2007

Enterprise network convergence: path to cost optimization

Vidhumana Sridharan

Rochester Institute of Technology

College of Applied Science and Technology

Enterprise Network Convergence: Path to Cost Optimization

By

Vidhumana Sridharan

Thesis submitted in partial fulfillment of the requirements for the degree of **Master of Science** in **Telecommunications Engineering Technology.**

May 2007

Vidhumana Sridharan Thesis approved by:

Professor Ronald G. Fulle (Chairman of the Thesis Committee)

Dr. Warren L. G. Koontz (TET Program Chair)

Dr. Anthony Trippe

Date: _____

Acknowledgements

I would like to thank my parents, Mr. Sridharan and Mrs. Anuradha, for their love and support. Without their help this degree would not have been possible.

I am deeply indebted to my thesis advisor Professor Ronald Fulle. His wide knowledge in the field of telecommunications has been of great value for me. His detailed constructive comments, stimulating suggestions, diligent support, encouragement and guidance have helped me in all the time of research for and writing of this thesis.

I would like to extend my sincere gratitude to Dr. Warren Koontz, Chair of Telecommunications Engineering Technology program, who made the BS/MS degree a dream that came true. I am deeply appreciative of all the support and assistance that he has provided ever since I started my degree at R.I.T.

I would like to express my warm and sincere thanks to Dr. Chance Glenn and Professor Mark Indelicato who have been a constant source of encouragement. Working with them gave me an opportunity to learn from their experience and gain a better understanding of various fields in the telecommunications industry.

Lastly, I am grateful to all the faculty and staff of the ECT ET department for providing me with resources and support to complete my thesis. I would like to express my appreciation to the staff at Wallace library for their assistance during my thesis research – especially Ms. Patricia Mull.

Purpose

During the past two decades, telecommunications has evolved a great deal. In the eighties, people were using television, radio and telephone as their communication systems. Eventually, the introduction of the Internet and the WWW immensely transformed the telecommunications industry. This internet revolution brought about a huge change in the way businesses communicated and operated. Enterprise networks now had an increasing demand for more bandwidth as they started to embrace newer technologies. The requirements of the enterprise networks grew as the applications and services that were used in the network expanded. This stipulation for fast and high performance communication systems has now led to the emergence of converged network solutions.

Enterprises across the globe are investigating new ways to implement voice, video, and data over a single network for various reasons – to optimize network costs, to restructure their communication system, to extend next generation networking abilities, or to bridge the gap between their corporate network and the existing technological progress. To date, organizations had multiple network services to support a range of communication needs. Investing in this type of multiple communication infrastructures limits the networks ability to provide resourceful bandwidth optimization services throughout the system. Thus, as the requirements for the corporate networks to handle dynamic traffic grow day by day, the need for a more effective and efficient network arises. A converged network is the solution for enterprises aspiring to employ advanced applications and innovative services.

This thesis will emphasize the importance of converging network infrastructure and prove that it leads to cost savings. It discusses the characteristics, architecture, and relevant protocols of the voice, data and video traffic over both traditional infrastructure and converged architecture. While IP-based networks present excellent quality for non real-time data networking, the network by itself is not capable of providing reliable, quality and secure services for real-time traffic. In order for IP networks to perform reliable and timely transmission of real-time data, additional mechanisms to reduce delay, jitter and packet loss are required. Therefore, this thesis will also discuss the important

mechanisms for running real-time traffic like voice and video over an IP network. Lastly, it will also provide an example of an enterprise network specifications (voice, video and data), and present an in depth cost analysis of a typical network vs. a converged network to prove that converged infrastructures provide significant savings.

Table of Contents

List of Tables

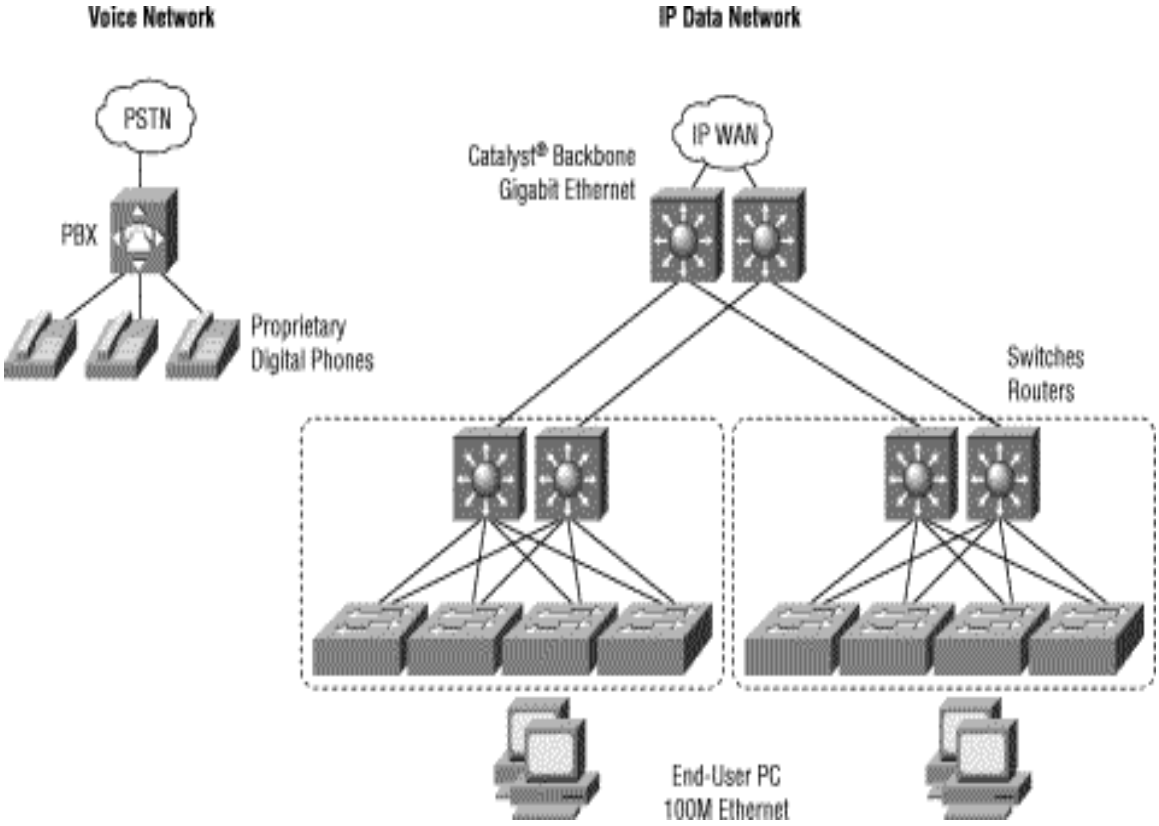
List of Figures

1 IP Convergence

1.1 Introduction

Enterprises across the globe are investigating the implementation of voice, video, and data over a single network for various reasons – to optimize network costs, to restructure their communication system to extend next generation networking abilities, or to bridge the gap between their corporate network and the current technological progress. To date, organizations had multiple network services to support a range of communication needs. Figure 1.1 illustrates the network infrastructure that has separate voice and data communications systems.

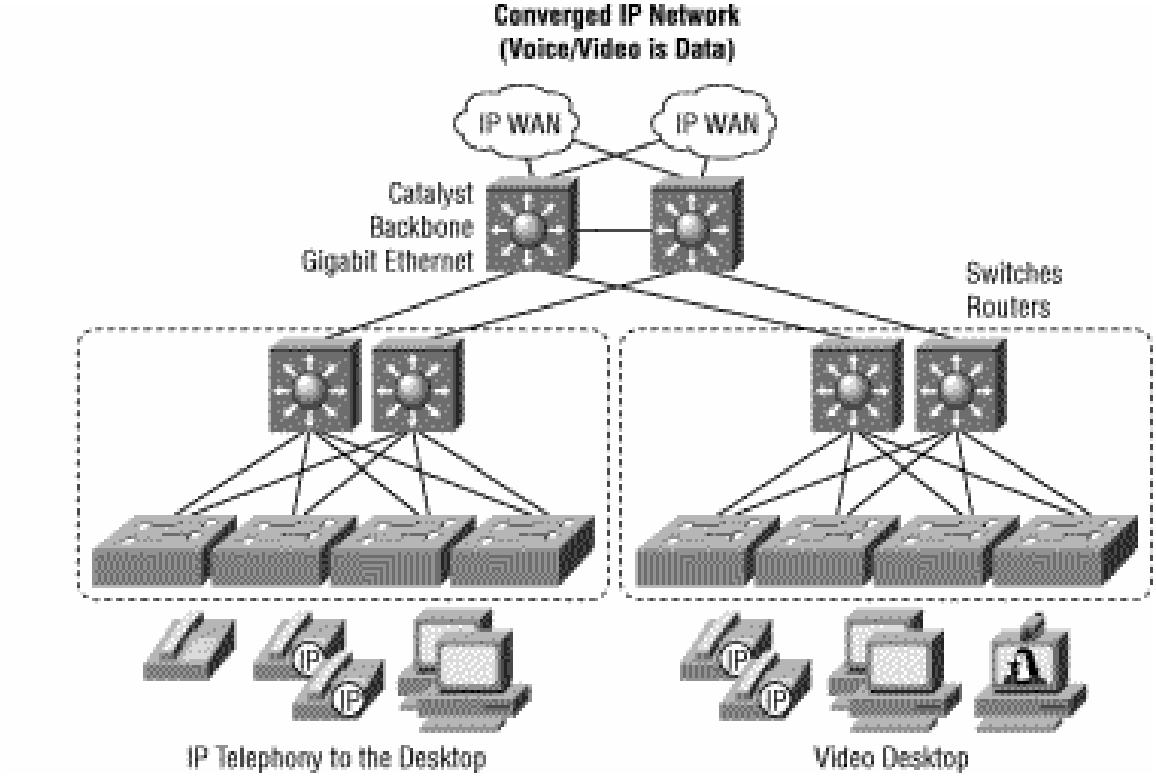
Figure 1.1 – Separate Voice and Data Network Infrastructure[1]



[1] *Cisco – Technical Considerations for Voice, Data and Video Networks.* Cisco Systems. http://www.itworld.com/WhitePapers/Cisco_AVVID_TechCon/

As the requirements for the network to handle dynamic traffic grows day by day, the need for a more effective and efficient network arises. A converged network is a feasible solution for enterprises aspiring to employ advanced applications and innovative services, while minimizing costs. This will be discussed in detail in sections to come. Figure 1.2 displays the layout of a typical converged network.

Figure 1.2 – Converged Network Layout[2]



1.2 Definition

Network convergence is defined as the integration of all traffic types – voice, data and video solutions – onto a single IP network[3]. Multiple data and traffic types are integrated to coexist seamlessly. A converged network must sufficiently handle different

[2] *Cisco – Technical Considerations for Voice, Data and Video Networks.* Cisco Systems. http://www.itworld.com/WhitePapers/Cisco_AVVID_TechCon/

[3] *The Coming of True Convergence: Why Service Providers Can Finally Turn Out the Lights on the Old Public Switched Telephone Network (PSTN).* International Engineering Consortium. http://www.iec.org/online/tutorials/true_converge/

traffic types and deliver consistent quality and reliability for the end-user. That is, both voice and data packets that are routed through the network should have comparable performance schemes. The voice packets must have less delay, and enhanced QoS, while the data packets must have high reliability and performance ratings throughout the network.

Convergence deployments have generated distinguished results in areas such as quality of service, bandwidth management, stability of vendor solutions, and return on investment (ROI)[4]. All of these factors put together make a converged network the appropriate base model during the network design process. Therefore, while designing a network decision makers should avoid getting fixed into a one dimensional solution that might not be able to provide and accommodate all application traffic simultaneously.

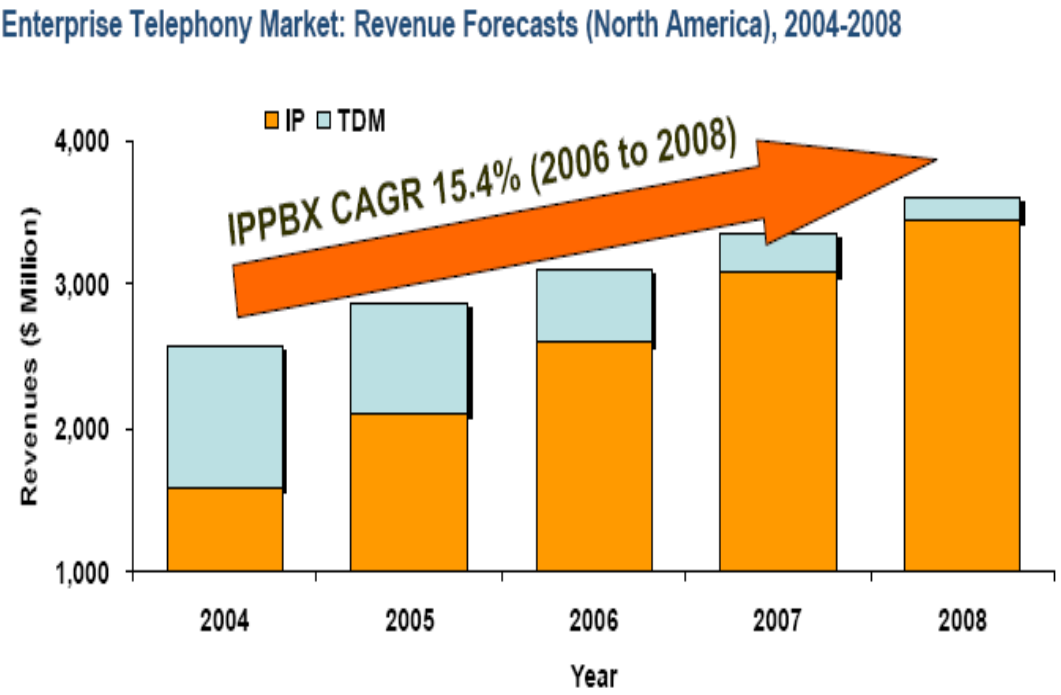
1.3 Significance

With adaptive network architecture, enterprises can successfully create a powerful converged network solution that efficiently handles several traffic types. With the introduction of advanced telecommunications technologies has come the great potential of having instantaneous access to business information from any part of the world in a synchronous fashion, and the urge to optimize cost and network resources. Enterprise networks are all about secure data communication, quality voice services and reliable video streaming capabilities. These three technologies, so distinct from one another, add their own value and have their own significance in the performance of any given telecommunications network. Independent infrastructure has forced them to remain as discrete methods of information-sharing. All information flowing through a corporate network can be categorized in one of the three forms: voice, data or video. Unifying these three important modes of communication media from separate network infrastructures onto a single IP-based network has the power to change how corporations operate, both internally and externally. IP convergence is moving into the corporate mainstream. By 2010, according to the Gartner Group, voice and data convergence based on IP telephony

[4] *Convergence: Preparing the Enterprise Network.* ProCurve Networking by HP. http://www.hp.com/rnd/pdfs/convergence_WP_june05.pdf

and VoIP will be under way in more than 95 percent of major companies[5]. Figure 1.3 shows how IP PBXs are being implemented in North America and how swift the revenue growth is for the enterprise telephony market. According to Frost & Sullivan, revenues are expected to grow from $2.87 billion in 2005, to $3.8 billion in 2008.

Figure 1.3 – Enterprise Telephony Market: Revenue Forecast[6]

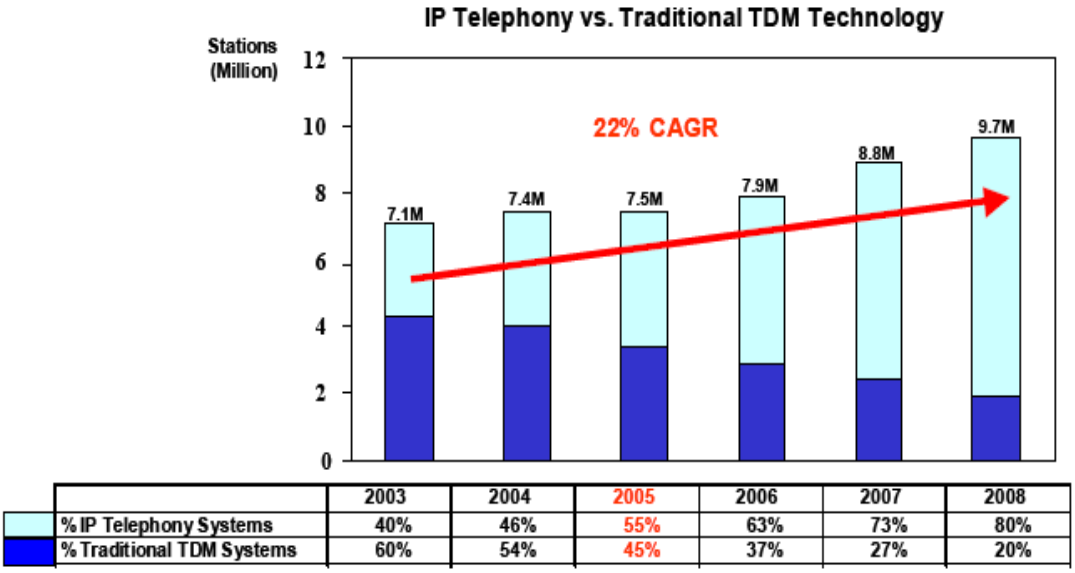


Based on a telecom market research by InfoTech, purchases of IPT platforms now exceed those of traditional TDM PBXs. By 2008, IPT solutions are expected to make up 80% of new enterprise system deployments. Figure 1.4 shows the statistics on how IPT is projected to take over the TDM world.

[5] *Gartner's Positions on the Five Hottest IT Topics and Trends in 2005*. Gartner Survey. May 12, 2005 http://www.gartner.com/resources/125800/125868/gartners_positi.pdf

[6] Manoj Menon. *IP Convergence in the Enterprise*. Cisco Systems. http://www.cisco.com/web/VN/voice/pdf/ip_convergence_in_the_enterprise_ver_3.0.pdf

Figure 1.4 – IP Telephony vs. TDM Telephony[7]

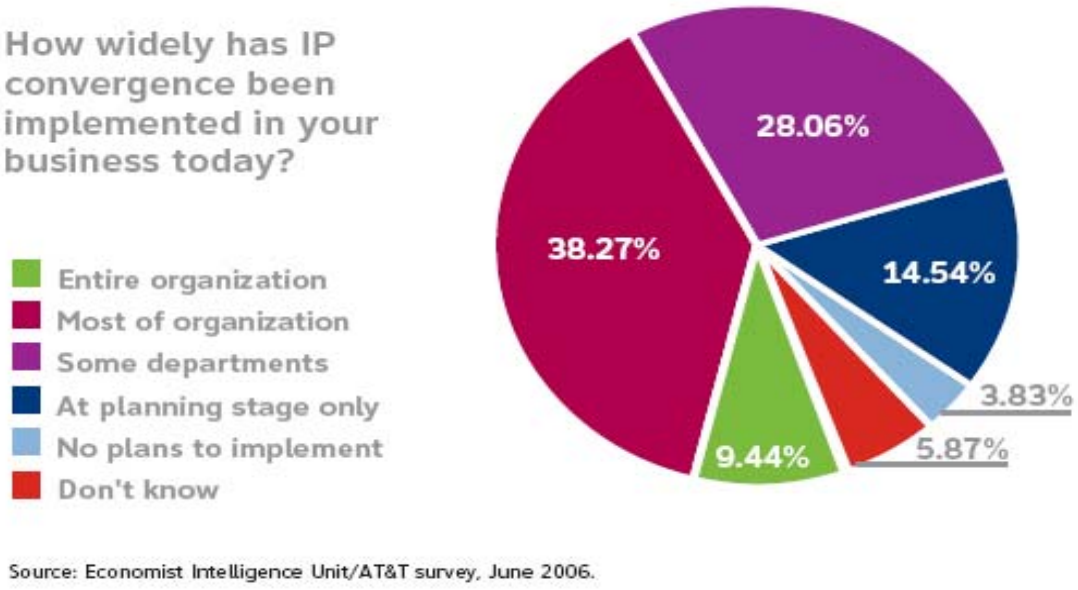


	2003	2004	2005	2006	2007	2008
% IP Telephony Systems	40%	46%	55%	63%	73%	80%
% Traditional TDM Systems	60%	54%	45%	37%	27%	20%

The Economist Intelligence Unit for AT&T conducted a global survey, which polled 395 senior executives across 51 countries and over 20 industries, with 63% hailing from large firms with annual revenue of more than US $500 million. Figure 1.5 and Figure 1.6 illustrate the results from that survey.

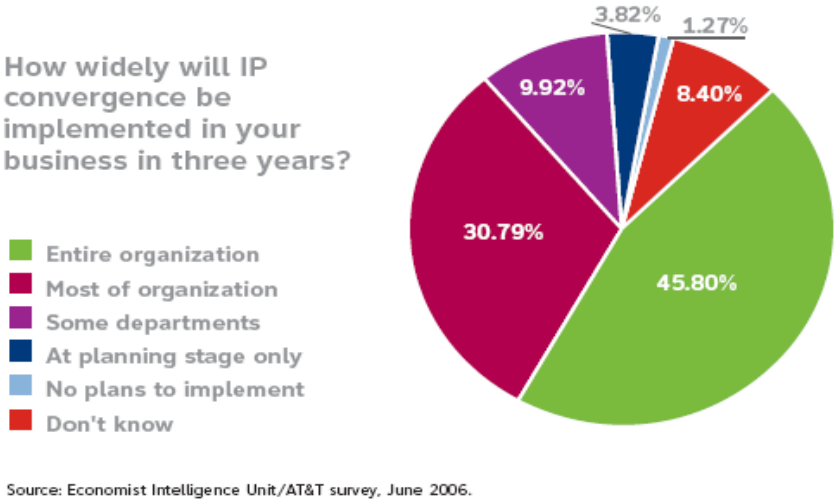
[7] *Strategies for IP Telephony Evaluation and Migration: Best Practice Considerations for Deploying IPT in the Enterprise.* InfoTech. April 2005. http://www.voicepro.com/_files/user/InfoTech%20Building%20Client%20Value1.pdf

Figure 1.5 – IP Converged Network Today[8]



Source: Economist Intelligence Unit/AT&T survey, June 2006.

Figure 1.6 – IP Converged Network in 3years[8]



Source: Economist Intelligence Unit/AT&T survey, June 2006.

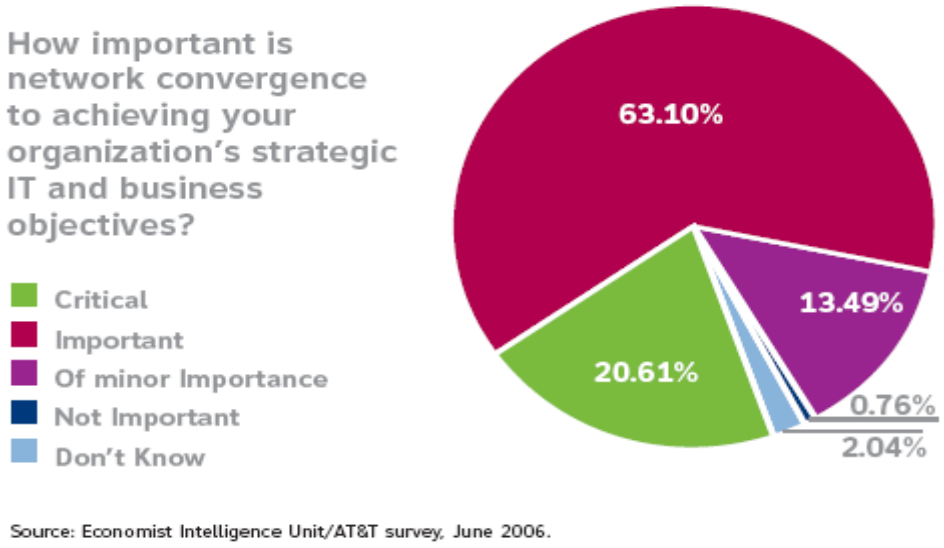
As evident from Figures 1.5 and 1.6, about 48% of the survey respondents say that IP convergence has been implemented in all or most of their business. Fully 77% of the same business owners say this will be the case in 3 years' time.

In the same survey, many executives also agreed that IP Convergence was important to achieving their organizations goals. Figure 1.7 illustrates that almost 63% of

[8] *Convergence Takes Hold in the Enterprise.* AT&T Corporation. http://www.corp.att.com/emea/docs/s4_convergence_eng.pdf

all the business owners that were surveyed agreed that network convergence played a big part in achieving their enterprise objectives.

Figure 1.7 – Importance of Converging Networks[9]



Source: Economist Intelligence Unit/AT&T survey, June 2006.

1.4 Incentives and Disincentives

IP Convergence is the union of data and voice communication protocols, using packet-based networks as the conventional infrastructure, to enhance the way enterprises implement communication networks. With data and voice IP protocols it is possible to improve the end-user information-sharing experience by combining voice, video, and data content seamlessly. Although interest in convergence was immediate, it took more than a half-decade for the overall market to reach a consensus that IP-based communications was fully enterprise-ready[10].

This section discusses the apparent incentives to convergence and analyses the various disincentives to the implementation process.

---

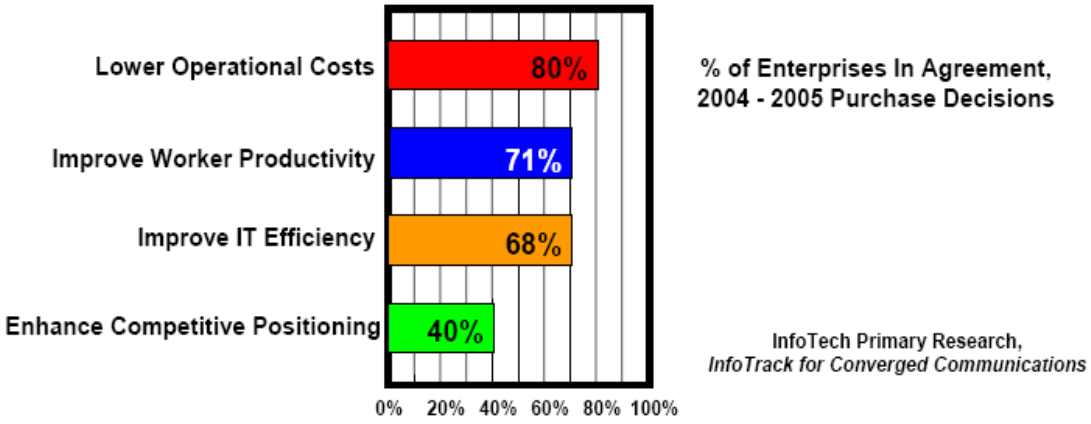
[9] *Convergence Takes Hold in the Enterprise.* AT&T Corporation. http://www.corp.att.com/emea/docs/s4_convergence_eng.pdf

[10] *Strategies for IP Telephony Evaluation and Migration: Best Practice Considerations for Deploying IPT in the Enterprise.* InfoTech. April 2005. http://www.voicepro.com/_files/user/InfoTech%20Building%20Client%20Value1.pdf

1.4.1 Incentives

There are many reasons why enterprises want to switch to converged network solutions. The two key driving forces behind corporate network convergence are cost and resource optimization. Convergence gives the enterprise workforce the ability to make effectual decisions and act in real time using whatever tools are available from any location. InfoTech found that while many companies vary in why they choose VoIP, most enterprises have found the most common anticipated benefits as lowering total operating costs, enhancing end-user productivity, improving IT organization efficiency, reinforcing market differentiation and brand image[11]. Figure 1.8 is the graphical representation of what companies say the initial interest is in deploying converged network infrastructure.

Figure 1.8 – Business Drivers for Deploying Converged Solution[12]



As it is made clear by Figure 1.8, cost cutting is one of the primary motivator and driving factors in the shift towards enterprise network convergence. But improving business efficacy and competence by revitalizing already available infrastructure, thus leading to resource optimization and enabling a user centric network design will be its

[11] *Strategies for IP Telephony Evaluation and Migration: Best Practice Considerations for Deploying IPT in the Enterprise.* InfoTech. April 2005. http://www.voicepro.com/_files/user/InfoTech%20Building%20Client%20Value1.pdf

[12] *Strategies for IP Telephony Evaluation and Migration: Best Practice Considerations for Deploying IPT in the Enterprise.* InfoTech. April 2005. http://www.voicepro.com/_files/user/InfoTech%20Building%20Client%20Value1.pdf

accurate motive. Cost optimization and will be discussed in detail in the following chapters.

1.4.2 Disincentives

There are reasons why there are critics of enterprise network convergence. Although enterprise wide network convergence promises cost savings in the long run, there is a huge amount of money that has to be invested for buying new equipment and implementing a single IP-based network. Cost of implementation and cost of purchasing equipments to support convergence is a main reason enterprises do not consider convergence solutions. Although long-run cost saving might be appealing to migrate conventional networks to converged solutions, enterprises might not have enough capital to spend on immediate implementation and equipment costs.

Another major reason why enterprises are not influenced by convergence options is that converged network is not easy to design and implement. Converged network planning and design involves learning multiple technological architectures. Implementing such complex technologies might not be an easy option without expert knowledge on the subject. This will be discussed further in the Network Design section that is to follow later in this chapter.

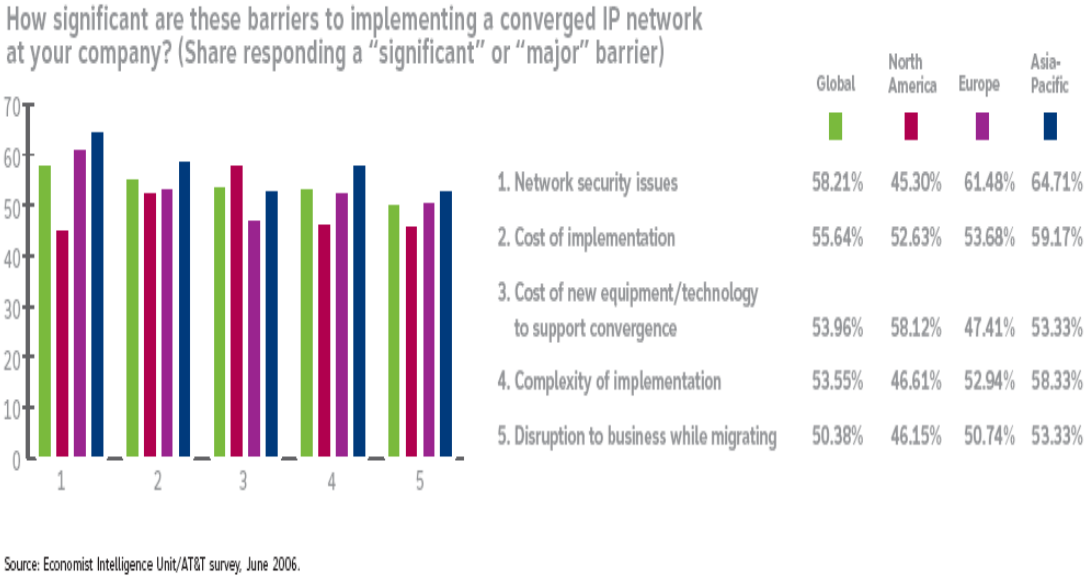
Also, network security becomes an issue when implementing converged solution. Although security can be seen as an advantage to converged network, it does have some drawbacks. This is discussed in detail in the Advantages and Challenges sections later in this chapter.

Finally, convergence does not occur overnight. It has to be carefully planned and implemented over a period of time. Some enterprises do not consider the option of convergence because they might not be able to afford the network downtime during the migration period. Today's businesses need connectivity to the network all the time. So the disruption to their network during migration poses as a negative outcome while they consider convergence.

The Economist Intelligence Unit for AT&T conducted a global survey polled 395 senior executives across 51 countries and over 20 industries with 63% hailing from large

firms with annual revenue of more than US $500 million. Figure 1.9 shows the results from the survey.

Figure 1.9 – Barriers to Implementing Converged Network[13]



Source: Economist Intelligence Unit/AT&T survey, June 2006.

As observed in Figure 1.9, North American companies consider the cost of new equipment to support convergence as their number one reason to not migrate their network, closely followed by the cost of implementation.

1.5 Advantages

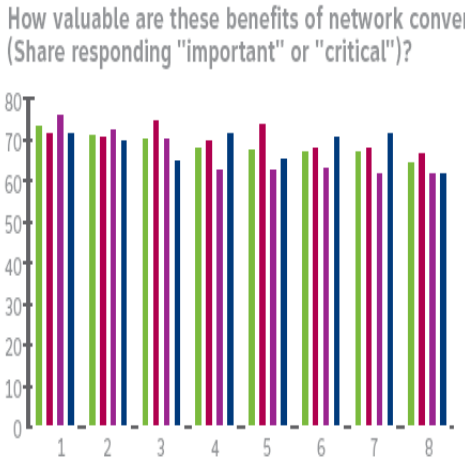
IP convergence is crucial to network modernization. Enterprise networks must be designed to constantly embrace change, be quick enough to adapt new technologies and be flexible to technological transformation. Over time, it will become apparent that enterprises that adapted network convergence have had an advantage over corporations that had been careless in the process.

The Economist Intelligence Unit for AT&T conducted a global survey that polled 395 senior executives across 51 countries and over 20 industries, with 63% hailing from large firms with annual revenue of more than US $500 million. With the results from that

[13] *Convergence Takes Hold in the Enterprise.* AT&T Corporation. http://www.corp.att.com/emea/docs/s4_convergence_eng.pdf

survey, it is apparent that enterprises in North America are moving towards converged solution, so the businesses can enhance their collaboration with customers, suppliers and partners among other important reasons. Figure 1.10 shows the results from the survey.

Figure 1.10 – Benefits of Network Convergence[14]



Source: Economist Intelligence Unit/AT&T survey, June 2006.

	Global	North America	Europe	Asia-Pacific
1. Simplified network management	73.67%	72.03%	76.47%	71.90%
2. Reduced cost of voice calls	71.65%	71.19%	72.79%	70.25%
3. Better collaboration with customers, suppliers and partners	70.74%	75.21%	70.59%	65.29%
4. Improved network scalability to accommodate business growth	68.35%	70.34%	63.24%	71.90%
5. Better customer service	68.19%	74.36%	63.24%	65.83%
6. Expanded number of access types by which employees can connect to the corporate network	67.68%	68.38%	63.70%	71.07%
7. Better network redundancy (for back-up in the event of an outage)	67.59%	68.64%	62.50%	71.90%
8. Better ability to launch and manage new applications	64.80%	67.24%	62.50%	62.50%

Some key advantages, as also mentioned in the survey above, in implementing a converged network are competence, user productivity, ease of network management, and scalability. Also, when implemented well, convergence can increase customer service. Some apparent advantages of convergence are discussed below: [14, 15]

1.5.1 Reduced Costs

A converged network solution moderates overall bandwidth costs by combining all kinds of telecommunications traffic over one single network. It also decreases management, maintenance and provisioning costs. It reduces the cost of integrating

[14] *Convergence Takes Hold in the Enterprise.* AT&T Corporation. http://www.corp.att.com/emea/docs/s4_convergence_eng.pdf

[15] *Network Services for Converged Communications.* IBM. http://www-935.ibm.com/services/us/gn/pdf/convservgd510-6388-00f.pdf

various applications. With such a simplification in network management, companies need fewer support staff and reduce capital and operational expenses to build, implement, maintain and manage the converged network.

A converged network brings down the number of vendors that a business enterprise deals with, thus reducing various costs that differ from vendor to vendor. This also increases the ability to associate invoices and services because enterprises trade with fewer dealers. Unlike traditional networks, converged solution reduces personnel costs to maintain different communication networks for voice, video, and data traffics. Overall, it reduces costs, optimizes network performance, and improves the elasticity of the network.

For an enterprise with 10,000 employees, saving just a ½ - 1 hour per employee during a week will save the corporation $8 – $25 million each year, according to Sage Research and Cisco Analysis[16]. Convergence brings about such time savings, which in turn can be equated to absolute dollar value. For example[16], innovative phone features and advanced voicemail[17] save up to 9 hours a week per employee. Also, extending system features to remote employees saves more than 4 hours per week per employee – this is an additional savings of 28 days per year. Extending system features to telecommuters leads to cost and time savings as well – on average, 5 hours per week. Besides, IT staffers would also save when end-users can use telephony features without additional assistance. Moreover, a Nemertes survey of 100 companies with average IT budgets of $10 million or more shows that employees move an average of .87 times per year (or almost once every year) at a cost of $100 per move[18]. For a company with 1,500

---

[16] *The ROI of Convergence.* Network World. June 14, 2004. http://www.networkworld.com/supp/2004/0621convergenceperspectives.html?page=1

[17] **Advanced Voicemail:** VoIP voicemail supports basic Class 5 voicemail features found in traditional phone systems. This includes the ability to: set-up mailboxes; record greetings and inbound messages; and play back, store, and delete inbound messages. VoIP voicemail also enables some advanced features such as voicemail forwarding to email as an audio attachment. Users may also choose to receive voicemail notification alerts via email. (http://www.voip-news.com/faq/voip-feature-faq/)

[18] *The ROI of Convergence.* Network World. June 14, 2004. http://www.networkworld.com/supp/2004/0621convergenceperspectives.html?page=1

employees, that amounts to $130,500 in annual costs that goes away with converged network implementation because users have the flexibility to move their own phones.

1.5.2 Productivity Enhancements

A converged network enhances the opportunities for intra- and inter-organizational cooperation. It enables distribution of a coherent and reliable user interface to the network regardless of location, application or business function. It enables a new level of integrated service. It provides a holistic rather than a disconnected view of business communications.

Employee productivity within the organization increases when they have better access to information regardless of their geographic location. By increasing information processing capabilities with a converged environment, employees will now have higher productivity through enhanced communication channels, as well as improvements for remote and telecommuters. IT staff productivity increases with support processes that improve service quality and reduce the time needed to perform day-to-day functions.

1.5.3 Simplicity

A converged network removes the complexity of managing multiple network infrastructures. It enables an understanding of the true cost of the network. As mentioned before, it improves invoice management. Another added advantage is that IP-based networks are able to translate any form of information they receive regardless of physical medium or service they may run on. This capability is built into the TCP/IP protocol suite – TCP/IP bridges the gap between dissimilar network environments, operating systems or applications[19].

[19] Tim Parker, and Karanjit S. Siyan. *TCP/IP Unleashed.* 3rd ed. Sams, 2002.

1.5.4 Reliability

Network reliability is another added advantage to converged IP-based infrastructure. If designed appropriately, a converged network will have the capabilities of providing both reliable voice and data services. Real-time applications such as video and audio have their own set of parameters that measure the network's reliability standards. This set of parameters, such as latency, jitter, throughput, etc, are addressed by resource management protocols such as RSVP. The IETF standards provided QoS standards for every network by making enough bandwidth available on a priority basis to support end-to-end quality control[20]. Furthermore, IP-based networks are able to offer higher quality voice services than the traditional TDM networks[21]. However, if not designed and implemented properly, the voice service might degrade compared to a traditional TDM service. Thus, while reliability is an advantage if the network is designed and implemented properly, it can turn around and pose as a challenge with a poor network analysis and design.

1.5.5 Scalability

A network that easily scales has become one of the fundamental requirements during the network design process. It enables the ability to expand the network as business needs change without having to invest money or time on different applications. A converged network makes the network absolutely adaptable and fully scalable. It supports evolving business applications and services like unified messaging and video/audio conferencing. In terms of scalability in converged networks, end-user scalability is achieved by the implementation of VPN services. The closed end-user grouping capability provided with the VPN like services lets enterprises limit incoming

---

[20] *Voice over IP Reliability: Architecture Matters.* ShoreTel. October 2004. http://www.infinitycomp.com/sbs/detail/shoretel_downloads/white_papers/Reliability_Whitepaper.pdf

[21] Carolyn R. Johnson, Yokov Kogan, Yonathan Levy, Farhad Saheban, Percy Parapore. *Voice over IP and Quality of Service.* January 2005. http://www.comsoc.org/tech_focus/pdfs/2005/jan/johnson.pdf

or outgoing calls to only members of the specified group. Also, VPN services have built-in security features, which allow for dedicated access without the need for deploying firewalls. For example, when a packet reaches a VPN access router, it is checked against a table for authorized source address. If there is no match found, then that particular packet is discarded. Furthermore, same filtering techniques can be implemented to enable enterprise to extend the access to their information to external users. This improves the network to go from being an intranet to an extranet while using the already existing infrastructure.

A converged infrastructure also makes it much easier to perform moves, adds or changes as organizational needs grow or reduce. Adding, dropping or moving phone numbers is all easily done on an IP-based network than on traditional TDM networks. DHCP, a standard protocol in data environment, can be used for managing IP phones on network. DHCP enables an IP phone to be moved to any part of the network without having to manually change the database configuration on a PBX or physical cabling/wiring alterations.

1.5.6 Security

All networks are vulnerable to attacks of one kind or another. For example, the voice communications industry has been faced with toll fraud for decades. In the data communications arena, security breaches have been prevalent for a longtime, especially with financial networks. A converged network reduces the number of access points, and hence the number of probable security threats. It facilitates real-time security policy management and enforcement. It enables a focused approach to dealing with security threats when they happen. Firms can implement and enforce a single set of standards across converged network to provide the optimal level of security while reducing safety breaches. Tunneling protocols such as IPSec, provide for authentication encryption and integrity in a packet switched network. Voice can be run over VPN’s, just like data traffic, with guaranteed improvement in performance by the QoS parameters that can be set for all applications on a VPN. Moreover, privacy issues can be handled during voice

transmission in a converged network that has VPN deployed because it inherits the properties from IPSec. The IPSec protocol will be discussed in detail in a later chapter.

Converged networks can offer a more stable, higher-performing alternative that helps increase your adaptability and reliability. With one network infrastructure, you can have greater visibility of the resources and monitor and address security threats in a straightforward fashion. A converged network environment also makes possible for quicker enterprise wide communications to coordinate the response to security issues.
Similar to reliability standards, although security is an advantage if the network is designed and implemented properly, it can turn around and pose as a challenge with a poor network risk assessment and design. The latter part will be discussed later in this chapter.

1.5.7 Potential

IP enterprise networks lead to boundless access to information. Every country in the world is tied to the public Internet. The magnitude of the Internet gives businesses worldwide reach. IP roaming capabilities enhance this overall reach by enabling end-user mobility. Roaming services give telecommuters, remote workers, and travelers access to the enterprise intranet from just about anywhere in the world. The Internet also offers unique opportunities for organizations to economically extend communication capabilities, applications and information to employees, remote offices, mobile workers, telecommuters, suppliers, partners, vendors, etc.

A VPN solution is the means by which global enterprises communicate securely via IP networks. VPNs allow corporations worldwide to carve out their own IP WAN within the service providers IP backbone. Thus, VPNs have the endless potential of being local, national or even global in geographical scope.

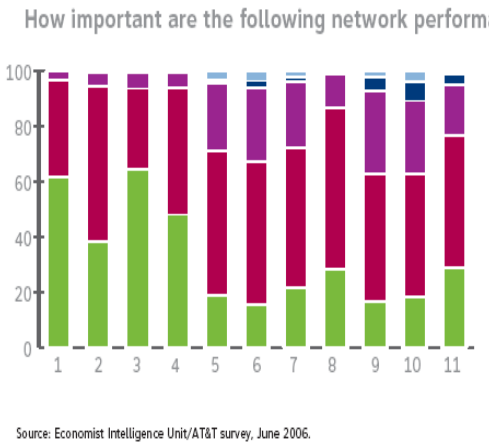
1.6 Challenges

Success with converged networks is not assured – rather it requires the understanding of the various challenges that shape it. Enterprise networks are global in

geographic scope – they connect corporate headquarters and branch offices to employees, customers, vendors, partners, etc. Such networks, when converged on a single IP-based backbone, clearly create challenges during the design and implementation process. Requirement for converged network performance raises questions such as quality of service, traffic analysis, and monitoring, performance management, etc. As communication needs increase, corporations need a more flexible architecture that promises hardware and applications interoperability and enables reliable delivery of voice, data and other multimedia applications. Although converged IP network implementation is the answer, the implementation itself can pose a challenge[22].

Figure 1.11 shows the results from the Economist Intelligence Unit for AT&T survey emphasizing the various important network performance attributes as seen by the company executives.

Figure 1.11 – Converged Network Performance Attributes[23]



	Critical			Not Important	Don't Know
1. Network availability/downtime	61.52%	35.19%	3.04%	0.25%	0.00%
2. Speed of network	38.23%	56.20%	5.32%	0.25%	0.00%
3. Security	64.72%	28.93%	5.58%	0.76%	0.00%
4. Reliability of connections	48.08%	46.04%	5.37%	0.00%	0.51%
5. Scalability	18.83%	52.16%	24.43%	1.27%	3.31%
6. Reach (number of end points)	15.74%	51.27%	26.90%	2.54%	3.55%
7. Efficiency/capacity utilization	21.52%	50.63%	23.80%	1.77%	2.28%
8. Applications performance	28.39%	58.31%	12.28%	0.51%	0.51%
9. Ability to prioritize applications	16.58%	46.17%	30.10%	4.85%	2.30%
10. Open, standards-based architecture	18.21%	44.62%	26.41%	6.92%	3.85%
11. Accessibility (ie, access to full suite of network applications from outside the office)	29.16%	47.31%	18.67%	3.84%	1.02%

Indicated in this chapter are few of those fundamental challenges to be considered during the planning and design process of a converged network.

---

[22] *Implementing VoIP Network.* Network General. http://i.i.com.com/cnwk.1d/html/itp/Network_General_Implement_VOIP.pdf

[23] *Convergence Takes Hold in the Enterprise.* AT&T Corporation. http://www.corp.att.com/emea/docs/s4_convergence_eng.pdf

1.6.1 Security

Figure 1.11 makes it evident that network security is the major concern while enterprises consider a converged network solution. About 65% of the executives that were surveyed clearly indicated that network security was their most important factor to be considered in view of their networks convergence. To date voice and data networks were implemented in different architectures, so the security needs were also kept separate. However, as enterprises are emerging towards implementing converged network solutions, designers are exposed to new voice and data susceptibilities. Security measures are required to be considered not only after being attacked but also while designing the network. All measures must be taken to protect the network from any kind of malicious attack and minimize the impact on data transfer. All three hierarchies that were discussed in the network design section must have individual security measures implemented. Only authorized users must be able to connect to the network. All the network users and devices that are connected to the network must have some limitation posed that might be stated in the enterprise security enforcement policies. Therefore, access control has become an essential security concern throughout the corporate world.

As security measures have taken priority while designing networks, so have the issues that complicate the measures. Thus, today's converged network solutions must have security policies that includes prevention measures while planning and designing the network and detection measures at each and every access point to the network. Converged network security issues must be in conjunction with providing QoS for all current and next generation applications.

1.6.2 Performance

Figure 1.11 makes it apparent that network performance – availability/downtime – is the second major concern while enterprises consider a converged network solution. A converged network carries both real-time and non real-time traffic. The real-time applications place operational limitation on the network that may not be present with traditional data network infrastructure. For example, voice traffic quality degrades when

the delay exceeds conventional standards – this is a constraint that typically would not affect email applications. Thus, there is a need for some technique to monitor and measure these real-time network statistics. Quality and performance measurements can be done in a network if real-time end-user experience data is readily available. So performance monitoring becomes an easier task on already functioning networks but complicated on newly designed networks.

Unlike data and voice traffic, video streaming and other multicast traffic can be harder to predict. Managing such real-time video traffic requires insight into the availability, bandwidth utilization and errors for all interfaces. Moreover, in any given enterprise network, the subscribers for multicast traffic can be dynamic. Thus, all interfaces must be aware of all user groups at any given time to be able to forward the traffic appropriately. Consequently, managing this user list itself adds more erratic overhead and degrades the performance of the network. And since these are so unpredictable, given ever changing user groups, predicting this type of traffic becomes a complex task.

Thus, while planning and designing a converged network, the network administrator has no other means but to gather data from other similar networks or simulate and obtain performance results with simulation software with projected traffic matrix.

1.6.3 Reliability

As evident from Figure 1.11, network reliability is the third major concern of enterprises while considering network convergence. With a converged solution, users now run all voice, data and multimedia applications on the same platform. Therefore, the expectations network performance is typically high. For example, users that have been using the PSTN network for a long time would have realized that it has 99.999% reliability[24]. Now, they expect the newly implemented VoIP network to be as reliable as

[24] *Implementing VoIP Network.* Network General. http://i.i.com.com/cnwk.1d/html/itp/Network_General_Implement_VOIP.pdf

that. If not properly designed to certain constrains, the network might not meet the users' need. Therefore, this is a challenge for the network designer to consider.

As enterprise networks go from different voice, video and data networks to a converged infrastructure, the QoS approach must undergo an immense transformation. As mentioned earlier, unlike most non real-time traffic such as asynchronous data packets over a network, audio and video traffic cannot afford to experience network delay, congestion or jitter. The quality is degraded if audio/video packets arrive late, or sometimes the packets are discarded. Therefore, converged networks require a QoS technique implemented that can distinguish and prioritize among the various types of traffic across the network. To accurately plan for desired QoS methodology, the network designer should understand the different traffic types as they relate to the enterprise's needs and necessities, analyze the traffic for service level requirements, and then implement an appropriate QoS policy. This way the approach will help prevent network congestion and delays and optimize network performance. A QoS strategy should be chosen that would ensure timely and reliable delivery of real-time traffic and reliable delivery of non real-time traffic across the network. The different types of QoS strategies and various methods of implementing QoS will be discussed in detail in later chapters.

1.6.4 Inexperienced Staff

A typical network uses different technologies for voice and data platforms the converged network encompasses elements from both voice and data communication. Consequently, IT and telecom professionals have to be knowledgeable about both of the technologies in order to integrate and incorporate them in the network design. Thus, to successfully manage a converged environment, a knowledgeable staff, with familiarity with both telephony and data networking expertise, is essential.

1.7 Network Analysis

Network convergence does not happen overnight. It requires careful analysis of current network infrastructure, hardware/software and other application requirements for

proposed infrastructure, thorough budgeting strategies, and knowledgeable staff to carry out tasks. Thus, enterprises have to put in immense effort in making sure the company is ready for convergence and can afford the initial costs integrated with the implementation process. This section will provide an insight on the facts that are to be considered when an enterprise plans to integrate its voice, data and video services into a single IP-based network.

The steps in transitioning a traditional network to a converged network might sound extremely simple – just analyze the current technology and business needs and plan and design a network that would meet the business needs and deploy next generation technologies. It is not, however, as easy as it might seem. Migrating to an effective functioning converged network solution is a complex task. Even after the network analysis and the business needs are identified, there is much work to be done. Below are some points that need to be carefully considered before making the big decision of migrating to a converged infrastructure.

- Globalization is a buzz word among big business owners. Enterprises are constantly looking to expand their empire globally to be able to reach a wider customer base. In such case, it is no longer good enough to have their communication systems just provide service in a specific area. Large enterprises worldwide are moving to service providers that can provide end-to-end connectivity and management services as a part of their service. If the corporation considering migration to converged network solution has a global presence, then it is time to move towards implementing network plans. This opens opportunities to enormous cost savings in the long run.
- Another eminent reason in the move towards converged network solution is that most of the equipment and software solutions running on today's separate voice and data networks will likely become obsolete with more and more next generation technologies being introduced. Given that new technologies are being developed and next generation technologies might not make full use of depreciated equipment, telecommunications equipment will need to be refreshed on a regular basis. Thus, as the need to replace any obsolete technological architecture arises, it gives the enterprise an opportunity to invest on converged

network solution to reduce cost, improve QoS, and simplify the network infrastructure.

Once an enterprise determines that there is a pushing need towards convergence, then the scheduling of the project plan must be done accordingly. Most of the time global corporations cannot afford to have their networks down for a long period of time, no matter how vital the reason for the downtime is. Convergence does not happen overnight. Therefore the migration process must be wisely scheduled.

1.8 Network Design

Whether it is designing a brand new converged network, or migrating from the traditional network to a converged infrastructure, the planning and design of a converged network has to take many points into consideration. Converged network planning and design involves learning multiple technological architectures. The vendors have to work with their clients and network architects to know and understand the business needs, application features, traffic requirements, and other services. Generally, a converged network just like a traditional network, has various design components such as the communication media, the signaling architecture, the hardware requirements, the service management facilities, etc[25].

It is a common misconception that designing a converged network is easy compared to designing traditional networks. But this is not always true[26]. For example, while designing traditional networks the designers look at voice and data as separate entities. Thus, the data network is designed to have good performance and reliability to keep the network running, while the voice network has to take GoS and traffic analysis at a certain level to have the voice system functional. Unlike that, for a converged network designers have to put both of these factors together to obtain a mutual agreement of

[25] *The ROI of Convergence.* Network World. June 14, 2004. http://www.networkworld.com/supp/2004/0621convergenceperspectives.html?page=1

[26] *The Coming of True Convergence: Why Service Providers Can Finally Turn Out the Lights on the Old Public Switched Telephone Network (PSTN).* International Engineering Consortium. http://www.iec.org/online/tutorials/true_converge/

network performance that is capable of handling all kinds of network traffic such as voice, video and data. Therefore, the way a converged network is designed and all the considerations that were taken into account while designing the network have a direct impact on the network performance and reliability.

Network designers usually take a hierarchical approach to designing a converged network[27, 28]. The hierarchy is divided into three layers. The first layer is the logical layer, where all the enterprise's applications and the hardware and software needed to access the network inhabit. The second layer is the transport layer, where the network's routing and switching takes place. This layer is where corporate networks all around the world are connected to each other for easy network access. This layer is also responsible for having the enterprise network connected to the outside world network, also known as the Internet and the PSTN. Lastly, there is the access layer where the network users reside. Figure 1.12 shows a layout of the three layers in the converged network design.

[27] *Implementing VoIP Network.* Network General. http://i.i.com.com/cnwk.1d/html/itp/Network_General_Implement_VOIP.pdf

[28] *Designing Converged Networks. Networks and Services.* December 2003. http://www.enterprisenetworksandservers.com/monthly/art.php?393

Figure 1.12 – Converged Network Hierarchy[29]

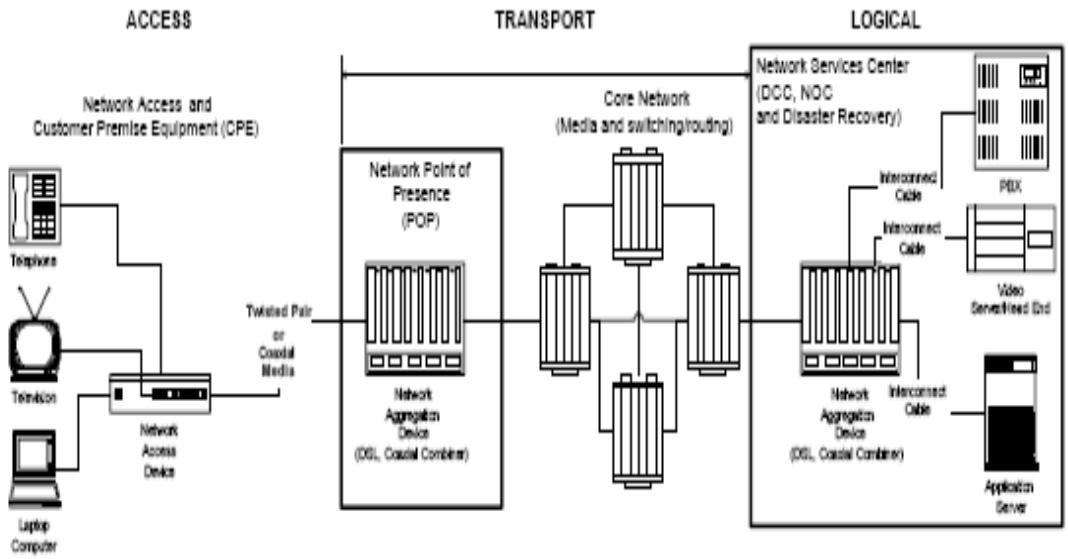


Table 1.1 illustrates the relationship between the OSI reference model and the converged network hierarchical model.

Table 1.1 – Converged Network Hierarchy Mapped to OSI[30]

OSI Model	Converged Network Hierarchy
Physical	Access
Data Link	Access and Transport
Network	Transport
Transport	Transport
Session	Logical
Presentation	Logical
Application	Logical

[29] Rick Allison. *Converged Networks Design Features.* Lucent Technologies. http://www.lucent.com/livelink/176405_Whitepaper.pdf

[30] Rick Allison. *Converged Networks Design Features.* Lucent Technologies. http://www.lucent.com/livelink/176405_Whitepaper.pdf

Similar to traditional network design, it is important to have network traffic analysis data for converged network design. A network designer should be able to determine the traffic flow across the network. This is an important factor so that a designer can determine where traffic is going to be most, at what time the traffic is going to be heavy, etc. Given that today's enterprises run various high bandwidth applications on a regular basis, having a real-time traffic data allows network architects to better design networks with redundancy and back ups incase a link goes down. QoS consideration is also another important factor in converged network design approach.

Therefore, the design of converged network is a complex process that takes many design principles and technologies into consideration. These features may not be clear to a traditional network designer[31]. Using the hierarchical approach to design the network will increase network efficiency and performance.

1.9 Conclusion

This chapter introduced the concept of convergence – a single IP-based infrastructure that runs voice, video and data traffic of enterprise networks. The advantages, incentives, disincentives, and challenges in implementing a converged network were also discussed. It is important to understand these various factors while analyzing a network to determine whether or not convergence might be the right solution. The following chapter will focus on the various technologies and protocols involved in a traditional network implementation.

[31] *Designing Converged Networks.* Networks and Services. December 2003. http://www.enterprisenetworksandservers.com/monthly/art.php?393

2 Network Technologies

2.1 Introduction

The concept of voice and data networks have been around since the theory of networking was developed. Enterprises across the globe have been implementing these technologies to improve their business communication process, develop their business opportunities and optimize costs. The legacy network model delivers voice, video and data on separate single purpose network infrastructure, while the converged model has a single network. This chapter will discuss the characteristics, architecture, and some relevant protocols of the voice, data and video traffic over a converged architecture.

2.2 Switching Characteristics

There are two types of predominant networks – circuit switched and packet switched. To understand the characteristics and behavior of converged networks, it is important to understand the fundamentals behind the voice and data traffics; mainly the difference between packet switched and circuit switched network technologies. This section provides an overview of packet switching and circuit switching.

2.2.1 Packet Switching

Packet switching is a type of technology in which data is broken into small "packets" that are routed through the network based on the destination address on the packet. Different packets might take different routes from the source, but they are reassembled in order at the destination. This method of data transmission gets hold of and frees bandwidth as and when it is needed. This mode of communication between the source and destination is viewed as connectionless service. Advantages of having a packet switched network include having the network ability to connect to simultaneous connectionless services, which in turn increases efficiency of the network. Packet switching might be a disadvantage when network congestion occurs; that is many users share the same network, and the bandwidth availability becomes low.

2.2.2 Circuit Switching

Circuit switching is a technique where a system establishes a physical connection between the source and the destination (callING party and the callED party) before the traffic transmission begins. Voice transmission via copper wires is a good example of circuit switched network. This method of transmission statically reserves the required bandwidth well in advance.

2.2.3 Circuit Switching vs. Packet Switching

Circuit switched networks are based on Time Division Multiplexing (TDM), in which various signals are combined for transmission on a single communication channel. Once a connection is established, it remains throughout the transmission session, whereas, in a packet switched environment, the packets are routed based on destination addresses contained in the header of each packet. Breaking down data into smaller packets allows the communication channel to be shared among users in the network.

Circuit switching, as explained above, is completely transparent. The source and the destination can use any bit rate, format, or framing methodologies they desire. In contrast, for packet switched networks, the carrier needs to determine the basic parameters before the transmission occurs. Compared to circuit switching, that statically reserves the required bandwidth, packet switched networks allocate bandwidth dynamically. Packets on a packet switched network can be routed through different routes, and still be organized back at the destination end. On a circuit switched network, all the data transmitted follow the same dedicated route.

The advantages of using a packet switched environment for voice and video communications will be discussed in the next chapter. Although packet switching for real-time applications seems easy to implement and has clear advantages, there is the QoS issue that needs to be taken care of. This will be discussed in the following chapter.

Table 2.1 shows a comparison between circuit switched and packet switched network characteristics.

Table 2.1 – Comparison Chart

	Circuit Switched Network	Packet Switched Network
Bandwidth Allocation	Static	Dynamic
Dedicated Path	Yes	No
Bandwidth Optimization	No	Yes
Call Setup	Required	Not Required
Congestion Probability	Only at setup	On every packet transmission

2.3 Network Models

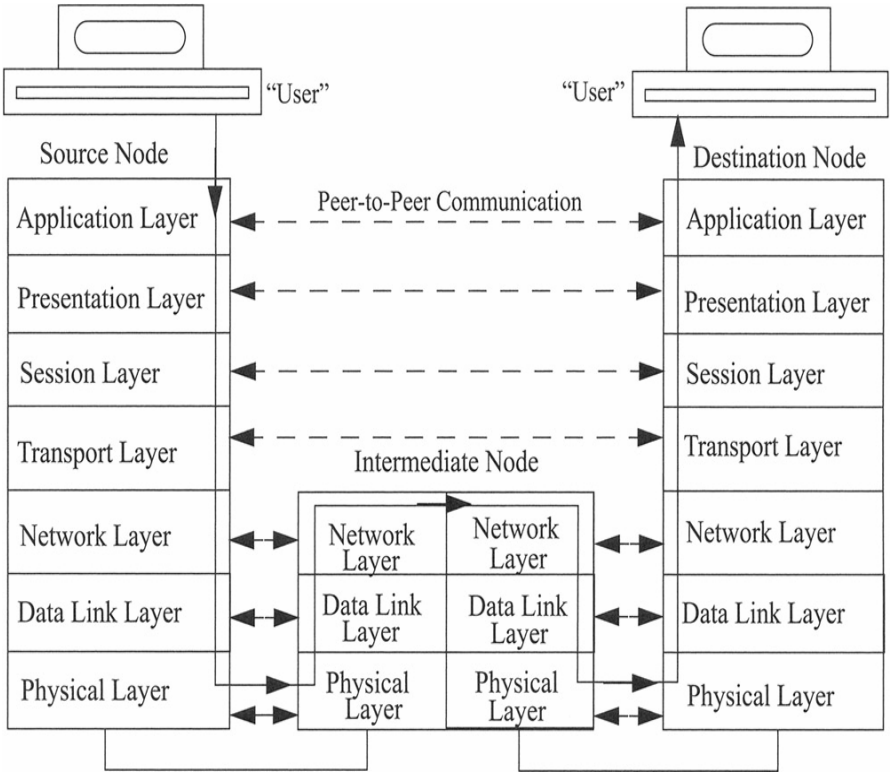
This section will discuss the two different reference models, Open Systems Interconnection (OSI) and the Transmission Control Protocol / Internet Protocol (TCP/IP) used for implementing data communication between interoperable systems.

2.3.1 Open System Interconnection

The first one, called the Open Systems Interconnection, proposed by the International Standards Organization (ISO) is a seven-layer architectural model. Each layer is responsible for specific, well-defined data communication function. Thus, the model attempts to decompose the complexity of information flow between communicating nodes into a set of functions that are independent of each other[32]. This is achieved by building the model so that the upper layer protocols depend on the services provided by the lower layers. Figure 2.1 represents the OSI reference model.

[32] Oliver C. Ibe. *Converged Network Architectures: Delivering Voice over IP, ATM, and Frame Relay.*

Figure 2.1 – OSI Reference Model[33]



The seven layers that are used in this model are: Application Layer, Presentation Layer, Session Layer, Transport Layer, Network layer, Data Link Layer, and the Physical Layer[33].

1. Physical layer is the lowest layer of the OSI model that is responsible to define the electrical and mechanical standards. It is concerned with transmitting raw bits over the given communication medium. That is, it is responsible for establishing, maintaining and terminating the required signaling. It deals with the physical medium of information exchange.
2. Data link layer is responsible for organizing the data in a specific format (called a frame) for transfer over the physical medium, and for detecting and correcting errors in a frame. This layer transforms raw transmission facility into a line that appears free of undetected transmission errors to the network layer.

[33] Oliver C. Ibe. *Converged Network Architectures: Delivering Voice over IP, ATM, and Frame Relay.*

3. Network layer is responsible for routing the data to its destination and for network addressing. It controls the operation of the subnet. This layer is also accountable for examining the packet to determine its destination and the required routing information.
4. Transport layer is responsible for the reliable transfer of data between the source and the destination, regardless of the performance and number of networks involved in the connection between the communicating networks. This layer accepts data and splits it up into smaller units, passes it on to the network layer, and ensures that the packets transmitted arrive at the destination node. Thus, it is responsible for end-to-end data integrity of data transmission[34].
5. Session layer is responsible for the establishment, maintenance, and termination of connections between applications. It controls data transfer by structuring data exchange into a series of dialog units. This facilitates restarting the exchange if service is interrupted. It is responsible for security during a connection and maintains the connection until data transmission is complete[33].
6. Presentation layer is responsible for translating the information to be exchanged, into formats that are understood by the destination. This layer is concerned with the syntax and semantics of the information transmitted. It ensures that information sent by the application layer of source is interpretable by the user of the application layer of the destination. It also performs data conversion, data encryption, and data formatting for display or printing.
7. Application layer is responsible for providing services to end-user applications that lie outside the scope of the OSI model. It defines the procedures by which end-user applications access network services.

The application layer, presentation layer, and session layer are concerned with application functions, while the lower four layers are concerned with data transport functions. Thus, the boundary between the transport layer and the session layer is the point of demarcation between application protocols and transport protocols.

[34] Oliver C. Ibe. *Converged Network Architectures: Delivering Voice over IP, ATM, and Frame Relay.*

2.3.2 TCP / IP

The second network model is the TCP/IP, which was developed for the Internet. Like the OSI model, the TCP/IP is also a layered architecture. The layers of the TCP/IP model includes the Network Access Layer, Internet Layer, Transport Layer and the Application Layer[35].

1. The Network Access layer is implementation specific. There are no specific protocols are defined for the network access layer. The TCP/IP model indicates that a host has to connect to the network using some protocol so it can send IP packets to it. This protocol is not defined clearly, and caries from one host to another, and from one network to another. It is expected that the network will rely on the upper layer protocols for appropriate proceedings.
2. The Internet layer is the key player in the TCP/IP reference model. The responsibilities at this layer include permitting hosts to bring in packets into any network and have those packets travel independently to its destination – a packet switched network based on a connectionless Internet layer. The Internet Protocol (IP) defined for this layer is a simple connectionless datagram protocol. It provides no error recovery, but performs error checking on each IP packet and discards any packet found to be in error without notifying the sender. Thus, there is no guarantee that the packet will be delivered.
3. The Transport layer provides a reliable data transfer between two communicating end systems. It provides mechanisms that establish, maintain, and carry out orderly termination of virtual circuits. The layer also provides mechanisms for error recovery and flow control. Two protocols are defined for the transport layer: Transmission Control Protocol (TCP), and User Datagram Protocol (UDP)[36].
    - TCP, detailed in RFC 793, defines a reliable mode for data transfer with error control mechanism[37]. TCP is a byte-oriented reliable

[35] *TCP / IP*. Cisco Systems. http://www.cisco.com/warp/public/535/4.html

[36] *Understanding TCP / IP*. Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/product/iaabu/centri4/user/scf4ap1.htm

[37] *Transmission Control Protocol. http://www.ibiblio.org/pub/docs/rfc/rfc793.txt*

protocol. It uses sequence numbers and acknowledgments to enable communication between two end users. The sequence numbers are used to determine the ordering of transmitted packets and to determine when a packet does not arrive at its destination. It is a connection-oriented protocol that first establishes a connection between source and destination, transfers the data, and then closes the connection.

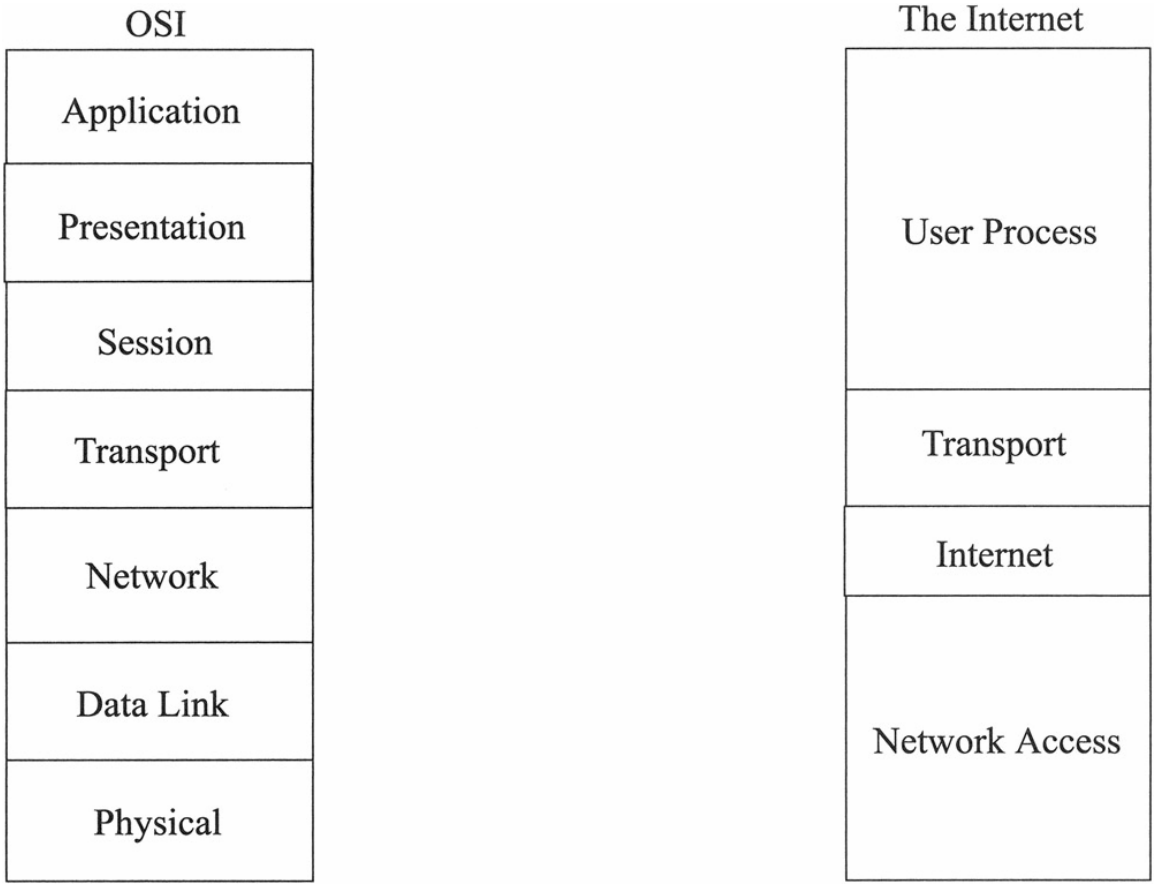
- UDP is a connectionless protocol that offers minimum reliability during data transfer, described in RFC 768, also provides minimal protocol overhead[38]. Therefore, it is usually considered an unreliable protocol. A connectionless path is one in which the communication channel is not established prior to the transmission of data. Unlike TCP, it does not require the connection between the source and destination to be established prior to data transfer. And it is an unreliable protocol because it does not issue an acknowledgment after the data has been received. UDP is designed to be simpler than TCP and is to be used by those applications that do not need the reliability and overhead of TCP. Instead, the network transmits the data in a package called a datagram. The datagram contains all of the addressing information necessary for that message to reach its intended destination.

4. The Application layer in the TCP/IP model is comparable to the abridgement of the Session, Presentation and the Application layers of the OSI standard based model. This layer contains all the higher level protocols such as TELNET, FTP, SMTP, DNS, etc.

Figure 2.2 illustrates the layer mapping of OSI to TCP/IP architectures.

[38] *User Datagram Protocol.* http://www.ietf.org/rfc/rfc0768.txt

Figure 2.2 – OSI and TCP/IP Mapping[39]



2.3.3 OSI vs. TCP / IP

As discussed above, in both the models, each layer has a distinct set of functions that it carries out; each layer also performs services for the layer above it. In both the above mentioned network models, the upper layer protocols provide end to end transport service to communicating applications. Also, the upper layers are used for application oriented protocols.

Although there are many similarities between the reference models, there are fundamental differences that make these architectural models apt for specific applications. The OSI model was developed before devising the corresponding protocols.

[39] Oliver C. Ibe. *Converged Network Architectures: Delivering Voice over IP, ATM, and Frame Relay.*

Therefore, this model is not biased toward any set of protocols[40]. Given this feature, the protocol developers did not have much exposure to the subject and did not have an idea of which functionality to put in which layer. On the contrary, the TCP/IP reference model was devised with the existing protocol stack. Therefore, there was no problem with the protocols fitting the different layers in the model. The drawback was that this model does not fit any other protocol stack. Hence, it is not useful for describing non-TCP/IP networks[40].

In addition, the OSI model supports both connection-oriented and connectionless protocols at the network layer level, but only connection-oriented services at the transport layer level. On the contrary, the TCP/IP model supports only connectionless services in the network layer level, but both connection-oriented and connectionless services in the transport level, giving the user a choice. Another major difference is that the TCP/IP model does not clearly distinguish between the physical and the data link layers. The differentiation between the transmission characteristics of the transmitting media and the framing characteristics is done only in the OSI reference model.

With all the major differences discussed above, it is clear that the OSI and the TCP/IP models have their own applications, advantages and disadvantages. Consequently, the OSI reference model has proven to be useful for discussing networks in the educational arena for beginners to comprehend computer networks[40]. In the contrary, the TCP/IP model has protocols that are widely used in the industry[40]. TCP/IP reference model is the one that is widely implemented in enterprise networks.

2.4 Legacy Networks

The goal of this section is to review data, voice and video communication systems and technologies that have been used in enterprise environments over the years. A clear understanding of these technologies and the demands that video and voice signals places over a network is vital for the study of network convergence and the expectations of legacy network users from a converged network solution.

[40] Andrew S. Tanenbaum. *Introduction: Reference Models.* Computer Networks.

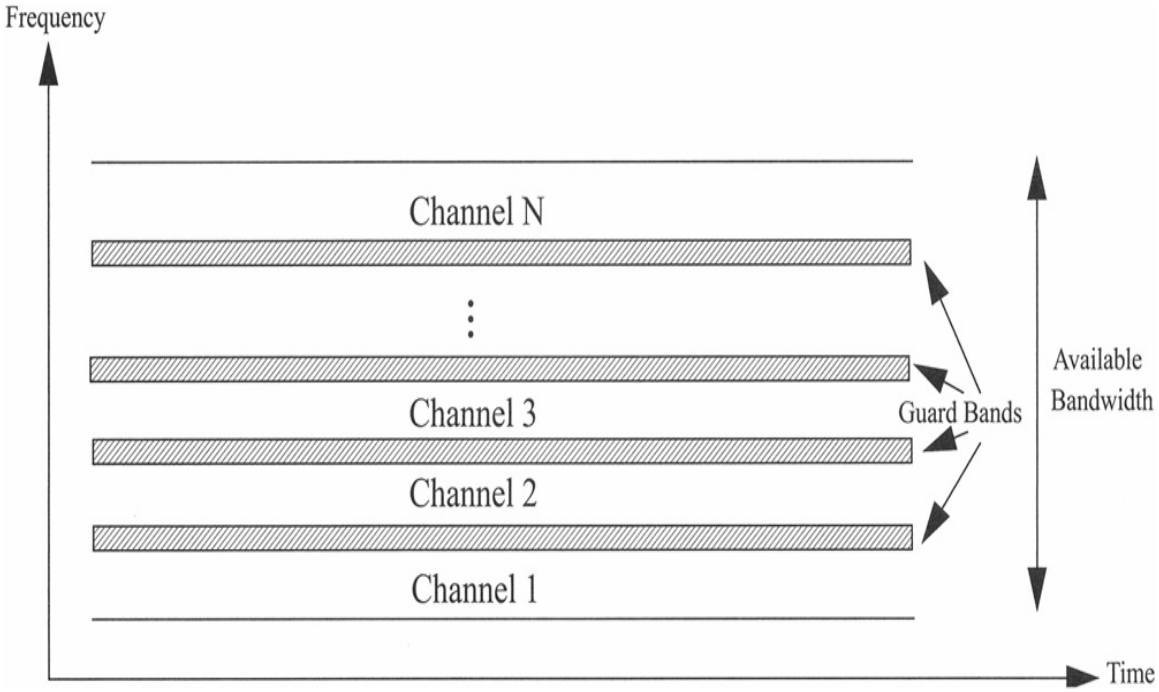
2.4.1 Voice Communications

Traditionally, voice networks have been deployed with two major multiplexing schemes – Frequency Division Multiplexing (FDM) and Time Division Multiplexing (TDM).

2.4.1.1 Frequency Division Multiplexing (FDM)

Early telephone systems multiplexed multiple calls into a single physical circuit using a technology called Frequency Division Multiplexing. In frequency division multiplexing, the available spectrum of frequencies for voice transmission is divided into independent channels. A user can transmit in one channel without affecting another user in another channel. Thus, all the channels generated in the link can be used simultaneously. FDM is used to partition the radio frequency spectrum, thereby making it possible to receive transmissions from different radio and television stations simultaneously. Figure 2.3 shows an FDM system with N channels.

Figure 2.3 – FDM System[41]



Although this technology was implemented in the earlier systems, FDM turned out to be insufficiently scalable for the demands of telephony. So in the mid 1900's the phone companies began digitizing voice signals and multiplexing them in the time domain using TDM.

2.4.1.2 Time Division Multiplexing (TDM)

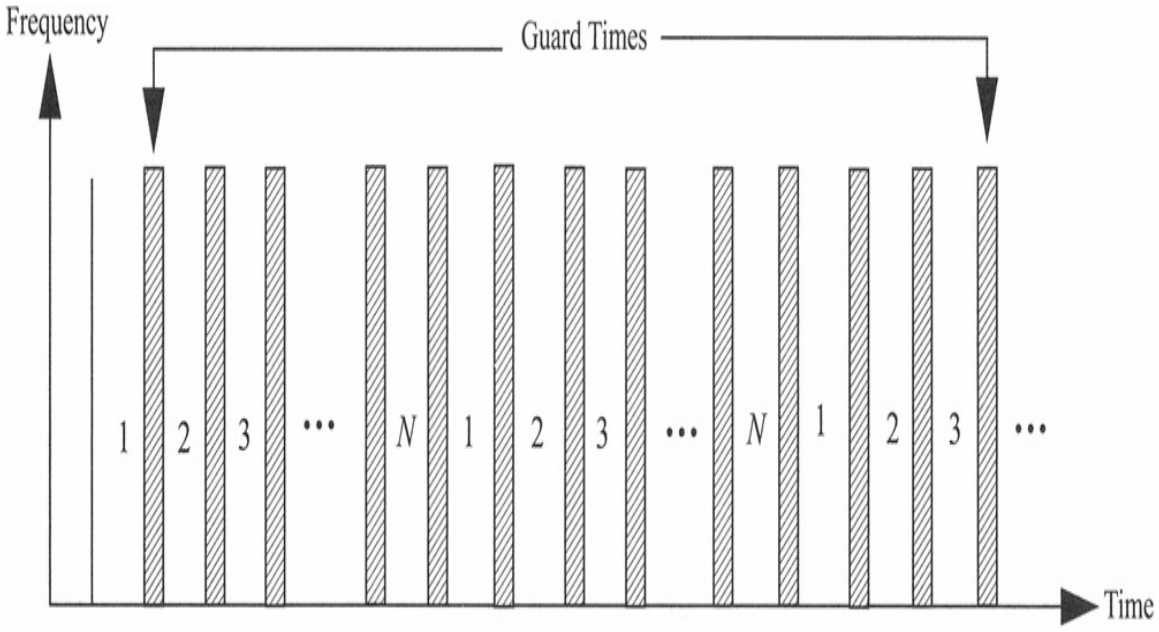
Voice signals needs to be delivered to the end users in real-time, which is a primary reason why voice networks have been implemented in a circuit switched environment. In any circuit switched network such as the Public Switched Telephone Network (PSTN) there exists the need to transmit multiple subscribers' calls along the same transmission medium. All traditional voice communications networks are based on the legacy Time Division Multiplexing architecture.

TDM is a scheme where various signals are combined for transmission on a single line or channel. Each signal from a lower bit-rate source is broken up into numerous

[41] Oliver C. Ibe. *Converged Network Architectures: Delivering Voice over IP, ATM, and Frame Relay.*

segments, each having very short duration and are multiplexed in rotating repeating sequence onto a high bit-rate transmission line[42]. The use of short duration pulses allows information to be transmitted at very high bit-rates. To do the actual switching, exchange of time-slot is done by a special type of switch called time-slot interchange (TSI) network[43]. TDM is accomplished by merely interleaving data from several bit streams. This can be done on a bit basis or on a byte basis. Figure 2.4 illustrates TDM architecture for *N* users.

Figure 2.4 – TDM system for *N* users[44]



TDM multiplexers interleave the output of 12 codecs into a multiplex frame. There are three standards for multiplexing – North American, European and Japanese. All three are based on a DS0 channel that is the pulse code modulation (PCM) output of a

[42] A. Als, F. Z. Ghassemlooy, G. Swift, P. Ball, and J. Chi, "Performance of passive re-circulating fiber loop buffer within an OTDM transmission link," *Optics Communications*, vol. 209, pp. 137–147, 2002.

[43] Majumder, S. P., Kabir, S. M. Raiyan, Rehman, Rezwanur., Imtiaz, Farrukh., Moniruzzaman. *A New Architecture of TDM Switch.* Bangladesh University of Engineering and Technology. http://ieeexplore.ieee.org/iel5/10170/32495/01517303.pdf?arnumber=1517303

[44] Oliver C. Ibe. *Converged Network Architectures: Delivering Voice over IP, ATM, and Frame Relay.*

codec, and define 5 or 6 levels of successive multiplexing. One main difference between the three multiplexing standards is that they use different number of voice channels multiplexed onto any given level. A standard voice signal has a bandwidth of 64Kbps, determined using Nyquist's sampling criterion[45]. TDM takes frames of the voice signals and multiplexes them into a TDM frame which runs at a higher bandwidth.

As mentioned above, this type of multiplexing is a way for many slow communication channels to share a high bit rate channel. The advantage of this scheme is that the cost per bit transmitted on a single fast channel is lower than on slower channels. Also, stations/lines are allocated the entire bandwidth of frequencies for use, but only for a small percentage of time.

Although TDM technologies has advantages of its own, implementing voice over an IP network has comparable compensation and beyond as mentioned in the earlier chapter.

2.4.2 Video Communications

Global enterprises make use of video conferencing technology for meetings that take place and require face to face communication. The advantages to using video conferencing include reduced travel costs as well as travel time for the meeting participants. This section will describe the protocol used for supporting video streaming on a LAN environment.

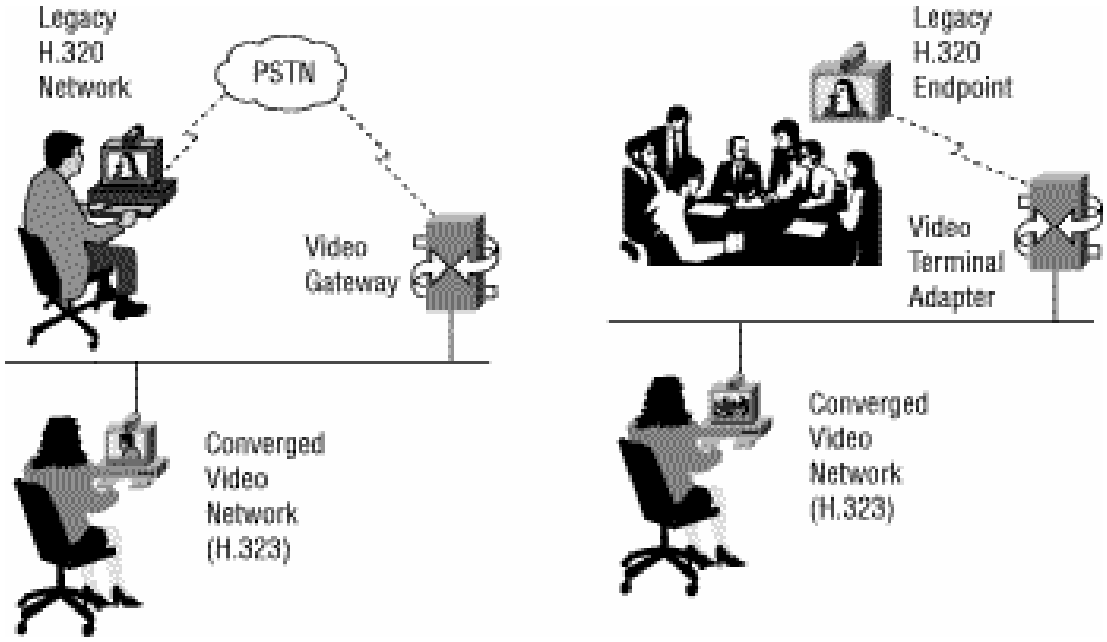
During video streaming, the information is divided into segments by the application, encoded and compressed, put into a series of data packets, and sent from the source to the destination network. The data packets that arrive at the destination might be slightly delayed, and possibly out of order. But to keep the real-time characteristic of video conferencing, the packets must arrive and be delivered on time.

---

[45] Ericsson Ltd, *Understanding Telecommunications*, http://web.archive.org/web/20040306215105/www.ericsson.com/support/telecom/part-a/a-2-7.shtml

The International Telecommunication Union (ITU) H.32x family of standards handles multimedia communications. This includes the H.320 protocol, which is used for communication over Integrated Services Digital Network (ISDN), and H.323 protocol, which serves as a communication standard aimed at the multimedia communication over LANs. Prior to H.323, the H.320 was used as a common protocol for video conference systems. Each system had its own Public Switched Telephone Network (PSTN) connection. Figure 2.5 illustrates the difference between the legacy video conferencing with H.320 and video conferencing with H.323.

Figure 2.5 – H.320 vs. H.323[46]



H.323 protocol was initiated in late 1996. The H.323 protocol has since been revised to include VoIP, and various other data communications that involve packet based networks. The H.323 standard specifies a great deal of information about the properties and components that interact within an H.323 environment[46]. It specifies the pieces that combine to provide a complete communication service, as follows[46]:

- Terminals: Computers or stand-alone devices that serve as the end points of communication lines.

[46] *Implementing QoS Solutions for H.323 Video Conferencing over IP*. Cisco Systems. http://www.cisco.com/warp/public/105/video-qos.html

- Gatekeepers: These serve as the brains of the network. They provide services such as addressing, identification, authorization and bandwidth management.
- Gateways: These are the devices that act as translators when connecting to a dissimilar network.
- Multipoint Control Units: MCUs allow multipoint conferencing, or communication between more than two parties at once.

It also described protocol standards, permissible audio and video codecs, RAS (registration, admission and status), call signaling and control signaling. H.323 specifies a mandatory level of compliance and support for these specifications for all terminals on the network[47]**Error! Bookmark not defined.**.

H.323 standards based video conferencing was engineered for the video streaming to take place on a data network without any QoS standard, such as the Internet. Such networks were not originally intended for delivery of sensitive real-time applications.

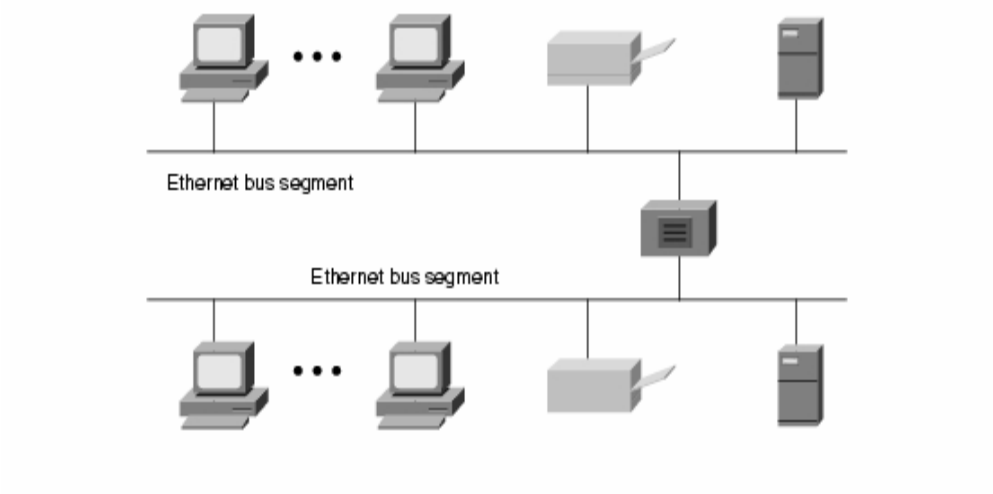
2.4.3 Data Communications

Traditionally, enterprise LAN data networks have been mainly implemented with the Ethernet. Ethernet is the most widely used LAN technology in today's enterprise networks. The Ethernet technology suite is defined by the IEEE 802.3 family of standards[48]. The initial Ethernet networks were implemented with a bus structure. Segment lengths were limited to 500 meters, and up to 100 stations could be connected to a single segment. Figure 2.6 illustrates an Ethernet network using a bus topology.

[47] *Implementing QoS Solutions for H.323 Video Conferencing over IP*. Cisco Systems. http://www.cisco.com/warp/public/105/video-qos.html

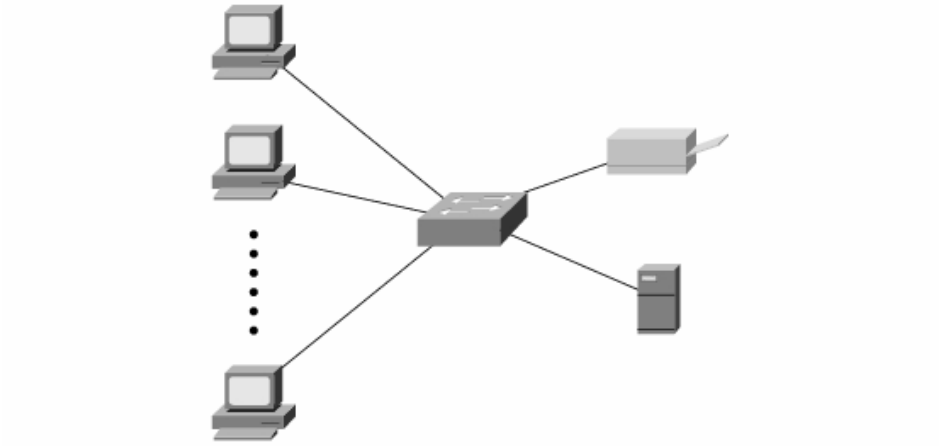
[48] *Ethernet Technologies.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/ethernet.htm

Figure 2.6 – Ethernet Bus Topology[49]



Most of the newly designed Ethernet networks are not designed for bus topology; rather the star topology has been adapted since the early 1990s. Figure 2.7 illustrates a star topological Ethernet network.

Figure 2.7 – Ethernet Star Topology[49]

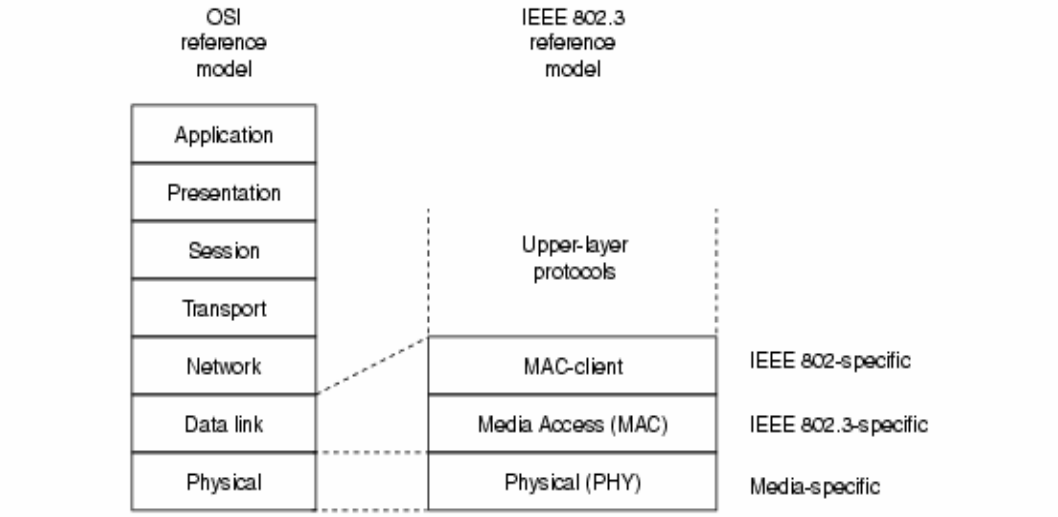


As illustrated in Figure 2.7, a hub or a switch is used as the central unit in the star network.

The IEEE 802.3 logical layers, used for Ethernet networks, correspond with the OSI seven layer model. Ethernet operates at the lower layers of the OSI reference model. Figure 2.8 illustrates the relationship between the OSI reference model and the IEEE 802.3 Logical Layers.

[49] *Ethernet Technologies.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/ethernet.htm

Figure 2.8 – OSI and IEEE 802.3 Layers Mapping[50]



As evident from Figure 2.8, the OSI data link layer is divided into two IEEE 802 sub-layers – the Media Access Control (MAC) sub-layer and the MAC-client sub-layer[Error! Bookmark not defined.]. The IEEE 802.3 physical layer corresponds to the ISO physical layer.

Ethernet can also be classified according to the speed of operation. At present Ethernet systems are capable of running at speeds of 10Mbps, 100Mbps, 1Gbps, and 10Gbps via various media[Error! Bookmark not defined.].

2.5 Converged Networks

The goal of this section is to provide an overview of one of the main technologies that is considered suitable for using in a converged environment.

2.5.1 Gigabit Ethernet

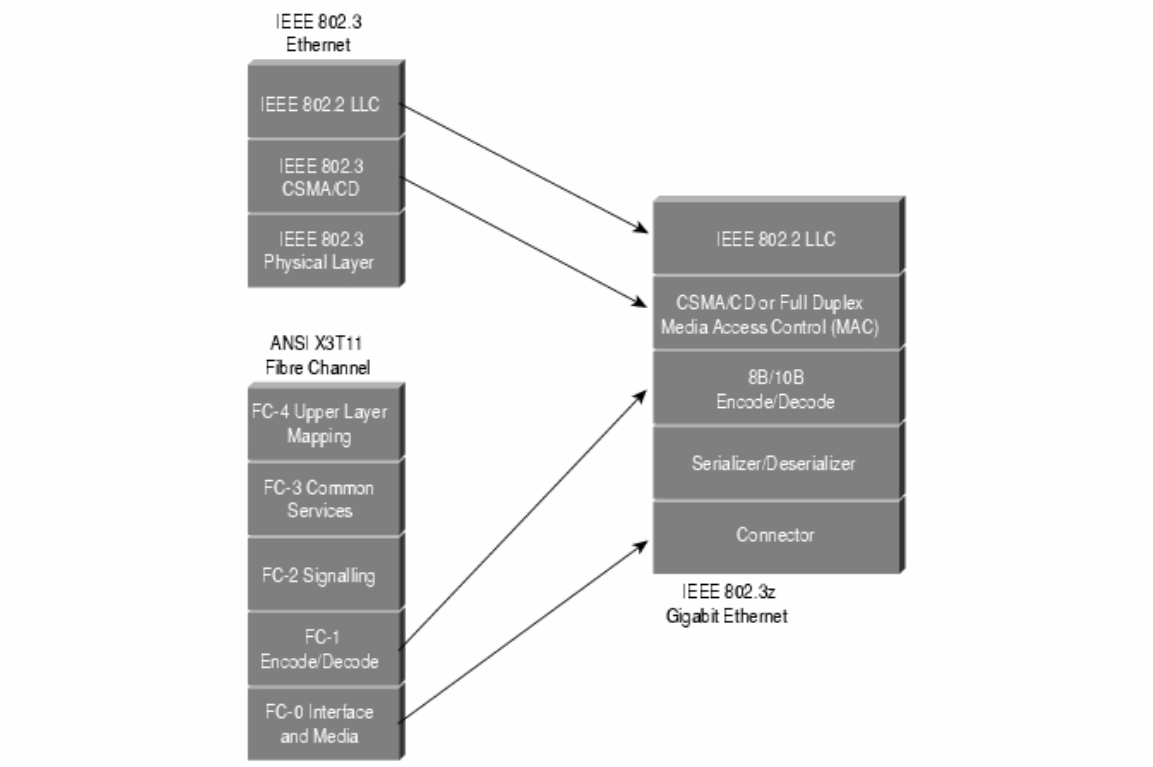
Gigabit Ethernet was developed by the IEEE 802.3 standard committee, which created the IEEE 802.3z task force. This 802.3z task force was assigned to develop a standard that would address the need for a high-speed data transfer technology. The IEEE

[50] *Ethernet Technologies.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/ethernet.htm

802.3z Gigabit Ethernet standard is an extension of the base IEEE 802.3 standard. Gigabit Ethernet has a lot in common with 802.3 Ethernet standards with respect to MAC layer characteristics and framing, but has a physical layer and data link layer that enables it to operate at a considerably higher speed[51].

The physical layer of the Gigabit Ethernet standard is modified from that of the standard Ethernet technology. Figure 2.9 depicts the physical layer of the Gigabit Ethernet. As seen in the figure below, in order to achieve 1Gbps throughput, a modified version of the ANSI X3.230 Fiber Channel standard physical layer is added to the established 802.3 Ethernet Standard.

Figure 2.9 – Gigabit Ethernet Physical Layer[51]

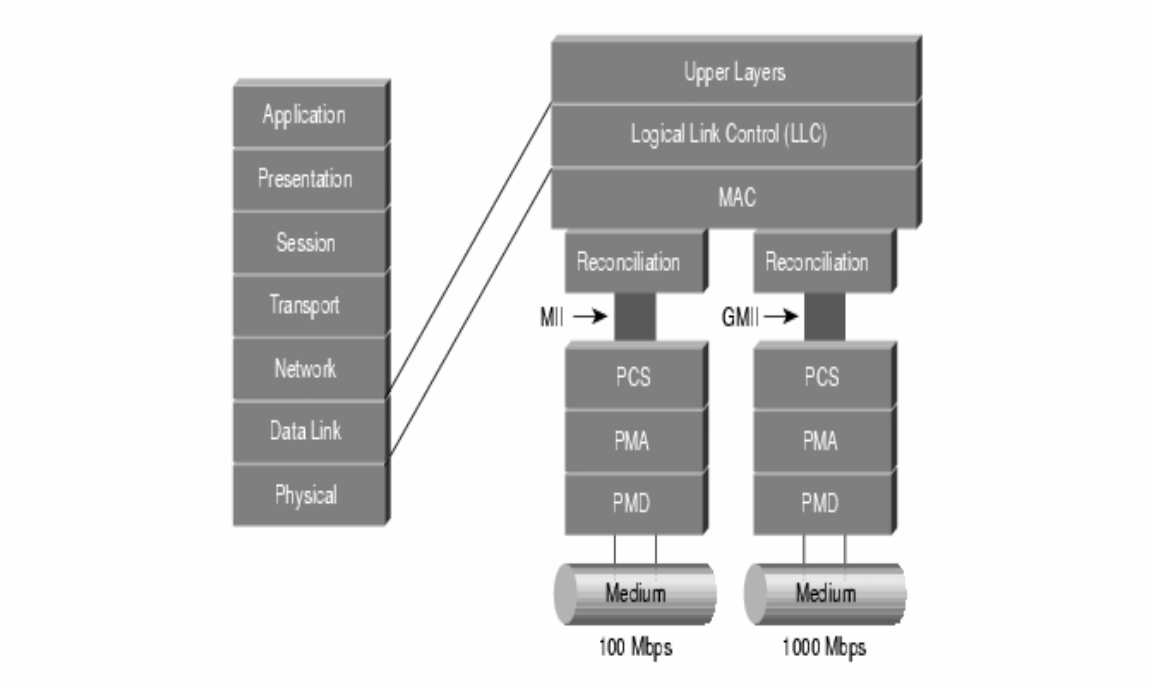


Fiber Channel was adopted because it was a technically advanced, readily available and commercially proven technology. The Fiber Channel technology uses long wavelength lasers to transmit data over fiber optic cable. Also, as per the standard specification, the Gigabit Ethernet uses IEEE 802.3 Ethernet frame format, employs the

[51] *Introduction to Gigabit Ethernet.* Cisco Systems. http://www.cisco.com/en/US/tech/tk389/tk214/tech_brief09186a0080091a8a.html

same half- and full-duplex MAC operation schemes as the 802.3 family, is backward compatible with 10Mbps and 100Mbps Ethernet standards, and supports all the protocols that are used in conjunction with the 802.3 Ethernet family. Figure 2.10 demonstrates the actual architectural model of the Gigabit Ethernet standard.

Figure 2.10 – Architectural Model of 802.3z Standard[52]



As mentioned earlier, the Gigabit Ethernet MAC layer operation can be in either half- or full-duplex mode. That is, in a half-duplex channel can receive and transmit, but not at the same time. With full-duplex transmission, it is possible to transmit and receive data at the same time. Another difference between the half- and full-duplex modes is that in full-duplex mode, the 802.3z MAC layer uses the IEEE 802.3x flow-control specification, whereas in half-duplex mode, the MAC layer uses the CSMA/CD access method. The full-duplex operation does not use the CSMA/CD for collision control because full-duplex operation eliminates congestion in the transmission media.

As seen in Figure 2.10, the Gigabit Media Independent Interface (GMII) is the interface between the MAC layer and physical layer. GMII allows the implementation of any of the physical layers with the same MAC layer. That is, the GMII supports 10Mbps,

[52] *Introduction to Gigabit Ethernet.* Cisco Systems. http://www.cisco.com/en/US/tech/tk389/tk214/tech_brief09186a0080091a8a.html

100Mbps, and 1Gbps data rates[53]. Also, it can support both full-and half-duplex modes of operation. The GMII provides two media status signals – one indicating the presence of the carrier and the other indicating the absence of a collision[54]. As evident from Figure 2.10 above, the GMII is positioned above three main sub-layers – Physical Coding Sub-layer (PCS), Physical Medium Attachment (PMA) and Physical Medium Dependant (PMD)[55].

- PCS provides a uniform interface to the Reconciliation layer for all of the possible physical media. The PCS sub-layer also generates the CSMA/CD indications for half-duplex operation.
- PMA provides a medium independent means for the PCS to support various serial bit-oriented physical media.
- PMD maps the physical medium to the PCS. It defines the physical layer signaling used for various media.

In conclusion, 802.3z is an Ethernet standard offering speeds in the gigabit per second range. It is backward compatible with its predecessors. Enterprises that opt to go for Gigabit Ethernet implementation when migrating to converged solution can make use of their existing Ethernet networks by upgrading the network performance without having to change wiring, protocols or applications.

2.6 Protocols

A basic understanding of the principles of Internet Protocol (IP) and other associated networking protocols is necessary to comprehend the concepts of real-time traffic over IP. As discussed in RFC 791, the Internet Protocol is designed for use in

---

[53] *Introduction to Gigabit Ethernet.* Cisco Systems. http://www.cisco.com/en/US/tech/tk389/tk214/tech_brief09186a0080091a8a.html

[54] *Gigabit Ethernet Technology and Solutions.* Intel Corporation. http://www.intel.com/network/connectivity/resources/doc_library/white_papers/gigabit_ethernet/gigabit_ethernet.pdf

[55] Vijay Moorthy. *Gigabit Ethernet.* Washington University. http://www.cs.wustl.edu/~jain/cis788-97/ftp/gigabit_ethernet/index.htm#PCS

interconnected systems of packet-switched communication networks[56]. With reference to the TCP/IP reference model discussed above, the Internet Layer transfers packets from one host to another host. The protocol that operates the Internet Layer is known as the Internet Protocol (IP). IP provides for transmitting blocks of data called datagram from sources to destinations. IP, along with TCP, can provide a reliable end to end communication service. Enterprise LANs use TCP/IP as the transmission protocol over Ethernet (IEEE 802.3 standard) physical medium. Other protocols, however, are necessary for data transmission from one host to another. This section will discuss few protocols that make up the essence of real-time communication over IP.

2.6.1 Multi Protocol Label Switching (MPLS)

MPLS is an IETF specification that provides for switching and routing to forward packets using fixed-length labels[57]. Although it uses Layer 2 switching and Layer 3 routing mechanisms, it remains independent of Layer 2 and Layer 3 protocols. A label in a MPLS network identifies the path a packet should traverse, and also identifies the underlying protocol. These labels are distributed across the MPLS network using the Label Distribution Protocol (LDP)[58]. A reliable protocol should be used as the signaling protocol – LDP uses TCP.

In MPLS, Label Switched Paths (LSPs) are used for data transmission. LSPs are a sequence of labels at every networking node along the path of transmission. The two methods by which LSPs can be established are: control-driven and data-driven. If the LSP is established prior to exchanging/transmitting data, then it is know as a control-driven LSP, whereas if the LSP is established upon detection of data flow, then it is known as data-driven LSP[57].

---

[56] *Internet Protocol*. IETF. http://www.ietf.org/rfc/rfc0791.txt

[57] *MPLS / Tag Switching.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/mpls_tsw.htm

[58] *Multi Protocol Label Switching.* IEC. http://www.iec.org/online/tutorials/mpls/topic03.html

The devices that participate in the MPLS transmission mechanisms are classified into Label Edge Routers (LERs) and Label Switching Routers (LSRs). LSR is a network device in the core of an MPLS network that participates in the establishment of LSPs using the appropriate signaling protocol and high-speed switching of the data traffic based on the established paths. LER is a device that operates at the edge of the access network and MPLS network. LERs support multiple ports connected to dissimilar networks and forwards this traffic on to the MPLS network after establishing LSPs. The LER plays a very important role in the assignment and removal of labels, as traffic enters or exits an MPLS network.

The Forward Equivalence Class (FEC) is a representation of a group of data packets that share the same requirements for their transmission. The assignment of a packet to a specific FEC is done when the packet enters the network. FECs are based on service requirements for the given packet[59]. Each LSR in the MPLS network builds a table to classify how a packet must be forwarded. This table is called a Label Information Base (LIB), which comprises of the FEC to label mapping[59]. A packet traversing through a MPLS network goes through the following steps[60]:

1. Label Creation and Distribution: When the packet arrives at the source networking device and the device labels the packet and assigns the packet to a specific FEC. At this stage, traffic related characteristics and MPLS capabilities are also negotiated using the LDP.
2. Table Creation at Routers: Once the labels are created for a packet, the LSR creates entries in its LIB. The contents o this table will map the label to a specific FEC. These entries are updated whenever negotiations of the labels occur.
3. LSP Creation: After the LIB creation, LSPs are created in the reverse direction to the creation of entries in the LERs LIB.
4. Label Insertion and Table Lookup: The source router uses the information in the LIB table to find the next hop and also requests a label for the specified FEC.

[59] Harry G. Perros. *Multi Protocol Label Switching Architecture.* Connection Oriented Networks: SONET/SDH, ATM, MPLS and Optical Networks.

[60] *Multi Protocol Label Switching.* IEC. http://www.iec.org/online/tutorials/mpls/topic03.html

Routers in the path of transmission just use the label to find the next hop. At the destination router, the label is removed and the packet is transmitted to the destination device.

MPLS is one of the widely used solutions to address the problems of QoS, bandwidth management, and traffic engineering. IP MPLS VPN will be discussed in the following chapters to give an overview of how QoS goals and security measures can be met using MPLS VPN solution.

2.6.2 Session Initiation Protocol (SIP)

SIP is a protocol used to establish IP calls. SIP is an application-layer control and signaling protocol for creating, modifying and terminating sessions with one or more participants. SIP is being developed by the SIP Working Group, within the IETF. RFC 2543 describes the basic operation of the SIP protocol[61]. Besides this RFC that describes the basic specifications, a number of extensions to SIP have been defined in other RFCs.
SIP is used to establish, modify, and terminate IP multimedia sessions. SIP is independent of the type of multimedia session handled and of the mechanism used to describe the session[62]. It can very well be used for videoconferencing, and IP calls.

SIP protocol also defines several network elements, and the following is a brief discussion of the important elements in a SIP session.

1. User Agents: User Agent (UA) is the SIP entity that interacts with the user. It usually has an interface towards the user. User Agent is usually the endpoint entity. User Agents initiate and terminate sessions by exchanging SIP requests and responses. RFC 2543 defines the User Agent as an application that controls both a User Agent Client (a client application initiating SIP requests) and User Agent Server (a server application that responds to SIP requests from clients)[61]. Figure 2.11 illustrates some common devices identified as SIP User Agents.

---

[61] *SIP: Session Initiation Protocol.* http://www.ietf.org/rfc/rfc2543.txt

[62] Mark. A. Miller. *Protocols for Converged Networks.* Voice over IP Strategies for Converged Networks.

Figure 2.11 – SIP User Agents[63]

2. Redirect Servers: Redirect servers help locate SIP User Agents by providing alternative locations where the user can be reachable. It accepts a SIP request from User Agents, maps the SIP address of the called party and responds with possible called party location information to the requesting client. The redirect server does not actually locate a user, but merely returns a list of possible locations where the user might be.
3. Proxy Servers: Proxy servers act as a third party entity that acts as both a server and a client. Requests are examined either internally or by passing them on to other appropriate servers. A proxy server interprets, and rewrites all the request messages before forwarding it[63].
4. Registrar: A Registrar is a server that accepts register requests from clients for the purpose of updating the location database with the contact information of the user specified in the request. A registrar is usually co-located with a redirect server or a proxy server[64].

SIP is a text-encoded protocol based on elements from the Hyper Text Transport Protocol (HTTP). SIP network entities exchange SIP messages, which can be categorized into requests and responses. The following section explains the message categories in detail:

[63] Gonzalo Camarillo. *The Session Initiation Protocol.* SIP Demystified.

[64] *SIP: Session Initiation Protocol.* http://www.ietf.org/rfc/rfc2543.txt

- Requests: Table 2.2 illustrates sample SIP requests. The first six methods are defined in RFC 2543, the base SIP specification[65][64]. The other SIP requests are described in detail in various other RFCs or Internet Drafts related to SIP.

Table 2.2 – SIP Requests Sample

SIP Request	Explanation
INVITE	Session setup
ACK	Acknowledgment of final response to INVITE
BYE	Session termination
CANCEL	Pending session cancellation
REGISTER	Registration of a user's URL
OPTIONS	Query of options and capabilities
INFO	Midcall signaling transport
PRACK	Provisional response acknowledgment
COMET	Preconditions met notification
REFER	Transfer user to a URL
SUBSCRIBE	Request notification of an event
UNSUBSCRIBE	Cancel notification of an event
NOTIFY	Transport of subscribed event notification
MESSAGE	Transport of an instant message body

- Response: SIP Response messages contain numeric codes. As mentioned earlier, SIP is partly based on elements from HTTP; this is evident in the response codes as it is based on HTTP response codes. There are two primary types of responses and six different response classes. The two types are[66]:
    a. Provisional – This type of response is used by the server to indicate progress of a SIP session, but it does not terminate SIP connections that

[65] *SIP: Session Initiation Protocol.* http://www.ietf.org/rfc/rfc2543.txt

[66] Henry Sinnreich and Alan B. Johnston. *SIP Overview.* Internet Communications Using SIP: Delivering VoIP and Multimedia Services with Session Initiation Protocol.

are up. The first class (discussed below) fall under this type of SIP response.

b. Final – This type of response is used to terminate SIP sessions. Classes 2 – 6 (discussed below) fall under this type of SIP response.

The six different classes of SIP responses are[6766]:

1. 1xx – Responses that follow this class indicate ringing, queuing, searching, etc. For example, SIP response 180 indicates ringing state.
2. 2xx – Responses that follow this class indicate any kind of SIP session success message. For example, SIP response 200 indicates OK.
3. 3xx – Responses that follow this class indicate redirection or forwarding of SIP requests. For example, SIP response 301 indicates party having moved permanently.
4. 4xx – Responses that follow this class indicate SIP request failure due to error in the client side. For example, SIP response message 408 indicates request time-out.
5. 5xx – Responses that follow this class indicate failure due to error in the server side.
6. 6xx – Responses that follow this class indicate failure due to global error. For example, SIP response message 606 indicates not acceptable.

SIP messages are organized in either the request method or response code, followed by a list of message fields. The three main parts of a SIP message include start field, header field and the body field[67].

- Start Field: This indicates the beginning of any SIP message. It contains the information about the message type (request or response) and the protocol version.
- Header Field: This field is used to share message attributes information and to modify the message meaning. It might include data such as route request, contact information, signaling information, etc.

[67] Henry Sinnreich and Alan B. Johnston. *SIP Overview*. Internet Communications Using SIP: Delivering VoIP and Multimedia Services with Session Initiation Protocol.

- Body Field: This field describes information regarding the session type (audio, video, multimedia), codec types being used, and any other information relevant to the session.

As indicated by the IETF, end-to-end protocols are better for providing end-to-end services[68]. Accordingly, IP was developed to be an end-to-end protocol. Similarly, SIP provides end-to-end connectivity between users with SIP servers. In conclusion, SIP can be called an efficient protocol because all intelligence in a SIP network is located in the User Agents. The messages sent by SIP User Agents contain all the routing information that the SIP server need; so the servers are not required to maintain transaction information.

2.6.3 Session Description Protocol (SDP)

The Session Description Protocol is defined in RFC 2327 developed by the IETF[69]. SDP is used in describing the multimedia sessions (audio or video) establishes using SIP. Generally, SDP contains the following information about the media session:

- IP Address
- Port number (TCP or UDP)
- Media Type (audio, video or multimedia)
- Codec Information
- Timers

Similar to SIP, SDP uses text coding. An SDP message is composed of different lines called fields. The field names are specified by a single lower-case letter, and are in a pre-defined order to simplify parsing. Table 2.3 provides description about few of the SDP fields.

[68] Gonzalo Camarillo. *The Session Initiation Protocol.* SIP Demystified.

[69] *Session Description Protocol.* IETF. http://www.ietf.org/rfc/rfc2327.txt

Table 2.3 – SDP Fields[70]

Field	Description
v=	Protocol version number
o=	Owner/creator and session identifier
s=	Session name
i=	Session information
u=	Uniform Resource Identifier
e=	Email address
p=	Phone number
c=	Connection information
b=	Bandwidth information
t=	Time session starts and stops
r=	Repeat times
z=	Time zone corrections
k=	Encryption key
a=	Attribute lines
m=	Media information
a=	Media attributes

As mentioned earlier, SDP is widely used with SIP. SDP was originally developed for scheduled multicast sessions. Therefore, many of the SDP fields have little or no significance in the perspective of multimedia sessions established using SIP. But, in order to maintain compatibility with SDP all required fields are to be included. For example, a typical SIP use of SDP includes the version, origin, subject, time, connection, media, and attribute fields. Although the origin, subject, and time fields are not used by SIP it is included for compatibility.

[70] *SDP*. Telecom Paris. http://www.infres.enst.fr/~dax/polys/multicast/sdp.html

2.6.4 Real-time Transport Protocol (RTP) and RTP Control Protocol (RTCP)

IP-based networks provide connectionless data transport. Thus, there is an absence of delay characteristics. This might function fine with non real-time data services, but real-time applications such as voice or video cannot be sent through a network that does not guarantee reliable delivery. Real-time Transport Protocol was developed to make reliable transport of real-time voice or video packets. RTP is defined in RFC 1889[71].

RTP is an application layer protocol that uses UDP for transport over IP network. RTP uses a bit-oriented header similar to UDP and IP. RTP provides end-to-end delivery services for data that requires real-time support, such as VoIP or Video over IP services. According to RFC 1889, the services provided by RTP include payload type identification, sequence numbering, time stamping, and delivery monitoring[71]. Applications typically run RTP on top of UDP to make use of UDP's multiplexing and checksum services, and as such both RTP and UDP contribute parts of the transport protocol functionality. Provisions are defined, however, to use RTP with other underlying network or transport protocols.

RTP, to some extent, allows for the detection of some drawbacks introduced by IP, such as packet loss, jitter, out of sequence packet arrival, etc. It is also important to note that RTP does not provide certain features and functions. For example, RTP, by itself, does not provide any mechanism to ensure timely delivery or provide other QoS guarantees, but relies on the lower layer services for these functions.

RTP allows detection of a missing or lost packet by a difference in the sequence number. RTP allows detection of these transport related problems but leaves it up to the codec to deal with the problem. RTP does not guarantee packet delivery or prevent out-of-order packet delivery. The sequence numbers included in RTP allow the destination network device to reconstruct the received packet sequence; sequence numbers might also be used to determine the proper location of a packet.

---

[71] *RTP: A Transport Protocol for Real-time Applications.* http://www.ietf.org/rfc/rfc1889.txt

There are two parts of RTP as defined in RFC 1889 – the RTP stream, which carries data that has real-time properties, and the RTCP stream that monitors the QoS and conveys information about the participants in a current data session. In addition to the protocol specification given in RFC 1889, an accompanying document, RFC 1890[72], describes a profile for the use of the RTP and RTCP within audio and video multi-participant conferences with minimal control. It provides a profile specification that defines a set of payload type codes and their mapping to payload formats, such as various media encodings.

RTCP allows participants in an RTP session to send each other QoS reports and statistics, and exchange profiling information. The following are the different types of RTCP packets[73]:

1. Sender Report (SR): Sent by a participant that both sends and receives RTP packets.
2. Receiver Report (RR): Sent by a participant that only receives RTP packets.
3. Source Description (SDES): Contain information about the participant in the session including e-mail address, phone number, and host.
4. Bye (BYE): Sent to terminate the RTP session.
5. Application Specific (APP): Defined by a particular profile.

The use of RTCP packet types allows feedback on the quality of the real-time connection; for example, number of packets sent and received; number of packets lost; packet jitter; etc. RTP, along with RTCP, provide some level of QoS to real-time packets in an IP network.

2.7 Conclusion

This chapter covered the various technologies, topologies and characteristics of voice, data and video on a legacy and converged network. It was important to understand

[72] *RTP Profile for Audio and Video Conferences with Minimal Control.* IETF. http://www.ietf.org/rfc/rfc1890.txt

[73] *RTP Packet Types.* Freesoft.org. http://www.freesoft.org/CIE/RFC/1889/47.htm

the operation of an IP network and the different protocols involved, to appreciate the scope of convergence. The following chapter will focus on the QoS and VPN requirements and implementation on converged IP network environment.

3 QoS and VPN Solutions

3.1 Introduction

As discussed in an earlier chapter, converging networks can result in both cost savings and productivity enhancements as there is only one network that is implemented, maintained, and managed. However, while IP-based networks present excellent quality for non real-time data networking, the network by itself is not capable of providing reliable, quality and secure services for real-time traffic. In order for IP networks to perform reliable and timely transmission of real-time data, additional mechanisms to reduce delay, jitter and packet loss are required. This chapter will discuss the important mechanisms for running real-time traffic like voice and video over an IP network.

3.2 Voice over TDM vs. Voice over IP

In order to understand the concepts of packet loss, jitter, and delay in the VoIP world it is important to understand how IP telephony network differs from the traditional time division multiplexing network. For the most part, a traditional voice call placed over the PSTN is transported over a TDM network. In a TDM network, a voice call uses a fixed amount of bandwidth, usually 64Kbps, throughout the duration of that call. This bandwidth is independent of whether or not the called or calling part is actually speaking. This technique is referred to as circuit switching.

Contrary to the TDM call in an IP network, fixed bandwidth is not assigned to any application. IP networks are packet switched networks. Packet switching is inherently more efficient than circuit switching. That is because in a packet switched network, there is never a time when there is an application needing bandwidth and yet bandwidth being idle.

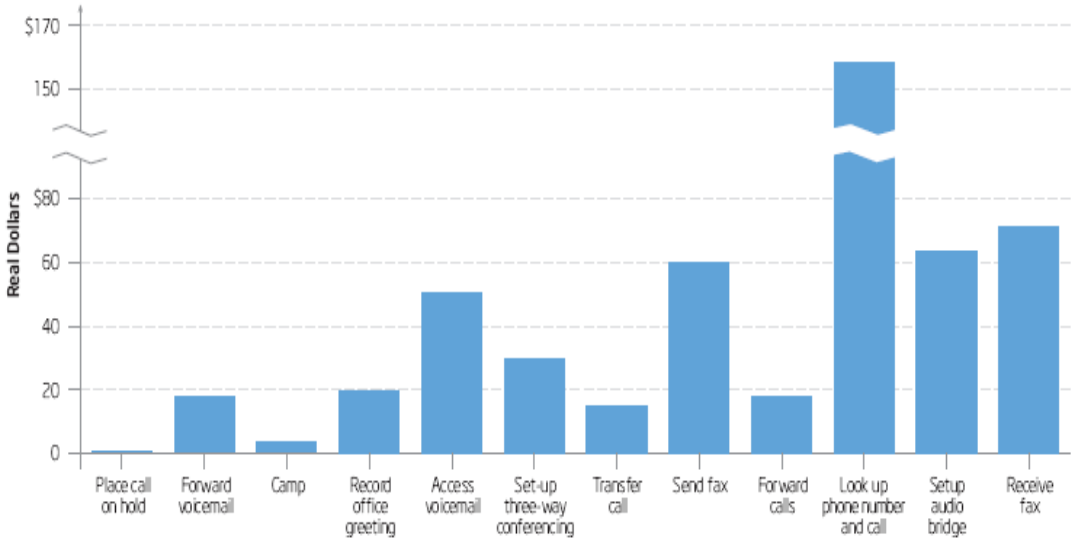
However, packet switching also means that traffic from one application can impact the performance of another application. This is not possible in a TDM network. Therefore, in the world of IP telephony, when someone makes a call at the same time a computer is attempting to transfer a large file through the network, the IP call is bound to

experience delay because of the contention for the transmission facility. As the delay of the IP Telephony packets increases, so does the jitter. This would not be a problem in a phone call being placed through the TDM network because the bandwidth would be dedicated to the voice call and additional bandwidth would be dedicated to the file transfers. As such, there would be no contention for the bandwidth, and hence no added delay.

In an IP network, each packet has to traverse through various routers before it reaches the destination. Because routers process each packet, each router in the network becomes a bottleneck. That is, if many of the devices that terminate on a particular router decide to use the network at the same time, then the router becomes loaded with data packets. The router might not be able to process all the packets at the same time. Therefore, the router stores each packet for a brief period of time before it can process it. However, if the network transmission continues past that brief duration, the router might start dropping packets. This process accounts for packet loss in the VoIP network, which is not acceptable for real-time traffic. This cannot happen in a TDM voice network. Since a dedicated channel is provided for the duration of each call, there is no need for the network to worry about dropping packets due to increased traffic.

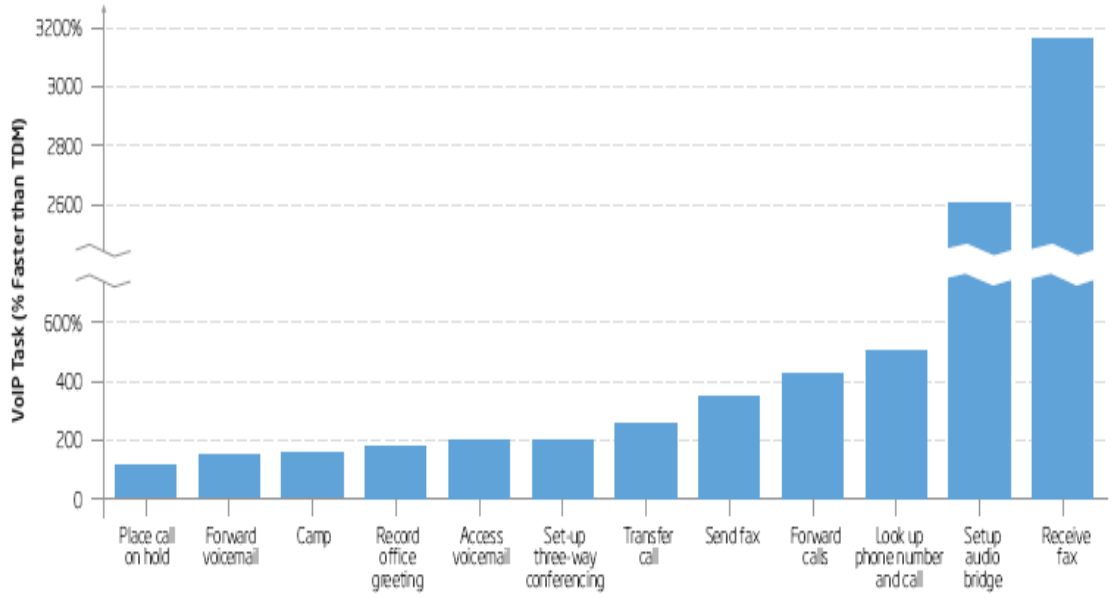
The problems with delay, packet loss and jitter can be overcome by implementing proper QoS measures on the IP network. This will make the IP network more reliable for voice traffic. Implementing IP telephony in an organization has obvious cost advantages – in the areas of moves/adds/changes, cabling/wiring, audio conferencing, etc. Figure 3.1 illustrates the various areas where an enterprise can save monthly costs by migrating to IP telephony.

Figure 3.1 – Cost Savings (IP Telephony vs. TDM)[74]



Along with the cost saving, deploying IP telephony also increases productivity. As seen in Figure 3.2, performance generally increased between 134 to 500 percent using the IP telephony.

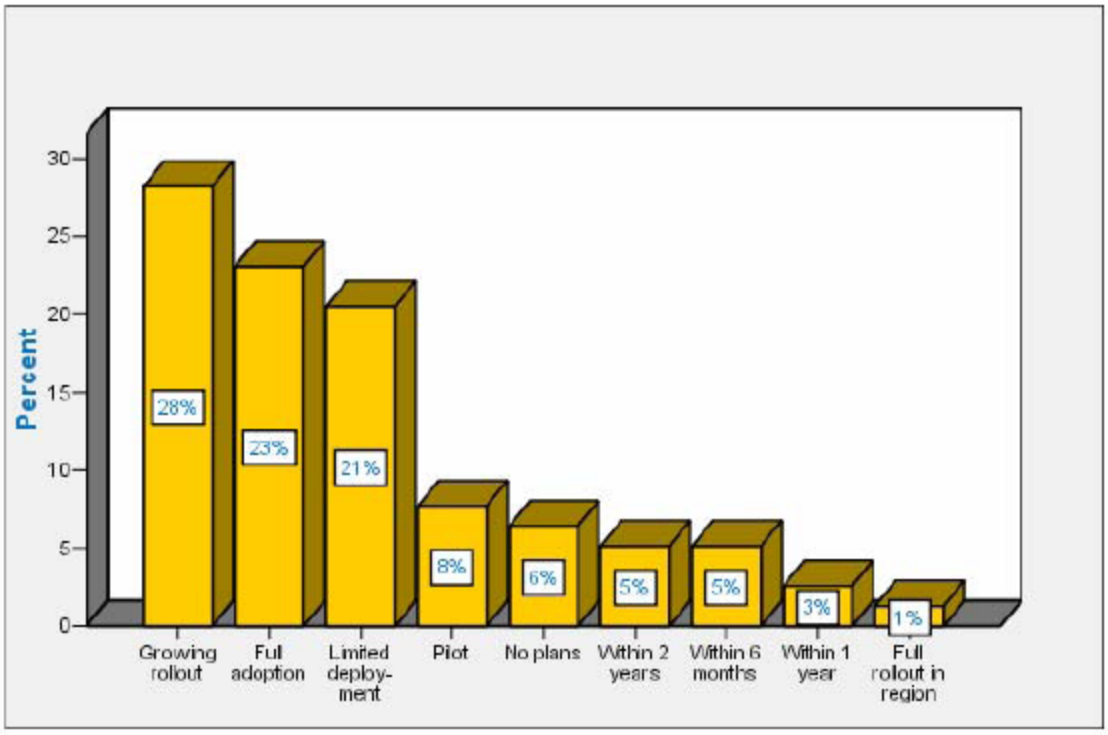
Figure 3.2 – Increased Productivity with IP Telephony[74]



[74] *The Business Case for Enterprise VoIP*. Intel Information Technology. http://www.intel.com/it/pdf/parsippany-voip.pdf

Therefore, it is evident that IP telephony is the answer to enterprises looking to implement next generation solutions, and save costs. According to a survey by Nemertes only about 6% of the companies are doing nothing when it comes to deploying IP telephony systems. About 80% of the companies that took part in the survey are using IP telephony in someway today, and the rest are planning to deploy it in the following years. Figure 3.3 illustrates the VoIP state of deployment.

Figure 3.3 – VoIP State of Deployment[75]



The differences between a TDM voice network and a VoIP network has to be measured carefully before implementing a converged network.

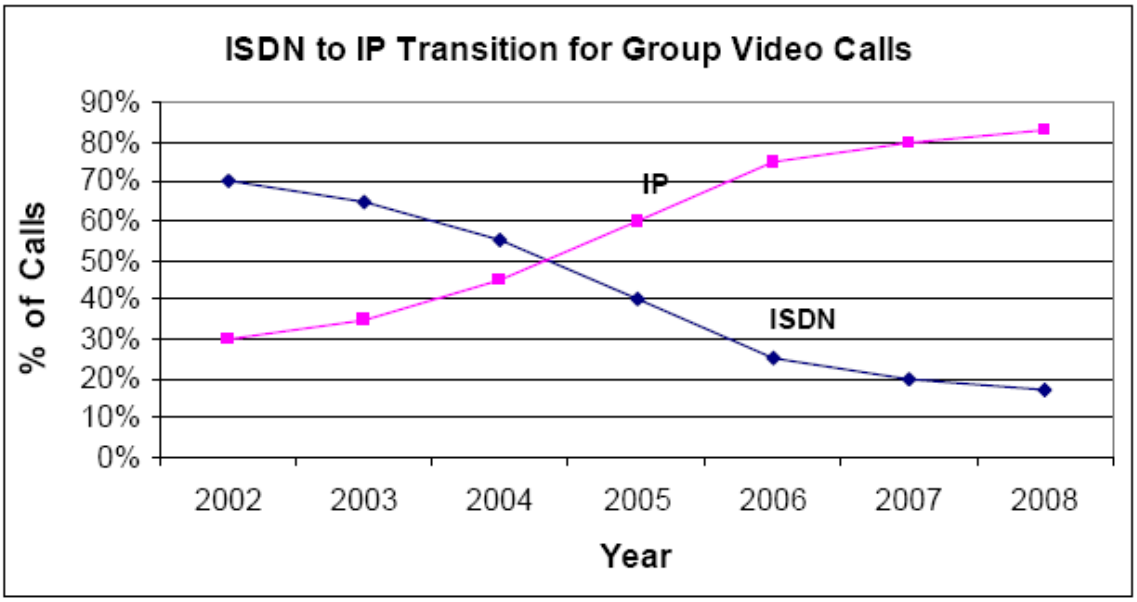
3.3 Video over ISDN vs. Video over IP

Traditionally, video needs of businesses have been taken care of by implementing video over ISDN using the H.323 protocol. As the concept of converged solution is

[75] Melanie Turek. *Voice and Video over IP: Leveraging Network Convergence for Collaboration.* http://www.polycom.com/common/pw_cmp_updateDocKeywords/0,1687,5742,00.pdf

emerging, the need to design robust IP networks to accommodate video traffic increases. IP is advantageous to use for videoconferencing purposes compared to ISDN in areas such as cost, reliability and scalability, if implemented properly. Figure 3.4 illustrates the shift from ISDN to IP in the area of videoconferencing. It is evident from the figure that since 2004 IP became the most common network used for hosting videoconference calls.

Figure 3.4 – ISDN to IP Shift in Videoconferencing[76]



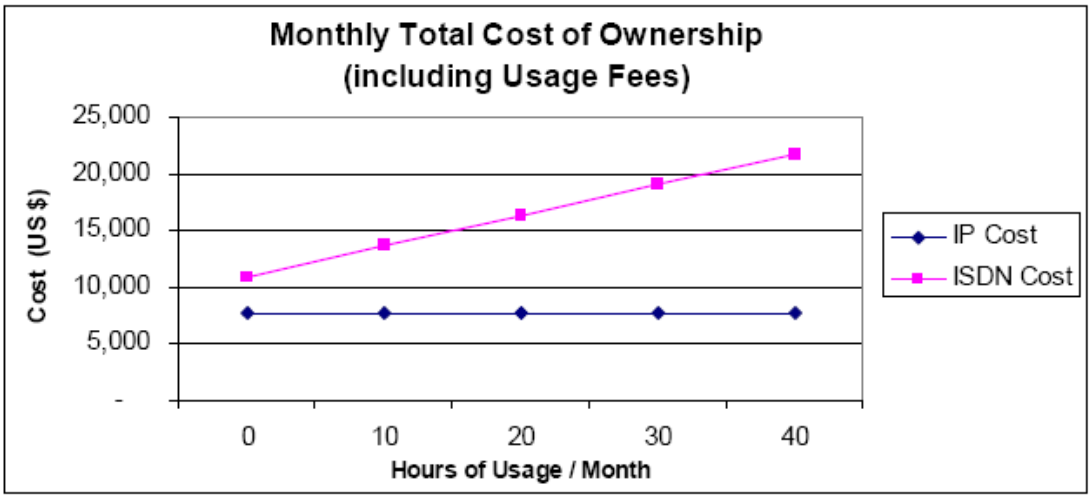
ISDN might be easy and inexpensive to implement, but it is costly to use. Apart from the initial capital to implement the ISDN network for videoconferencing, bandwidth costs for ISDN channels add to the expenses. Videoconferencing over ISDN usually requires 384Kbps of bandwidth – which comes to 6 ISDN B channels put together. Usually, enterprises pay for ISDN services on per-channel basis, but the cost might also vary upon distance. This makes the use of ISDN costly. These expensive ISDN usage costs often prohibit the adoption of ISDN for videoconferencing by any enterprise. In contrast, IP is based on flat rate pricing, which makes it a reasonable replacement for ISDN[77]. Since, IP is affordable and is one of the most adapted emerging technologies, IP

[76] *The ISDN to IP Migration for Videoconferencing.* Wainhouse Research. http://www.wrplatinum.com/Downloads/6128.aspx

[77] *Why IP based Videoconferencing?* Hitachi Software Engineering America, Ltd. http://host271.ipowerweb.com/~hitachi-/tsg/solutions/video/ipbased.html

videoconferencing can be deployed across the enterprise economically. Figure 3.5 illustrates the cost analysis for ISDN vs. IP videoconferencing.

Figure 3.5 – ISDN vs. IP Videoconferencing Cost Comparison[78]



As mentioned before, while using ISDN for videoconferencing a minimum of 6 ISDN B channels are bonded together. This can be unreliable because if one the channels is dropped during the conference, the entire conference call is torn down. IP networks do not have channels similar to ISDN, so there is increased reliability. Enterprises that use ISDN are delighted to achieve a 92-94% success rate whereas companies using IP videoconferencing often achieve greater than 99% reliability[79].

When video is deployed in a converged IP-based network, it makes the management of the network easier. Video over IP systems have constant connectivity to the enterprises' backbone network. This stable connectivity allows these systems to be remotely controlled and managed. Enterprise IP-based conferencing environments use a software product called the gatekeeper, which allows easier performance management over the network.

The differences between an ISDN video network and IP video network has to be measured carefully before implementing a converged network. The following sections

[78] *The ISDN to IP Migration for Videoconferencing.* Wainhouse Research. http://www.wrplatinum.com/Downloads/6128.aspx

[79] Saqib Jang, Brent Kelly, Andrew W. Davis. *A Technical FAQ.* http://www.wainhouse.com/files/papers/wr-faq-ip-conf.pdf

will discuss in detail some of the processes that need to be considered to determine if an IP network would be robust enough to support real-time traffic.

3.4 Quality of Service

Now that the differences between TDM vs. IP for voice and ISDN vs. IP for video traffic have been discussed, it is important to understand the concepts behind implementing a network that would provide comparable performance and reliability measures in the IP world.

A network must be able to handle bandwidth requirements in order to accommodate different types of traffic. Even when a network is designed for optimal bandwidth management, there could be contention for bandwidth during short periods of time. The three main concerns while implementing real-time traffic through an IP-based network are jitter, delay and packet loss. Delay is defined as the time it takes for a data packet to reach its destination from its source. Jitter is the variation in the time between packets arriving, caused by network congestion, timing drift, or route changes[80]. Packet loss occurs when networking devices drop data packets. When traffic arrives at a network device such as a router, the device processes and forwards the traffic. The amount of traffic that the router can forward is limited by the capacity of its interfaces. Therefore, when too many packets arrive at the router at the same time, it cannot forward all the packets immediately. This leads to the router dropping a packet, which is referred to as packet loss.

A major obstacle in deploying a converged network solution has been the inability to provide the required network quality needed by voice and video applications. As enterprises increase the capacity of their networks to accommodate voice, video and data traffic over IP, the network must be designed such that the performance is not degraded in any way. Applications such as email and file transfer are not sensitive to delay and jitter. On the other hand applications such as voice and video are susceptible to loss, delay and jitter. To provide conventional performance measures, real-time applications

[80] *Jitter*. Search Networking. com http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci213534,00.html

require significant bandwidth with minimal delay, jitter and loss. These factors are not permissible while transferring real-time data across a network. Therefore, there arises a need to either over provision the network with extra bandwidth to ensure all applications can be satisfied at all time or implements QoS techniques to make optimal use of the available resources. The first option, although ideal, is not feasible. Bandwidth is not available in abundance. Therefore, applying the latter option of defined QoS measures will lead enterprises to implement a reliable network that can carry both real-time and non real-time traffic.

In any network, with both real-time and non real-time traffic, a measure should be implemented such that the network devices are able to differentiate among the arriving traffic and satisfy their requirements on an individual basis. QoS mechanisms provide a means that can be used to differentiate traffic types, and fulfill requirement needs. They enable the devices to recognize traffic type and provide preferential services to it. Therefore, QoS is the best way to handle network congestion and provide resources to the various applications. QoS is only used to manage the resources in a network according to predefined policies; it does not create any additional network capacity[81]. This section discusses different Quality of Service (QoS) building blocks, the most popular QoS architectures, and solutions for obtaining the required QoS from the network for real-time traffic.

3.4.1 Classification Mechanisms

There are different QoS mechanisms that are used to help manage the utilization of network resources. These means provide the functionality to implement QoS in the network. This section will discuss in detail the QoS mechanisms.

[81] *Quality of Service Networking.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/qos.pdf

3.4.1.1 Packet Classification

This is a QoS classification tool in a network device that is capable of separating the arriving traffic into different classes such that each class of traffic maybe provided with distinct preferences and services. This classifier is usually present in the IP header. The network devices look at the explicit markings in the type of service field. Once the traffic is classified into multiple classes, it is possible for each class of packets to receive different treatment from the device[82].

3.4.1.2 Queuing

Packets are assigned to different queues based on their classification. This mechanism helps meet diverse service needs. That is, voice or video packets that have low delay requirements can be provided by examining and servicing their queue more frequently. Based on this type classification, the queuing application on the network device determines which queue packets enter and when packets exit the queues. The queuing application configured in the network device also executes the scheduling. It determines how the packets in each of the queues are scheduled for transmission[82]. A few queuing models that provide QoS are priority queuing, round robin queuing and weighted fair queuing.

Priority queuing is a simple queuing model. In this model, packets are forwarded based strictly on the priority of the queue that they are in. The packets in the high priority queue are always forwarded first. If there are no packets in the high priority queue then the packets in the medium priority queue are forwarded and so on. High priority packets are never scheduled behind lower priority packets. Therefore, this model provides minimal latency service to the packets in the high priority queue. This might lead to starvation of traffic in lower priority queues. Hence, it is recommended to use priority queuing with effective policies of how much traffic can enter the high priority queue.

---

[82] *Quality of Service Networking.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/qos.pdf

With round robin scheduling the queues are given turns at sending packets. There are many different algorithms within the round robin scheduling category, such as simple round robin, weighted round robin, deficit weighted round robin, and self-clocked fair queuing[83]. Unlike priority queuing, round robin scheduling gives every queue a chance to forward packets, so no queue is ever totally starved. However, round robin scheduling can be disadvantageous because the packets belonging to real-time traffic can be delayed as they wait for their transmission turn.

Weighted Fair Queuing (WFQ) is another scheduling technique that allows guaranteed bandwidth services. The purpose of WFQ is to let several sessions share the same link. WFQ can also manage duplex data streams such as those between pairs of applications and simplex data streams such as voice or video. There are two categories of WFQ sessions: high bandwidth and low bandwidth[84]. Low bandwidth traffic has effective priority over high bandwidth traffic, and high bandwidth traffic shares the transmission service proportionally according to pre-assigned weights[84].

With standard WFQ, packets are classified by flow. Packets with the same source IP address, destination IP address, source TCP or UDP port, or destination TCP or UDP port belong to the same flow. WFQ allocates an equal share of the bandwidth to each flow. Flow-based WFQ is also called fair queuing because all flows are equally weighted[8484].

3.4.1.3 Admission Control

Admission control consists of bandwidth control and policy control. Applications can request a particular QoS mechanism for their traffic. The network device will then examine the network for factors like capacity, load, policies, etc, and either grant or deny the request. If the request is granted then the application has a contract for that service. If

---

[83] *Advanced QoS.* Allied Telesis. http://www.alliedtelesyn.com/media/pdf/adv-qos_wp.pdf

[84] *Configuring Weighted Fair Queuing.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt2/qcfwfq.pdf

the request is denied then the network device communicates the denial message back to the application. At this point, the application can then either stop sending traffic into the network or request a different service[85].

In the absence of admission control, the network application may be unaware of the resource availability in the network. Therefore, the application may not get the service that it expects from the network.

3.4.1.4 Policing

Policing is the mechanism that ensures that network traffic measures up to the network policy standards. The policing function in the network device must be able to monitor, control and enforce the use of network resources with respect to its profile and configured policies. Packets that do not confirm with configured policies are either dropped or shaped into the profile. Such policing helps the network device maintain its side of the agreed upon policies.

3.4.2 QoS Architectures

The QoS mechanisms described above are used in the network to create different QoS architectures. This section will discuss the important QoS architectures explained by the IETF: Differentiated Services (DiffServ) and Integrated Services (IntServ). These two architectures can also be used together to attain the most practical QoS objectives.

3.4.2.1 Differentiated Services

Differentiated Services (DiffServ), as discussed in IETFs RFC 2475, define a set of QoS mechanisms for implementing scalable service differentiation in the Internet[86].

---

[85] Kenichi Mase, Yuichiro Toyama, Abdulkhalog A. Bilhaj, Yosuke Suda. *QoS Management for VoIP Networks with Edge-to-Edge Admission Control.* IEEE. http://ieeexplore.ieee.org/iel5/7633/20835/00966237.pdf

[86] *An Architecture for Differentiated Services.* IETF. http://www.ietf.org/rfc/rfc2475.txt

DiffServ is a service model that offers more than one kind of service. It uses in-band signaling to differentiate various classes of traffic. Therefore, the signaling is carried in the data packet itself.

DiffServ defines a field in the IP header called the DiffServ Code Point[87] (DSCP). Hosts on the network sending traffic requiring QoS into a network supporting DiffServ mark each packet with a DSCP value. Network devices within the network use these values to classify the traffic into distinct service classes. Based on the configuration of the service class, the packets are queued and scheduled.

Expedited Forwarding (EF) is one of the most popular DSCP values (value: 101110). It is described in RFC 2598[88]. It is mainly used to provide guaranteed end-to-end service across a network. It gives traffic a low-loss, low-jitter, end-to-end service by assuring bandwidth availability across networks. This is done by reserving bandwidth before the packet is sent. The goal is to limit delay and deliver the packet on a timely basis. If a packet arrives at a router with an EF marking and passes the administered network policies, then the router schedules it into the highest priority queue. This allows the packet to be forwarded with minimum queuing latency thus incurring the least delay and jitter measures. EF marking is commonly used by latency and loss sensitive real-time applications as those applications usually require a very stringent service from the network. EF is commonly used for voice over IP services[88].

Assured Forwarding (AF) is another means of providing better than best effort handling for an IP packet[89]. AF is suggested for applications that require a better reliability than the best-effort service. There are four classes of service, and within each class, there are three different drop precedence. A drop precedence level determines how likely it is for the packets to be dropped. The classes each have their own bandwidth specifications. Class 1, high, gives the policy the lowest priority and Class 4, low,

[87] The DSCP is a six-bit field, spanning the fields formerly known as the type-of-service (TOS) fields and the IP precedence fields.

[88] *An Expedited Forwarding PHB.* IETF. http://www.ietf.org/rfc/rfc2598.txt

[89] *An Assured Forwarding PHB.* IETF. http://www.ietf.org/rfc/rfc2597.txt

describes the policy the highest priority[90]. A low drop level means the packets in this policy have the lowest chance of being dropped in this particular class level. Table 3.1 shows all the code points in an AF class.

Table 3.1 – AF Classes and Code Points[91]

	Class 1	Class 2	Class 3	Class 4
Low Drop	001010 AF11 DSCP 10	010010 AF21 DSCP 18	011010 AF31 DSCP 26	100010 AF41 DSCP 34
Medium Drop	001100 AF12 DSCP 12	010100 AF22 DSCP 20	011100 AF32 DSCP 28	100100 AF42 DSCP 36
High Drop	001110 AF13 DSCP 14	010110 AF23 DSCP 22	011110 AF33 DSCP 30	100110 AF43 DSCP 38

AF41 is considered the most suitable for video traffic[92].

The main advantage of using DSCP markings is that it enables the network devices to classify and prioritize certain traffic. When all the network devices along the path of data transfer supported this mechanism, then it results in an end-to-end QoS solution[93].

The advantages of DiffServ are simplicity and scalability. Implementing QoS using DiffServ is comparatively simple to accomplish. That is because DiffServ does not introduce any additional signaling messages to the packets it just marks the packets with

---

[90] *Use Code Points to Assign Pre-Hop Behaviors.* IBM. http://publib.boulder.ibm.com/infocenter/iseries/v5r3/index.jsp?topic=/rzak8/rzak8phb.htm

[91] *Implementing Quality of Service Policies with DSCP.* Cisco Systems. http://www.cisco.com/warp/public/105/dscpvalues.html

[92] *An Assured Forwarding PHB.* IETF. http://www.ietf.org/rfc/rfc2597.txt

[93] *Implementing Quality of Service Policies with DSCP.* Cisco Systems. http://www.cisco.com/warp/public/105/dscpvalues.html

code points. Also, in the DiffServ architecture, the network device does not process the network traffic flow-by-flow. It simply accumulates the traffic at the arriving node into different traffic classes that are defined using the DSCP marking. The device using DiffServ architecture is not required to uphold any elaborate state information to identify the traffic flow, hence reducing the processing of huge overheads. Consequently, the network device can handle a lot of traffic without encountering any drop in the network performance. Hence, it can be implemented without the scalability concerns.

Although DiffServ architecture is scalable and simple, this method does not have the facility for admission control. That is, although the network device is able to classify and prioritize incoming traffic given the DSCP marking, there is no methodology implemented in DiffServ for the device to know whether this traffic would cause network overloads or not.

3.4.2.2 Integrated Services

Integrated Services (IntServ) is a QoS architecture that uses explicit signaling to request QoS mechanisms. This architecture is described in RFC 1633[94]. In this architecture, the QoS message carries information that enables the network device to identify the traffic type and notifies it of the specific service requirements for that traffic. Unlike the DiffServ architecture, the traffic using this model is classified, queued and scheduled based on the flow at the network device.

IntServ uses the Resource Reservation Protocol (RSVP) to represent traffic, request network resources and achieve admission control[95]. This protocol enables applications to share their traffic profile with the networking device and to request certain QoS measures from the network based on the its bandwidth, packet loss tolerance and delay constraints. RSVP signaling messages are required to take the same route through

---

[94] *Integrated Services in the Internet Architecture: an Overview.* IETF. http://www.ietf.org/rfc/rfc1633.txt

[95] *Resrouce Reservation Protocol.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/rsvp.htm

the network as that of the applications traffic; this way, RSVP can reserve resources in the same networking devices that will process the traffic flow.

Host devices on a network generate RSVP messages. This message contains the source and destination address information, port data, the protocol being used, profile of the traffic, and a request for a specific traffic class for that profile. When a network device receives the RSVP signaling messages, it compares it against the available network resources. The network device bases the grant or denial of service decision on the application type, the user, and network resource availability. If the network is unable to provide requested resources at that time, then the network device denies the request and sends a failure message back to the requesting application. If the network device decides to grant the services requested then the information in the request message is stored in the network device, and the RSVP message is passed to the next network device. Once all the devices on the path to destination have granted the resource reservation request, the application can be guaranteed that the resources it asked for are set aside for its use, and that the services that were requested will be received.

There are various advantages to use RSVP. RSVP ensures that higher priority traffic gets the required services by not overloading the high priority queue. Unlike DiffServ in which when requests are denied the denial message is not communicated to the application requesting the service, RSVP makes sure the denial is communicated back. This way, if a request is denied by a network device, then the requesting application may take alternative measures to transfer data again. Similarly, if a particular request is granted, then that message is passed on to the requesting application.

One major disadvantage of using RSVP for QoS measures is scalability. In contrast to DiffServ, RSVP is a flow based protocol. Therefore, network devices are required to perform extra processing to identify traffic flows. For example, if the traffic flow at the network device is large, then the performance of the device will be affected. Therefore, this issue must be addressed before implementing RSVP for QoS measures.

3.4.2.3 IntServ + DiffServ

As discussed in the above sections, DiffServ and IntServ with RSVP have their own advantages and disadvantages. Therefore, combining both the architectures to give optimal utilization can be practical. Measures were taken to make use of the best of both architectures, which is described in RFC 2998[96]. The RFC describes how end-to-end IntServ QoS maybe supported over DiffServ networks. It discusses the methods by which network traffic can get end-to-end QoS with admission control without requiring the entire network to support the IntServ architecture[9696]. This RFC also addresses the scaling concerns in the network.

3.5 IP VPN

Enterprises that have dedicated private line infrastructure for their data networks often rely on a limited number of point-to-point links along with the Virtual Private Network (VPN) tunnels through the public network. While this option can be satisfactory for data applications that are not time-critical, it typically lacks the performance, accessibility, reliability and security features required for real-time voice and video. Enterprises usually have Layer 2 VPN services, such as ATM VPN, FR VPN, to support their existing data infrastructure. But these Layer 2 VPN services do not provide clear and reliable solution for real-time applications such as voice and video. This suggests a need to migrate to Layer 3 IP VPN solutions, which is optimized to run such real-time data traffic[97]. This section will provide an overview of IP VPN solution for enterprises.

VPNs are deployed over the shared public network infrastructure that uses various technologies to help ensure reliability, security, and privacy[97]. Therefore, VPNs offer businesses the same security, quality of service, reliability, and privacy of private networks. To run real-time applications, enterprises need a service provider who is capable of providing real-time traffic sufficient priority throughout the ISP backbone

[96] *A Framework for Integrated Services Operation over DiffServ Networks.* IETF. www.ietf.org/rfc/rfc2998.txt

[97] Andrew Mason. *VPNs and VPN Technologies.* Cisco Secure Virtual Private Networks.

network to meet enterprises' quality, security and reliability requirements. A Layer 3 IP VPN optimized for real-time applications is an efficient way to provide this level of service. High-speed connectivity, reliability, and security make IP VPNs viable for supporting services, such as voice over IP and video over IP. Typically, IP VPNs can be categorized in two different models[98]:

- ISP Network IP VPN: In this model, VPN intelligence is in the service provider end, which makes it transparent to end-users. This architecture enables the service providers to reduce the complexity of network implementation and lower the cost of delivering VPN services to businesses.
- Customer Network IP VPN: In this architecture, the intelligence is in the Customer Premise Equipment (CPE) at the customer's sites. This enables the enterprise network administrators to implement different classes of service based on their needs and depending on the ISP's network infrastructure.

The basis behind any kind of VPN implementation is the encapsulation or tunneling algorithms that it supports. In the IP VPN field, there are two major types of technologies, namely Multi Protocol Label Switching (MPLS) based IP VPN, and IP Security (IPSec) based IP VPN. These different protocols provide various benefits and serve different purposes.

3.5.1 MPLS based IP VPN

Figure 3.6 provides a graphical view of a MPLS based IP VPN service.

[98] Andrew Mason. *VPNs and VPN Technologies.* Cisco Secure Virtual Private Networks.

Figure 3.6 – MPLS based IP VPN[99]

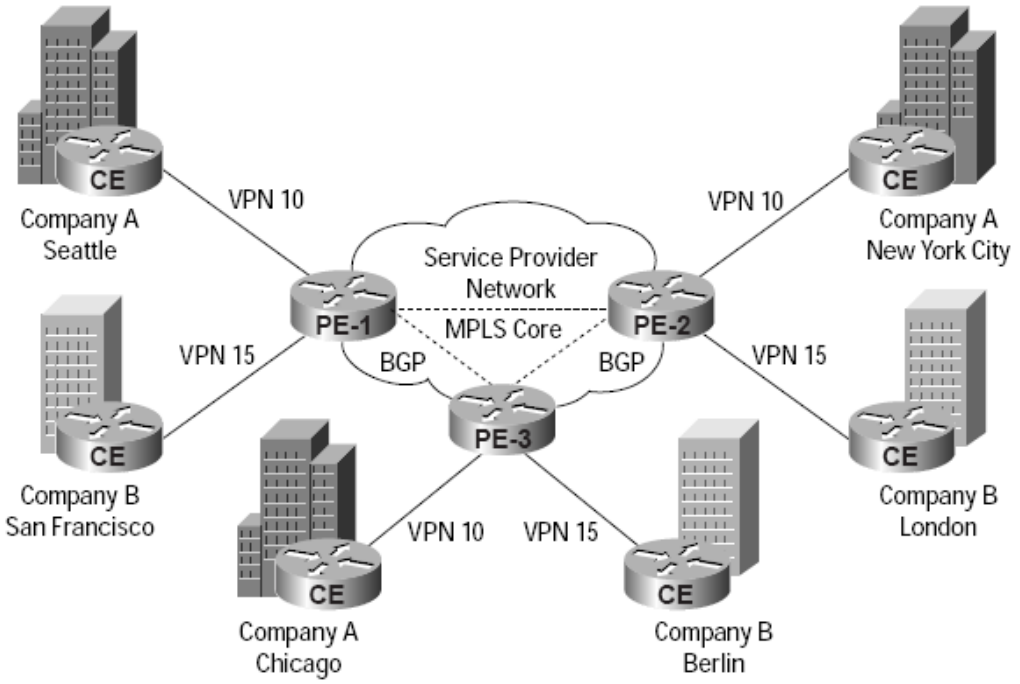


Figure 3.6 illustrates that MPLS technology enables ISPs to deliver differentiated VPN services to numerous enterprise customers over a single, shared network infrastructure. MPLS based IP VPNs uses multi protocol labeling algorithms and signaling protocols to encapsulate IP packets and distribute network information. MPLS based IP VPN capable of seamlessly interfacing with traditional Layer 2 VPN technologies. Enterprises that have a VPN solution deployed can use MPLS based IP VPNs as an alternative or a complementary to their existing solution.

MPLS technology integrates capabilities of Layer 2 switching performance with the flexibility of Layer 3 routing. At the network end, routers apply simple labels to IP packets or data frames[99]. Then MPLS enabled switches or routers in the core network switches the received packets based on those labels with minimal lookup overhead.

---

[99] *MPLS based IP VPN Service.* Cisco Systems. http://www.paetec.com/downloads/mpls_overview.pdf

Traffic Engineering and QoS are two key features of MPLS technology[100]. Traffic engineering is enabled by MPLS algorithms that route traffic through a specific path, even if it is not the least-cost route. By using these techniques in the core network, policies to ensure optimal traffic distribution can be achieved, which in turn would improve network resource utilization. The QoS features enable network administrators to provide priority services across the network by marking packets with specific DSCPs. With this technique, MPLS QoS supports traffic classifications and improves capabilities for congestion management.

MPLS IP VPN solutions support end-to-end quality of service requirements, rapid fault correction of link or node failure, bandwidth protection, etc. MPLS technology by itself simplifies network configuration, administration, and provisioning, helping ISPs to deliver highly scalable, differentiated, end-to-end IP based services. This is a network based VPN technology for site-to-site VPN communications only.

3.5.2 IPSec Based IP VPN

IPSec protocol is a standards-based method that functions at the network layer used for providing confidentiality, integrity, and authenticity to data transferred across IP networks. Although IPSec is described in a series of RFCs , the main RFC that defines the protocol is RFC 2401[101]. This protocol provides the framework for CPE based Layer 3 VPNs[102]. IPSec based IP VPNs is a replacement technology to the traditional Layer 2 VPN solutions.

IPSec suite defines two main parts to the protocol – Authentication Header (AH) and Encapsulating Security Payload (ESP). AH, as the label suggests, provides authentication and integrity to the data packets passed between network devices.

---

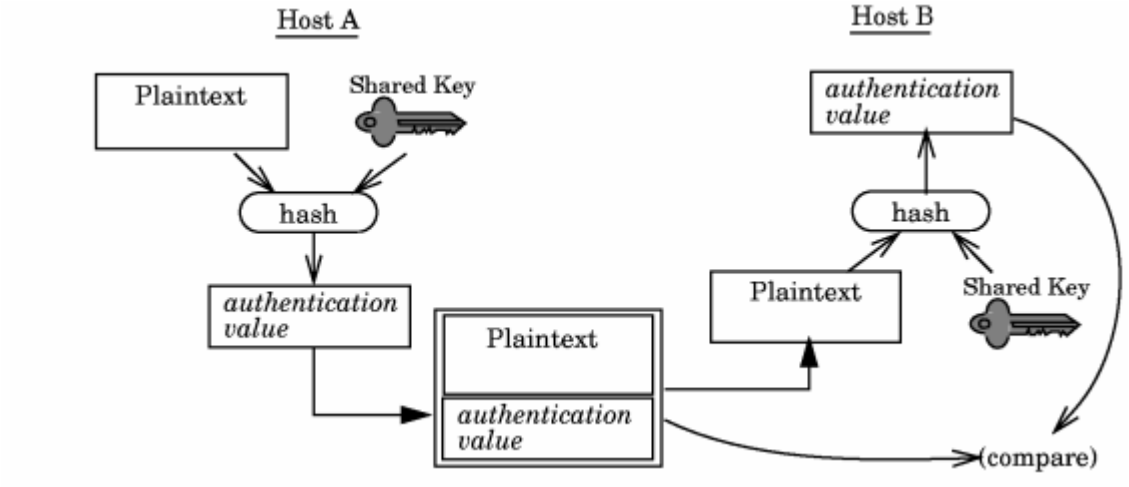
[100] *MPLS based IP VPN Service.* Cisco Systems. http://www.paetec.com/downloads/mpls_overview.pdf

[101] *Security Architecture for the Internet Protocol.* IETF. http://www.ietf.org/rfc/rfc2401.txt

[102] Andrew Mason. *VPNs and VPN Technologies.* Cisco Secure Virtual Private Networks.

Authentication Header is discussed in detail in RFC 2402[103]. The AH contains an authentication value based on a symmetric-key hash function[104]. Figure 3.7 illustrates the AH authentication process[104].

Figure 3.7 – AH Authentication Process[104]



A keyed one-way hash function is applied to the data packets to create a message digest. If any part of the packet is changed during data transfer period, it will be detected by the receiving network device when it performs the same one-way hash function on the data packets and compares the value of the message digest that the sender has supplied. Since this methodology involves the use of a secret key shared between the two systems, authenticity of data transmission is ensured. One main disadvantage of AH is that it does not provide data privacy because it does not encrypt the actual data that is sent it simply adds an authentication value to the plain text.

ESP is another security protocol used by IPSec to provide data privacy, authentication, and integrity. This protocol is discussed in RFC 2406[105]. ESP provides confidentiality by performing encryption at the IP packet layer. It supports a variety of symmetric encryption algorithms. ESP takes the data packets carried by IP and encrypts

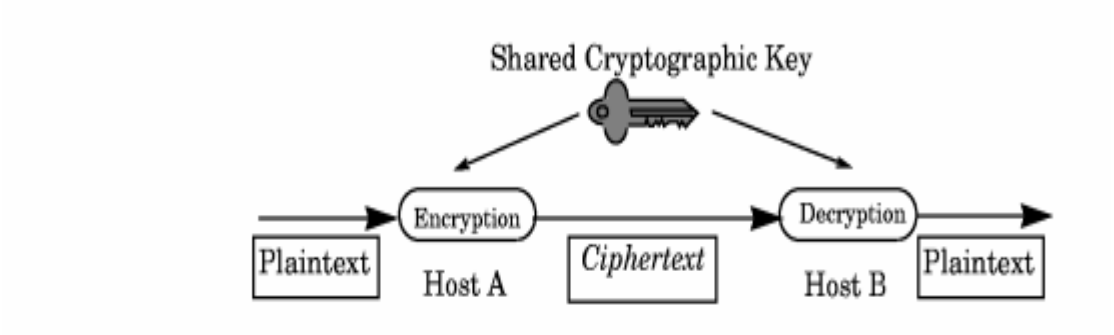
[103] *IP Authentication Header.* IETF. http://www.ietf.org/rfc/rfc2402.txt

[104] *Authentication Header.* HP. http://docs.hp.com/en/J4256-90003/ch01s02.html

[105] *IP Encapsulating Security Payload.* IETF. http://www.ietf.org/rfc/rfc2406.txt

the packets using an encryption algorithm and cryptographic key[106]. The output is in the form of a cipher-text that is difficult to decode without knowing the key that is shared by the sender and the receiver. The receiving IPSec ESP network device uses a decryption algorithm and the same key to extract data from the cipher-text. Figure 3.8 provides a visual representation of the ESP process.

Figure 3.8 – ESP Encryption Process[106]



The IPSec protocol provides protection for data packets that are transmitted through the IP network by allowing network administrators to identify the traffic that needs protection, define the protocols using which the data traffic will be assured authentication and privacy, and control destination routes. IPSec is suitable for both site-to-site and remote-access IP VPNs[107].

3.6 Conclusion

In conclusion, QoS is an important component for successful real-time IP network deployment. With QoS, the network administrator is able to control the resource consumption of real-time application as well as provide acceptable network service to the end users. In addition to QoS solutions, IP VPN solutions provide authenticity, confidentiality, security, and reliability to real-time applications running on the converged network. Now that all the QoS and security measures have been discussed, the

[106] *Encapsulating Security Payload.* HP. http://docs.hp.com/en/J4256-90003/ch01s03.html?btnNext=next%A0%BB

[107] Andrew Mason. *VPNs and VPN Technologies.* Cisco Secure Virtual Private Networks.

next chapter will focus on demonstrating how a converged solution leads to cost optimization.

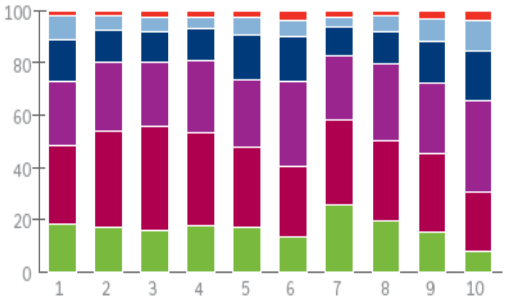
4 Cost Optimization

4.1 Introduction

During the past decade, modern communication technologies have become more and more pervasive in enterprise networks. There is growing pressure to use these technologies to reduce costs, while making enterprise networks more adaptable to newer technologies. Convergence addresses many of the problems faced by enterprise networks, providing a holistic solution that meets business needs now and will be more adaptable for any future innovation. Figure 4.1 shows the key barriers to investing in new technology, identified in a study completed by AT&T in 2006.

Figure 4.1 – Barriers to Implementing Converged Solution[108]

How significant are these barriers to implementing a converged IP network at your company?



Source: Economist Intelligence Unit/AT&T survey, June 2006.

	Major Barrier				No Barrier	Don't Know
1. Lack of in-house skills and experience in relevant technologies	18.02%	30.20%	24.62%	15.99%	9.39%	1.78%
2. Cost of new equipment/ technology to support convergence	17.39%	36.57%	26.09%	12.53%	5.12%	2.30%
3. Cost of implementation	16.15%	39.49%	24.36%	12.05%	5.13%	2.82%
4. Complexity of implementation	17.51%	36.04%	27.16%	12.44%	4.06%	2.79%
5. Concern about quality (eg, of voice communications)	16.88%	30.95%	25.83%	17.14%	6.65%	2.56%
6. Time scales for return on investment	13.55%	27.11%	32.23%	17.39%	6.14%	3.58%
7. Network security issues	25.90%	32.31%	24.36%	11.28%	3.59%	2.56%
8. Disruption to business while migrating	19.34%	31.04%	29.52%	11.96%	6.11%	2.04%
9. Concern about "putting all eggs in one basket" on a single network	15.52%	30.03%	26.72%	16.03%	8.65%	3.05%
10. Satisfied with present system	7.89%	22.65%	35.11%	19.08%	11.20%	4.07%

Although the concept of converged IP networks has bee around for years, the adoption of IP convergence has been low. The fact that the financial benefits of converging network infrastructures have not been clearly stated has been the principal barrier to migrating to converged solution. As seen in Figure 4.1, it is evident that cost of

[108] *Convergence Takes Hold in the Enterprise.* AT&T. http://graphics.eiu.com/ebf/PDFs/WP_2006_Convergence_ENG.pdf

new equipment/technology and cost of converged network implementation are two of the main factors that are keeping enterprises from implementing a converged solution. This chapter aims to demonstrate the cost savings that are possible when a company decides to switch to an IP-based network solution. Also, this chapter will discuss the resource optimization possibilities with a converged network solution.

4.2 Overview

As mentioned before, this chapter is provided to demonstrate the financial savings possible when an enterprise migrates to converged network infrastructure. The analysis is about an enterprise that has five sites across the United States. The estimated traffic volume will be based on the number of employees, and the usage of voice and data networks. The study will investigate the operating costs for both the legacy network solution and a converged network solution. Once the financial analysis of the two network designs is done, the data will be compared to establish that the converged model provides cost savings in the long run.

4.2.1 Network Infrastructure

The company being analyzed here is APX Corporation, which has 5 major sites across the United States. These sites are located in Washington D.C., Miami, Boston, Denver and Rochester. Number of employees in each site is shown in Table 4.1.

Table 4.1 – Site Location and Number of Employees

Site Location	Number of Employees
Washington D.C.	300
Miami	200
Boston	130
Denver	150
Rochester	120

4.2.2 Voice Traffic Volume

Each employee at each site makes 25 intra-site calls per day that lasts for 10 minutes each. Intra-site calls costs nothing, because there are PBXs in each site. There are no direct private trunks between sites. Sites are connected to each other via PSTN links. Table 4.2 shows the inter-site voice traffic table.

Table 4.2 – Inter-site Voice Traffic

Site	# of Employees	# of Calls per Employee	Duration of one Call (minutes)
Washington D.C.	300	15	5
Miami	200	10	8
Boston	130	7	10
Denver	150	8	12
Rochester	120	6	15

Trunking calculation is done assuming 20% of the calls go through during busy hours, and the company allows for 1% blocking.

4.2.3 Data Traffic Volume

As per the specified requirements the inter-city network is a fully meshed design. The data network volume of traffic out of the each of the 5 corporate sites is assumed to be 100Kbps.

4.2.4 Video Traffic Volume

Each site has to be capable of videoconferencing with one other site at any given time. The bandwidth requirement for the videoconferencing is 384Kbps.

4.2.5 Assumptions

This analysis does not apply to any particular business operation. However, the company represented here is logically suitable for analytical purposes. The basic

assumption for the scenario being considered here are based on the employee levels and construction of each of the business sites, data and voice network utilization, and the potential business growth rate. The following are the critical assumptions that were made during the analysis:

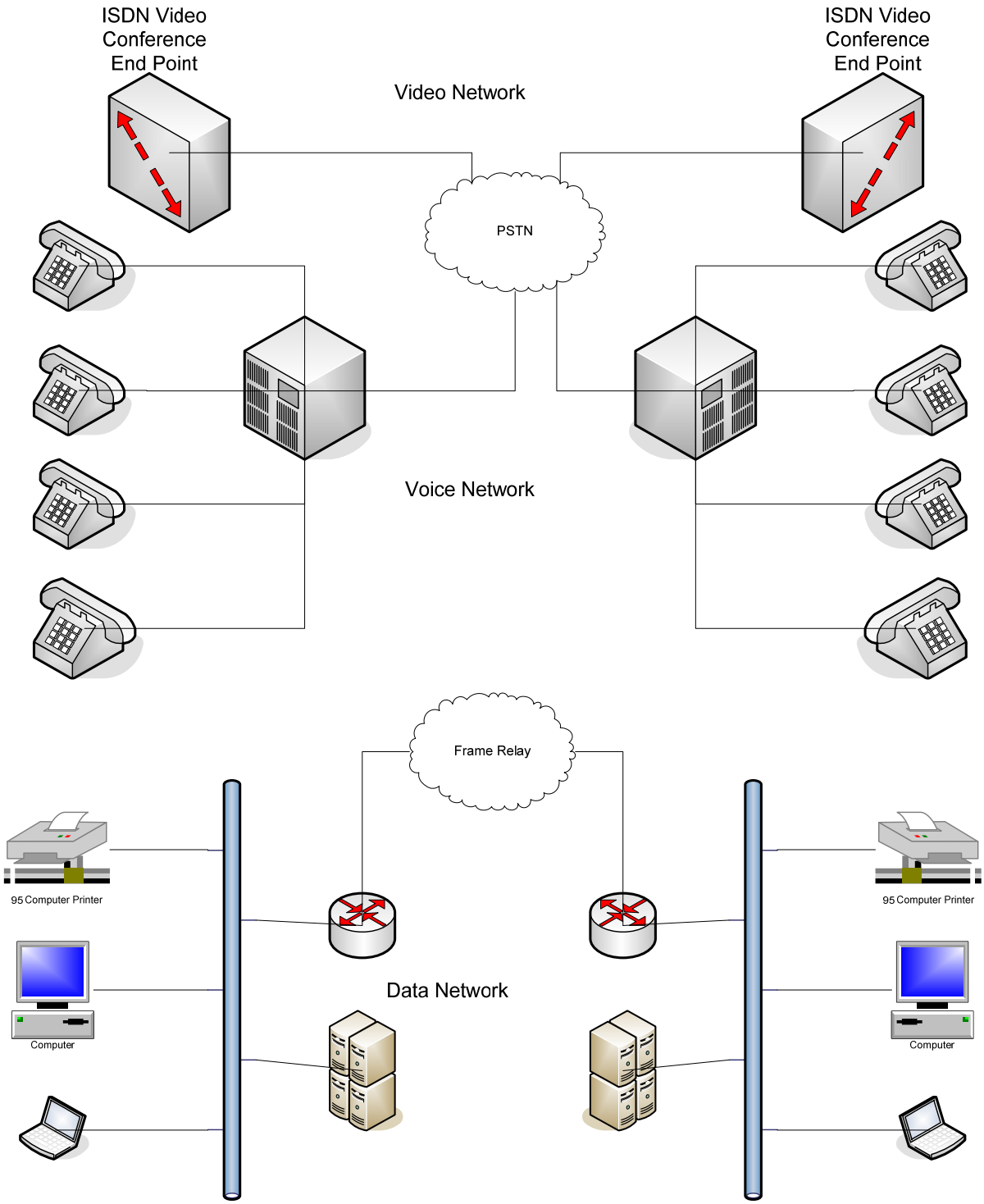
- The installation and equipment charges are not considered in this analysis. This study only investigates the long run cost savings when migrating to a converged solution.
- Employees at all sites are uniformly distributed throughout the premises.
- No productivity gains have been considered.
- All employees have equal access to phones, PCs and servers.
- All employees are site-based; there are no telecommuters or remote workers.
- The cost for domestic long distance voice calls over PSTN is 5 cents / minute.
- The customer already has Internet connection, which will not be changed.
- There is the same access requirement for data regardless of the WAN transport type.

NOTE: The rates included in this study are dated April/May 2007. These were obtained from AT&T website and from www.shopforT1.com Voice-Data representative.

4.3 Traditional Network Solution

The traditional network infrastructure is depicted in Figure 4.2 (only two sites are shown for illustrations purposes, but there are 5 sites being considered for this study).

Figure 4.2 – Traditional Network Infrastructure



With this solution voice, video and data services are provided on separate networks. For the voice network, there is a PBX at each site. Therefore, intra-site calling incurs no extra monthly costs. Inter-site voice connections use trunks to connect to PSTN – there is no private voice network. The data LAN at each site is using Ethernet with

Cisco equipment. The data WAN infrastructure consists of fully-meshed FR, with a 256 Kbit/s access circuit at the five sites. This is calculated using the data network volume mentioned above and assuming 50% utilization on each site. The video network uses 6 Basic Rate ISDN lines to get to the minimum bandwidth requirement of 384Kbps. Given the voice, video and data traffic information, this section will calculate the equipment and operational costs for the traditional network design.

4.3.1 Voice Network

The voice network trunking analysis for the traditional network infrastructure is shown below in Table 4.3. The intra-site calls do not incur any cost because there is a PBX in each site, and thus, the intra-site calls do not go through the PSTN. The voice traffic analysis assumes that the 30 inter-site calls can be made to any of the sites at the same rate. This is assuming that the long distance charge for inter-site calls is 5 cents/minute. PSTN trunks are rented at $32 / trunk at each site.

Table 4.3 – Number of Trunks Calculated with Erlang B Table[109]

Site	# of Employees	# of Calls per Employee	Duration of one Call (minutes)	Total Duration of Calls (minutes)	BH Call Minutes	BH # of Erlangs	Trunks Required
Washington D.C.	300	15	5	22500	4500	75.00	91
Miami	200	10	8	16000	3200	53.33	67
Boston	130	7	10	9100	1820	30.33	42
Denver	150	8	12	14400	2880	48.00	62
Rochester	120	6	15	10800	2160	36.00	48

Table 4.4 illustrates the total cost for implementing the above mentioned PSTN trunks at the rate of $32 / month / trunk.

109 *Erlang B Traffic Table.* http://www.stttelkom.ac.id/staf/UKU/Buku%20Referensi%20Cellular/Table%20ERLANG.pdf

Table 4.4 – PSTN Voice Trunk Costs

Site	Trunks Required	Cost Per PSTN Trunk	Total Trunk Cost Per Site
Washington D.C.	91	$32.00	$2,912.00
Miami	67	$32.00	$2,144.00
Boston	42	$32.00	$1,344.00
Denver	62	$32.00	$1,984.00
Rochester	48	$32.00	$1,536.00

Table 4.5 depicts the cost for call minutes used by the employees at each site, assuming $0.05 / minute.

Table 4.5 – Per Minute Call Costs

Site	Number of Employees	# of Calls per Employee	Duration of one Call (minutes)	Total Duration of Calls (minutes)	Cost Per Minute	Total Cost
Washington D.C.	300	15	5	22500	$0.05	$1,125.00
Miami	200	10	8	16000	$0.05	$800.00
Boston	130	7	10	9100	$0.05	$455.00
Denver	150	8	12	14400	$0.05	$720.00
Rochester	120	6	15	10800	$0.05	$540.00

The total voice network charge on a traditional infrastructure comes to $13,560.00 per month. This will later be compared with the converged network solution.

4.3.2 Data Network

As per the specified requirements the inter city network is a fully meshed design. That is, each major city is interconnected with all the other 4 major sites. This provides for maximum redundancy; if one link goes down for some reason, there is always other ways to reach the desired destination. As mentioned above, the volume of traffic out of each major city is 100Kbps. The required network utilization is assumed to be 50%; therefore, the link speed required for all the inter city links must be 256Kbps links or

above. Thus, cost calculations for inter city traffic using frame relay network will use 256Kbps links.

Frame Relay is a fixed rate service. The 3 main factors that are used while calculating frame relay link costs are access costs, CIR, and port charges. Table 4.6 shows the frame relay CIR charges, while Table 4.7 represents the frame relay port charges.

Table 4.6 – CIR Charges[110]

CIR Charges	
Link Types (Kbps)	**Two Way Charge ($)**
4	21
8	26
16	36
32	67
48	96
56/64	112
128	254
192	384
256	511
320	639
384	766
448	969
512	1149
576	1309
640	1465
704	1624

110 *AT&T Packet Services.* AT&T. http://new.serviceguide.att.com/portals/sgportal.portal?_nfpb=true&_pageLabel=aps_page

Table 4.7 – FR Port Charges[111]

Port Charges		
Link Type (Kbps)	**Domestic ($)**	**Global ($)**
56/64	305	1395
128	540	2620
192	680	3135
256	795	3645
320	925	4110
384	1040	4575
448	1135	5217
512	1255	5860
576	1355	6010
640	1445	6165
704	1520	6320

The following is the calculation for domestic FR charge:

FR Port Charge for 256Kbps Link = $795 / month

FR CIR Charge (two way) for 256Kbps Link = $511 / month

Access Charge @ each major site = $500 / month

Total Domestic FR Charge = $1806 / month

Table 4.8 illustrates the total FR charges to form a fully meshed network as mentioned earlier.

[111] *AT&T Packet Services.* AT&T. http://new.serviceguide.att.com/portals/sgportal.portal?_nfpb=true&_pageLabel=aps_page

Table 4.8 – FR Charges

CITIES	Washington DC	Miami	Boston	Denver	Rochester	Total
Washington DC	$0.00	$1,806.00	$1,806.00	$1,806.00	$1,806.00	$7,224.00
Miami	$1,806.00	$0.00	$1,806.00	$1,806.00	$1,806.00	$7,224.00
Boston	$1,806.00	$1,806.00	$0.00	$1,806.00	$1,806.00	$7,224.00
Denver	$1,806.00	$1,806.00	$1,806.00	$0.00	$1,806.00	$7,224.00
Rochester	$1,806.00	$1,806.00	$1,806.00	$1,806.00	$0.00	$7,224.00
	Total Data Network Cost					**$18,060.00**

The total data network charge on a traditional infrastructure comes to $18,060.00 per month using FR for WAN access. This will later be compared with the converged network solution.

4.3.3 Video Network

The bandwidth required for videoconferencing was mentioned to be 384Kbps. This would require 6 BRI links. The monthly cost of a BRI link to any corporate site is $45 / link. Since we require 6 links at each of the 5 sites, the total video network cost comes to: $45 / link * 6 links / site * 5 sites = $1,350.00 / month. Therefore, the total video network charge on a traditional infrastructure comes to $1,350.00 per month using 6 BRI link access per site. This will later be compared with the converged network solution.

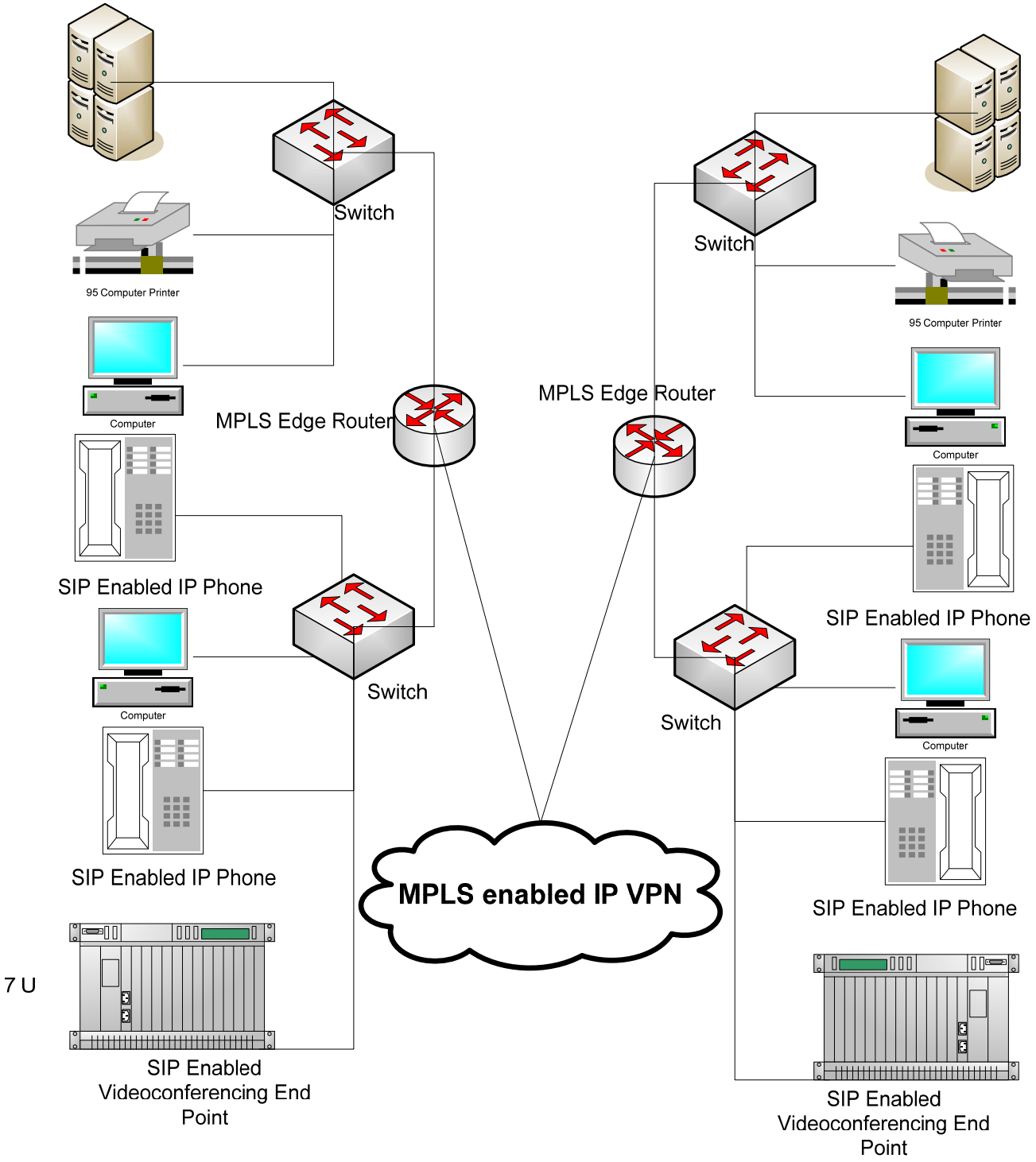
4.3.4 Total Costs

The total monthly charge for the traditional network solution is: $13,560.00 + $18,060.00 + $1,350.00 = $32,970 / month. This will later be compared with the converged network solution.

4.4 Converged Network Solution

The converged network infrastructure is depicted in Figure 4.3 (only two sites are shown for illustrations purposes, but there are 5 sites being considered for this study). All sites are interconnected in a VPN environment IP VPN QoS service. Each site is connected to a local point of presence by a point-to-point circuit running IP. Real-time traffic, such as voice and video, must be given the highest priority to keep jitter, delay and packet loss at acceptable levels and maintain quality and performance.

Figure 4.3 – Converged Network Infrastructure



4.4.1 Voice Network

The inter- and intra-site voice calls would now incur no charges over the PSTN network because they are carried over the private IP network. Therefore, instead of paying for PSTN trunks, calculations have to be done to accommodate the required voice bandwidth over the IP network.

PCM voice packets are now to be digitized using voice codecs to transport real time audio through an IP network. For this study the G.711 codec will be used to digitize voice, and obtain voice network bandwidth requirements. G.711 digitizes the voice signal at uncompressed 64 Kbps and creates a payload of 160 bytes for the IP audio packet.

Since the traffic is run on an IP network, the IP, UDP, and RTP headers are to be included in the bandwidth calculation. The layer 3, 4 and 5 headers (IP, UDP and RTP respectively) sum up to 40 bytes uncompressed, but by using compressed RTP, the headers can be reduced to 2 bytes[112].

G.711 codec will digitize voice at 50 packets per second. Therefore, the 64Kbps calculation for 50pps:160 bytes / packet * 8 bits / byte * 50 packets / second = 64Kbps. However, since the layers 3, 4, and 5 are 40 bytes as described above, then the amount of bandwidth needed for G.711 is greater when the headers are included. The total bandwidth required with headers included would be (using uncompressed RTP): [(160 bytes / packet + 40 bytes / packet) / (160 bytes / packet)] * 64Kbps = 80Kbps. Therefore, one voice call requires 80Kbps of bandwidth when it becomes a VoIP packet to run on a converged network. With this information, the voice network can now be designed to calculate required bandwidth. Table 4.9 illustrates the bandwidth requirements for voice calls over the converged network.

[112] *RTP Header Compressions.* Cisco Systems. http://www.cisco.com/univercd/cc/td/doc/product/software/ios112/rtphead.htm

Table 4.9 – Voice Calls Bandwidth Requirement

Site	# of Employees	# of Calls per Employee	Bandwidth Per Call (Kbps)	Total Bandwidth (Kbps)	BH Bandwidth (Kbps)
Washington D.C.	300	15	80	1200	240
Miami	200	10	80	800	160
Boston	130	7	80	560	112
Denver	150	8	80	640	128
Rochester	120	6	80	480	96

4.4.2 Data Network

The data network volume of traffic out of the each of the 5 corporate sites is given to be 100Kbps.

4.4.3 Video Network

The real-time video streaming bandwidth requirement was given to be 384Kbps.

4.4.4 Total Bandwidth Calculation

Table 4.10 depicts the bandwidth requirement at each site.

Table 4.10 – Bandwidth Requirement

Site	Voice Bandwidth (Kbps)	Data Bandwidth (Kbps)	Video Bandwidth (Kbps)	Total Bandwidth (Kbps)
Washington D.C.	240	100	384	724
Miami	160	100	384	644
Boston	112	100	384	596
Denver	128	100	384	612
Rochester	96	100	384	580

4.4.5 Total Cost Calculations

As illustrated in Table 4.7, each site has specific bandwidth requirements. But keeping in mind the 50% utilization factor, it is fair to designate a DS-1 point-to-point circuit from each of the sites to the ISPs network. DS-1 circuits operate at a speed of 1.544Mpbs, which will fulfill the 50% utilization factor. Now, the cost for implementing a DS-1 circuit would be $600 / month per circuit. Therefore, for five DS-1 circuits, the cost adds up to $3,000 / month. In addition, the business must pay a monthly fee to have a private IP VPN setup for secure and reliable data transfer. This monthly fee would be anywhere from $600 / month to $1,000 / month depending upon the type of IP VPN and the other service level agreements with the ISP. If the enterprise chooses to go with IP MPLS VPN service then the rate charges can be obtained from AT&T[113]. As illustrated in the converged network diagram, this analysis will use the MPLS VPN solution. Therefore, the VPN cost would be $3,458 / month per port. Therefore, the total for IP MPLS VPN is $17,290 / month.

Adding all the above mentioned costs, the total monthly expenditure comes to: $3,000 + $17,290 = $20,290 / month.

4.5 Comparison

The traditional network solution costs $32,970 / month, while the converged solution costs $20,290 / month. It is a saving of about 38% per month. Calculating annual costs, a business of the size described in the analysis could save up to $152,160 per annum.

This study demonstrates the cost savings available through deploying a converged network solution is noticeable. The capital costs of the two solutions might be similar when starting to design and implement a network. The considerable operational savings associated with the converged network solution, however, strongly support the case for early replacement of existing infrastructure.

---

[113] *AT&T Packet Services.* AT&T. http://new.serviceguide.att.com/portals/sgportal.portal?_nfpb=true&_pageLabel=aps_page

4.6 Conclusion

This analysis proves that an IP-based network solution gives the potential for significant savings on annual network operational costs. The operational expenses of maintaining a converged network is approximately 38% less than the traditional solution. The analysis presented in this chapter evaluated and compared the costs incurred by an organization when deploying an IP-based converged solution, as opposed to a traditional network infrastructure.

Thesis Conclusion

A converged network infrastructure can be flexible and easy to manage. There are many convergence solutions provided by different vendors suiting different business needs. Converged solution allows an enterprise network to adjust to all communication needs of the business and also ensures that all current and future next generation applications will function properly. Converged solutions provide proper functionality desired to deliver data, voice and multimedia connectivity over single network infrastructure.

This thesis introduced the concept of convergence, discussed the advantages, challenges, incentives and disincentives in implementing a converged network solution. It is important to understand and analyze these points before making a decision to migrate to a converged solution.

Understanding the underlying technologies and protocols is important to comprehend the operation of an IP network and to appreciate the scope of convergence. Along with understanding the protocols, enterprise network administrators need to have knowledge of quality of service measures that need to be taken in the network when it carries real-time traffic like voice and video. In addition to discussing the various technological architectures and protocols that are involved in implementing a converged network infrastructure, this thesis also discussed the various QoS options available to enterprises to properly implement IP network to carry real-time traffic.

As previously stated in this paper, network security is one of the major concerns when it comes to implementing a converged infrastructure. One of the solutions to implementing a secure IP-based network is deploying Layer 3 IP VPN services. This paper discussed the various IP VPN solutions available that provide authenticity, confidentiality, security, and reliability to real-time applications running on the converged network.

Cost is the other main point of concern in a converged network solution. The analysis provided in this thesis proved that converged networks can increase the efficacy and productivity of the overall network, while reducing the cost significantly. It proved that an IP-based network solution gives the potential for significant savings on annual

network operational costs. The operational expenses of maintaining a converged network is approximately 38% less than the traditional solution. The analysis presented in this paper evaluated and compared the costs incurred by an organization when deploying an IP-based converged solution, as opposed to a traditional network infrastructure.

In conclusion, this thesis emphasized the importance of a converged network infrastructure and proved that it leads to significant cost savings. It provided an example of an enterprise network specifications (voice, video and data), and presented an in depth cost analysis of a typical network vs. a converged network to emphasize that converged infrastructures provide significant savings.

References

Als and F. Z. Ghassemlooy; G. Swift; P. Ball; J. Chi. Optics Communications. 2002.

"Advanced QoS White Paper." Allied Telesis. www.alliedtelesyn.com. <http://www.alliedtelesyn.com/media/pdf/adv-qos_wp.pdf>.

Andrew Mason. Cisco Secure Virtual Private Networks. 1st ed. Cisco Press, 2001.

Andrew S. Tanenbaum. Computer Networks. 4th ed. USA: Prentice Hall PTR, 2002.

"Authentication Header." HP. www.hp.com. <http://docs.hp.com/en/J4256-90003/ch01s02.html>.

"The Business Case for Enterprise VoIP." Intel Corporation. www.intel.com. <http://www.intel.com/it/pdf/parsippany-voip.pdf>.

Carolyn R. Johnson, Yakov Kogan, Yonatan Levy, Farhad Saheban and Percy Tarapore. "VoIP Reliability: A Service Provider's Perspective." IEEE Communications Magazine. July 2004 2004. www.comsoc.org. <http://www.comsoc.org/tech_focus/pdfs/2005/jan/johnson.pdf>.

"Cisco - Technical Considerations for Converging Voice, Data, and Video Networks." IT World. www.itworld.com. <http://www.itworld.com/WhitePapers/Cisco_AVVID_TechCon/>.

"The Coming of True Convergence: Why Service Providers Can Finally Turn Out the Lights on the Old Public Switched Telephone Network (PSTN)." www.IEC.org. International Engineering Consortium. <http://www.iec.org/online/tutorials/true_converge/>.

"Configuring Weighted Fair Queuing." Cisco Systems. www.cisco.com. <http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt2/qcfwfq.pdf>.

"Converged Networks." Computer Sciences Corporation. www.csc.com. <http://www.csc.com/solutions/itinfrastructureoutsourcing/offerings/uploads/1632_1.pdf>.

"Convergence Takes Hold In The Enterprise." AT&T. October 2006. www.corp.att.com. <http://www.corp.att.com/emea/docs/s4_convergence_eng.pdf>.

"Convergence: Preparing the Enterprise Network." ProCurve Networking by HP. June 2005. www.hp.com. <http://www.hp.com/rnd/pdfs/convergence_WP_june05.pdf>.

David W. Cearley, Jackie Fenn, and Daryl C. Plummer. "Gartner's Positions on the Five Hottest IT Topics and Trends in 2005." Gartner. 12 May, 2005. http://www.gartner.com/. <http://www.gartner.com/DisplayDocument?doc_cd=125868>.

---. "Gartner's Positions on the Five Hottest IT Topics and Trends in 2005." Gartner. May 12, 2005. www.gartner.com. <http://www.gartner.com/resources/125800/125868/gartners_positi.pdf>.

"Encapsulating Security Payload." HP. www.hp.com. <http://docs.hp.com/en/J4256-90003/ch01s03.html?btnNext=next%A0%BB>.

"Erlang B Traffic Table." <http://www.stttelkom.ac.id/staf/UKU/Buku%20Referensi%20Cellular/Table%20ERLANG.pdf>.

"Ethernet Technologies." Cisco Systems. October 12, 2006. www.cisco.com. <http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/ethernet.htm>.

"Fiber Distributed Data Interface." Cisco Systems. October 12, 2006. www.cisco.com. <http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/fddi.htm>.

G. Hudson Gilmer. "The Real-Time IP Network: Moving IP Networks Beyond Best Effort to Deliver Real-Time Applications." Nortel Networks. 2003. www.nortel.com. <http://www.nortel.com/products/01/cr_rtr/collateral/the_realtime_network.pdf>.

Gonzalo Camarillo. SIP Demystified. 1st ed. McGraw-Hill Professional, 2001.

H. Schulzrinne. "RTP: A Transport Protocol for Real-Time Applications." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc1889.txt>.

H. Schulzrinne. "RTP Profile for Audio and Video Conferences with Minimal Control." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc1890.txt>.

Harry G. Perros. Connection Oriented Networks: SONET/SDH, ATM, MPLS and Optical Networks. Wiley, 2005.

Henry Sinnreich, and Alan B. Johnston. Internet Communications using SIP. 1st ed. Wiley, 2001.

"Implementing QoS Solutions for H.323 Video Conferencing over IP." Cisco Systems. 6 February 2006. www.cisco.com. <http://www.cisco.com/warp/public/105/video-qos.html>.

"Implementing Quality of Service Policies with DSCP." Cisco Systems. May 24 2006. www.cisco.com. <http://www.cisco.com/warp/public/105/dscpvalues.html>.

"Internet Protocol." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc0791.txt>.

"Introduction to Gigabit Ethernet." Cisco System. www.cisco.com. <http://www.cisco.com/en/US/tech/tk389/tk214/tech_brief09186a0080091a8a.html>.

Ira M. Weinstein. "The ISDN to IP Migration for Videoconferencing." Wainhouse Research. 2006. www.wrplatinum.com. <http://www.wrplatinum.com/Downloads/6128.aspx>.

"ISB Policies and Requirements." Information Services Board. <http://isb.wa.gov/policies.aspx>.

J. Heinanen, F. Baker, W. Weiss and J. Wroclawski. "Assured Forwarding PHB Group." IETF. June 1999. www.ietf.org. <http://www.ietf.org/rfc/rfc2597.txt>.

J. Postel. "User Datagram Protocol." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc0768.txt>.

---. "User Datagram Protocol." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc0768.txt>.

Kenichi Mase, Yuichiro Toyama, Abdulkhalog A. Bilhaj and Yosuke Suda. "QoS Management for VoIP Networks with Edge-to-Edge Admission Control." IEEE. 2001. www.ieee.org. <http://ieeexplore.ieee.org/iel5/7633/20835/00966237.pdf>.

M. Handley and V. Jacobson. "Session Description Protocol." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc2327.txt>.

Manoj Menon. "IP Convergence in the Enterprise." Frost & Sullivan. July 13, 2006. www.cisco.com. <http://www.cisco.com/web/VN/voice/pdf/ip_convergence_in_the_enterprise_ver_3.0.pdf>.

Mark A. Miller. "Implementing the VoIP Network." Network General. August 2005. www.networkgeneral.com. <http://i.i.com.com/cnwk.1d/html/itp/Network_General_Implement_VOIP.pdf>.

Mark Miller. Voice Over IP: Strategies for the Converged Network. Hungry Minds, 2000.

Mark Santkuyl. "Jitter." SearchNetworking.com. January 12 2005. www.searchnetworking.com. <http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci213534,00.html>.

Melanie Turek. "Voice and Video over IP: Leveraging Network Convergence for Collaboration." Nemertes Research. 2006. www.polycom.com.

<http://www.polycom.com/common/pw_cmp_updateDocKeywords/0,1687,5742,00.pdf>
.

"MPLS / Tag Switching." Cisco Systems. www.cisco.com.
<http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/mpls_tsw.htm>.

"MPLS based IP VPN Service." Cisco Systems. www.paetec.com.
<http://www.paetec.com/downloads/mpls_overview.pdf>.

"Multi Protocol Label Switching." International Engineering Consortium. www.iec.org.
<http://www.iec.org/online/tutorials/mpls/topic03.html>.

"Networking services for converged communications." IBM Global Services.
www.ibm.com. <http://www-935.ibm.com/services/us/gn/pdf/convservgd510-6388-00f.pdf>.

Oliver C. Ibe. Converged Network Architectures: Delivering Voice Over IP, ATM, and Frame Relay. New York: Wiley, 2002.

Paul Desmond. "The ROI of Convergence." Network World. June 14, 2004.
www.networkworld.com.
<http://www.networkworld.com/supp/2004/0621convergenceperspectives.html?page=1>.

Paul Ruppert. "Videoconferencing at UofT - Methods, Means and Modes." University of Toranto. May 05, 2006. www.utoronto.ca.
<http://content.library.utoronto.ca/rcat/services/presentationnoteslunch/pdf/videoconferencing.pdf>.

"Quality of Service Networking." Cisco Systems. www.cisco.com.
<http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/qos.pdf>.

R. Braden, D. Clark and S. Shenker. "Integrated Services in the Internet Architecture: an Overview." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc1633.txt>.

"Resource Reservation Protocol." Cisco Systems. October 2006. www.cisco.com.
<http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/rsvp.htm>.

Rick Allison. "Converged Networks Design Features." Alcatel-Lucent. www.alcatel-lucent.com. <http://www.lucent.com/livelink/176405_Whitepaper.pdf>.

"RTCP Packet Types." Freesoft.org. www.freesoft.org.
<http://www.freesoft.org/CIE/RFC/1889/47.htm>.

"RTP Header Compressions." Cisco Systems. www.cisco.com.
<http://www.cisco.com/univercd/cc/td/doc/product/software/ios112/rtphead.htm>.

S P Majumder, S M Raiyan Kabir, Rexwanur Rahman, Md. Farrukh Imtiaz and Md. Moniruzzaman. "A New Architecture of TDM Switch." IEEE. www.ieeeexplore.org. <http://ieeexplore.ieee.org/iel5/10170/32495/01517303.pdf?arnumber=1517303>.

S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss. "An Architecture for Differentiated Services." IETF. December 1998. www.ietf.org. <http://www.ietf.org/rfc/rfc2475.txt>.

S. Kent and R. Atkinson. "IP Authentication Header." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc2402.txt>.

---. "IP Encapsulating Security Payload (ESP)." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc2406.txt>.

---. "Security Architecture for the Internet Protocol." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc2401.txt>.

Saqib Jang, E. Brent Kelly and Andrew W. Davis. "A Technical FAQ: Frequently Asked Questions About Voice and Video Over IP." Wainhouse Research. January 2003. www.wainhouse.com. <http://www.wainhouse.com/files/papers/wr-faq-ip-conf.pdf>.

"Session Description Protocol." Telecom Paris. www.infres.enst.fr. <http://www.infres.enst.fr/~dax/polys/multicast/sdp.html>.

"Session Initiation Protocol." IETF. www.ietf.org. <http://www.ietf.org/rfc/rfc2543.txt>.

Steve Murphy and Steve Brodson. "Designing Converged Networks." Enterprise Networks and Servers. December 2003. www.enterprisenetworksandservers.com. <http://www.enterprisenetworksandservers.com/monthly/art.php?393>.

"Strategies for IP Telephony Evaluation and Migration: Best Practice Considerations for Deploying IPT in the Enterprise." InfoTech. April 2005. www.accessintel-infotech.com. <http://www.voicepro.com/_files/user/InfoTech%20Building%20Client%20Value1.pdf>.

"TCP/IP." Cisco Systems. 6 February, 1996. www.cisco.com. <http://www.cisco.com/warp/public/535/4.html>.

Tim Parker and Karanjit S. Siyan. TCP/IP Unleashed. 3rd ed. Sams, 2002.

"Token Ring / IEEE 802.5." Cisco Systems. October 12, 2006. www.cisco.com. <http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/tokenrng.htm>.

"Transmission Control Protocol." IETF. www.ibiblio.org. <http://www.ibiblio.org/pub/docs/rfc/rfc793.txt>.

"Understanding TCP/IP." Cisco Systems. 28 September, 2002. www.cisco.com. <http://www.cisco.com/univercd/cc/td/doc/product/iaabu/centri4/user/scf4ap1.htm>.

"Understanding Telecommunications." Ericsson. www.ericsson.com. <http://web.archive.org/web/20040306215105/www.ericsson.com/support/telecom/part-a/a-2-7.shtml>.

"Use Codepoints to assign per hop behaviors." IBM. www.ibm.com. <http://publib.boulder.ibm.com/infocenter/iseries/v5r3/index.jsp?topic=/rzak8/rzak8phb.htm>.

V. Jacobson, K. Nichols and K. Poduri. "An Expedited Forwarding PHB." IETF. June 1999. www.ietf.org. <http://www.ietf.org/rfc/rfc2598.txt>.

"Voice over IP Reliability." ShoreTel Intelligent Phone Systems. October 26, 2004. www.infinitycomp.com. <http://www.infinitycomp.com/sbs/detail/shoretel_downloads/white_papers/Reliability_Whitepaper.pdf>.

"White Paper: IP Convergence Based On SIP - Enhanced Person-To-Person Communications." Forum Nokia. June 24, 2004. forum.nokia.com. <sw.nokia.com/id/d9589d7d-ee9d-4d16-8419-b339c01ad37a/White_Paper_IP_Convergence_Based_On_SIP_v1_0_en.pdf>.

"Why IP-based Videoconferencing?" Hitachi Software Engineering. www.ipoerweb.com. <http://host271.ipowerweb.com/~hitachi-/tsg/solutions/video/ipbased.html>.

Y. Bernet, P. Ford, R. Yavatkar, F. Baker, L. Zhang, M. Speer, R. Braden, B. Davie, J. Wroclawski and E. Felstaine. "A Framework for Integrated Services Operation over Diffserv Networks." IETF. www.ietf.org. <www.ietf.org/rfc/rfc2998.txt>.

Appendix A – List of Abbreviations

AF – Assured Forwarding
AH – Authentication Header
ANSI – American National Standards Institute
ATM – Asynchronous Transfer Mode
BRI – Basic Rate Interface
CPE – Customer Premise Equipment
CSMA/CD – Carrier Sense Multiple Access with Collision Detection
DHCP – Dynamic Host Control Protocol
DiffServ – Differentiated Services
DSCP – Differentiated Services Code Point
EF – Expedited Forwarding
ESP – Encapsulating Security Payload
FDDI – Fiber Distributed Data Interface
FDM – Frequency Division Multiplexing
FEC – Forward Equivalence Class
FR – Frame Relay
FTP – File Transfer Protocol
GMII – Gigabit Media Independent Interface
GoS – Grade of Service
HTTP – Hyper Text Transport Protocol
IEEE – Institute of Electrical and Electronic Engineers
IETF – Internet Engineering Task Force
IntServ – Integrated Services
IP – Internet Protocol
IPT – IP Telephony
ISDN – Integrated Services Digital Network
ISO – International Standards Organization
ISP – Internet Service Provider
ITU – International Telecommunication Union
LAN – Local Area Network
LDP – Label Distribution Protocol
LER – Label Edge Router
LIB – Label Information Base
LSP – Label Switched Path
LSR – Label Switching Router
MAC – Media Access Control
MCU – Multipoint Control Unit
MPLS – Multi Protocol Label Switching
OSI – Open Systems Interconnection
PBX – Private Branch Exchange
PCM – Pulse Code Modulation
PCS – Physical Coding Sub-layer
PMA – Physical Medium Attachment
PMD – Physical Medium Dependant

POTS – Plain Old Telephone Service
PSTN – Public Switched Telephone Network
QoS – Quality of Service
RFC – Request For Comments
ROI – Return Of Investment
RR – Receiver Report
RSVP – Resource reSerVation Protocol
RTCP – Real-time Transport Control Protocol
RTP – Real-time Transport Protocol
SDES – Source Description
SDP – Session Description Protocol
SIP – Session Initiation Protocol
SMTP – Simple Mail Transfer Protocol
SR – Sender Report
TCP – Transmission Control Protocol
TCP/IP – Transmission Control Protocol / Internet Protocol
TDM – Time Division Multiplexing
UA – User Agents
UDP – User Datagram Protocol
VoIP – Voice over IP
VPN – Virtual Private Network
WAN – Wide Area Network
WFQ – Weighted Fair Queuing

Appendix B – List of RFCs

RFC 768 – User Datagram Protocol
RFC 791 – Internet Protocol
RFC 793 – Transmission Control Protocol
RFC 1633 – Integrated Services in the Internet Architecture: an Overview
RFC 1889 – RTP: A Transport Protocol for Real-time Applications
RFC 1890 – RTP Profile for Audio and Video Conferences with Minimal Control
RFC 2327 – Session Description Protocol
RFC 2401 – Security Architecture for the Internet Protocol
RFC 2402 – IP Authentication Header
RFC 2406 – IP Encapsulating Security Payload
RFC 2475 – An Architecture for Differentiated Services
RFC 2543 – SIP: Session Initiation Protocol
RFC 2597 – An Assured Forwarding PHB
RFC 2598 – An Expedited Forwarding PHB
RFC 2998 – A Framework for Integrated Services Operation over DiffServ Networks