12-1999

The design and analysis of a corporate data network supporting a real-time clinical data application

Craig M. Longo

To the Graduate Council:

I am submitting herewith a thesis written by Craig M. Longo entitled "The design and analysis of a corporate data network supporting a real-time clinical data application." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Electrical Engineering.

Daniel Koch, Major Professor

We have read this thesis and recommend its acceptance:

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a thesis written by Craig M. Longo entitled "The Design and Analysis of a Corporate Data Network Supporting a Real-Time Clinical Data Application." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Electrical Engineering.

Daniel B. Koch

Daniel Koch, Major Professor

We have read this thesis
And recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor and
Dean of The Graduate School

THE DESIGN AND ANALYSIS OF A CORPORATE DATA NETWORK SUPPORTING A REAL-TIME CLINICAL DATA APPLICATION

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

C. M. Longo
December, 1999

DEDICATION

This Thesis is dedicated to my parents

Mr. Joseph Gitano Longo

And

Mrs. Linda Louise Longo

Who have provided me with the proper guidance and structure and afforded me invaluable educational opportunities.

And to my wife

Mrs. Jeanne Longo

Whose support and encouragement assisted me in the completion of this work.

Abstract

In this study a design is proposed for a corporate data network supporting real-time data applications. The proposed network incorporates both Local Area Network and Wide Area Network technologies to form a system capable of supporting a variety of applications. Multimedia software, like desktop video conferencing, IP telephony, and video streaming are becoming more pervasive. Since multimedia applications depend on active human involvement and perception, they are commonly referred to as real-time. The content of real-time applications relies on the timely and consistent delivery of information. If real-time applications experience any variation in information delivery, usually referred to as jitter, the result is unacceptable application performance. However, real-time applications are not solely limited to traditional multimedia. Interactive client-server based data applications also fall into this category. This project will specifically focus on the performance of a real-time clinical application, which has become predominant in the healthcare industry.

To support the implementation of the proposed network, empirical data was gathered from system testing. Testing involved comparing the performance of a real-time application on the proposed design, against the current architecture. The result found that the proposed data network design reduced transport latency, allowing the real-time application to perform more efficiently.

Table of Contents

List of Tables

Table | Page

List of Figures

Figure Page

1.0 Introduction

This thesis explores the design associated with a corporate data network supporting real-time data applications. The proposed network incorporates both Local Area Network (LAN) and Wide Area Network (WAN) technologies to form a system capable of supporting a variety of such applications. Multimedia software, like desktop video conferencing, IP telephony, and video streaming are becoming more pervasive. However, real-time applications are not solely limited to traditional multimedia. Interactive client-server based data applications also fall into this category. Since multimedia applications depend on active human involvement and perception, they are commonly referred to as real-time.

The content of real-time applications relies on the timely and consistent delivery of information. If real-time applications experience any variation in information delivery, usually referred to as jitter, the result is unacceptable application performance. This thesis will specifically focus on the performance of a real-time clinical data application, which has become predominant in the healthcare industry.

Meditech [1] sells healthcare companies an application suite that provides an electronic means of patient accounting, faxing, billing, and

processing. A proprietary protocol called Magic is used for communications between the Meditech processor and host nodes. Nodes may take the form of personal computers, terminals, nursing hand-held units, and lab instrumentation. Meditech and other real-time applications place a new set of demands on legacy data networks. To support these applications legacy networks often require a redesign.

1.1 Network Requirements for Real-time Clinical Application Protocols

The Magic protocol was developed for use on Local Area Networks. Magic can be categorized as a non-windowing protocol which is bound by strict delay characteristics. Since the application relies on interactive data entered over the network, the protocol will time-out causing the application to fail when the latency between the processor and the node becomes too great, roughly 400-milliseconds.

Real-time applications such as Meditech impose a new set of demands on corporate telecommunication networks. Specifically, real-time applications are very data intensive; these applications require an increase in network bandwidth. Without sufficient bandwidth, the network will constantly experience congestion. If congestion continues, performance associated with the application will suffer.

However, bandwidth is not the only network requirement. The interactive nature of these time-sensitive applications requires deterministic transmission. Non-deterministic transmission will result in jitter or variable delay between application packets. Jitter and delay are acceptable for store and forward applications like e-mail, and file transfers. Real-time applications are intolerant of jitter and variable delay. Non-deterministic access to the network will cause real-time applications to time-out. Time-outs affect user perception and result in poor application performance.

Finally, real-time applications share the network with other non real-time applications. During periods of network congestion, it is imperative that these applications are differentiated via a prioritization mechanism. Without prioritization, all packets are treated equally. Store and forward applications that accommodate delay and retransmissions should receive a lower priority than Meditech. With the ability to assign priorities during congestion, the network can choose to discard file transfer packets while allowing delay-sensitive Meditech packets to reach their destination on time. When network prioritization is applied, both types of applications, real-time and non real-time, may co-exist on a single network infrastructure.

1.2 Shortcomings of Current Network Design

The current network design for the case study to be considered here was constructed in 1992. Although the Meditech application constraints mentioned above existed when the original network design was proposed, a decision was made to separate the clients from the application processors. This decision was based solely on economic factors regarding the resources involved in the management of the processors. This left the network designers responsible to solve the latency problem created by the data network. The approach was to construct two data centers housing the application processors.

Each of the original twenty facilities was connected to the data center, located in close geographical proximity. The network design consisted of three major components. The first component used shared Ethernet hubs to connect low-bandwidth, highly delay-sensitive clients to the facility LAN. Dedicated T-1 transmission lines were chosen for the second component, constructing the WAN. Previous analysis indicates that the latency induced by the WAN accounts for sixty percent of the overall delay in the system. Dedicated T-1 lines were utilized, because they have fixed latency based on their physical distance (Figure 1.1).

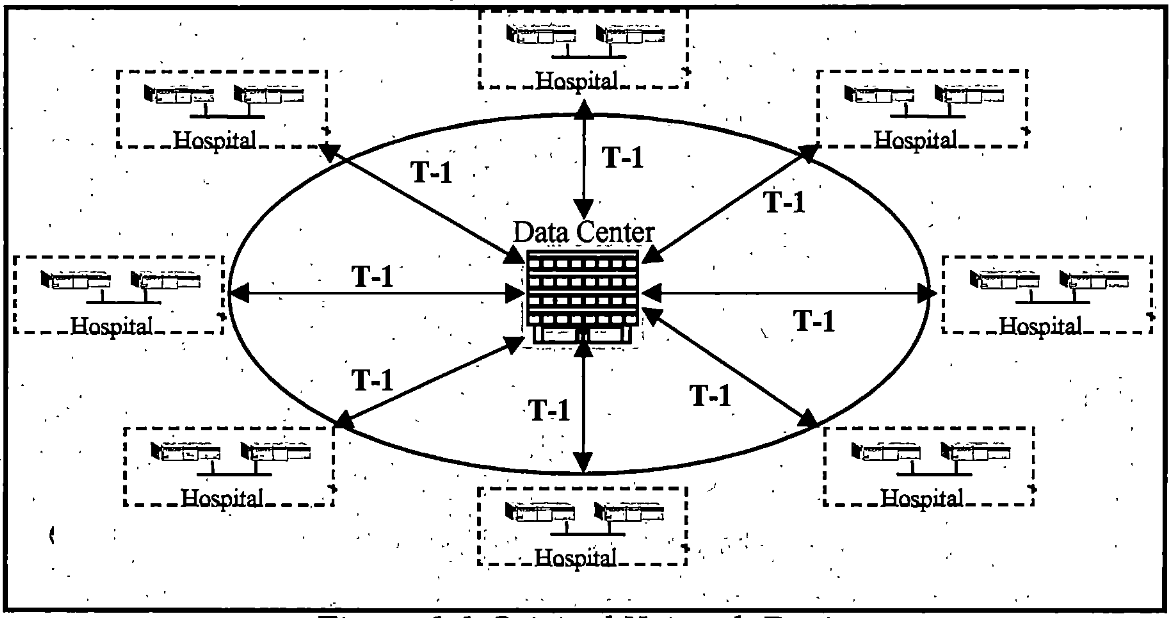

Figure 1.1 Original Network Design

By connecting a facility to the closest data center, distance, and therefore latency, could be minimized. However, these T-1 lines require the purchase of fixed bandwidth; in this case, more than necessary to operate all the applications in the facility. This wasted bandwidth translated into wasted money. Finally, the third component utilized switched Ethernet to connect the application processors to the data center LAN.

During the early and mid 1990s the corporation grew from twenty facilities to over five hundred. This rapid growth through acquisition caused the data network to increase to its current size of eight regionalized data centers with over one thousand T-1 transmission lines.

Rapid growth has uncovered several scaling issues with the original network design. Latency experienced when traversing the network has become so large that many facilities have suffered application time-outs. This latency is induced by the non-deterministic nature of the shared Ethernet LANs, and the inability of the WAN to distinguish the clinical application and prioritize it. A new data network design has been proposed, and will be explored in this thesis. The new design has the following goals:

- Reduce all latency components in the network, providing the best possible application performance.
- Accommodate future growth, and the convergence of voice, video and data multimedia applications onto a single network.
- Take advantage of lower cost alternatives in the Wide Area Network.

1.3 Proposed Network Design Providing Real-time Application Support

The new design has accomplished these goals by incorporating a variety of recently developed technologies. On both the facility LANs and Data Center LANs, Ethernet switching was introduced with a modification to the Ethernet back-off algorithm. The modification technique used is trademarked from 3Com Corporation under the name Priority Access

Control Enabled (PACE™). PACE™ technology allows for deterministic transmissions and provides a mechanism to prioritize certain applications over the LAN. Field measurements are used to determine the effect this algorithm has on reducing jitter and delay. The second technology investigated in this thesis is Asynchronous Transfer Mode (ATM). ATM was chosen as part of the new design because of the ability it possesses to distinguish between protocols, assign priorities, and define delay characteristics to specific flows.

This thesis will use empirical data from field measurements to determine the effects of implementing the proposed data network. Each of the three components in the design will be tested separately; comparisons against the performance of the current network component will be analyzed. Finally, each component will be assembled and the system will be tested as a whole.

2.0 WAN Current Network Design

The current WAN design consists of leased lines deployed as a collection of hierarchical trees. At the root of each tree are the regional data centers. The hierarchical tree model was chosen to accommodate the traffic flow of the predominantly time-sensitive Meditech application. Meditech clients, which reside in the facilities, communicate exclusively with servers located in the regional data centers. Since clients do not communicate with clients in other facilities, there is no need to directly connect facilities together. If applications were deployed in facilities that required any-to-any traffic flows, such as wide world web and file transfer, the tree model would be inefficient. Regional Data Centers (RDC) are inter-connected via a leased line mesh. Mesh architectures accommodate applications that require any-to-any data flows. Each RDC is directly connected to every other RDC, supporting web, financial and backup file transfer applications.

The hierarchy of the design consists of three levels: Regional Data Centers, Tier-1, and Tier-2 facilities. Groups of Tier-1 and Tier-2 facilities, which share the same servers in the data centers, are called *markets*. Markets are constructed by grouping facilities according to their physical location. Tier-1 facilities are the largest, and have a

greater number of clients. Tier-1 facilities have direct connectivity to the data center via their primary circuit. A second redundant circuit is connected to a peer Tier-1 facility in the market (Figure 2.1). Tier-2 facilities have a smaller number of clients, and do not have direct connections to the data center. Instead, Tier-2 facilities connect through Tier-1 facilities. Data from Tier-2 facilities must transit Tier-1 facilities incurring additional delay in route to the data center. Each Tier-2 facility has both a primary and a secondary circuit. The primary and secondary circuits are connected via different Tier-1 facilities within the market for redundancy. Since leased lines are priced by bandwidth and distance metrics, this hierarchical model allows leased line costs to be minimized across the enterprise.

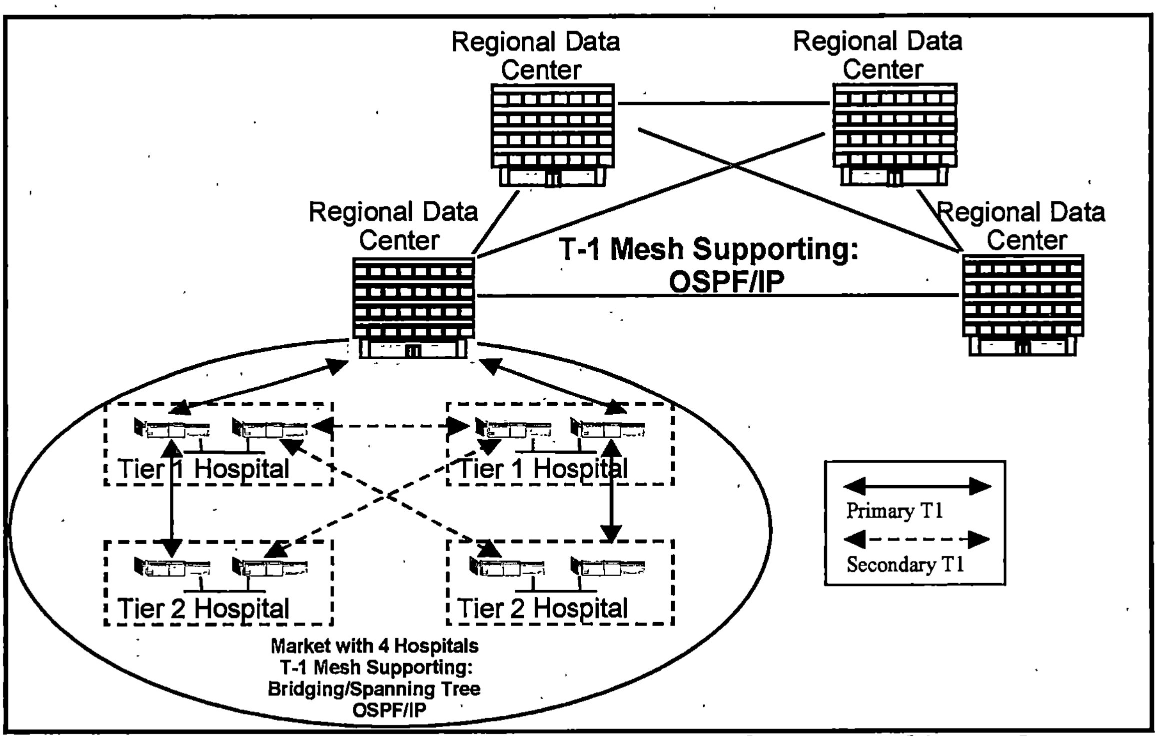

Figure 2.1 Current WAN Design Tier-1, Tier2, Regional Data Centers

2.1 Leased Lines

Leased lines are dedicated digital transmission circuits connected between two terminating points. The fundamental building block of the North American Digital Hierarchy is a single transmission channel know as Digital Signal Zero, or DS-0. Each DS-0 consists of eight bits per frame, which are transmitted at 8000 frames/second, yielding a bandwidth of 64-Kilobits/second. The current network design utilizes leased lines with a fixed bandwidth of 1.544 Megabits/second. This type of leased line is commonly referred to as a T-1 or DS-1, and is comprised of 24 DS-0s. At the time of the design, only DS-0s and T-1s were available for purchase. T-1s were chosen based on the requirement that each facility must have a sustainable 768Kilobits/sec of bandwidth to the data center. This left the additional bandwidth purchased unused. Each T-1 line was terminated between a router at the facility and the data center. At each facility, two routers were deployed. The primary and secondary T-1s were connected to separate routers for redundancy. The routers are configured to route certain protocols like the Internet Protocol (IP), and bridge non-routable protocols like Magic over the same physical T-1 circuits. Data from the routers are transmitted serially to the T-1 line. This data is time division multiplexed into the 24 channels of the T-1. Leased lines do not posses any capability to prioritize traffic, and are unable to distinguish between Meditech and IP data packets in this

design. Ethernet frames are directly encapsulated in the point-to-point protocol (PPP) for transmission across the leased line. The PPP protocol introduces minimal overhead. This protocol is utilized on leased line WAN links to provide additional error checking for greater reliability.

At the data center, two routers were deployed for each market. All primary T-1s from Tier-1 hospitals terminate in the market routers. Although both circuits from each hospital represent a unique path to the data center, they can not be used simultaneously for Meditech which is a bridged protocol. A logical structure must be placed on the bridged network to ensure there are no loops in the design.

2.2 Logical Topology Layer 2

The logical structure evolved in computer networking separates the communication functions from the application processing. This separation of networking functions is called layering. The 7-layer technique standardized by the International Organization for Standardization (ISO) committee is referred to as the Open Systems Interconnect (OSI) [4] reference model. Each layer uses its own layer protocol to communicate with its peer layer between communicating stations. The functions of each layer are outlined in figure 2.2.

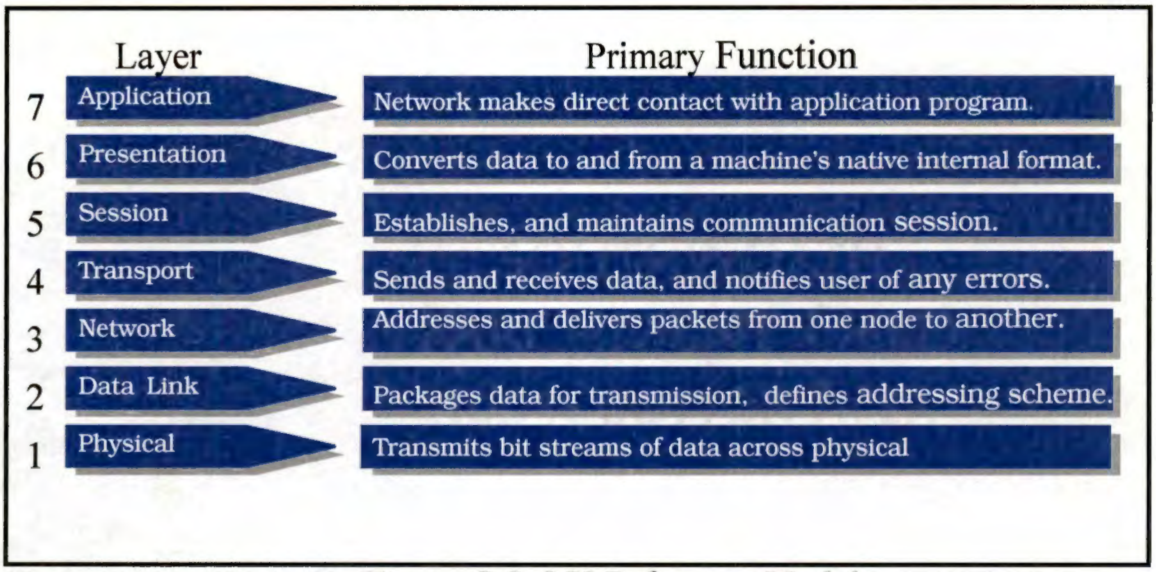

Figure 2.2 OSI Reference Model

The application Meditech runs on a proprietary protocol called Magic, which is non-routable. If a protocol is non-routable, meaning it does not posses a network layer in the OSI reference model, it must be bridged at the data link layer. Bridging uses broadcasts as a means of communication. Since broadcasts are forwarded by every bridge, loops must be prevented in the network design. If a loop existed, a broadcast could be infinitely forwarded. A protocol was developed called Spanning Tree [8] to logically detect loops. Once Spanning Tree detects a loop, it communicates its existence among all bridges on the network. The protocol then instructs the bridges to activate connections that do not cause loops in the network. Spanning trees are created on a per market basis. Any change in circuit status affects the spanning tree for the

entire market. Magic traffic is purposely confined to individual markets. The only protocol allowed to traverse markets is the routable protocol IP.

2.3 Logical Topology Layer 3

Similar to the Layer 2 protocol Spanning Tree, Layer 3 protocols such as the Internet Protocol (IP) exchange reachablility information between routers. A significant different exists in the Layer 3 treatment of network loops. In a Layer 3 environment broadcasting does not occur. Consequently, loops may exist. Redundant circuits always remain active and capable for IP data transmission. All applications except Meditech utilize IP, and operate concurrently on the network.

The original design incorporated the use of a link state IP routing protocol-Open Shortest Path First (OSPF) [3]. OSPF, a link state protocol, was chosen over distance vector-based protocols for its scalability, reliability, and administrative ability to prioritize circuits with a configurable cost metric. Distance vector protocols, such as Routing Information Protocol (RIP) [3], provide reachability information based solely on the number of hops between endstations. Each router broadcasts reachability information to other routers every 30-seconds. This type of protocol does not take into account the bandwidth of each

hop, and consumes a great deal of bandwidth by broadcasting frequently. Link-state protocols on the other hand take into account the bandwidth associated with every hop and will prioritize the path with the most bandwidth or lowest administrative cost to the destination. Additionally, routers communicate via multicasts only when a link is dropped or added to the network. Multicasting, or sending a single packet to a group of stations, represents a more efficient means of communications compared to broadcasting. Broadcasting requires every station on the network to process the packet and decide whether it has an application that uses the information. Finally, an OSPF enabled networks only send updates to existing information when needed. If a link fails or returns a single OSPF, advertisement will be sent to the remaining routers indicating that a change has occurred in network topology.

OSPF was designed for large networks and possesses a hierarchy in the form of areas. A common backbone area is created with sub-areas connecting to the backbone. Topology information is shared internally with every sub-area and summarized for the backbone area. This form of distributed processing per area allows OSPF enabled networks to scale to large implementations. The original design organized sub-areas by markets. By constructing OSPF sub-areas based on markets and

controlling the number of routers per area, the routing table held by each router is minimized. This allows for efficient operation providing quick route table look-ups and fast topology convergence.

At the data center the market routers serve as area border routers, connecting the market sub-area to the backbone OSPF area. The backbone area is referred to as area zero. Routers in area zero posses large routing tables, containing connectivity information for all facilities connected to the data center.

In order to facilitate inter-data center traffic, a highly complex yet scalable link-state protocol called Border Gateway Protocol (BGP) was used [5]. BGP has the capability to summarize routes based on contiguous assignment. Similar to OSPF, BGP dynamically summarizes routing information, However BGP uses a process called Classless Inter-domain Routing (CIDR) [5]. This form of summarization is the most powerful, condensing a large number of routes to form a single route advertisement. Additionally, BGP communicates through reliable Transmission Control Protocol (TCP) connection to peers. Since routes were assigned contiguously to each data center, and a leased line mesh existed between regional data centers, BGP was the most efficient routing protocol to employ in this section of the design.

2.4 LAN

The facility and the data center both use Ethernet hubs to provide Local Area Network connectivity for end-station devices. In the facilities these devices consist of personal computers, terminal servers connecting multiple serial devices, hand-held devices, and lab equipment. At the data center, end-stations are comprised of processors that contain the application databases. Ethernet hubs operate at Layer-1 on the OSI reference model. This layer, called the Physical Layer, allows electrical signal propagation between ports. All hubs connected to a switch-port, and all stations connected to the hubs, form what is called an Ethernet segment. Ethernet uses Carrier Sense Multiple Access with Collision Detection (CSMA/CD), for network access [6]. Each segment forms a collision domain, and operates at 10-Megabits/second. All the end-stations with Ethernet controllers participate in the collision domain on their segment. Excessive collisions result in delayed network access for the end-stations. This variable delay can significantly degrade the performance, and responsiveness, of real-time applications.

The only time that an Ethernet controller detects a collision, is when the controller's state machine is in the transmitting mode. After transmission, the Ethernet controller's state machine switches to receive

mode and listens for the 64-bit preamble that signals the start of a frame (Figure 2.3).

2.5 Truncated Binary Back-off Algorithm

When packets from two or more stations overlap in time, the stations transmissions are said to have collided. When a station detects a collision has occurred, the sending station's controller generates a random retransmission interval, which is determined by the truncated Binary Exponential Back-off algorithm (BEB). The BEB algorithm is based on the concept of slots. One slot corresponds to the maximum round trip delay value of 51.2 microseconds or 512-bit times [6].

An Ethernet controller begins transmission of each new packet with a mean retransmission interval of one slot. Each time the controller records a collision, the retransmission delay interval is increased by a random length. This random length is twice the mean of the previous

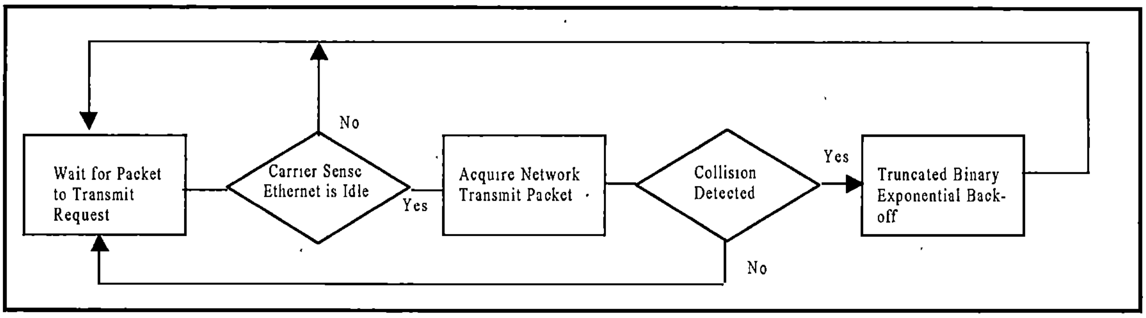

Figure 2.3 Flow Chart for Collision Detection [7]

interval, up to a maximum truncated value of 16. This can be expressed mathematically as follows: The number of slot times to delay before the n^{th} retransmission attempt is chosen as a uniformly distributed random value r in the range:

$$0 < r \leq 2^k,$$

Figure 2.4 Equation for Slot time Calculation

Where $k=\min(n,10)$ and $0 < n \leq 15$ (where $n=0$ indicates the initial attempt). [2] After 16 consecutive attempts the controller drops the packet from the transmission queue and indicates a deferred attempt.

2.6 Capture Effect

The truncated Binary Exponential Back-off algorithm was developed to arbitrate station contention to minimize delay under light loads, while stabilizing under heavy loading conditions. A significant problem with the algorithm appeared to be caused by an increase in utilization from real-time applications. Under heavy loading, it is possible for a station to capture the channel. This is known as channel capture, or packet starvation, depending on whose perspective is taken. The station gaining

almost exclusive use of the network has captured it, while the station locked out, potentially aborting future transmission is starved.

As an illustration of the capture effect, consider the following example: A pair of stations each with an infinite transmit-queue attempt to gain access to an initially quiet Ethernet. Both stations observing a quiet channel will transmit and collide with each other. After the collision occurs each controller will abort the transmission, jam the line, and enter back-off.

Since it is the first retransmission for each station, both controllers pick a random delay interval between (0,1). If both stations choose the same delay value, each station will repeat the process until different values are chosen. Once this occurs, the station that chose the lowest value (Station A) will win, gaining access to the channel (Figure 2.4). After successful transmission, both stations still having data to send will attempt to transmit. A collision will occur, this time station A having reset its collision counter after the successful transmission will enter the back-off algorithm with a delay range between (0,1). Station B suffering its second collision for this transmission will choose a delay range between (0,3). The probability that Station B will win over Station A is only 1 in 4 (A picks 1 and B picks 0). More likely, Station A will win and transmit its

second frame. The process repeats with great probability that Station A will transmit sixteen frames, until Station B discards its packet, and resets it algorithm.

As the result of the capture effect, one Ethernet station is capable of dominating the network. Other stations are essentially blocked from sending data for long periods. This inconsistent network access causes latency and jitter effects to real-time data applications, often rendering the application useless to its user.

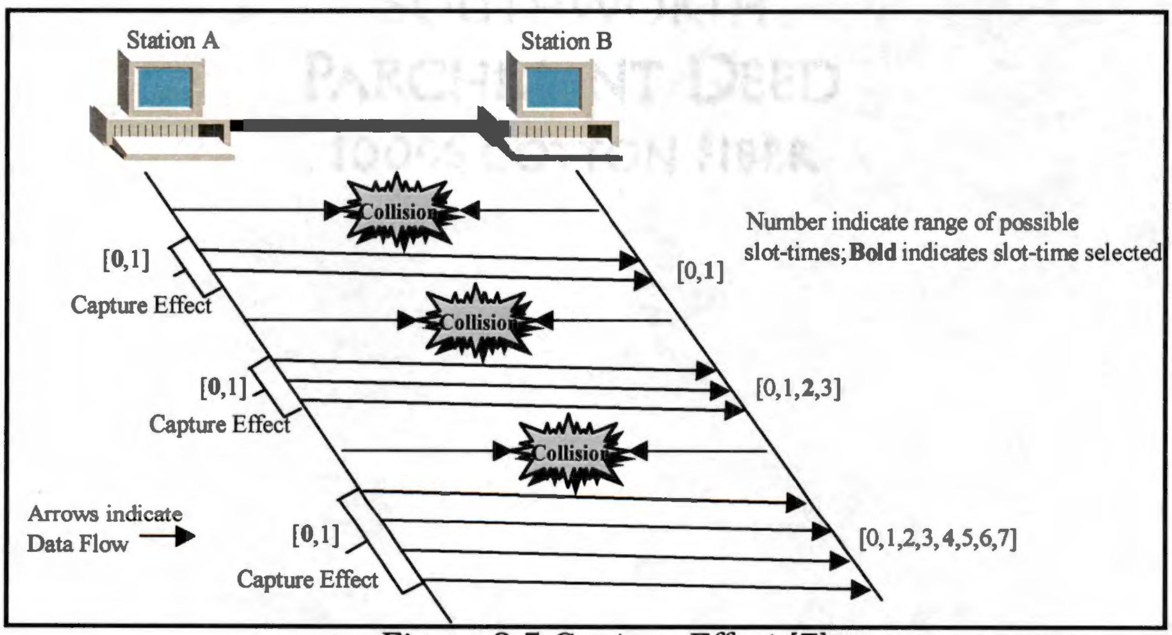

Figure 2.5 Capture Effect [7]

3.0 Modified Network Design

The modified network design replaces the hierarchical leased line tree with a public Asynchronous Transfer Mode (ATM) cloud. Public service providers offer cloud technologies to customers by provisioning physical bandwidth with the creation of Virtual Circuits (VC). Connection identifiers on each virtual circuit, ensure that information is routed correctly across the network. Aggregating multiple customers onto public cloud technologies allows service providers to offer less expensive circuits in comparison to leased lines. Cloud technologies also offer connection flexibility. A single facility may be simultaneously connected to many locations utilizing multiple point-to-point virtual circuits over a single physical connection (Figure 3.1).

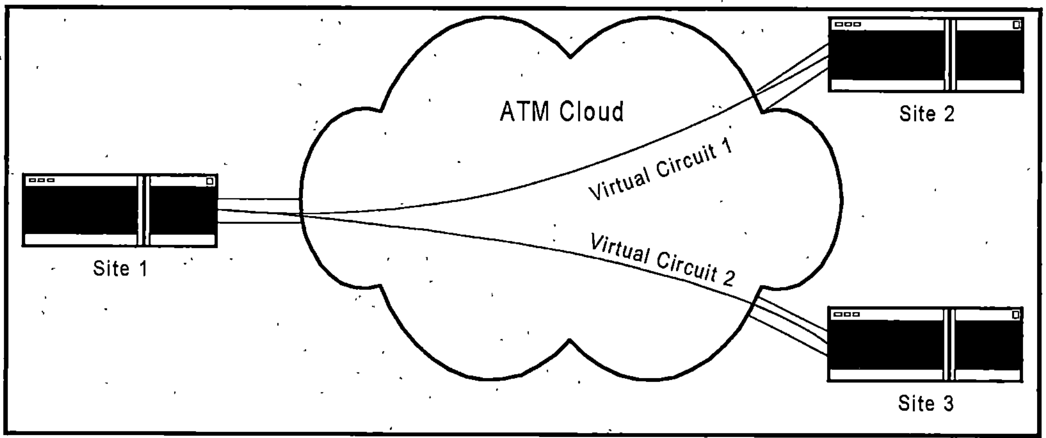

Figure 3.1 Virtual Circuits in Cloud Technologies

The most popular cloud technologies are Frame Relay [9], ATM, and SMDS [9]. Of the three, only ATM has the ability to support the combined services of voice, video, and real-time data over a single infrastructure. To accomplish this convergence, ATM has the capability to prioritize data and to guarantee service level parameters such as: variable delay, burst sizing, and sustainable bandwidth. These parameters are required for reliable transportation of real-time data from the hospital clients to the regional data centers. Since ATM satisfies the requirements, it was chosen as the WAN architecture for the modified network design.

3.1 ATM WAN

Public ATM Wide Area Networks are comprised of a series of ATM switches that execute highly complex routing protocols, distributing reachability information along with the available capacity of each route [10]. Edge devices, such as routers, connect to the switches and terminate virtual circuit connections (Figure 3.2). Fixed length frames called ATM cells are switched across the network based on virtual circuit identification numbers.

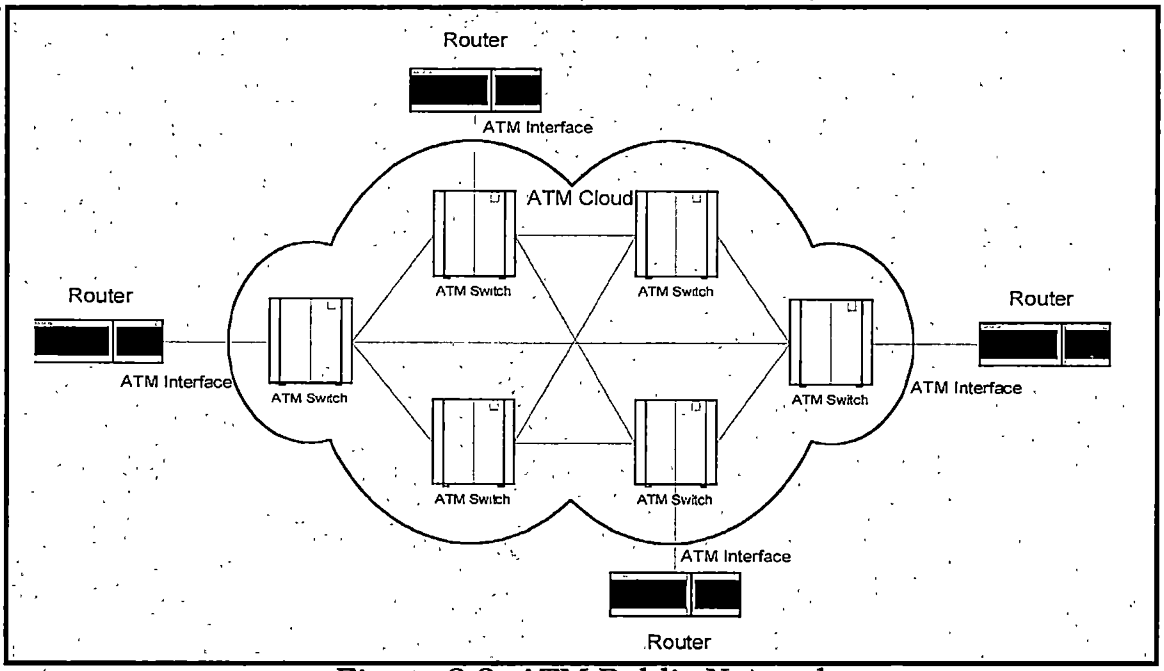

Figure 3.2 ATM Public Network

ATM virtual circuits are identified by the combination of two fields. Virtual Path Identifier (VPI) fields are concatenated with Virtual Channel Identifiers (VCI) to form a VC. Logically a virtual path is a bundle of virtual channels (Figure 3.3). It is important to keep in mind that each virtual circuit has only local significance across a particular connection. Therefore, each switch will re-map, as appropriate, to construct the virtual circuit across the ATM network. Often many virtual circuits are constructed over a physical link to partition the physical bandwidth. This allows disparate application requirements to be met by separate VCs over the same physical link.

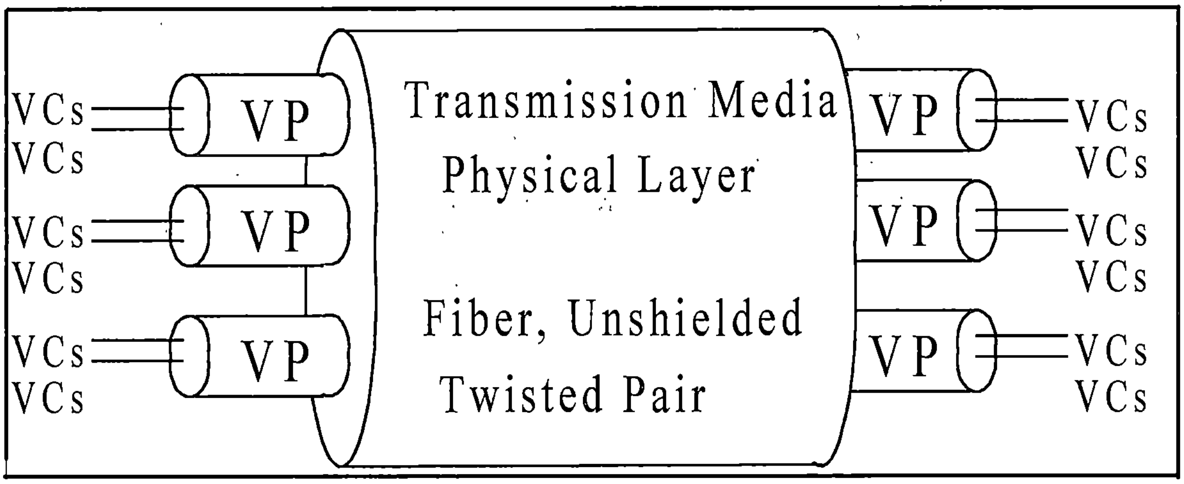

Figure 3.3 Virtual Circuits Comprised of VPI/VCI Identifiers

By using a fixed length cell, ATM has greater control over service-level parameters. Other technologies, such as frame relay, utilize variable length frames and cannot control variations in delay associated with transmitting interleaved variable sized frames. An example of this would be the when small real-time data packets and large file transfer packets need to be transmitted across the WAN simultaneously. Regardless of the queuing mechanism used, the small packets will experience variable delay in their transmission, due to the time required for the larger packets to be sent. ATM avoids this scenario by using fixed length cells. However, a conversion process must be implemented at the LAN/WAN boundary to accomplish this task.

Since ATM is used in the modified network design as a WAN technology, a standard is employed to convert Ethernet frames on the LAN to ATM

cells for transmission across the WAN. Although different adaptation techniques are defined, the most commonly available for data transmission, and the one used exclusively in this design is ATM Adaptation Layer 5 (AAL-5). The AAL-5 functionality is performed by the ATM interface card on the router.

One function of the adaptation layer is known as Segmentation and Reassemble (SAR). The primary function of the SAR sublayer is to format data received by higher layer protocols into 48-byte SAR Protocol Data Units (SAR-PDUs). The amount of overhead associated with the creation of the SAR-PDU is AAL type dependent. AAL-5 has the least amount of overhead and has therefore gained industry predominance. Data that has been prepared by the adaptation sublayer into SAR-PDUs will now fit into ATM cell payloads. ATM is the next layer in the protocol stack, and will append the ATM header prior to transmission (Figure 3.4). The ATM header has the control information necessary such as, VPI/VCI priority, and congestion information, all necessary to route the cell throughout the ATM network. After the 5-byte ATM header is appended, the total length of the ATM PDU is 53 bytes. Now in standard cell format, the 53-byte ATM cell is passed to the physical layer for transmission. This layer is media dependent, and will transmit the ATM cells based on the type of media and technology being used. There are many physical layer

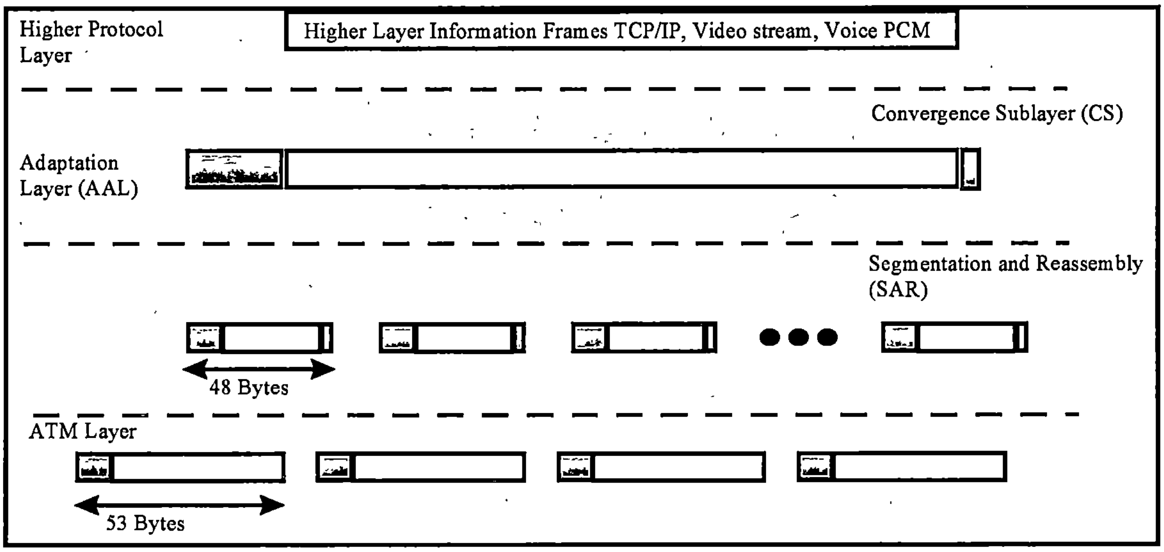

Figure 3.4 ATM Sublayers

technologies available including standard formats like DS-1, DS-3, OC-3c, and OC-12c.

To realize the benefits of ATM the process of SAR-ing (converting packets to cells) must be carried out. A term called *cell-tax* is often used to describe the tradeoff this operation produces. The price paid for the conversion process itself, and overhead of ATM VPI/VCI information carried in each cell, contributes to end-to-end latency. This added latency component has been evaluated in the result section of this thesis.

ATM edge devices signal the network with parameters indicating the type of connection required. If these connections are temporary, they are called Switched Virtual Circuits or (SVC). Signaling information is

carried out of band, on a separate VC. The current specification for SVC signaling between host devices and public ATM switches is UNI 4.0. If connections do not vary over time, they are referred to as permanent virtual circuits (PVC). PVCs do not require host devices to have an ATM address since all calls are set up statically and signaling is not required between the host and the switch. Service level parameters for the connection are specified during the PVC configuration. PVCs are used exclusively in the modified network design. Two reasons contribute to their use: Fist SVCs are not readily available today. Secondly, all WAN connections for the modified design are required to be static. Since connection requirements, from the facilities to the data centers, do not vary with time, PVCs are an ideal fit for the modified network.

Three types of service levels are commonly available from carriers. Constant Bit Rate (CBR) virtual circuits guarantee a sustainable cell rate and variable delay constant. CBR circuits are purchased for voice applications requiring the transport of voice pulse code modulated (PCM) data. Variable Bit Rate (VBR) virtual circuits are separated into real-time (RT) and non real-time (nRT). VBR-RT circuits specify cell delay variation, sustainable bit rate and maximum cell rate. VBR-nRT circuits are used in data applications, such as Meditech, which require latency characteristics to be satisfied. VBR-nRT specifies only sustainable and

maximum cell rates. These circuits are used to carry non-time sensitive data such as e-mail, and file transfer.

3.2 Logical Topology

In the modified design the Tier-1/Tier-2 concept is eliminated. Every facility has equal access to the data center. This topology removes the performance degradation experienced by Tier-2 facilities traversing a Tier-1 in reaching the regional data center. Leased line T-1s are replaced by T-1 User to Network Interface (UNI) connections to the public ATM network. The conversion from T-1 leased line to T-1 UNI does not require a hardware change on the router. The same serial interface module is used by the router for both technologies, since the physical transmission media remains a constant as a T-1. A software change on the router is required to support the UNI signaling and SAR functions required for ATM communications. Additionally, the service provider re-terminates the original T-1 circuit on a public ATM switch. Virtual circuits created over the physical T-1 UNIs connect facilities to the data center (Figure 3.5). Separate VCs are used to transport IP and Meditech back to the data center. At the data center two T-3s are used to aggregate all the facilitie's virtual circuits. T-3 connections are able to aggregate

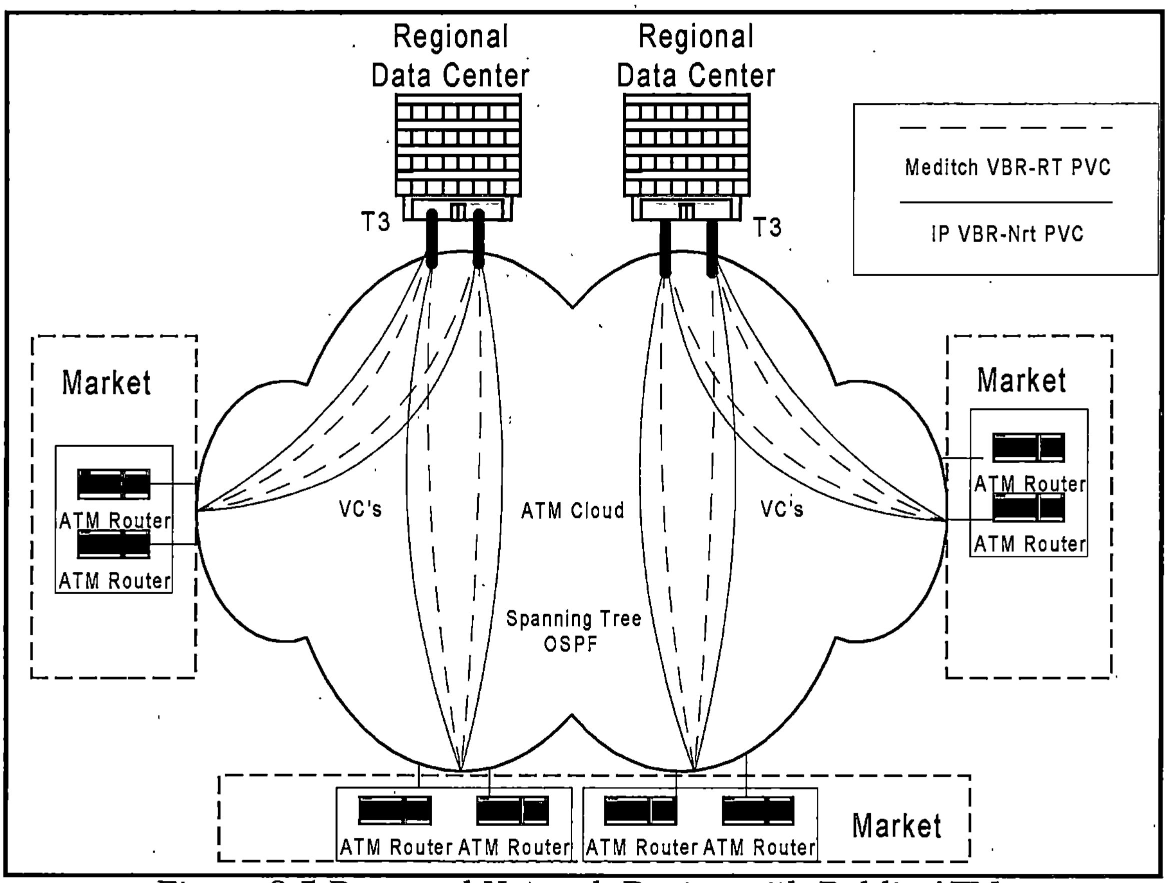

Figure 3.5 Proposed Network Design with Public ATM

bandwidth up to 45-Megabits/second. Two T-3s are utilized in each data center, allowing for the same level of redundancy used in the previously design with the simplification on the number of physical circuits required. Redundancy is achieved by distributing the four virtual circuits at each facility, one for IP and one for Meditech, to both T-3s at the data center evenly. Therefore, the primary Meditech and IP virtual circuits go to the primary T-3 at the data center, while the secondary virtual circuits connect to the redundant T-3.

VBR-RT PVCs are purchased for Meditech data, while VBR-nRT PVCs are used for IP data. By isolating application data to individual PVCs,

custom requirements can be assigned to varying application requirements. The Meditech PVC carries strict cell delay variation tolerance parameters along with a sizing of 512-Kilobits/second. On circuits configured to transport Meditech data cell delay variation is restricted to 25-milliseconds. Each ATM switch along the path of the virtual circuit guarantees that this requirement is satisfied. The IP PVC carries best-effort service characteristics that allow for fluctuations in network latencies. E-mail and file transfer applications, which are transported across these PVCs, do not require real-time delivery. Sizing on the IP PVC is 384-Kilobits/second, and since delay characteristics are not controlled, this PVC is less expensive. Another benefit from application separation is that any fluctuation in traffic levels from one application will not effect the operation of the other data.

3.3 Logical Topology Layers 2 and 3

The logical operation of Spanning Tree and IP routing protocols like OSPF and BGP are not affected by the conversion to ATM. Virtual circuit aggregation is performed at the data center routers allowing more connections from the facilities to share the same physical access at the data center. Although physical aggregation occurs with the use of virtual circuits, the same numbers of connections exist in comparison to the

original design. Since the number and location of the connections are not modified, logical operation of Spanning Tree and OSPF continue to execute transparently with the underlying physical media. With the hierarchical facility-to-data center design maintained, changes to OSPF and Spanning Tree configurations are not required.

3.4 LAN

The modified network design replaces the shared Ethernet hubs at the facility and at the Regional Data Centers with Ethernet switches. Ethernet switches follow the same standards-based protocol definitions for Ethernet and the underlying CSMA/CD algorithm; however switched Ethernet reduces the collision domain to two nodes: the switch-port and the endstation. Since each endstation is attached to its own segment, a common terminology for Ethernet switching is called Private Ethernet. Obviously, Ethernet switching allows for higher throughput by reducing the number of nodes on a shared segment to a single user. However, at periods of high utilization the collision effect mentioned previously still adversely affects throughput. To correct this situation, modification techniques were developed to the back-off algorithm. One technique called Priority Access Control Enabled (PACE™) took advantage of the two-node collision domain association of Ethernet switching.

3.5 PACE™

Modifications to the CSMA/CD back-off algorithm were constructed to solve problems associated with the latency and jitter caused by the capture effect. Several competing algorithms were developed. The two most popular are 3Com's PACE™ used exclusively in switched Ethernet environments, and Dr. Mart L. Molle's [11] Binary Logarithmic Arbitration Method for Ethernet (BLAM) used in both shared and switched environments. PACE™ was the only modification algorithm that changed only the network switch-port side of the communication link. By eliminating the requirement for endstation reconfiguration, yet remaining standards compliant, PACE™ gained market dominance. This can easily be explained since the administrative complexity in modifying existing endstations is just too impractical for network managers to implement.

A functional representation of PACE™ interactive access technology is shown in Figure 3.6. When the switch-port has data to send it first checks to see if it has received a packet from the endstation, or if the delay timer it operates has expired. If either of these two criterion are satisfied the switch-port attempts to transmit. If successful, the packet

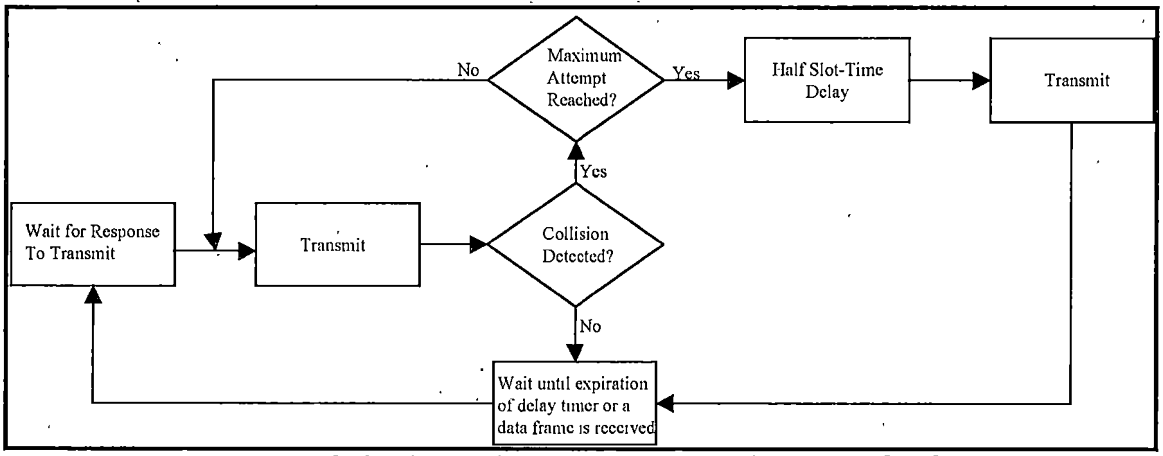

Figure 3.6 Flow Chart PACE™ Technology [12]

is transmitted to the endstation. However if a collision occurs, the switch-port attempts to retransmit immediately, bypassing the back-off algorithm. With each attempt the probability of successful transmission increases. This can easily be shown since the endstation is more likely to experience longer delays after each attempt by following the standard back-off algorithm. The switch-port will continue this behavior for a maximum of six attempts. In the unlikely event that the maximum number of attempts is reached, the switch-port will delay for half a slot-time. By waiting exactly half a slot-time, the switch-port allows just enough time for the endstation to transmit data. After receiving the data, the switch-port immediately transmits, guaranteeing a successful transmission.

Once the switch-port successfully transmits, the port enters wait-mode and starts a delay timer. The delay timer is set in proportion to the

number of collisions experienced in the previous stage. By doing so, the switch-port allows sufficient time for the endstation to come out of back-off and begin transmission. Once the switch-port receives the data from the endstation it clears the delay timer and is ready for the next transmission.

Regardless of how much data needs to be sent, the capture effect will never occur under this mode of operation. PACE™ allows the reduction of latency and jitter by assuring fair and equal network access to both the endstation and switch-port. The algorithm causes data to be interleaved more efficiently and is graphically represented in figure 3.7. Access latency is used to measure jitter on Ethernet networks. For Meditech and other real-time applications to run effectively, access latencies of less then 5-10 milliseconds are required. With PACE interactive access, the maximum access latency for six attempts is 3.23 milliseconds [12], well under the requirement.

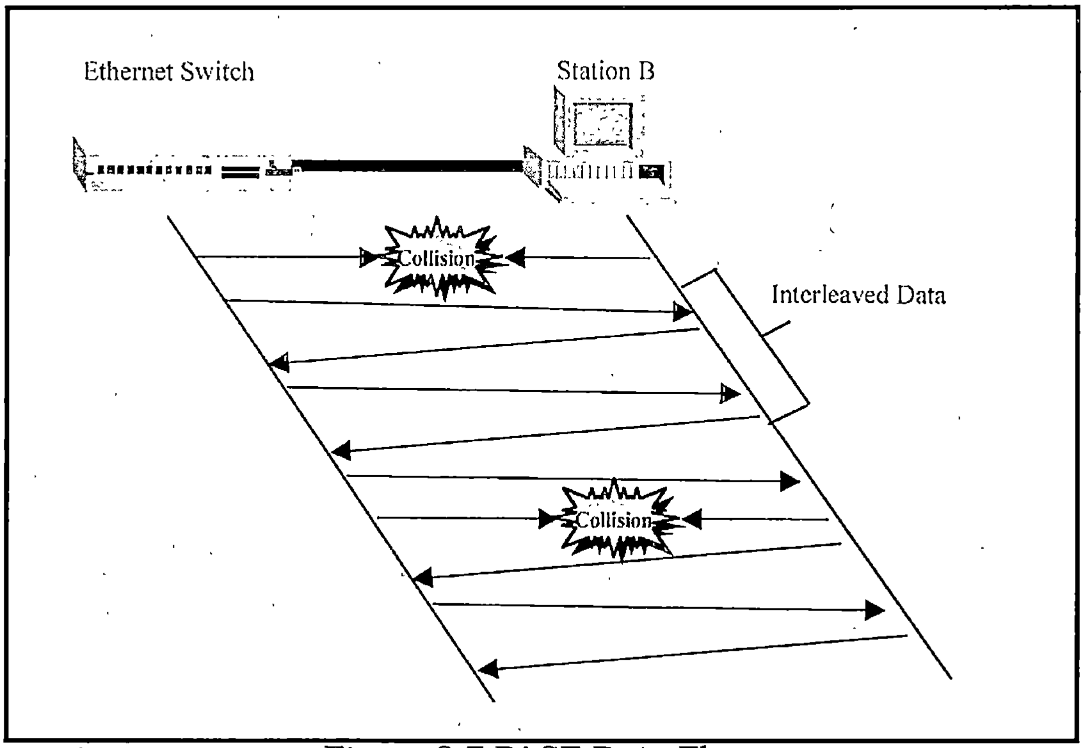

Figure 3.7 PACE Data Flow

4.0 Hardware Simulation

Simulation centered on the three components introduced by the proposed design. The first component analyzed was the ATM access equipment required to convert Ethernet data from the LAN to ATM cells for transport over the WAN. The remaining tests separated the proposed design into WAN and LAN elements. This allowed for a comparative evaluation of the new design with the respective legacy technology currently employed by the network. The WAN simulation compared the latency experienced by the Meditech protocol utilizing an ATM network with the latency of a leased line T-1 infrastructure. LAN simulation compared the latency experienced by the Meditech protocol utilizing the modified backoff algorithm (PACE ™) with traditional Ethernet switching. The last simulation evaluated the proposed design as a complete system with LAN and WAN elements against the legacy network.

4.1 Intra-Nodal Delay Introduced by ATM Access Equipment

The first test conducted determined the amount of intra-nodal delay introduced by the ATM access equipment utilized in the proposed design. The term "cell-tax", used in the previous chapter, describes the performance penalty introduced by adding the conversion process

necessary to translate Ethernet frames into ATM cells. In the proposed design a specialized piece of hardware performs the conversion process. This hardware, called the ATM access equipment, replaces the Channel Service Unit / Data Service Unit (CSU/DSU) used in the legacy design. Figure 4.1 indicates how this equipment interfaces with the existing router. In both cases the equipment connects the router to the service provider. The service provider terminates the ATM T-1 circuit in the ATM access equipment.

The inter-nodal delay introduced by the ATM access equipment is a function of the processing capability of the ATM access hardware. To measure this inter-nodal delay, or latency, a simulation was conducted illustrated by Figure 4.1.

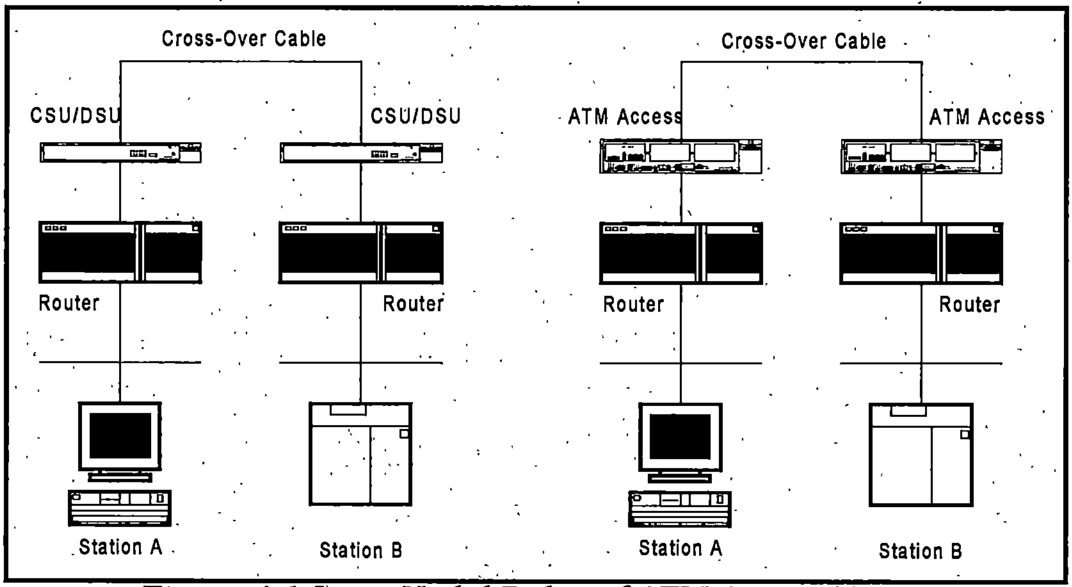

Figure 4.1 Inter-Nodal Delay of ATM Access Equipment

A 3-foot crossover T-1 cable was used to emulate a T-1 circuit that would normally be issued by the service provider. The crossover cable removed any latency that would normally be associated with the length of the T-1 circuit. The latency introduced by the circuit is directly proportional to the circuit length, and is simply cumulative for the overall system. Since each circuit in the network has a unique latency measure based on the circuit mileage, utilizing 3-foot crossover cable effectively removed circuit latency from the simulation.

Test equipment located on each LAN performed a Meditech echo test where the sender (Station A) would generate an echo packet to the receiver (Station B). The receiver would then reply to the echo, with an echo reply. The sender displays the round-trip latency of the system with an accuracy of ± 10 microseconds (μs). A 3-minute test was conducted on each configuration providing a sustainable 256Kbs of Meditech data traffic. A 3-minute test was used to eliminate any transient effects.

The legacy system yielded a total round-trip latency of 5.15 milliseconds (ms). The modified design produced a round-trip system delay of 9.17

ms. Therefore, the intra-nodal delay introduced by each ATM access device is 2.1 ms with an accuracy of ± 10 μs.

4.2 Latency of Meditech Protocol as a Function of Bandwidth Utilization

The focus of the second simulation was evaluating ATM as the wide area transport technology with a leased line. The test environment consisted of a simulated regional data center and a simulated hospital facility (Figure 4.2). At the regional data center a router was connected to two Local Area Networks. LAN-1 consisted of a traffic generator.

This device was used to control the amount of low priority "background" traffic injected during the simulation.

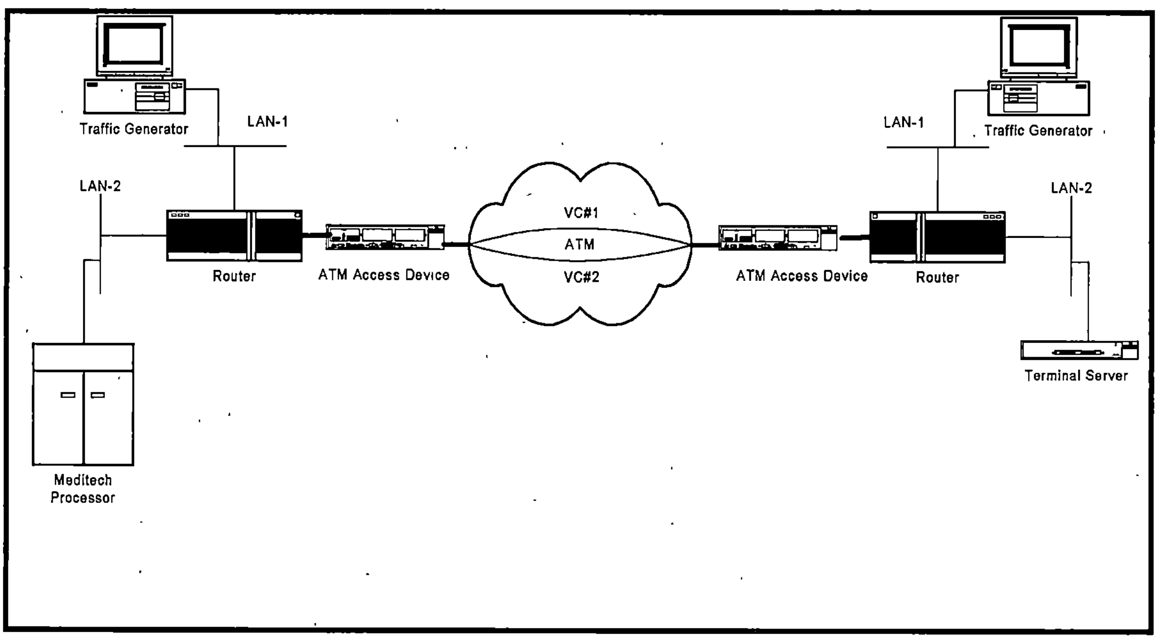

Figure 4.2 WAN Simulation

LAN-2 consisted of a Meditech processor. Two different LAN segments were utilized to eliminate interference between sources. Both devices source traffic through the same WAN interface on the router. The router was configured to interleave data in a round-robin fashion from both LAN ports onto the WAN interface. Although the WAN interface consisted of a single T-1 ATM User to Network Interface (UNI), two virtual circuits were created on the interface to carry the data streams to the hospital facility. Low priority or background traffic was configured to traverse VC-1, while high priority Meditech data used VC-2. VC-1 was configured as a 768Kbs Variable Bit Rate non Real Time (VBR-nRT) connection. VC-2 was configured as a 768Kbs Variable Bit Rate Real Time (VBR-RT) circuit. The VBR-RT circuit was provisioned for a maximum Cell Delay Variation (CDV) of 25 *ms*. The ATM network from the service provider was responsible for maintaining CDV. This guaranteed that each cell traveling across VC-2 would arrive at the destination within 25 *ms* of the previous cell.

At the hospital side the configuration was duplicated. A traffic generator was placed on LAN-1. A Meditech client, or terminal-server, was attached to LAN-2. An Ethernet testing utility was used to generate Meditech data between the processor and the client. The utility was run on the Meditech processor and allowed the construction of multiple

master-slave tasks. Each task emulated an individual interactive Meditech session across the WAN. The processor was always configured as the Master while the terminal-server became the slave. By controlling the number of tasks for each test the amount of WAN bandwidth utilization could vary. The results of each test produced values for the average round trip latency and the number of Meditech frames per second transmitted.

For testing which required IP loading a traffic generator was used to simulate "background" low priority IP traffic. The traffic generator was placed on the LAN segment in the simulated data center (see Figure 4.2). By controlling the number of frames per second transmitted the traffic generator simulated IP loading over the WAN as a percentage of total available bandwidth on the virtual circuit.

First Test :

For this test, the Meditech processor was used to generate varying amounts of Meditech uni-cast data over the WAN. The amount of data generated was related to how many master/slave tasks were run which varied from 20 to 120 per terminal-server. A graph of latency and frames/second vs. the number of tasks, can be viewed in Figure 4.3.

It is clear from Figure 4.3 that latency and frames/second have an inversely proportional relationship. A "knee" in the graph can be seen at approximately 80 tasks. At this point the performance of Meditech applications would begin to degrade rapidly due to the increase in latency experienced by each frame. An analysis quickly revels that using 64-Byte Meditech packets with a frame per second rate of 1500 completely fills the available bandwidth on the configured virtual circuit. The packet experiences additional delay when greater than 80 tasks are configured since packets must be queued in the buffers of the router when the virtual circuit has reached capacity. Buffering in the router will contribute to fewer lost packets at the expense of additional latency.

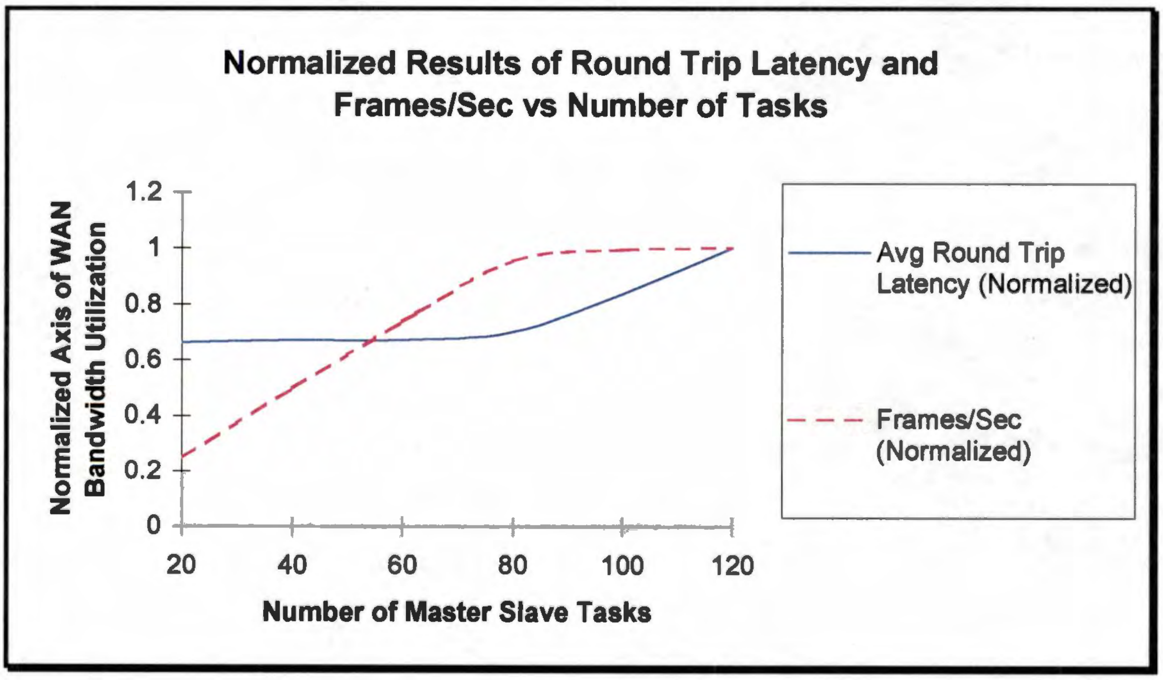

Figure 4.3 Normalized Round Tip Meditech Latency and Frames per Second as a function of Bandwidth

Although the packets may experience additional delay, as long as they reach the destination before the protocol time out, they will not be considered lost. In all simulations buffering on the routers remained constant.

Second Test:

For this test the Ethernet utility was used again to generate variable amounts of Meditech uni-cast data. In addition to the Meditech data, varying amounts of IP data was injected over the same WAN connection. The amount of IP or background data that was generated was incremented at fixed levels. Initially no IP loading was injected on the system, then IP loading was increased to 25%, 50%, and finally 80% of the maximum WAN bandwidth. Simulations were conducted three times for each level of IP loading. The three simulations were then averaged and plotted in Figure 4.4. For comparison, data from the current leased line environment with 80% IP load was included in Figure 4.4. Values lower than 80% loading for the leased line were omitted for clarity since they were identical with the ATM testing.

It is clear from Figure 4.4 that until 50% or more background IP loading occurs, results are consistent with the baseline measurements. Once

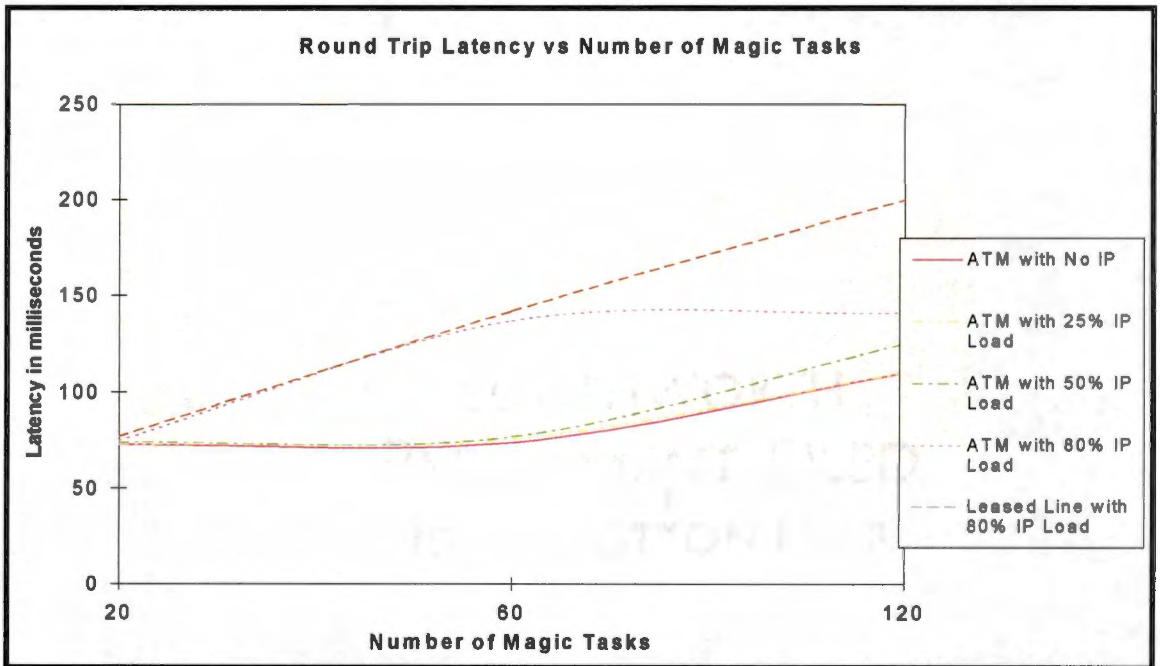

Figure 4.4 Round Trip Meditech Latency Subjected to Congestion with Low Priority IP Traffic

50% or more loading occurs, IP background data begins to interfere with Meditech. At this point the total aggregate data on the WAN reaches 100% of the maximum available bandwidth. Analysis reveals an interesting result. Under heavy loading of 80% or more the ATM design experiences less latency than a leased line. This can be attributed to the logical virtual circuit reserved in the ATM design designated to transport Meditech via VBR-RT. In the ATM design the ATM access equipment empties Meditech data onto VC-2 under a fixed delay threshold of 25 ms. This occurs at the expense of the IP background data which remains in the buffer longer. With the legacy T-1 model both data streams share the same logical pipe, and were interleaved evenly for transportation across

the WAN contributing to a uniform delay experienced by both protocols. This explains the linear response under heavy loading with a leased line.

Third Test:

For the third test the same simulation environment was utilized. This time instead of measuring latency through the WAN, frames per second of the Meditech protocol were measured. Again IP loading was introduced at the same levels of the second test. Leased line measurements were included for comparison at 80% loading. Figure 4.5 depicts the relationship between Meditech frames/second and the amount of IP load. Clearly, as the IP load increased, the number of Meditech frames per second decreased across each operating value. The linear response of a leased line under heavy background loading corresponds to the increase in latency evaluated in the second test. The ATM virtual circuit maintained a saturation level under heavy loading due to the VBR-RT virtual circuit configured.

4.3 LAN Simulation

The next simulation measured the effect that the modified back-off algorithm had on the performance of the LANs residing in the regional

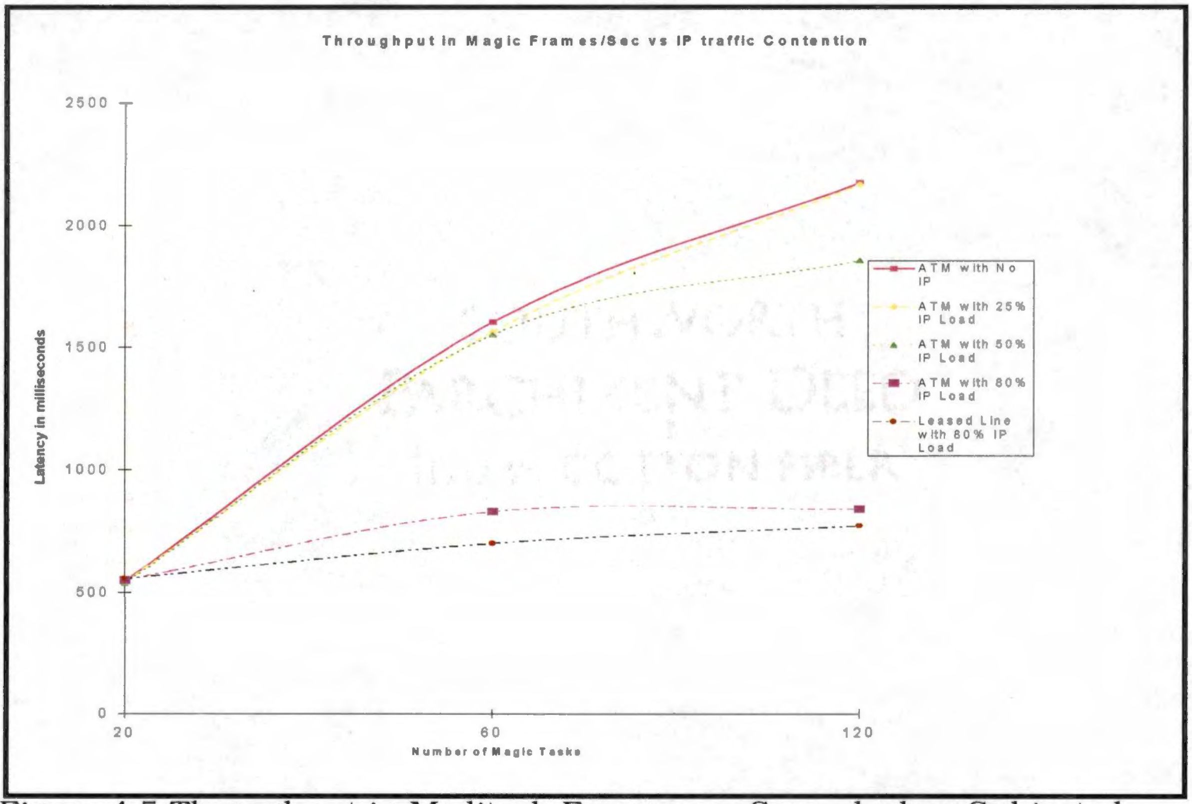

Figure 4.5 Throughput in Meditech Frames per Second when Subjected to Congestion with Low Priority IP Traffic

data centers and at each hospital facility. The simulation included all the hardware elements for the LANs in the proposed design (Figure 4.6). The data center LAN was comprised of a single Ethernet switch. Attached to the switch were a Meditech processor and a router. The data center router was connected to another router via an Ethernet crossover cable.

The hospital portion of the simulation consisted of a single Ethernet switch. Connected to the switch at the hospital were a router, four

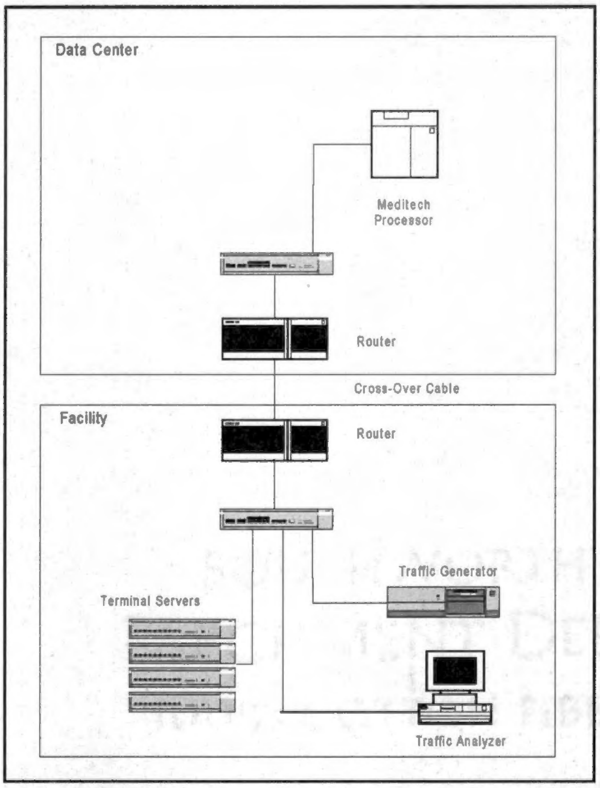

Figure 4.6 LAN Simulation Test Environment

terminal servers, a traffic analyzer, and a traffic generator.

LAN testing utilized the same Ethernet utility previously described in the WAN simulation to generate Meditech uni-cast traffic. Traffic generators and traffic analyzers were employed to generate IP background data and collect information.

The testing methodology for the LAN simulation included establishment

of a baseline. The baseline was calculated without the back-off modification applied to the Ethernet switches and without any additional background traffic generated. The Ethernet utility was run utilizing four terminal-severs on the hospital LAN as Meditech clients. Four terminal-servers were used to generate enough data capable of saturating the LAN. Each test was conducted five-times for a duration of two minutes per simulation. An average was calculated eliminating any transient effects experienced during the simulation process. After a baseline was established in each of the subsequent tests, the traffic generator was used to inject IP broadcast traffic onto the hospital LAN. Again in this simulation IP data represented non-Meditech background data such as e-mail or HTTP. By varying the amount of back-ground traffic as a percentage of maximum hospital LAN bandwidth measurements were conducted for different operating environments. A traffic analyzer was used to measure the amount of Meditech data frames transmitted and received on the hospital LAN. This measurement represented the aggregate throughput of Meditech data to and from the facility. This measurement was made to compare the amount of throughput achieved under varying loads with and without the modified back-off algorithm invoked.

Test 1 Baseline

From Table 4.3.1 the average number of packets transmitted and received from each of the four terminal-server clients was 143,467. This represented an average of 33% utilization of the LAN and established a baseline for the remaining simulations.

Tests 2-8 Modified Back-off Algorithm under Varying Loads

The remaining tests were conducted by using the traffic generator to inject varying amounts of background data simulating an actual network's traffic loading experienced on hospital LANs. The traffic generator injected IP frames with a MAC layer broadcast. Since these frames were broadcasts, each packet was forwarded to every switch-port. This meant that every terminal-server transmission had to contend for access to the Ethernet with every packet injected from the traffic generator.

Table 4.3.1 Baseline Measurement

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average	Baseline	
TS #1	35229	35984	36566	36105	35444	35866	Pace	Disabled
TS #2	36444	36485	36011	34498	36063	35900	% Util.	0
TS #3	34604	35883	35988	34527	36509	35502	F/S	0
TS #4	36923	36254	35804	36018	35996	36199	Size	0
Totals	143200	144606	144369	141148	144012	143467	Meditech %	33

There were six test points for background network utilization selected: 20%, 35%, 50%, 60%, 75%, 89% (Note: This does not include the bandwidth utilization of Meditech packets generated by the terminal-server clients). These six points were selected as the most common operational levels of the LAN.

Table 4.3.2 indicates that an average of 121,583 packets were successfully transmitted under a 20 % loading environment with PACE™ enabled.

Table 4.3.2 PACE™ Analysis with 20% Background Data Generated

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	30743	29725	30973	31153	28996	30318	PACE	Enabled
TS #2	29693	30579	30580	30700	30483	30407	% Util.	20
TS #3	29672	31121	30261	29396	30799	30250	F/S	170
TS #4	30748	30913	30138	30930	30310	30608	Size	1500
Totals	120856	122338	121952	122179	120588	121583	% Incre.	3.10%

Device	Test1	Test2	Test3	Test4	Test5	Average		
TS #1	29240	29139	28980	29006	29104	29094	PACE	Disabled
TS #2	29357	29438	29323	29126	29779	29405	% Util.	20
TS #3	29818	29503	29481	29519	30154	29695	F/S	170
TS #4	30258	30047	29953	29097	29302	29731	Size	1500
Totals	118673	118127	117737	116748	118339	117925		

As a comparison, with PACE™ disabled an average of only 118,339 packets were successfully transmitted. This represents a relative improvement, with PACE™ enabled, of 3.10%.

Table 4.3.3 indicates that an average of 110,110 packets were successful transmitted under a 35 % loading environment with PACE™ enabled. As a comparison with PACE™ disabled an average of only 100,813 packets were successfully transmitted. This represents a relative improvement in transmission with PACE™ enabled of 9.22%.

Table 4.3.3 PACE™ Analysis with 35% Background Data Generated

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	26097	27898	26957	26816	26875	26929	PACE	Enabled
TS #2	28756	27911	26576	27831	28421	27899	% Util.	35
TS #3	27837	26401	28194	28222	28097	27750	F/S	295
TS #4	27796	27579	28069	26985	27230	27532	Size	1500
Totals	110486	109789	109796	109854	110623	110110	% Incre.	9.22%

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	25061	25344	25333	24987	24641	25073	PACE	Disabled
TS #2	24973	25700	25545	25279	25244	25348	% Util.	35
TS #3	25424	25353	24722	25699	24918	25223	F/S	295
TS #4	25399	25386	25000	25065	24993	25169	Size	1500
Totals	100857	101783	100600	101030	99796	100813		

Increasing the background utilization even further, Table 4.3.4 indicates that an average of 100,312 packets were successfully transmitted under a 50 % loading environment with PACE™ enabled. As a comparison with PACE™ disabled an average of only 84,005 packets were successfully transmitted. This represents a relative improvement in transmission with PACE™ enabled of 19.41%.

Table 4.3.5 data indicates that an average of 95,951 packets were successfully transmitted under a 60 % loading environment with PACE™

Table 4.3.4 PACE™ Analysis with 50% Background Data Generated

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	24947	25402	25783	24490	24875	25099	PACE	Enabled
TS #2	26195	25009	24829	24744	25647	25285	% Util.	50
TS #3	24558	24134	25038	25596	25910	25047	F/S	420
TS #4	24721	25141	25561	24826	24156	24881	Size	1500
Totals	100421	99686	101211	99656	100588	100312	% Incre.	19.41%

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	20976	21198	20849	20775	20974	20954	PACE	Disabled
TS #2	20613	20931	21412	21530	21104	21118	% Util.	50
TS #3	21207	21311	21044	20669	21269	21100	F/S	420
TS #4	20473	21098	20475	20863	21255	20833	Size	1500
Totals	83269	84538	83780	83837	84602	84005		

Table 4.3.5 PACE™ Analysis with 60% Background Data Generated

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	24528	23343	22621	23152	23730	23475	PACE	Enabled
TS #2	24542	24466	23693	24310	24136	24229	% Util.	60
TS #3	23526	24755	24101	24109	24719	24242	F/S	500
TS #4	22814	23696	24100	24033	23579	23644	Size	1500
Totals	95410	96260	94515	95604	96164	95591	% Incre.	31.35%

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	18022	17858	18178	17989	17705	17950	PACE	Disabled
TS #2	18724	18488	18559	18283	18422	18495	% Util.	60
TS #3	17996	18317	18024	18225	17972	18107	F/S	500
TS #4	18192	18226	18248	18076	18363	18221	Size	1500
Totals	72934	72889	73009	72573	72462	72773		

enabled. As a comparison with PACE™ disabled an average of only 72,773 packets were successfully transmitted. This represents a relative improvement in transmission with PACE™ enabled of 31.35%.

Table 4.3.6 indicates that an average of 74,583 packets were successfully transmitted under a 75 % loading environment with PACE™ enabled. As a comparison with PACE™ disabled an average of only 54,803 packets were successfully transmitted. This represents a relative improvement in transmission with PACE™ enabled of 36.09%.

Table 4.3.6 PACE™ Analysis with 75% Background Data Generated

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	18445	18013	18297	18308	18831	18379	PACE	Enabled
TS #2	19551	19183	18565	17890	19372	18912	% Util.	75
TS #3	19581	17640	18451	18721	19303	18739	F/S	625
TS #4	18481	18824	18341	18665	18451	18552	Size	1500
Totals	76058	73660	73654	73584	75957	74583	% Incre.	36.09%

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	13313	13202	13597	13999	13170	13456	PACE	Disabled
TS #2	13735	13823	13567	14831	13829	13957	% Util.	75
TS #3	13969	13542	13520	14208	13245	13697	F/S	625
TS #4	13327	13397	13695	14543	13504	13693	Size	1500
Totals	54344	53964	54379	57581	53748	54803		

To complete LAN testing one last data point was observed. Table 4.3.7 indicates that an average of 55,266 packets were successfully transmitted under a 89 % loading environment with PACE™ enabled. As a comparison with PACE™ disabled an average of only 41,529 packets were successfully transmitted. This represents a relative improvement in transmission with PACE™ enabled of 33.08%.

From each of the LAN simulations it is clear that enabling PACE™ under any loading condition improves the overall throughput performance of

Table 4.3.7 PACE™ Analysis with 89% Background Data Generated

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	13908	13678	13699	13797	13695	13755	PACE	Enabled
TS #2	13728	13873	13921	14179	14021	13944	% Util.	89
TS #3	13661	13554	13673	14111	13799	13760	F/S	738
TS #4	13513	13477	13839	13877	14325	13806	Size	1500
Totals	54810	54582	55132	55964	55840	55266	% Incre.	33.08%

Device	Test 1	Test 2	Test 3	Test 4	Test 5	Average		
TS #1	10377	10257	10840	10015	10174	10333	PACE	Disabled
TS #2	10554	10548	10874	10416	10143	10507	% Util.	89
TS #3	10472	10230	10409	9885	10130	10225	F/S	738
TS #4	10498	10401	10804	10107	10513	10465	Size	1500
Totals	41901	41436	42927	40423	40960	41529		

the LAN. The greatest relative improvement was measured while 75% of background data was injected onto the LAN. At this point maximum interleaving is achieved between Meditech and non-Meditech data. Periods below this level simply do not contribute enough data for the algorithm to be completely utilized, while levels above 75% saturate the theoretical capacity of the system (Figure 4.7). When saturation of the LAN occurs the switch must buffer packets in a First-In-First-Out (FIFO) queue. The depth of the queue and the length of the saturation determines how many packets the switch will discard. For each test

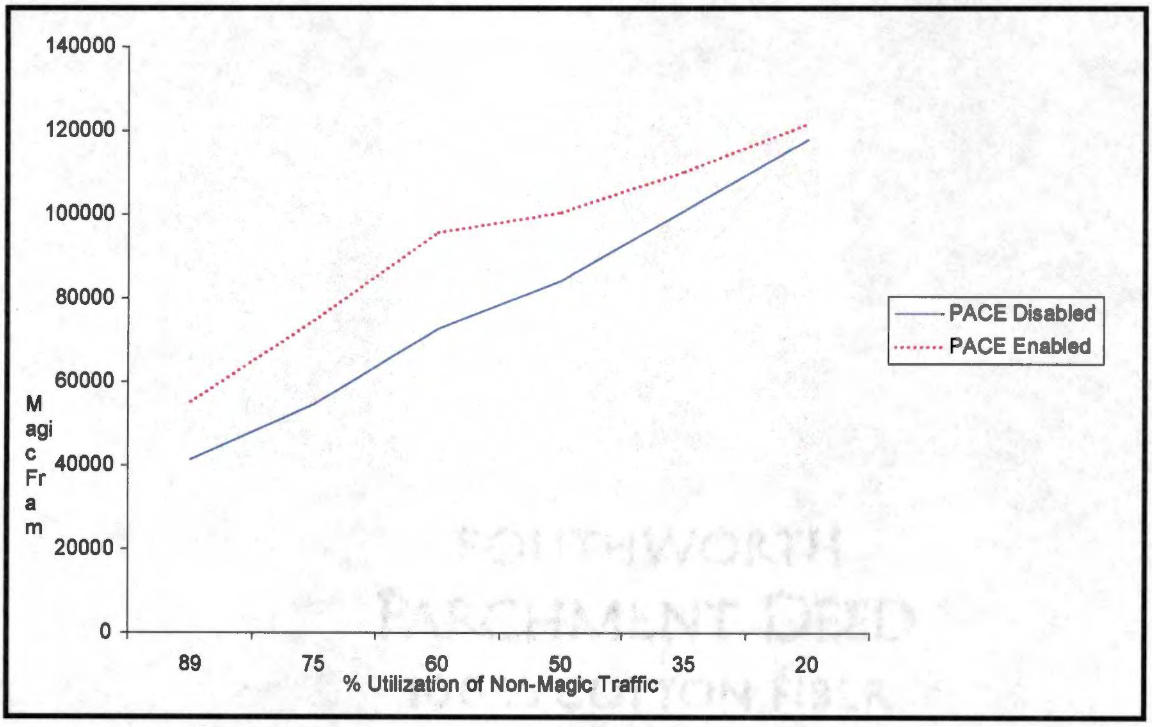

Figure 4.7 Summarization of LAN simulation

switch buffering remained constant. Each hospital facility and data center operates well under the saturation point of the LAN.

4.4 System Testing

A final test was conducted simulating both the WAN and the LAN components of the proposed design operating simultaneously forming the complete system. The WAN simulation in Figure 4.1 was run again, this time with PACE™ enabled on the Ethernet switches in the data center and the hospital facility. An operating value of 80% non-Meditech data was chosen. This level represented a realistic value observed frequently

on the legacy system, and previous testing indicated this level was well suited for comparisons.

Conducting multiple tasks generated various levels of Meditech uni-cast data and represented a complete operating spectrum. The results presented in Figure 4.8 indicate system improvement above both the current leased line design and the ATM only model. Consistently throughout each level of operation the proposed design with ATM and PACE™ enabled produced a greater value of frames/second through the network.

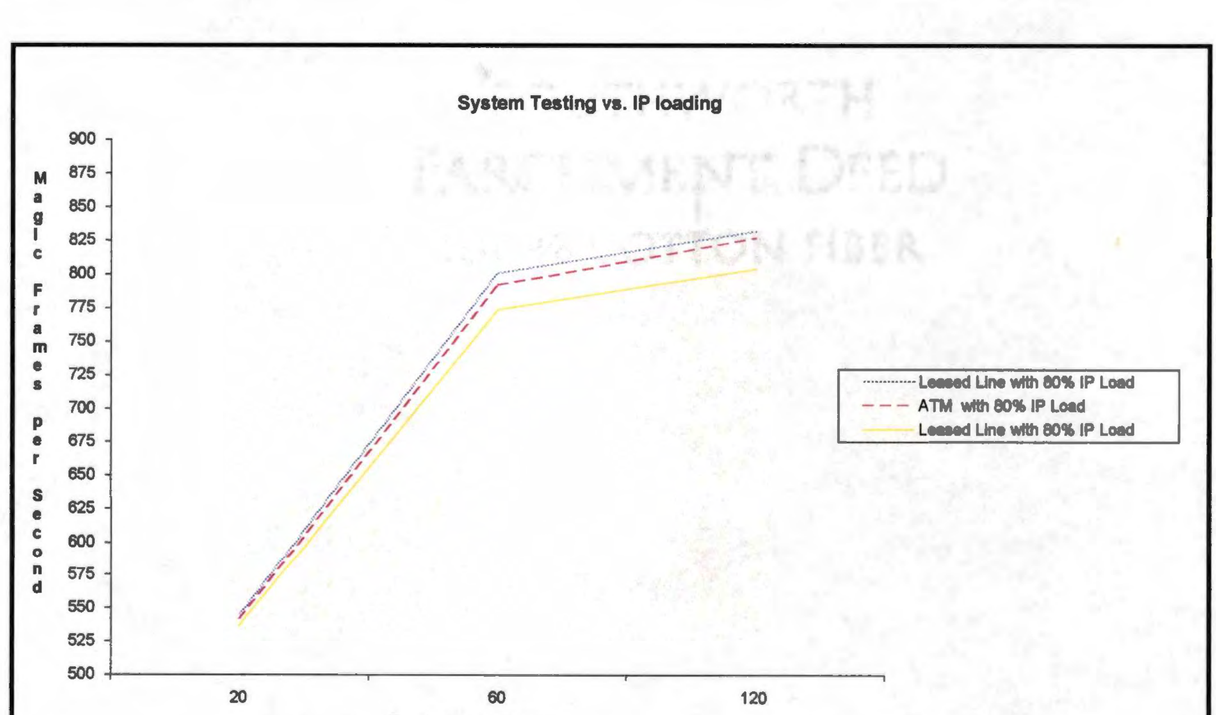

Figure 4.8 System Testing both LAN & WAN of the Proposed Design

5.0 Conclusion

This thesis explored the effects associated with the proposed design of a corporate data network supporting real-time data applications. Specifically, two enhancements were proposed in the design. The first enhancement converted a Wide Area Network from leased lines to ATM. The second enhancement focused on a modification to the Ethernet back-off algorithm present on the LAN. This chapter will discuss the results presented in the previous chapter illustrating the effects these two modifications had on the network. Results were measured as a function of performance in a real-time healthcare data application called Meditech. Finally, this chapter will offer insight into emerging technologies that may complement the proposed design in the future.

5.1 Enhancements in the WAN

The first enhancement instituted ATM on the WAN. ATM was chosen to allow flexibility in design and control over bandwidth allocation to specific applications. Logically separating the bandwidth through ATM required instituting VBR-RT PVCs for real-time Meditech data, and VBR-nRT PVCs for the remaining background traffic. The sizing of the PVCs,

512Kbs for Meditech and 384Kbs for the remaining background data, were determined through an analysis of current circuit utilization. Background traffic was limited to 384Kbs so that it could not monopolize network bandwidth and cause interference with Meditech. The real benefit in using ATM however, lay in the ability to assign a 25 *ms* cell delay variance to the VBR-RT PVC. This ensured that every successive cell entering the WAN would transverse the ATM network with a maximum of 25 *ms* of variance. By instituting such tight constraints on the PVC, Meditech was transported across the WAN successfully within the protocol timeout window and without congestion due to excess background traffic. Since leased lines do not provide a similar mechanism, this guarantee was not present in the existing environment.

The results of the previous chapter indicated that through each simulation the Meditech protocol operated more efficiently in the ATM environment. The most noticeable effects were seen when ATM was compared to a leased line subjected to a background utilization of 80% of the theoretical maximum bandwidth of the T1. Testing conducted on latency demonstrated that the average latency of each Meditech packet was reduced from 142 ms using a leased lined to 137ms with ATM, while Meditech frames/second were increased from 800 to 830. The results are consistent with the original hypothesis. Logically separating and

controlling the bandwidth via ATM resulted in increased performance and consistency of the real-time data protocol Meditech. Even under heavy loading conditions Meditech's performance on the network was increased through the adoption of ATM.

The conversion to ATM does come with a price. Each facility must introduce ATM access equipment in between the router and the circuit provider. This equipment was tested in the previous chapter to understand what effects it would have on the system. The application of converting Ethernet frames into ATM cells, called segmentation and re-assembly, performed by this hardware introduced an additional delay of 2ms. The simulation process used factored in this delay in all of the results presented. Even with this delay included a net increase in Meditech's performance occurred. Clearly, the outcome of utilizing ATM justified the penalty associated with the conversion process.

5.2 Enhancements in the LAN

The second enhancement occurred on the LAN. Since technology improvements and economics have made Ethernet switching widely available, a simple modification to the original Ethernet back-off algorithm produced considerable performance improvements on the LAN.

The PACE™ algorithm was chosen over several competing schemes since it did not require additional modification to standard compliant endstations deployed in the network. PACE™ modified the backoff algorithm on the switch side only and allowed traffic under heavy utilization periods to be interleaved between the switch and the endstation more efficiently.

Through the testing presented in the previous chapter, performance increases were observed over the entire operational environment ranging from 20-89% background loading. The algorithm presented the largest increase in performance, a 36.09% increase in frames/second during 75% background utilization. At this point the bi-directional communication of Meditech data and background data between the endstation and the switch-port reach maximum efficiency. Below this level the amount of data present is insufficient to maximize the communication. Above this point the amount of data is too much and begins to be queued by the switch-port for transmission. However, it was determined that over any level of operation, PACE™ enabled switches produced a positive effect on Meditech performance on the LAN.

5.3 Overall System Improvements

Performance improvements realized individually in the LAN and WAN were tested together as a system. As anticipated a performance increase was observed across the entire operational spectrum. The most interesting results are obtained when compared to a leased line with 80% background utilization. At this point the overall system increased the Meditech throughput from 756 to 812 frames per second. This operational level of 80% background traffic also represented the average operational level of the system.

From the results obtained through each simulation a definitive statement can be made.. Namely, the implementation of the proposed design produces improvements in performance of the network supporting the real-time data application Meditech throughout the operational spectrum of the network. With this conclusion reached the final issues to explore are the system implementation requirements, and the future migration possibilities which exist for the system.

5.4 Migration to Future Technologies

This section explores how new evolving technologies may complement the design in the future. Although beyond the scope of this work, two

evolving standards are making their way into the commercial market. A new IEEE standard 802.1p allows for prioritization of Ethernet frames on the LAN. This standard is dependent upon 802.1Q which appends a Virtual LAN (VLAN) tag onto standard Ethernet frames. VLAN's logically associate endstations into a bridge domain. Each station in a VLAN communicates at the Data Link Layer. Inter-VLAN traffic is accomplished through a router. By grouping Meditech endstations on the LAN into a single VLAN, other protocol traffic confined to the Data Link Layer will not interfere with Meditech traffic on the LAN. Furthermore, by implementing 802.1p capable switches, Meditech Ethernet frames may be given a higher priority tag. The priority tag, which is interpreted by the switch, allows Meditech packets to be transmitted out congested switch-ports based on the vendor queue algorithm used. This Class of Service (CoS) technique is used on the LAN to interrogate every packet and ensure each packet is transmitted in accordance to the priority level assigned.

In order to implement this technology future software modifications to both the switches and endstations would be required. The endstations would have the responsibility for tagging each packet a priority level while the switch-port forwards packets based on the priority tag.

Another evolving standard called Resource reSerVation Protocol (RSVP) provides a quality of service to an IP associated flow. This protocol could be added to each router in the network through a software upgrade in the future to support IP applications such as Voice Over IP (VoIP) and desktop video conferencing. Routers implementing the RSVP algorithm detect delay sensitive flows through messaging frames and allocate bandwidth from the ingress router to the egress router across the network to support the flow. This technique is similar to ATM except that instead of implementing the bandwidth reservation at the Data Link Layer which ATM provides, the reservation is made at the Network Layer for IP. Since both layers are mandatory for this network, implementing a reservation scheme at both layers give the most predictable results.

To implement this technique software modification to all routers in the network would be required. Each router would be responsible for identifying and reserving resources to process each flow. Additional memory may also be required based on the large number of simultaneous connections of flows present in the system.

These techniques and others could be added to the proposed design in the future. This ensures a stable migration path accommodating new technology and increasing existing performance.

REFERENCES

[1] Http://www.meditech.com

[2] J. Shoch, Y. Dalal, and D. Redell, Evolution of the Ethernet Local Computer Network, Computer, Aug. 1982, pp.10

[3] D. Comer, Internetworking with TCP/IP Vol. 1: Principle, Protocols, and Architecture, Prentice-Hall (1991). (Second Edition)

[4] A. Tanenbaum, Computer Networks

[5] International Telegraph and Telephone Consultative Committee (CCITT), Recommendation X.200, Reference Model of Open Systems Interconnection for CCITT Applications, Fascicle VIII.5-Rec. X.200 (1984), p.8

[6] RFC 1771 A Border Gateway Protocol 4 (BGP-4), 1995.

[7] "Carrier Sense Multiple Access with Collision Detection (CSMA/CD)," IEEE Std 802.3-1990 Edition (ISO/DIS 8802-3), IEEE, New York (1990).

[8] K. Chang, PACE Technology: Making Multimedia and Real-Time Networks Possible Today. 3Com Corporation Technology White Paper, 1996

[8] "Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges ," IEEE Std 802.1D-1990 Edition (ISO/IEC 10038), IEEE, New York (1990).

[9] Mark A. Miller, P.E, Analyzing Broadband Networks Frame Relay, ATM, SMDS, 1994.

[10] Raif O. Onvural, Asynchronous Transfer Mode Networks: Performance Issues, Artech House, Inc (1994).

[11] Mart. L. Molle, A New Binary Logarithmic Arbitration Method for Ethernet, Technical Report CSRI-298, University of Toronto, April 1994.

[12] K. Chang, PACE Technology: Making Multimedia and Real-Time Networks Possible Today. 3Com Corporation Technology White Paper, 1996.

VITA

Craig Longo was born in Harrisburg Pennsylvania on March 28th 1971. He entered the Pennsylvania State University during June of 1989 where in January 1993 he received a Bachelor of Science in Electrical Engineering. While working at the Oak Ridge National Laboratory, in the Telecommunications department, he entered the Master's program in Electrical Engineering at the University of Tennessee, Knoxville. After completing the required course work he joined 3Com Corporation where he was responsible for the design and architecture of the network proposed in this thesis. After completing the design he moved on to join and become one of the original members of Neteffect Corporation. While employed at Neteffect he completed his thesis and received a Masters of Science degree in Electrical Engineering from the University of Tennessee, Knoxville.

He is presently working with several Fortune 500 companies, as a managing consultant for Neteffect Corporation in Atlanta GA.