. The vision includes extensive support of multiple classes of service for native ATM, IP–based, frame-relay-based, and circuit-based applications. ATM accommodates the inherently bursty nature of data, voice, and video applications and the compressibility of these traffic types for increased storage and bandwidth effectiveness. Nortel also believes that frame relay and ATM, being both virtual-circuit based, provide a service continuum supporting the broadest sets of speeds from sub-64 kbps all the way to Gbps. Finally, Nortel envisages a family of application servers around the periphery of this network to provide a range of data, image, video and voice services that take advantage of increasing insensitivity of the network to distance (see Figure 1.9).



A flexible scalable architecture:

· Service and media independent access
· Multiservice high performance transport
· Application servers on industry standard platforms
· Bandwidth on demand eliminating network as bottleneck

Figure 1.9 Nortel's ATM Architectural Vision

CHAPTER 2
ATM FIREWALLS

2.1 Abstract

Here are many differences between ATM and today's most commonly used network 2.1 technologies. New firewall architectures are required to exploit the advantages of ATM technology and to support the high throughput available in ATM networks.

This chapter begins with a discussion of the impact of ATM on firewalls and then introduces the idea of parallelized firewalls, which may be used in order to achieve the high performance necessary for ATM networks.

2.2 Introduction

Firewalls are a widely used security mechanism in the Internet today. They are mostly used to provide access control and audit at the border between the public Internet and private networks, but are also used to secure critical subnets within private networks.

ATM is another somewhat newer trend in networking today. ATM provides a scalable high-speed network infrastructure, based on the concepts of fixed-length cells and virtual circuits. These conceptual differences to "legacy" networks and the high throughput of ATM networks present both challenges and new opportunities for firewall concepts.

This chapter discusses ATM specific topics of firewall design for ATM networks. General firewall issues such as security policies or implementation of firewalls are not discussed.

It presents performance measurements of the two most important firewall components: packet screens and proxy servers. It will be shown that the high processing requirements in both packet screens and proxy servers are the source of a severe throughput bottleneck of firewalls in ATM networks.

Also parallel protocol processing is introduced in chapter, this is one promising solution to the need for increased firewall performance resulting from the high scalability of ATM networks. Several concepts for parallel firewalls are discussed.

The following section gives a short introduction into ATM before discussing the consequences of using ATM in conjunction with firewalls. Different approaches to integrate packet screens into "Classical IP over ATM" networks are considered.

2.3 ATM as a challenge for firewalls

Firewalls are widely deployed to protect critical subnetworks from public networks. While today firewalls are mostly used in networks not exceeding throughputs of 10 Mbit/s, most sites are currently upgrading to high-speed networks (HSN) like Fast-Ethernet, Gigabit-Ethernet or ATM. As firewalls are, by design, "choke-points", firewall performance is a major concern in HSNs. In addition to the performance requirements, ATM networks also introduce new networking concepts, which require a revision of current firewall concepts.

2.3.1 Implications of ATM on firewalls

ATM networks introduce four major challenges to firewalls:

2.3.1.1 Performance

Firewalls, as already stated, are a bottleneck by design. In order to increase security, all traffic is channeled through a small number of firewall systems. Current increase in workstation performance cannot cope with the easy scalability of ATM networks. 622 Mbit/s or even 1.2 Gbit/s can easily be achieved with ATM networks. Workstations cannot perform even simple filtering at these speeds. In addition to the lack of firewalls to transfer legitimate traffic at high speeds, various attacks can be performed much more efficiently in high-speed networks. This is especially true for various "denial of service" attacks, such as SYN-flooding and ICMP attacks resulting in packet storms. Audit files created during an attack can easily grow by some megabytes within minutes, preventing the machine collecting further audit data after all the available audit data storage space has been filled.

2.3.1.2 New requirements

ATM has a number of features not available in "legacy" networks, most notably, ATM supports various "quality of service" requirements; a certain bandwidth or a fixed maximum delay during transmission can be specified individually for virtual connections in ATM networks. No currently available firewall supports resource reservation in order to keep track of these qualities of service requirements. This is currently an active research area.

2.3.1.3 New risks

ATM networks require a number of new protocols (e.g. PNNI - "Private Network-Network Interface" and ILMI - "Integrated Local Management Interface"). Even more services are necessary to support CLIP or LANE. The security implications of these protocols are not fully understood. Before firewalls can be integrated into such an environment, the risks associated with these new protocols must be identified; this requires extensive research.

2.3.1.4 Technical problems

As already described, ATM networks differ from "legacy" networks in many ways. Most firewall concepts have implied assumptions about the underlying network. Application layer firewalls (proxies) are on a high level of abstraction and are therefore more loosely coupled with the underlying network.

Packet screens, on the other hand, are usually based on the assumptions that every packet sent contains complete address information and that also the services accessed can be identified in every packet. Both assumptions are no longer valid in ATM networks. These aspects are discussed in the following section.

2.3.2 Integration of packet screens into ATM networks

Packet screens filter packets based on information in the packet headers. In "legacy" networks the address information available allows packet screens to restrict access to certain IP addresses and to TCP or UDP services.

As ATM cells contain only 5 bytes of header and 48 bytes of payload, a packet screen operating on every cell has only very limited information available. IP datagrams, usually a few hundred bytes long, must be "segmented" into multiple cells by the sender and "reassembled" at the destination.

2.3.2.1 Classical packet screens

As packet screens operate on IP datagrams, they have to reassemble IP datagrams from cells before filters can be applied. Datagrams, which are allowed to be forwarded, must be segmented into cells once again after filtering. Segmentation and reassembly is performed by hardware on the ATM interfaces and therefore does not increase the packet screens processing load. After cells are reassembled to IP datagrams, the further processing of these IP datagrams does not differ from packet screens used in "legacy" networks.

2.3.2.2 Cell screens

The reassembly and segmentation in classical packet screens increases the transmission time, as all cells must arrive before the original datagram can be recovered and filters can be applied. The delay can be reduced, if the packet screen could extract the information required for filtering from cells, thus avoiding reassembly.

To understand how this could be implemented a short description of the transmission mechanisms for IP datagrams over ATM networks is required. The CLIP protocol stack is shown in figure 1.1. First, a SNAP header is prepended to an IP datagram. The SNAP header identifies the transmitted payload as IP. SNAP header (8 bytes) and IP datagram are then encapsulated in an AAL-5 frame. The AAL-5 frame has a trailer of 8 bytes. It also contains a variable number of padding bytes to match the frame exactly into multiple 48 bytes cells. This AAL-5 frame is segmented into cells, where the last cell of the AAL-5 frame is marked. All cells are then sent on the same virtual circuit across the ATM network. ATM guarantees the ordered delivery of cells.

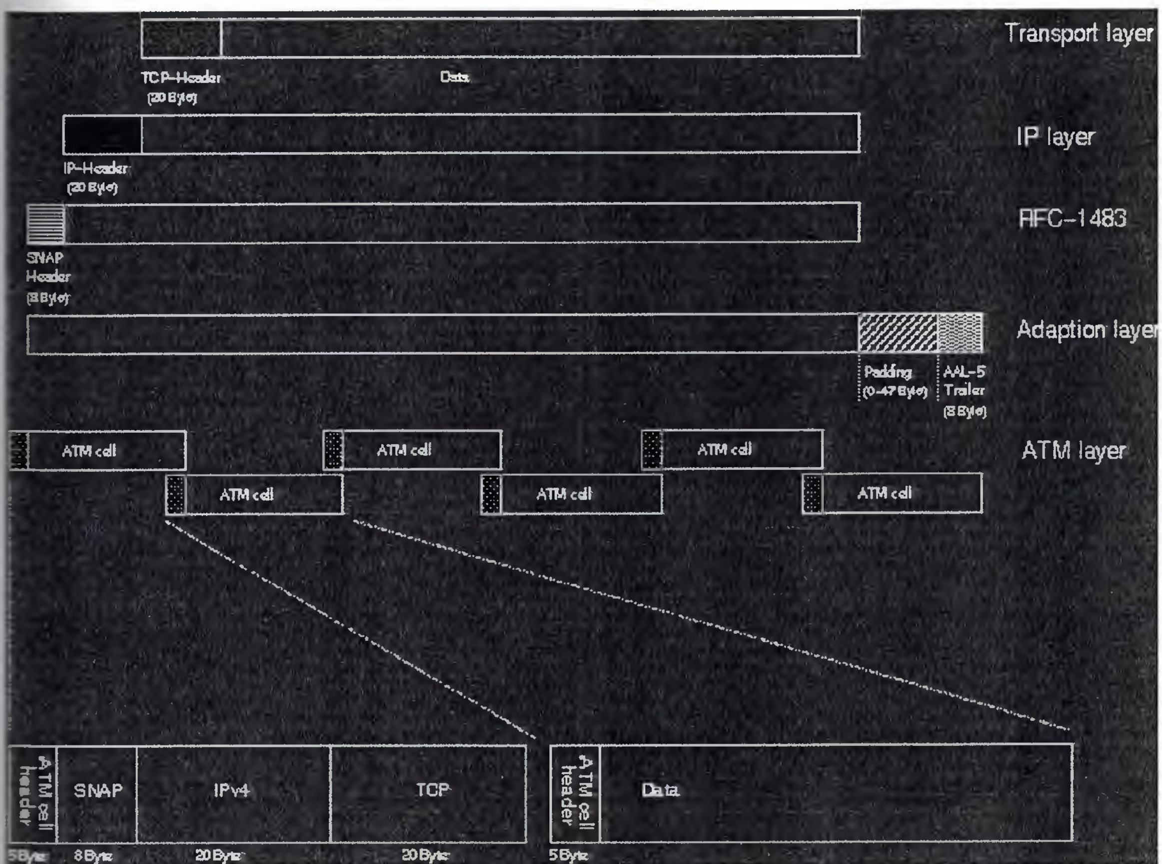


Figure 2.1 Cell Screen: CLIP protocol stack

A cell screen can identify the last cell of an AAL-5 frame. As all cells are delivered in order, the next cell will be the first cell of the next datagram. This first cell contains 8 bytes of SNAP header, 20 bytes of IP header and 20 bytes of TCP header. With the complete IP and TCP (alternatively UDP) headers in the first cell, the packet screen has all information that is required for filtering. If the forwarding of the datagram is allowed by the filtering rules, the first cell and all following cells on the same virtual circuit are forwarded until the last cell of an AAL-5 frame is found. If the datagram must be blocked, the packet screen discards all cells up to and including the last cell of the AAL-5 frame.

A cell screen can be implemented mainly in hardware and installed between an external link and an internal switch. But, as most parts of a cell screen are already required in ATM switches (forwarding of cells, recognition of the end of an AAL-5 frame and selective discard of cells), the extension to support the missing cells screen features is a natural one.

A cell screen imposes a shorter delay, as screening can occur after the first cell has been received. The copy operations performed by classical packet screens, which move whole IP datagrams are also avoided; cell screens only have to copy the first cell for screening, subsequent cells can be forwarded or dropped efficiently by the switching hardware. The screening overhead for evaluation of the filter rules is, however, the same for cell screens and classical packet screens.

2.3.2.3 Signaling Screens

Most firewall concepts rely on a combination of one or more packet screens and one or more bastion hosts. The bastion hosts perform connection authentication on an application layer level. The packet screens function is to allow the proxy servers to communicate, while preventing all other communication. If another mechanism is available, which ensures that only this legitimate communication can take place, no packet screen is needed. For example a gateway firewall does not require a packet screen, as it is the only machine that is connected to both internal and external networks.

In ATM networks routing and forwarding are separate tasks. All routing decisions are made during the setup of a virtual circuit. All data sent is forwarded along this virtual circuit. The end systems of a virtual circuit (sender and receiver) can be identified during connection setup before any data is sent. By specifying rules which define which circuits may be setup between which end systems, all traffic can be forced to be processed by a bastion host before it enters a network on the other side of the firewall.

Current ATM switches already support a simple filtering language; its structure is similar to the filter rules of packet screens in routers. Rules can be defined in an ATM switch to expressively allow or deny the establishment of virtual circuits to a list of ATM addresses. It requires only moderate effort to define rules that forbid the establishment of virtual circuits between internal and external end systems except for the bastion host. The filter rules only have to be examined during the setup of a new virtual circuit; there is no impact on the performance of the following communication. Obviously the bastion host must be powerful enough to support the high bandwidth available or the traffic must be distributed among several parallel bastion hosts.

2.4 Performance of firewalls in ATM networks

An ATM test-network was setup for performance measurements of different firewall concepts in high-speed networks. The following discussion summarizes the results of performance measurements for the two most important firewall components - packet screens and proxy servers.

2.5 Performance of packet screens

A workstation equipped with two ATM interfaces was used as a packet screen for the performance measurements. The software "IP-filter" (version 3.2) used on the packet screen allows the specification of filter rules. The tool "Netperf" repeats write calls on an already opened TCP connection for 10 seconds. The throughput is calculated by the amount of data transferred.

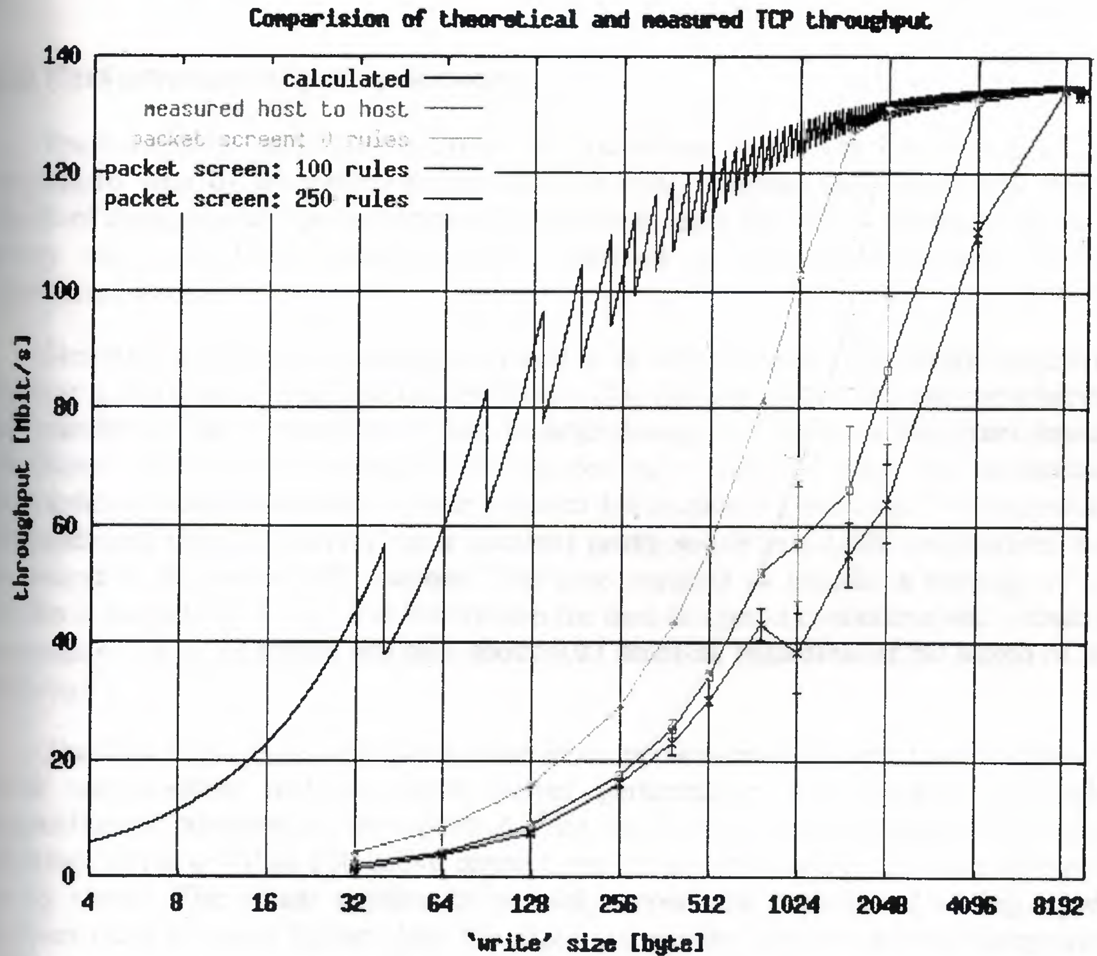


Figure 2.2 Throughput over a packet screen

The figure 2.2 shows the achieved throughputs for three selected filter configurations with 0, 100 and 250 rules. As expected the performance depends primarily on the number of filter rules configured. The calculated theoretical maximum throughput can only be reached with `write` calls longer than 2048 bytes. The reason for the sharp drop of throughput for shorter `write` calls is the limited packet throughput of the packet screen. In an OC-3c ATM network (155 Mbit/s) almost 180,000 datagrams per second are necessary in order to achieve the calculated theoretical maximum throughput with message sizes below 40 bytes. While the workstations in our environment were able to generate about 16,000 datagrams per second, the tested packet screen reaches only about 8.000 datagrams per second. Adding more filter rules will further reduce this value. As typical message sizes rarely exceed 500 bytes the expected throughput of the packet screen in a real environment will be limited to 30-40 Mbit/s. In order to reach the calculated theoretical maximum throughput of 120 Mbit/s for this message size, the packet throughput of the packet screen must be four times higher. A packet throughput of approximately 30,000 datagrams per second is necessary to reach 120 Mbit/s with a message size of 500 bytes. These results show that the actual data throughput is not a bottleneck, the packet throughput of the packet screen limits the maximum throughput instead.

2.6 Performance of proxy servers

Proxy servers control connections at application level. The processing of the transferred data by an application process obviously requires more resources than a check of datagrams at a packet screen. Nevertheless most firewall concepts are based on proxy servers as better security can be achieved by doing access control on the application level.

Despite the higher processing overhead it is also possible with proxy servers to achieve a maximum throughput of 134 Mbit/s. But this throughput can only be achieved for transfers of large quantities of data in large datagrams. For more important smaller quantities of data, for instance the transfer of a HTML page, the connection establishment time dominates the time required for transferring the data. The connection establishment time to a server via a standard proxy server in a LAN environment was measured to be about 0.03 seconds. The time required to transfer a message of 16 kbytes is magnitudes lower. For that reason the time to open a connection and transfer a message of 16 or 32 kbytes will take about 0.03 seconds, regardless of the length of the message.

The data throughput and the increase in connection establishment time designate a more user-oriented view on proxy server performance. The number of parallel connections is however just as important. First results show that dependent on the type of proxy server less than 100 active connections can be processed at the same time on a proxy server. The actual number of parallel connections experienced in high-speed networks can be much higher. Also the rising complexity of proxy servers (integration of virus scanner, encryption etc.) will require a distribution among several bastion hosts.

2.7 Concepts for Parallel Firewalls

The throughput measurements for packet screens and proxy servers have shown that the performance of these classical firewall concepts is not sufficient for high-speed networks. The packet throughput of workstation-based packet screens is too low to result in an adequate throughput for typical IP datagram sizes. The high processing overhead of application layer firewalls such as proxy servers result in low maximum throughputs, so that proxy servers are perceived as a bottleneck for communication. In the following section we start with an overview of parallel protocol processing in order to introduce parallel firewall concepts later on.

2.8 Parallel Protocol Processing

It turned out that a typical workstation is unable to provide the available throughput of a high-speed network at transport level or application level due to a bottleneck in the protocol processing in higher layer protocols such as TCP or IP. For this reason many parallel processing approaches have been suggested.

While the static methods all introduce some kind of pipelining in the protocol stack which differ in the achievable granularity of parallel processing, the dynamic methods handle either incoming packets (packet parallelism) or whole connections (connection parallelism).

While packet parallelism fits very well for parallel packet screens the connection parallelism is the better choice for concepts for parallel application level firewalls.

2.8.1 Parallel Bastion Hosts

As most proxy servers support TCP based services and TCP is a connection-oriented protocol the connection parallelism is a straightforward choice for parallel application level firewalls. The load that has to be distributed among parallel processes is the accumulated number of parallel connections a proxy server has to handle. We will now discuss different basic approaches for distributing the load. This will lead to parallel application level firewalls.

2.9 Static Distribution of Connections

The "load" denotes the number of open connections to a proxy server. If the load has to be shared we need solutions for distributing these connections.

The easiest way to distribute the load is to provide a separate proxy server for each service (e.g. HTTP, FTP...) that has to be supported. As traffic is statically mapped to dedicated proxy servers, measurements have to show which proxy servers can be mapped together on a single processor (e.g. bastion host) and which proxy servers should be mapped onto separate processors or hosts.

By distributing the proxy servers among different hosts the security can also be improved. If an intruder succeeds in attacking one bastion host, he still has no access to other proxy servers. If on the other hand all proxy servers are concentrated on a single bastion host, all these services can be used by an intruder who succeeds in attacking this bastion host to proceed attacking the guarded net. The major disadvantage of this solution is the static mapping of all connections to a certain service to a dedicated proxy server on a dedicated processor. If the current traffic differs from the expected traffic (e.g. more FTP requests than HTTP requests) the forejudged mapping may be inefficient. This may lead to situation where a single bastion host is under heavy load while other parallel bastion hosts are idle.

2.10 Dynamic Distribution of Connections

The throughput can be improved by replicating a proxy server on multiple processors, so that connections can be dynamically mapped to replicated proxy servers.

Example: Round-Robin DNS

A well-known example for dynamically distributed connections is a "Round Robin" extension to DNS. All names of replicated WWW servers which shall share the connections are registered with a CNAME for the virtual "WWW" server. After each lookup the DNS server rotates the list of CNAMEs. The next client requesting the name of the "WWW" server will receive a different answer from the DNS server and the connections to the "WWW" server will be distributed among the parallel servers.

We have a dynamic distribution of connections, but we still can not make sure that this load is balanced, as the distributing process does not get any feedback about the load of the parallel proxy servers.

Example: Distribution by "meta" Proxy

A "Meta" proxy can improve the distribution of connections to the proxy servers. All requests are sent to this "meta" proxy who chooses one of the parallel proxy servers to process the request. As the parallel proxy servers may send status information to the Meta proxy this choice can be made load dependent (e.g. the proxy server with the lowest load gets the request).

The main problem of the dynamic distribution is to find an inexpensive (fast) algorithm that distributes the incoming connections among the parallel bastion host. This distribution can either be centralized or decentralized.

2.11 Centralized vs. Decentralized Distribution

A single (Meta) proxy who distributes all incoming connections among the pool of proxy servers, which actually serve the requests, may realize a centralized distribution. The advantage of this solution is that the meta proxy may gather status and load statistics from the proxy servers that enables a fair and balanced distribution of incoming connections. The meta proxy may also be able to redirect requests if it detects an intrusion or failure of a bastion host. On the other hand the meta proxy has to handle all incoming connections. It must be fast enough so that the distribution of connections is not a bottleneck itself.

Another way to improve the throughput of centralized distribution is to make the decision in the kernel (e.g. on the network (IP) layer). The low throughput of proxy servers results from the fact that the proxy servers are application level processes which receive the request via one connection and forward it via another one. If we just want to distribute the incoming connections a kernel level process could forward the datagrams to the bastion hosts. This mechanism is a special case of "Network Address Translation" (NAT).

Example: Packet Screen as a Central Distributor

An example for a centralized distribution at the network level is a combination of packet screen and parallel bastion hosts.

Modern packet screens are able to map one IP address onto another ("Network Address Translation" (NAT)) while they are filtering datagrams. This can be used to distribute the datagrams to parallel bastion hosts depending on the load of the proxy servers (see figure 2.3).

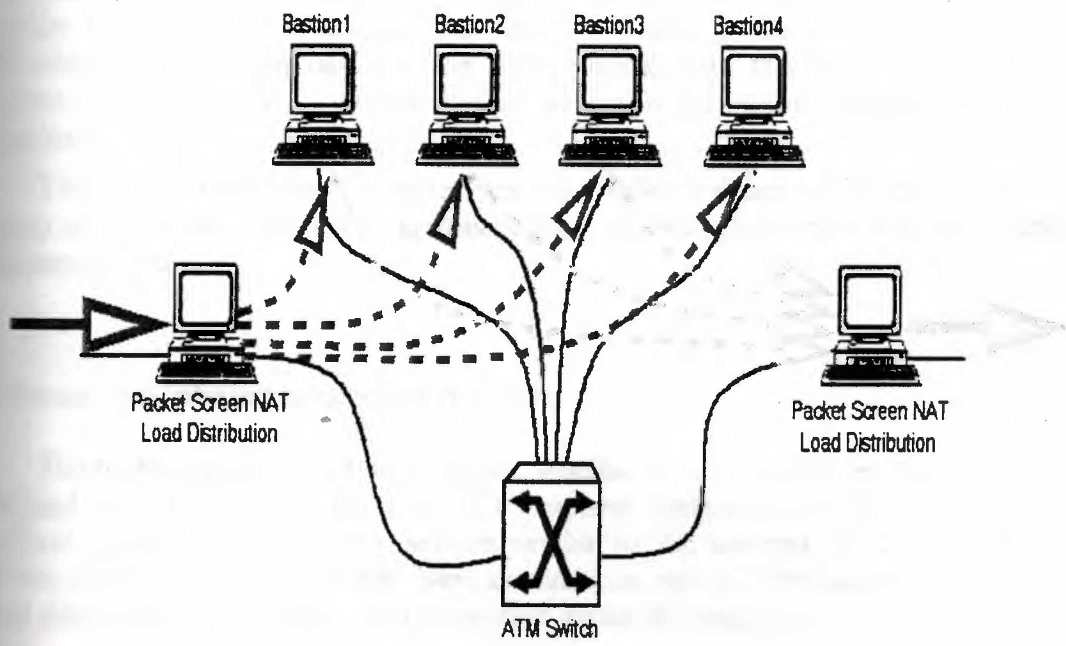


Figure 2.3 Distribution by "Network Address Translation" (NAT)

Measurements for a packet screen with enabled NAT show that the performance impact of NAT is almost equal to the impact of 10-20 filter rules. Depending on the type of traffic and the processing overhead of the proxy servers there is a risk that the packet screen used for distribution may become the bottleneck in this setup.

There are, however, some points to be considered. First all datagrams of one connection have to be forwarded to the same proxy server. Secondly there may also exist inter-connection dependencies. For example the data stream and the control stream of an FTP session should be mapped to the same proxy server. This may be implemented by forwarding all connections of a client (represented by its IP address) to the same proxy server. On the other hand mapping all incoming connections of the same client onto the same proxy server may be too restrictive in some environments. For each kind of proxy server the context information that have to be shared in order to resolve inter-connection dependencies must be specified. The propagation of context information enables a higher degree of parallelism among the proxy servers.

2.12 Distribution over "native" ATM

The previous setup (figure 2.3) may be improved by using native ATM connections between the packet screens and the parallel bastion hosts. All parallel bastion hosts respond to the same IP address. The packet screen does not need to use time consuming

NAT transformation on every IP datagram, it acts as a router and simply forwards an IP datagram over one of the native ATM connections to a bastion host instead. The route for an IP datagram cannot be a function of the destination IP address because the IP addresses of the bastion hosts are all the same. The load distribution algorithm calculates the path of a datagram. The use of the same IP address for multiple systems usually has disastrous effects on the network. These problems do not occur in the described setup as the bastions can only communicate through the packet screen. Further studies have to show quantitative aspects of performance improvements of this approach.

The parallel proxy servers themselves may realize a decentralized distribution. Each proxy server inspects all incoming connections and decides whether it is responsible for this connection.

Example: Distributed connection response

The decentralized algorithm to decide whether or not to serve a connection may be realized as a function of the first TCP segment (indicated by the SYN flag). This method requires that all proxy servers similar to the concept of the parallel packet screen can receive every packet. New connections can be distributed either randomly, load dependent, or dependent on information inside the datagram.

Example: Redirecting HTTP requests

The HTTP protocol enables servers to redirect requests by sending the clients an alternative URL. Parallel servers to balance their load can use this feature. Whenever a server under heavy load receives a request it may decide to redirect the request to a replicated server. An obvious problem with this approach is a spoofing of redirect messages which increases the risk of "man in middle" attacks.

Example: Transparent redirecting connections

Another disadvantage of the solution for HTTP is the need for an explicit cooperation between client and server. This cooperation may be hidden by providing a transparent `connect()` system call in a shared library that replaces the standard library function. The new `connect()` system call tries to open a connection. A server may accept the connection or supply the address of an alternative server. As this redirection is hidden by the `connect()` call, there is no need to change the client software. Another advantage is that it is a generic solution. It works for all TCP based applications that use the library replacement. Note that this solution can be built on the simple protocol used by SOCKS.

As a redirection of a connection increases the connection setup time, the tradeoff between increased throughput and connection setup time has to be taken into account. A redirection is usually only worth the increased setup time, if a large amount of data has to be transferred. For small quantities of data it is more efficient to process the connection without redirection and notify the client to use an alternate proxy server for subsequent connections. A prototype that uses the described transparent redirection for a distributed load balancing is currently developed.

2.13 Parallel packet screening

Measurements for the performance of packet screens (figure 3.1) have shown that a typical workstation is able to perform screening in a 155 Mbit/s ATM network for large packet sizes only. Unfortunately the average packet size in the Internet is much smaller. The packet throughput of the investigated packet screen has to be increased about four times to reach acceptable throughputs with smaller packet sizes.

Parallel packet screens based on the paradigm of "packet parallelism" provide a scalable solution for high-speed networks. "Packet parallelism" fits very well for parallel packet screens that do not care about connection contexts. As there is no need to update any connection contexts, any packet of any connection may be screened in parallel. Of course this is the case for connectionless (UDP) traffic anyway.

A distributed decision about which packet screen is responsible for the filtering has to be made. The additional costs for this decision must be very low compared to the total costs of filtering. Every packet screen in the parallel setup inspects every packet and immediately discards packets that another screen is responsible for. The decision which packet screen is responsible for a packet may be a function of information elements in the packets. For example the hash value of the IP checksum may be used as an index to the packet screen that has to examine the packet. All other packet screens may discard the packet. Because the IP checksum is likely to differ for successive packets this algorithm should assure an almost balanced distribution.

Example: Broadcast LAN implementation

The implementation is very simple for broadcast LANs such as Fast-Ethernet, Gigabit-Ethernet, or FDDI. In the case of Fast-Ethernet the parallel packet screens can be placed between two Hubs. The Hubs ensure that all packets are distributed to all parallel packet screens.

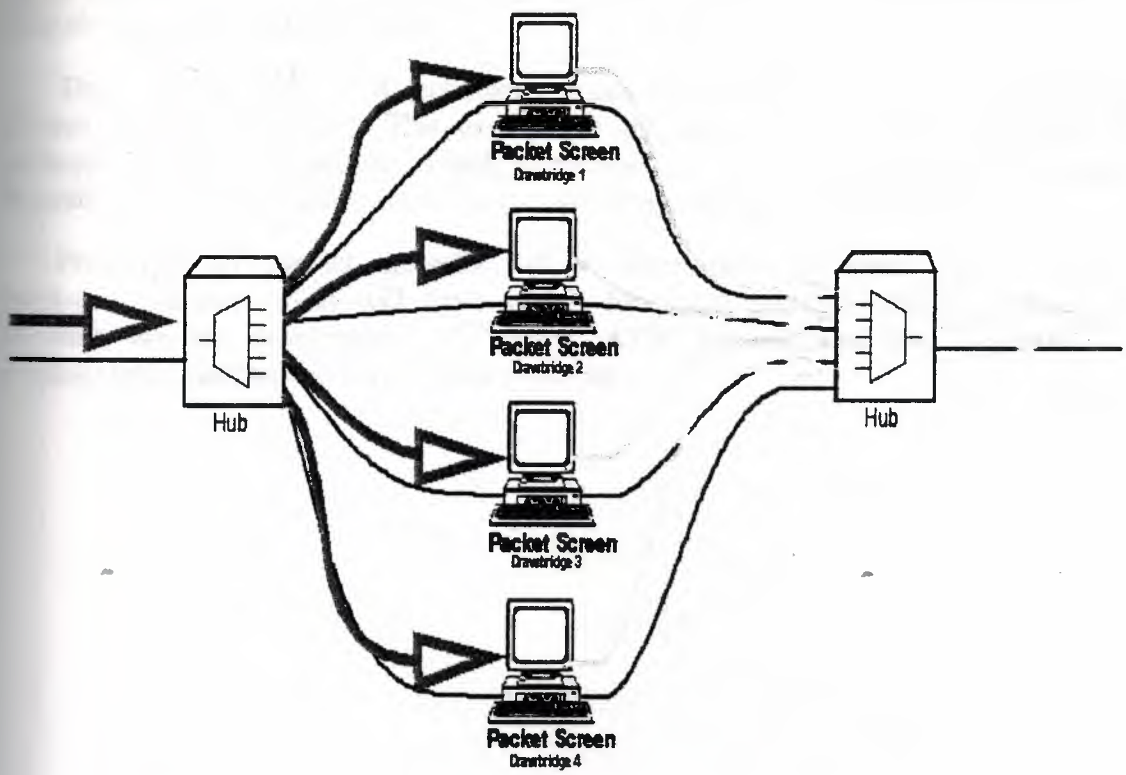


Figure 2.4 Parallel Packet Screen (Fast-Ethernet/Gigabit-Ethernet)

Example: ATM Implementation

As ATM networks are connection-oriented the required broadcast functionality must be emulated. An ATM switch can be used for a very efficient implementation. It is possible to configure a point-to-multipoint connection so that the switch copies all incoming cells to multiple outgoing virtual channels. This mechanism can be used to assure that all packet screens in an ATM network receive the incoming packets.

A problem arises by the increased possibility of failure due to the parallel packet screens. By monitoring or status propagation one should make sure that a failure is detectable so that the distributed filtering algorithm may be adjusted to the new number of parallel packet screens. On the other hand with failure detection there is no longer a single point of failure. If one of the packet screens breaks down the others take over.

2.14 Conclusions

This chapter discussed the impact of ATM on firewalls. The functional differences between ATM and “legacy” network means that classical firewall concepts (for "legacy" networks) cannot be applied to ATM technology without modification. These differences have a greater impact on packet screens than application level firewalls, as they are more dependent on the underlying network than application level firewalls.

Three different approaches for packet screens were introduced: classical packet screens, cells screens and signaling screens.

The highly scalable ATM technology raises throughput problems for both packet screens and proxy servers. This emphasizes the necessity for the parallel firewall concepts we have introduced, which overcome the throughput bottleneck. Parallel firewalls can provide scalable solutions for upcoming high-speed networks.

Prototypes of parallel firewalls will be developed in further research. Other important aspects, which will have strong functional and performance impact on firewalls, are the development of "native" ATM firewalls and the integration of cryptographic mechanisms into firewall concepts.

CHAPTER 3
WIRELESS ATM

3.1 Introduction

As the deployment of wireless LANs grows, there is a need for higher data rates. As a result, spectrum has been allocate for high performance LANs (HIPERLAN) and SUPERNET activities at 5GHz, supporting connectivity of 20 to 25 Mbit/s. Moving to even higher frequencies (40 and 60 GHz) with connectivity of 100Mbit/s is the subject of current research, although these higher frequencies are more suited to fixed links applications.

Because of the wide range of services supported by ATM networks, ATM technology is expected to become the dominant networking technology for both public infrastructure networks and LANs. ATM infrastructure can support all types of services, from time-sensitive voice communications and multimedia conferencing to bursty transaction processing and LAN traffic. Extending the ATM infrastructure with wireless access meets the needs of users and customers who want a unified end-to-end networking infrastructure with high performance and consistent service. Wireless ATM adds the advantages of mobility to the already great service advantages of ATM networks.

3.2 Background of Wireless Technologies

Wireless technologies and systems are fairly new to telecommunications and are still emerging on the scene. Currently, wireless LAN technologies are comprised of infrared, UHF radio, spread spectrum, and microwave radio. These technologies can range from frequencies in the MHz (US), GHz (Europe), to infrared frequencies. The personal communication network (PCN) can either use code-division multiple access (CDMA), or time-division multiple access (TDMA). There is a considerable controversy among experts in the field regarding the relative merits of spread spectrum (CDMA) and narrow-band (TDMA) for private communication network (PCN). The preferred technique may actually vary with the specific PCN application scenario to be addressed. Below is a brief description of CDMA and TDMA.

- Spread Spectrum (CDMA): The term spread spectrum defines a class of digital radio systems in which the occupied bandwidth is considerably greater than the information rate. The term code-division multiple access (CDMA) is often used in reference to spread spectrum systems and refers to the possibility of transmitting several such signals in the same portion of spectrum by using pseudo random codes for each one. This can be achieved by either frequency hopping (A series of pulses of carrier at different frequencies, in a predetermined pattern), or direct sequence (A pseudo random modulating binary waveform whose symbol rate is a large multiple of the bit rate of the original bit stream) spread spectrum.

- Time Division Multiple Access (TDMA): TDMA divides the radio carriers into an endlessly repeated sequence of small time slots (channels). Each conversation occupies just one of these time slots. So instead of just one conversation, each radio carrier carry's a number of conversations at once. With the development of digital systems TDMA is being more widely used.

3.3 Reasons for Wireless ATM

Since the beginning the concept of ATM is for end-to-end communications (i.e. in a WAN environment). The communication protocol will be the same (i.e. ATM), and companies will no longer have to buy extra equipment (like routers or gateways) to interconnect their networks. Also, ATM is considered to reduce the complexity of the network and improve the flexibility while providing end-to-end consideration of traffic performance. That is why researchers have been pushing for an ATM cell-relay paradigm to be adopted as the basis for next generation wireless transport architectures.

There are several factors that tend to favor the use of ATM cell transport for a personal communication network. These are:

- Flexible bandwidth allocation and service type selection for a range of applications
- Efficient multiplexing of traffic from bursty data/multimedia sources
- End-to-end provisioning of broadband services over wireless and wired networks
- Suitability of available ATM switching equipment for inter-cell switching
- Improved service reliability with packet switching techniques
- Ease of interfacing with wired B-ISDN systems that will form the telecommunications backbone

In general, internetworking may always be seen, as a solution to achieve wireless access to any popular backbone network but the consequence, in this case, is a loss of the ATM quality of service characteristics and original bearer connections. The more internetworking there is in a network, the less harmonized the services provided will be.

Therefore, it is important to be able to offer appropriate wireless extension to the ATM network infrastructure.

One of the fundamental ideas of ATM is to provide bandwidth on demand. Bandwidth has traditionally been an expensive and scarce resource. This has affected the application development and even the user expectations. So far, application development has been constrained because data transmission pipes cannot support various quality of service parameters, and the maximum data transmission bandwidth that the applications have to interface with is relatively small. Finally, ATM has removed these constraints. Bandwidth has become truly cheap and there is good support for various traffic classes. A new way of thinking may evolve in application development.

The progress towards ATM transport in fixed networks has already started and the market push is strong. It can be expected that new applications will evolve that fully exploit all the capabilities of the ATM transport technology. The users will get used to this new service level and require that the same applications be able to run over wireless links. To make this possible the wireless access interface has to be developed to support ATM quality of service parameters.

The benefits of a wireless ATM access technology should be observed by a user as improved service and improved accessibility. By preserving the essential characteristics of ATM transmission, wireless ATM offers the promise of improved performance and quality of service, not attainable by other wireless communications systems like cellular systems, cordless networks or wireless LANs. In addition, wireless ATM access provides location independence that removes a major limiting factor in the use of computers and powerful telecom equipment over wired networks.

The following diagram shows a typical ATM Network:

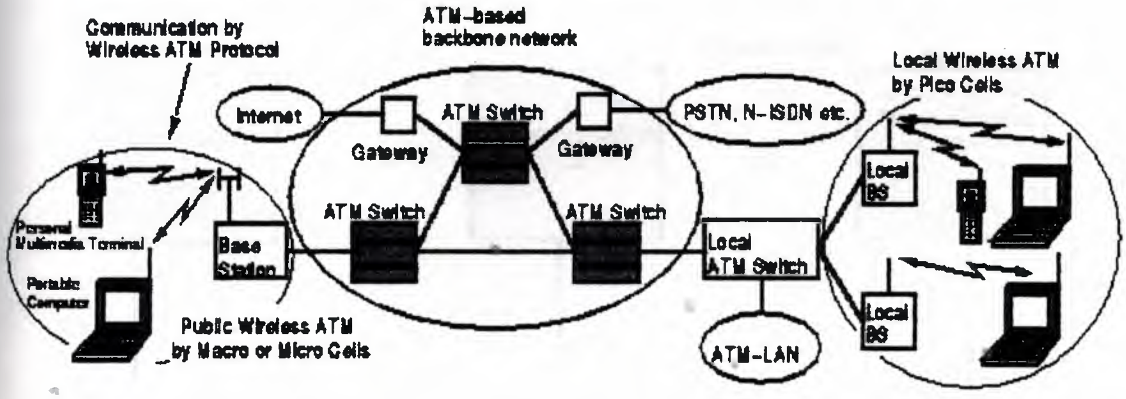


Figure 3.2 A Typical ATM Network

3.4 Wireless ATM Architecture

The architecture proposed for wireless ATM is composed of a large number of small transmission cells called Pico cells. A base station serves each Pico cell. All the base stations in the network are connected via the wired ATM network. The use of

ATM switching for intercell traffic also avoids the crucial problem of developing a new backbone network with sufficient throughput to support intercommunication among large number of small cells. To avoid hard boundaries between Pico-cells, the base stations can operate on the same frequency.

Reducing the size of the Pico-cells has major advantages in mitigating some of the major problems associated with in-building wireless LANs. The main difficulty encountered is the delay due to multi-path effects and the lack of a line-of-sight path resulting in high attenuation. Pico-cells can also have some drawbacks as compared to larger cells. There are a small number of mobiles, on average, within range of any base-station, so base-station cost and connectivity is critical. As cell size is reduced, hand-over rate also increases. By using the same frequency, no hand-over will be required at the physical layer. The small cell sizes also give us the flexibility of reusing the same frequency, thus avoiding the problem of running out of bandwidth.

The mobile units in the cell communicate with only the base-station serving that particular cell, and not with other mobile units. The basic role of the base station is interconnection between the LAN or WAN and the wireless subnets, and also to transfer packets and converting them to the wired ATM network from the mobile units.

In traditional mobile networks, transmission-cells are "colored" using frequency division multiplexing or code division multiplexing to prevent interference between cells. Coloring is a wasteful of bandwidth because in order for it to be successful there must be areas between re-use which are idle. These inactive areas could potentially be used for transmission.

The following diagram shows a typical ATM to Base Station Connection:

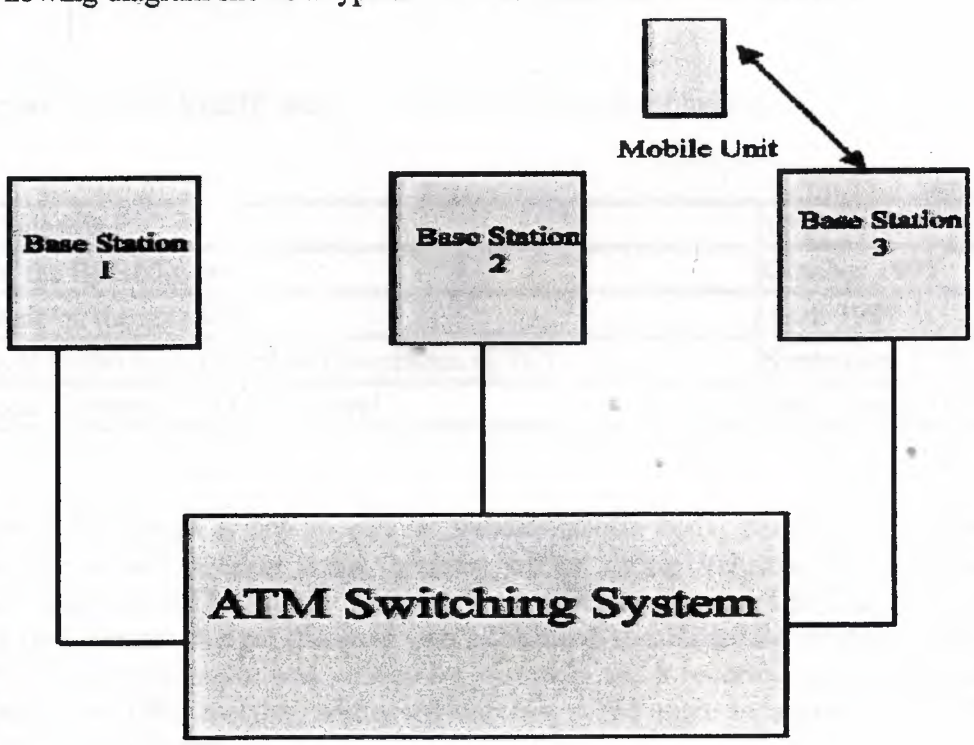


Figure 3.3 ATM to Base Station Connection

3.5 Standards

Wireless ATM research has been active for some time now. There are many papers written on wireless ATM and there are even announced wireless ATM prototypes such as RATM (Radio ATM) by Olivetti research laboratory. Yet, the most important type of activity has been missing from wireless ATM scene. For companies with business interests, the main objective is often to implement only equipment/systems conforming to standards. Therefore, October 1995 has been a significant milestone for wireless ATM. In October, the wireless ATM subject has been brought to two different standardization forums, namely the ETSI STC RES10 and ATM forum.

Currently, there are three standard bodies that have defined the physical layer in support of ATM: ANSI, CCITT/ITU-T, and the ATM Forum. None of these forums have considered the wireless ATM interface. ETSI RES10 sub-technical committee is the first standardization body to start working on wireless multimedia, ATM compatibility, and standardization. RES10 committee has already been engaged with the HIPERLAN (High Performance Radio Local Area Network) standardization and the wireless ATM group has just started to work on this new subject. The initial work will be concentrating on possible usage scenarios and specific requirements. Also, the search for available spectrum in the 5.2 GHz range for wireless ATM system is crucial and therefore one of the first tasks of RES10. The following table shows a tentative work schedule.

Workplan for ETSI RES10 work on wireless multimedia standard:

Work milestone	12
Start of the RES10 work	October 1995
First draft on the standard	April 1997
Approval by the Sub-Technical Committee (STC)	September 1997
Technical Committee (TC) approval	December 1997

The ATM forum is not an official standardization body, but it plays a significant role in the standardization arena because of its strong industrial participation and support. Wireless ATM activity has not been officially approved in The ATM forum, but the first contribution on this item was introduced in a Honolulu meeting, on October of 1995. This contribution was written by Motorola and it received significant attention. The December 1995 meeting addressed wireless ATM again to evaluate the relevance and role that it will play.

One solution would be to divide the standardization of wireless ATM between the ATM forum and RES10. It would probably be wise to let the ATM forum concentrate on the fixed network side and RES10 on the wireless interface. The main focus of The

ATM forum should be on the fact that the ATM physical layer is not necessarily always a reliable medium and that terminals may be mobile. Both of these facts are due to the fact that ATM/B-ISDN connections may be stretched over the wireless links in the future and should be independent of the specific wireless interface.

3.5.1 Wireless ATM Network Demonstrator

Objectives

- To specify a wireless, customer premises, access system for ATM networks that maintains the service characteristics and benefits of the ATM networks to the mobile user
- To promote the standardization of wireless ATM access
- To demonstrate and carry out user trials and test the feasibility of a radio based ATM access system

Summary

The Magic WAND project (Wireless ATM Network Demonstrator) covers the whole range of functionality from basic (wireless) data transmission to shared multi-media applications, in Europe. The primary goal of the project is to demonstrate that wireless access to ATM, capable of providing real multi-media services to mobile users, is technically feasible. The project partners have chosen to use the 5 GHz frequency band for the demonstrator and to perform studies on higher bit rate operation >50 Mb/s in the 17 GHz frequency band.

The aim of the user trials is to verify a wireless access system for ATM networks that maintains the service characteristics and benefits of ATM networks in the 5 GHz range allocated to wireless high speed data transmission. The feasibility of a radio based ATM access system will be demonstrated by the user trials with selected end-user groups in hospital (medical consultation) and office environments.

The medical consultation shows an advanced scenario, fully exploiting the wireless ATM service capabilities in the hospital environment. The JVTOS (Joint Video Telecommunication Operating System) will be used with an X-ray viewing application, using both native audio and video services over ATM. In this scenario, doctors will be equipped with a mobile terminal while visiting patients. With the help of a wireless ATM connection, doctors are able to retrieve patient information from the network, consult expert doctors, and share documents.

The setup can be seen in the following figure.

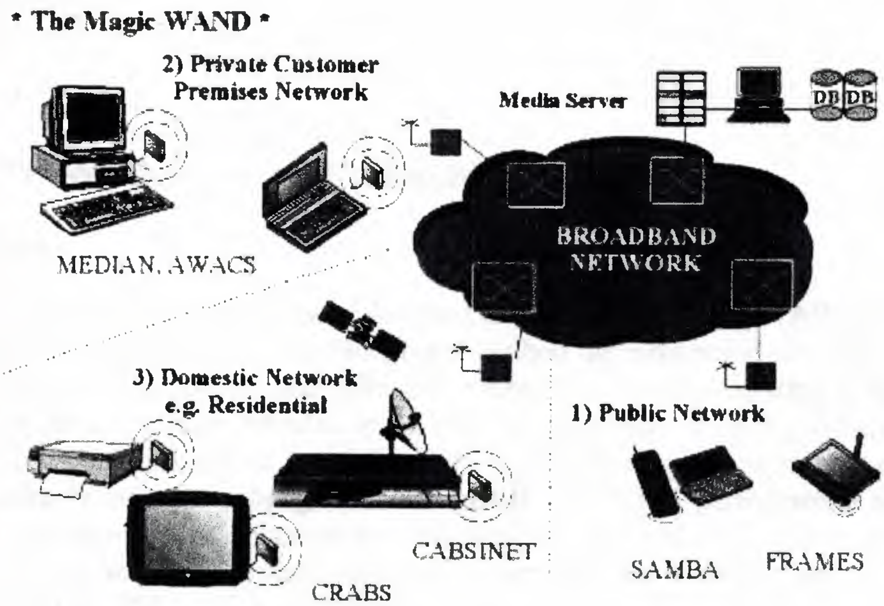


Figure 3.4 Setup of a Wireless ATM Network

Wireless ATM intends to extend all the benefits of the ATM and therefore also the ATM signaling and virtual channels/paths into the mobile terminal raising important issues that have to be solved both in the wireless access interface and in the supporting customer premises ATM network. In the air interface the wireless ATM transmission will be subject to the problems associated with the radio medium and therefore special radio design measures will be required in order to offer users an adequate level of service. These measures constitute some of the major technical challenges of this project.

The main result of the project will be a Wireless ATM Access Network Demonstration system that will serve as a proof of concept for the developed technology and help the wireless ATM standardization work. The current achievements of the project include the complete functional system specification on the demonstrator that has been specified with the SDL (Specification and Description Language) and verified with the simulation model. In addition, the project has defined the exact demo platform setup and therefore enabled the basis for the implementation work that has been started on all parts of the system.

Besides demonstrator work, the project has been active in its liaison and standardization activities. The stochastical radio channel model for channel simulations was developed and verified by measurements on 5 and 17 GHz frequency bands. The model has been given as an input e.g. for SIG1 work. Furthermore, the project has been active in the standardization forum by contributing and harmonizing the work between ATM forum and ETSI RES10.

The Magic WAND project will continue the work on gaining knowledge on the wireless ATM radio design and its medium access control functions as well as wireless

ATM specific Control and Signaling functions. These results will be further contributed to ETSI and ATM forum in order to influence all of the relevant standards for wireless ATM systems.

3.5.2 ATM Wireless Access Communication System

Objectives

The goals of the ATM Wireless Access Communication system (AWACS) project are the development of a system concept and testbed demonstration of public access to B-ISDN services. The system will offer low mobility terminals operating in the 19GHz band with a support of user bit rates up to 34Mbit/s with radio transmission ranges of up to 100m. The demonstrator of ATM Wireless Access (AWA) pre-prototype equipment will immediately provide propagation data, BER and ATM performance at 19GHz. Based on this information, enhancement techniques for AWACS will be investigated to support cellular, as well as spectrum and power efficient radio access technologies associated with HIPERLAN type 4 specifications.

Summary

AWACS technical approach will be centered around a testbed and associated trial campaign program. Trials will be conducted using the existing ATM Wireless access platform made available to the project by one of its partners. Associated program of work will then be directed on enhancing this current state-of-art system towards the final target features of the emerging ATM wireless specifications, in particular HIPERLAN type 4 as currently being defined by ETSI-RES10. These enhancements to the existing demonstrator will be considered in the following areas:

- application of source/channel coding and intelligent antennas
- optimization of link layer protocols to match ATM bearer types
- feasibility of 40GHz RF technology for ATM wireless LAN applications
- mobility management techniques together with the impact on the radio bearer appropriate for high bit rate communications

The AWACS field trial will cover the concept of 'virtual office' trials. This will include three potential cases, depending on the technical capabilities of the demonstrator:

- wireless multimedia communication link between an engineer at the production site and an expert at this office

- video communication in meetings between physically separated sites
- visual, wireless network access to virtual office facilities at one of the partner's location

The objectives of these trials can be summarized as follows:

- improvement of communication between physically separated offices by telepresence technologies
- reduce the need of traveling between the geographically separated offices
- improve the response time of expert advise in problem solving by visual communications
- free staff from fixed office hours

Key Issues

The key issues to be considered include:

- the performance evaluation of a 19GHz ATM compatible modem
- identification of the strengths and weaknesses of the existing ATM wireless experimental demonstrator
- investigation of possible enhancement to the ATM compatible modem
- AWACS field trials with the concept of 'virtual office,' which aims to improve the communication between physically separated offices by telepresence technologies

3.5.2.1 Expected Results

The AWACS demonstrator based on ATM in packet transmission schemes will support limited, slow speed mobility as it is in line with expected use of high data services. Therefore, the project generally covers the following directions that are open to developers of mobile communication systems for the future:

- construction of a wireless system providing seamless service in connections to hardwired systems (quality oriented system)
- services making the most of the excellent mobility and portability of mobile communication systems (mobility oriented system)

The AWACS trials will indicate the capacity of the available system in a real user environment. The trials results will contribute to the development of common specifications and standards such as ETSI-RES10 (for HIPERLAN type 4 specifications), ITU, TTC and ARIB in Japan.

3.5.3 International joint ventures

Wireless ATM is just barely getting started and there is a worldwide effort to unify and standardize its operation. The Public Communication Networks Group of Siemens AG, Newbridge Networks, and Broadband Networks Inc. announced an extensive joint research and development program to address the digital wireless broadband networks market. The three companies will focus on integrating BNI's broadband wireless technology with the Siemens /Newbridge Alliance's MainStreetXpress(TM) family of ATM switching products to develop wireless network base stations that are fully compatible with wireline services.

BNI has already deployed terrestrial wireless networks that provide wireless cable in a digitally compressed MPEG2 (Motion Pictures Experts Group) format, delivering laser disk quality transmissions with the capacity for hundreds of channels. The Siemens / Newbridge Alliance offers carriers the most comprehensive suite of ATM products and the largest ATM coreinfrastructure switch, scaleable up to 1 Terabit and beyond. The introduction of ATM into the broadband wireless environment will enable network operators to cost effectively deploy high capacity access services such as high speed data, broadcast (cable) distribution, and Internet access in the 28 GHz range. By incorporating both MPEG2 and ATM into the broadband wireless environment, the network solution provided by BNI and the Siemens/Newbridge Alliance ensures high speed, high quality, and high capacity video, voice, and data transmissions. It also represents an effective bandwidth allocation that ensures sufficient capacity for additional innovative residential and commercial services as they evolve.

3.6 Conclusion

While wireless communication is experiencing fast evolution, the fixed network has been going towards B-ISDN with ATM concept. ATM will offer data rates that are considerably higher than current fixed network services. Interworking with ATM will set extremely hard requirements on the wireless air interface, but hopefully continued development in technology will enable the industry to manufacture smaller and less power consuming terminals with increased performance and functionality.The prediction of future is always uncertain but it can be assumed that frequencies under 2 GHz remain mainly for mobile communications where only low bit rate services are offered (both data and speech). In this case, connections requiring close to 2 Mbit/s or more will need to be moved on, to the higher frequencies. The possible choices at the moment seem to be around 5.2 GHz and 17.1 GHz. The successful introduction of wireless ATM is strongly related to the success of ATM/B-ISDN in wired networks.

CHAPTER 4
ATM VS. ETHERNET

4.1 Abstract

ATM has been seen as the ultimate networking technology that will allow true broadband networking for the future. At the time ATM was introduced, it was generally estimated that it is just a matter of time when ATM replaces other network technologies. Today the estimates are no longer wholly in favor of ATM. The most popular LAN technology, Ethernet, is still a hot technology, despite its age. Switching, Fast Ethernet and Gigabit Ethernet have brought more bandwidth to the technology, quality of service is under development and the road goes towards company backbones. There is an ongoing battle between ATM and Ethernet and this paper studies the two technologies, their advantages and disadvantages compared to each other.

4.2 Introduction

ATM is a telecommunication technology with great promises. In 1988, ITU-T proposed ATM as the telecommunication standard for the broadband ISDN (BISDN). Many have seen ATM as the ultimate integrated-services network: unlimited bandwidth on demand; computer data, voice and video over one cost-effective infrastructure; seamless interconnect between the local area and the wide area. A technology with such promises does sound great.

Ethernet is a network technology that was designed to transfer computer data in local area networks. It didn't aim to provide a solution for all telecommunication needs. However, Ethernet has become a very successful and dominant technology. With the latest enhancements, it has turned out to be a feasible technology even beyond the LAN. Ethernet has not been replaced by ATM, which did seem inevitable at the time ATM was introduced.

4.3 Ethernet overview

Ethernet is the most popular LAN technology today. Approximately 80% of all LAN installations deploy Ethernet. This represents over 120 million hosts [10]. Ethernet is well understood, easy to install, cheap and the technology with the best support from manufacturers. The basic transfer rate of Ethernet is 10Mbps.

Traditional Ethernet is a shared access technology, based on a broadcast medium. All of the stations attached to the Ethernet share a single communications medium, a coaxial or a twisted pair cable. The broadcast nature of Ethernet is very different to the peer-to-peer networking of ATM.

Ethernet comes in two topologies: bus and star. The bus, utilizing coaxial cable, was the original topology that is rarely used anymore, due to difficulties of adding or moving users and troubleshooting. Today, by far the most common topology is a star, which uses twisted pair cables, shielded or unshielded. In a star topology all stations are wired to a central wiring concentrator who has a port for each station. The concentrator can be a hub or a switch, which is more common nowadays.

A hub is a "dump" physical layer device that broadcasts signals from the source port to all other ports. A switch is a smarter datalink layer device that forwards frames from the source port to the destination port only. This decreases the number of collisions in the whole network. In short, switching is a technique that divides a LAN to several smaller network segments, or collision domains, providing full bandwidth to each segment and diminishing the overall network congestion. Switching is a very popular and easy way to add capacity to Ethernet. To get a switched Ethernet, only the hub needs to be replaced into a switch. The equipment and software in the Ethernet hosts remain unaffected. Most current Ethernet networks are based on switching.

Stations access the shared medium of Ethernet using an access scheme called Carrier Sense Multiple Access with Collision Detection (CSMA/CD). It's a democratic scheme giving all stations an equal ability to transmit data. Before transmitting data, a station listens to the medium and starts transmitting only if the medium is free. A collision can happen if two stations start transmitting at the same time. When a collision is detected, all transmissions are damaged and stations stop transmitting. A station restarts transmit after a partially random period of time, determined by a backoff algorithm, and again only if the medium is free. CSMA/CD is a simple and viable access technology.

Ethernet uses variable-length packets called frames to carry data. The size of the frame is from 72 to 1526 bytes, of which 26 bytes is dedicated to the header [4].

Ethernet has been extended twice over the years to provide more bandwidth and to better compete with other broadband technologies. Fast Ethernet transports data at 100 Mbps and Gigabit Ethernet at a staggering 1000 Mbps. These extensions leverage the familiar Ethernet technology while retaining the CSMA/CD scheme of the original 10 Mbps Ethernet.

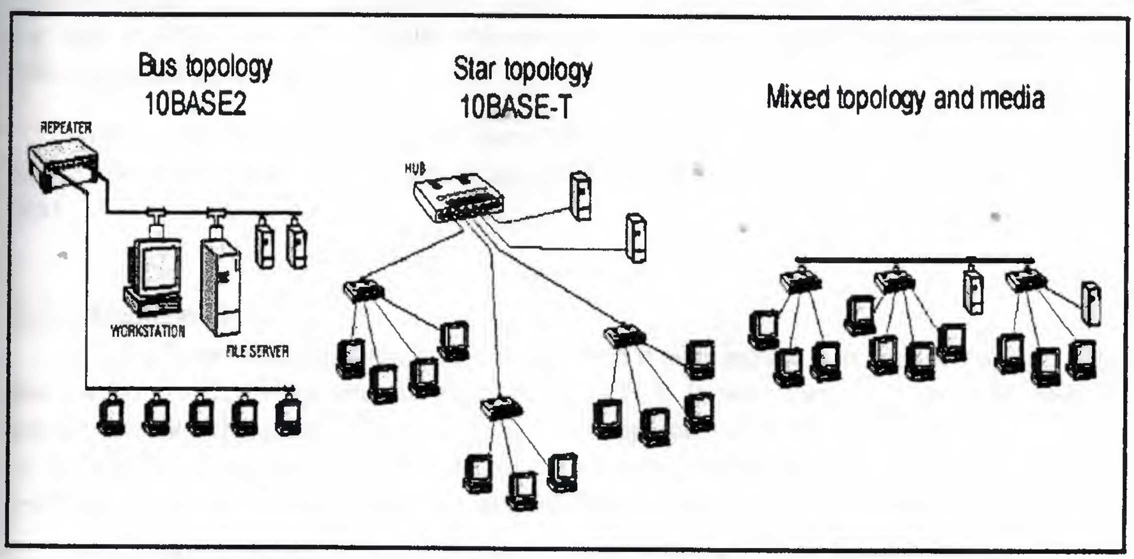


Figure 4.2 Ethernet network topologies

4.4 Technology comparison

4.4.1 General

ATM is a complex technology. There are tons of standards covering various aspects of ATM [2]. One can imagine that the complexity is pretty much due to the nature of ATM of trying to be one-solution-fits-all, technology from LAN to WAN, for all data types. The connection-orientation also largely contributes to the overall complexity, because it requires the existence of specific signaling and routing protocols. The complexity is driven by the powerful capabilities in ATM, but not all need them.

While there are many standards, some parts of the technology had to wait standardization for a long time, and this slowed down the adoption of the technology. The PNNI-interface was not standardized until 1996, and being a major interface between public ATM networks, the interoperability between public networks was not possible. Interoperability is still an issue.

In contrast, the beauty of Ethernet is its simplicity. Only a few standards cover the whole technology. The technology is easy to understand and deploy. This is the primary reason for the popularity and wide adoption of Ethernet.

A great benefit for ATM is that it is independent of the underlying transport mechanism. ATM does not define Media Access Control (MAC) mechanism (lower part of datalink layer, the other being LLC) or the physical layer, whereas Ethernet does define these. As a consequence, ATM can run on top of different transport mechanisms and can adapt to new transport technologies and greater speeds. Without physical independency, the final goal of ATM running everywhere would simply be impossible.

The original Ethernet technology has been available since 1976 [14]. ATM technology has been utilized since 1994, USA and Finland being the first countries [12]. By looking at these years, we can say that Ethernet technology is more mature than ATM. However, the comparison nowadays is done between Gigabit Ethernet and ATM, and Gigabit Ethernet is surely less mature of these two. Gigabit Ethernet Alliance was formed in 1996 to develop and standardize the technology. On the other hand, the standardization has been speedier for Ethernet, and people speaking for Gigabit Ethernet say that building on the original well-proven principles, Gigabit Ethernet should pose little technical problems.

Only a few applications have been developed for ATM. In addition of lacking native applications taking true advantage of ATM, there are few experts on ATM in the field.

4.4.2 Bandwidth

ATM offered massive bandwidth at the time it was introduced. Speeds of 155 Mbps and awesome 622 Mbps attracted people to ATM. It was seen that once the need for greater bandwidths arises, Ethernet LANs get replaced with ATM networks. The speed of 25 Mbps was proposed as the bandwidth to the desktop, but in practice it was not worth going from 10 Mbps Ethernet to 25 Mbps ATM, mostly for economical reasons.

After two upgrades, Ethernet fights back. ATM can no longer compete in pure speed. The routing and switching technology has improved and ATM alone can't take advantage of simple and fast hardware switching. The new gigabit speeds put burden on the back-end servers, and the server processing speed is becoming the bottleneck rather than the network.

An important point to make is that the actual bandwidth for the payload is always smaller than the full transmission speed of the medium. This is because the protocols and their headers eat up some of the total bandwidth, and there usually exists a couple layers of protocols.

4.4.3 Scalability

One important benefit of ATM is that ATM can be used as both LAN and WAN technologies. The original idea of ATM was the concept of spanning over LAN/MAN/WAN, and utilizing the same protocol over the entire network, eliminating the requirement for routers and gateways. This vision has not materialized. Instead, of these, the WAN has proven to be ATM's strong suit. "Despite the talk about Gigabit Ethernet displacing it, ATM continues to be a great fit in the WAN" [13].

The original Ethernet was purely a LAN technology, but Fast Ethernet took Ethernet to company backbones as well, and Gigabit Ethernet is striving itself to even bigger backbones, over campus areas. There has been criticism over Ethernet of being a technology that cannot scale, but its underlying transmission scheme continues to be one of the principal means of transporting data for contemporary campus applications.

Although Ethernet has not been considered to span over a WAN, research is being done to make it a reality. Ethernet is stepping on ATM's shoes in WAN too.

4.4.4 Overhead

One could argue that ATM is not optimized for any application. The technology holds a compromise. At the beginning of the standardization, the size of the cell generated heated discussion between USA and Europe. Europe wanted 32 bytes of payload and USA wanted 64 bytes. A compromise was agreed on 48 bytes for the payload [18]. The overhead from a 5 byte header of a total cell size of 53 bytes is 9.4%.

In Ethernet, the overhead is minimal as the frame size can be 1526 bytes, of which the header is 26 bytes. However, one should remember that this doesn't hold the whole truth: Ethernet frames do carry other protocol packets in them, and these packets have headers too. And Ethernet alone can't move data over WAN so a direct compare is not fully justified.

Another aspect creating overhead is the connection-orientation of ATM. A virtual circuit must be setup end-to-end before communication can take place. For exchanges of small amount of data, the connection setup can take considerably more time than the actual data exchange. A good example is the connectionless DNS-protocol of TCP/IP-networks. DNS-messages mainly consist of a single UDP-packet, because it's good enough and a TCP-connection would take way too much time to setup and produce extra congestion on the network.

In today's world, ATM is mostly used to transmit TCP/IP-packets. The larger TCP/IP-packets must be sliced into smaller ATM cells, which increases the overhead. Total overhead on ATM backbones typically comes in between 15% and 25%. On a 155 Mbps circuit, effective throughput can drop to 116 Mbps. That's 39 Mbps down the drain. [8]

4.4.5 Interoperability

Different Ethernet versions work well together. Gigabit Ethernet technology is based on the original technology: CSMA/CD media access with the original protocol frame format. The upgrade path is relative straightforward. Upgrade can be performed in segments and divide the costs over a period. And what's important: the old LAN applications will operate unchanged. The cabling needs updates though. Gigabit Ethernet practically needs fiber-optic cables. Unshielded twisted-pair copper runs only up to 25 meters.

Since Ethernet technology occupies the lowest two layers of the ISO-protocol stack, layer 3 protocols such as IP run happily over Ethernet. Ethernet does not get in the way.

Because large portions of all networking applications have been builds for traditional LANs, ATM has had hard time to pursue to the desktop. Applications would have to be changed for ATM. ATM LANE, LAN Emulation, has been created to accelerate adoption of ATM. In essence, LANE allows ATM technology to be used in traditional LANs without any change in applications at the workstations. LANE allows an easy migration from Ethernet to ATM, but also means that sophisticated ATM features are not exploited.

The goal of MPOA, Multiprotocol over ATM, is to make existing LANs and their protocols interoperate with an ATM backbone. MPOA will have a greater role than LANE since ATM has not replaced existing LANs, and interoperability with legacy networks is the issue. MPOA will gradually replace LANE.

4.4.6 Management

What comes to installation and configuration, Ethernet beats ATM. Ethernet is well understood and nearly plug and play. ATM network configuration is rather difficult with many arcane parameters at the switch and the workstation. ATM takes time to install and requires a bit of expertise. This in turn directly affects the total costs of the technology.

ATM network management is more difficult than Ethernet LAN networks, due to many parameters of ATM networks and interoperability issues. In an essential role is the Interim Local Management Interface (ILMI) which uses SNMP across UNI and NNI to access status and configuration information within each network node. ILMI and ATM networking in general are still evolving.

4.4.7 Price

The Ethernet family holds a clear economic advantage over ATM. This is true for both network interface cards and for the network infrastructure equipment. A transition to an ATM LAN costs more than an Ethernet upgrade. The price is always an important business decision factor. Very few administrators will use a more-expensive technology unless they get actual benefit from it.

When deciding between ATM and Ethernet, professor Raj Jain from the Ohio State University says that buyers face the old house versus new house-dilemma. Fixing the old house is cheaper initially, but whereas a new house is more expensive, does it pay back in the long run [11]? The question remains to be answered.

4.5 Quality of service comparison

4.5.1 ATM

ATM has established quality of service standards. This is the key strength of ATM. Currently ATM is the best technology for transmitting voice, video and computer data over a single line. ATM offers a choice of four different types of service:

- Constant Bit Rate, CBR, provides a fixed and steady bit rate for real-time data. Analogous to a circuit-switched line. This is the most simple service level.
- Variable Bit Rate, VBR, provides a service for real-time and bursty data where the bit rate varies. This class has been lately divided into real-time and non-real-time.
- Unspecified Bit Rate, UBR, does not guarantee any bit rate. Used for data that can tolerate delays, such as traditional computer data. This can be seen as an interpretation of the common term "best effort service".
- Available Bit Rate, ABR, is a service for applications that can negotiate the bit rate during the transfer. The available bit rate varies in the network and the applications must adjust to different bit rates based on feedback from the network.

While it seems that ATM can fill every need for quality of service, there are problems. The mechanisms for achieving quality of service are complex. In order to have quality of service from end to end, the ATM switches must implement quality based routing over the PNNI-interface. Since PNNI lacked a standard for a long time, and being a complex entity, there are still problems in getting quality of service features to work between equipment from different vendors.

Besides the fact that operators rarely support all classes of service [7], users found it difficult to specify a particular quality of service. The service types are associated with a wide range of parameters to choose from, like cell transfer delay, peak cell rate, cell loss rate and cell delay variation tolerance. What happens is that users end up ordering ATM connections as leased lines, rather than as ATM services [7]. The technology is not being utilized to its maximum.

Despite the few problems and high cost, currently ATM offers the best way to implement quality of service over a WAN. A company private ATM network should integrate reasonably well with a public ATM network in the WAN and provide good quality of service for those who need it now and can carry the costs.

4.5.2 Ethernet

There is active debate among the networking community about Ethernet and it's quality of service. The current family of Ethernet does not provide explicit quality of service capabilities. Ethernet LANs have usually provided enough bandwidth to make quality of service needless. Some say that increasing the bandwidth of Ethernet is the same as adding quality of service. "If you give them a fatter pipe, you've pretty much solved their problem", says one network vendor [13]. While this may be true for smaller LANs, quality of service starts to matter when the networks gets larger and the interoperability with other networks gets into consideration.

Quality of service mechanisms for Ethernet is on the way. Current implementations include policy servers, tag switching, intelligent queuing, various IP-based implementations and tools. The problem with current implementations is that they are not standards-based; each vendor has its own solution and their devices will not interoperate. There is a clear need for standards or Ethernet can end up in a chaos of nonintegrating proprietary networks.

An attempt that is expected to bring quality of service standard to Ethernet is standards 802.1q and 802.1p, proposed by IEEE. They both operate at layer 2.

802.1q is a standard for providing Virtual LAN identification and quality of service levels. A Virtual LAN is a logical subgroup within a LAN whose purpose is to isolate traffic within the subgroup. 802.1q uses 3 bits to allow eight priority levels and 12 to identify up to 4096 VLANs. 802.1p allows switches to reorder packets based on the priority level. 802.1p also defines means for stations to request a membership in a multicast domain and map it to a VLAN.

IEEE 802.1 proposals bring support for RSVP, which is a protocol to request and provide quality of service network connections at layer 3. RSVP is dependent on layer 2 to provide the quality of service over a data link. Since Ethernet has not been able to provide quality of service, RSVP has been run over unprioritized Ethernet links, which is clearly not the desired scenario. Taking advantage of IEEE 802.1, RSVP support can be achieved by mapping RSVP sessions into 802.1p priority levels. Whether RSVP will succeed or not is another matter.

A key question for Ethernet is how to be able to provide quality of service across a WAN and over heterogeneous network infrastructures. Ethernet alone cannot be used to transmit data from point A to point B anywhere in the world. In the middle, there lies a WAN, and before end to end quality of service is achieved, the quality of service requirements must be mapped and transmitted from one transport technology to another. Before we get there, technologies need to evolve.

4.6 Future Network Technology

4.6.1 IP

IP has a strong position of becoming the basis of the future networking technology. Today, IP is the de facto standard for transmitting data over the Internet. The family of TCP/IP-protocols is becoming increasingly successful. One of the elements contributing to the success of IP is that it is completely independent of the underlying network technology. IP can operate over heterogeneous network infrastructures. At the other side of the coin, the greatest weakness of IP is the lack of quality of service. IP provides only best effort delivery.

During the early years of ATM, it was seen as the unified choice for virtually every networking. IP was not a big player at the time and certainly not a threat to ATM. Now things have changed. IP has become the dominant networking technology, and we face the issue of how these two mesh together. The nature of the two collides with each other: IP is connectionless protocol while ATM is fundamentally connection-oriented. IP is a packet-routing technology routing variable length packets with no delay guarantees while ATM is a cell-switching technology with strict quality of service.

IP and ATM don't mix well together. Integrating IP and ATM to operate efficiently together is a challenging task. Today, IP is transmitted on top of ATM without taking advantage of the features of ATM. There exists overhead in encapsulation, routing, assembling and reassembling packets, and ATM's quality of service is left unexploited. The strengths of ATM are not being used. ATM could be the technology for providing quality of service to IP-networking, creating a happy marriage of these two. On the other hand, some see that ATM is not needed in the future TCP/IP network infrastructure [16].

Ethernet and IP operate well together. Ethernet is packet-oriented connectionless service just like IP. Operating at different layers, there are little conflicts with these technologies. Seeing the strong future of IP, it is easy to believe that the network technology that will win will be the one that is the simplest, fastest, cheapest, and easiest to use with IP. Ethernet fits well into this picture.

4.6.2 LAN Technology

It is expected that Ethernet remains the most popular LAN technology and the most widely deployed access network in companies in the foreseeable future. Ethernet has a huge base of installations, the technology is simple and well-understood. For those that want extra bandwidth, Ethernet provides switching and easy upgrade paths to the newer incarnations of Ethernet, Fast Ethernet and Gigabit Ethernet. Quality of service is being standardized. Ethernet is still hot technology.

To take ATM to the LAN and desktop it would require quite a large infrastructure change, which doesn't come free. Equipment has a high price and the interoperability issues which legacy LAN applications must be taken into account. While it is true that in order to take full benefit of ATM, it should reach all the way to the desktop, Ethernet with good-enough quality of service mechanisms can eliminate the need of ATM to the LAN.

4.6.3 Backbone Technology

ATM is widely employed in Internet WAN backbones. Almost all Internet operators run and offer ATM services. Ethernet has grown from pure LAN to backbones as well. The question from ATM's point of view is how far towards LAN can ATM infiltrate while Ethernet's point of view is how far towards backbones can Ethernet infiltrate.

ATM was designed to provide broadband networking and run efficiently from LAN to WAN. But ATM's problem is IP. There is a large overhead with IP over ATM, and alternatives to ATM as the transport technology are being researched. One such a technology is Packet over SONET, or IP over SONET, as it is also being called. The European counterpart to SONET is SDH.

On the bottom layer of a contemporary Internet backbone, there is SONET or SDH, which are layer 1 specifications for data transmission over optical fibers in the public network. ATM runs on top of SDH. In IP over SDH, ATM is eliminated from the transmission picture all together, and IP-packets are transmitted directly on top of SDH frames through the use of Point to Point Protocol (PPP). As a result, the bandwidth gets utilized more efficiently. IP over SDH can provide as much as 25% to 30% higher throughput than ATM [5]. IP over SDH is an ideal for transmitting IP. And when we talk about IP-packets, one doesn't have to stretch his mind far to start thinking about transferring Ethernet frames on top of SDH. Pretty nice picture.

But this is not the whole truth. The migration will not be that easy. By eliminating ATM, we loose the management, routing and other features of ATM. The management infrastructure required for SDH is completely different from ATM's. IP over SDH is best suited for high-volume point to point configurations. In a more complex and hierarchical network, it runs into trouble.

Looking even further into the future, some have questioned how SDH fits into the future broadband networks [6, 9]. A new technology called WDM, Wave Division Multiplexing, is an emerging technology that allows multiple optical signals to be transported over a single fiber, providing massive bandwidth for the 21st century. Each signal can carry a different channel, an SDH-channel, for example. But instead of using SDH frames in the channel, one could use Ethernet frames.

Recently Siemens conducted a pilot project where it transmitted Ethernet frames at a speed of 1Gbps over a full-duplex optical fiber with WDM [17]. The link distance was 1570 kilometers, a current world record. With such advancements in technology, the battle between ATM and Ethernet is strongly extending from LAN to WAN too.

4.7 Conclusion

The competition between ATM and Ethernet most likely continues many years to the next millennium. ATM has a well-established position in Internet backbones, and Ethernet dominates LANs. Neither is disappearing any time soon. Both technologies will co-exist for time to come. There are huge investments on both sides to the infrastructure and no matter how superior any technology is, migrations from one technology to another always takes time.

Looking at the sheer number of installations, price and easy of use reveal that Ethernet is the dominating technology in the LAN, and ATM can hardly change this. In the backbones, the competition seems to be harder. Ethernet is a not real threat but there are experimental transport technologies that ATM has to face.

Quality of service will have its effect in the development. There is a clear trend that Internet is increasingly being used to run real-time communication services such as voice over IP. Quality of service will matter in the future. ATM already has quality of service, though not fully implemented, that puts it ahead in this sector. Ethernet does not have quality of service, but development is active. While Ethernet may not achieve the "state of the art" quality of service of ATM, it may well provide good enough mechanisms to satisfy most of the needs.

The unified factor in the future networking seems to be IP rather than ATM. Marrying these two technologies could provide an answer, but due to the fundamental differences of the two, other alternatives are being seeked as well. Running IP directly over backbones is a viable technology, but at the current state it is somewhat limited and cannot provide as comprehensive solutions as ATM. On the whole, somehow it makes sense to believe that the winning technology will be the one that integrates most efficiently with IP.

CHAPTER 5
ATM AND THE 3RD GENERATION

There has always been a general need for more and cheaper bandwidth, but never more so than today. This need has been accelerated by the availability of high performance and relatively low-cost PCs and the workstations and capable business software applications. In the future public network operators (PNOs) and the service providers will be increasingly engaged in providing easy to manage Broadband services on demand at an affordable price.

The situation now taking shape holds out new opportunities and the prospect of more widely differentiated roles for all new and traditional operators in the sector. The network offers a wide range of services, including information transport on stationary, mobile and satellite network with national, international and global connectivity, supply of voice, data transmission, video distribution or integrated services and basic as well as value added services.

ATM development rests on a basis of well-established standards and on subsequent detail specifications. These specifications were formulated with the precise objective of providing detailed and widely accepted reference specifications to the international community of network and the service operators, manufacturers, system integrators and the users. They are used in developing introworkable commercial equipment capable of providing the advanced communication services required for the information society.

In July 1996, a major milestone in the development of ATM Technology was reached, then the ATM Forum announced agreement on the Anchorage Accord. The Anchorage Accord contains the foundation Specs for mission-critical ATM infrastructure , as well as expanded features Specs for migration to ATM multi-service networks. In future , the foundation Specs will be revised to correct the problem , align with ITU-Recommendations , and add Cohesiveness/Parity between specifications . Any new specifications will be backward compatible Anchorage Accord specifications. The introduction of ATM Layer services will increase the benefits of ATM , making the technology suitable for a virtually unlimited range of applications, since an ATM network is able to offer different levels of services . The corporate information technology and the communication market have flexibility and a capacity of rapid change that are far superior to those of conventional Telecommunications. The emerging ability to work together in an open environment (i.e. between solutions provided by different manufacturers) over a wide geographical area will increase the weight of this market.

In the corporate setting, one of the priority requirements is the interconnection of Local Area Network (LAN). Offering this service on a public or private Wide Area Network (WAN) allows high volume data and image transfer, e.g. for the Computer Added Design (CAD) and the manufacturing applications, workgroup sessions, or multimedia communications, where correlated processing of data, image and voice information is performed by the user terminal. In the area of corporate services, the most pressing need is to be able to make good use of the speed, availability quality and the security, which have become such market features of the Local Area Network (LAN). As far as future home applications are connected, multimedia will be an essential feature and particularly if associated with interactivity is destined to have a profound impact on our habits and social dealings.

This certainly aside, it is still too early to tell how quickly this market will widely develop, given its heavy dependence on mass phenomena. Nevertheless, it is now time to begin assessing the appeal of new services and new ways of providing services on sample group of customers. Given that the residential sector has not yet made much use of Narrow Band of ISDN services. One exception is the home entertainment sector. The demand for improved quality cable television and simultaneous growth in the installation of residential fiber access offers a chance for integration, atleast at the access level, of residential communication and entertainment needs. This will provide a starting point for future Broad Band services. Many studies based on questionnaires and analysis of cable TV services have shown that, in principle atleast , subscribers are also willing to pay for various video retrieval services and enhanced quality TV.

There are several reasons why ATM is the basic technological choice, endorsed internationally, is aimed at developing general, unifying solution for all application environments. Here the primary objective is integrated handling of signals produced by voice sources, data , images and video , regardless of bandwidth associated with each type of information. In principle a network based on ATM can thus satisfy the requirements of both the business environment and the household environment. The demand for broadband services is in its initial phase and growing. While just at the beginning of the commercial offer by network operator and service providers, inherent speed, response time and quality advantages will ensure the gradual uptake of services based on broadband network.

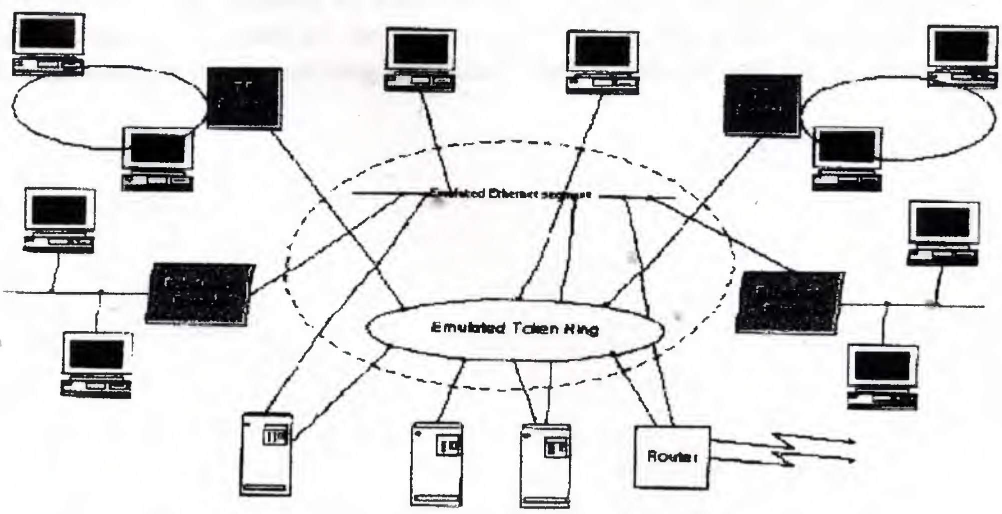


Figure 5.8 Logical views of integrated LAN - ATM environment

CONCLUSION

In this project, I have tried to analyze different issues pertaining to ATM, and discussed some of the ideas of different researches done on it. ATM is a relatively new field. The ATM networks play an important role in the broadband communication networks now and in the future.

This project has been compiled to make such an examination possible. It describes the wonderful resources and achievements that have been done on the field of networking and what it has become. It is hoped that this project will encourage many who are not yet online to realize that ATM networking is well worth the difficulty and trouble of studying and working on it. Also this project is intended to provide for courses for the student in education, communication, library sciences, computer sciences and engineering etc. The material contained in this project is intended to encourage students and teachers to begin to research and write about these developments so that there begins to be a substantial body of material documenting both where these developments have come from and the potential that they promise for a better world .

Most importantly, the information in this project are offered as a way of providing the public whose money and labor made these achievements possible, with a way of evaluating proposals to change the course of development for this network. These information about ATM are a contribution toward evaluating what has been created, and what its social and scientific potential is The past few years have seen people in Eastern Europe and around the world demonstrate that they need better need and working conditions. These cries for change mean that the methods that have achieved this Global network needs to be applied to other aspects of society so that the well being of the people become the concern of the government. Finally, this collection of the ATM is for those who want to see the coming millenium bring a better world for everyone.

REFERENCES

[1] "ATM User Network Interface Specification Version 3.0", ATM Forum, PTR Prentice Hall, September 10, 1993.

[2] "Asynchronous Transfer Mode (ATM) and ATM Adaptation Layer (AAL) Protocols Generic Requirements", Bellcore, Issue 1, August 1992.

[3] "Synchronous Optical Networks (SONET) Transport System: Common. Generic Criteria", Bellcore, Issue 2, December 1991.

[4] "Broadband ISDN ATM Layer Functionality and Specification", January 1993.

[5] "Fiber Channel Physical Layer (FC-PH)", Revision 2.1, Working Draft, May 25,1991.

[6] "Broadband ISDN Transport Network Elements Framework Generic

[7] "Broadband ISDN Switching System Generic Requirements", August 1992.
Criteria", Dec 1990.

[8] "B-ISDN User Network Interface-Physical Layer Specification", June 1990.

[9] "OAM principles of B-ISDN access", Geneva, June 1992.

[10] “General Aspects of Quality of Service and Network Performance in Digital Networks including ISDN.

Web References :

[11] http://www.ieee.org

[12] http://www.cisco.com/univercd/cc/td/doc/product/core/cis7505.htm

[13] http://www.cisco.com/univercd/cc/td/doc/cisintwk/idg4/nd2008.htm

[14] http://www.computer.org/