

# *Quality of Service for Conferencing – Reality or Hype?*

*A guide to understanding  
and successfully deploying  
the appropriate level of QoS  
for conferencing*

# Quality of Service for Conferencing – Reality or Hype?

---

A guide to understanding and successfully deploying the appropriate level of QoS for conferencing.

December 2003



# Contents

<b>Executive Summary</b> .....	<b>4</b>
<b>What Is QoS?</b> .....	<b>5</b>
<b>Why Do I Need QoS for Conferencing?</b> .....	<b>6</b>
What generates contention? .....	6
<b>QoS - An End-To-End Concept</b> .....	<b>8</b>
QoS on the Local Area Network (LAN).....	8
QoS on the Last Mile.....	9
QoS on the Wide Area Network (WAN).....	9
<b>WAN Bandwidth Options</b> .....	<b>11</b>
<b>Flavors of QoS</b> .....	<b>12</b>
Option 1 - Best Effort (Lack of QoS) .....	12
Option 2 – Intelligent Queuing .....	12
Option 3 – MPLS (Multi-Protocol Label Switching) .....	13
Option 4 - ATM.....	14
Option 5 – Non-Standards Based Solutions.....	16
<b>The Impact on the IT Department</b> .....	<b>17</b>
<b>Real World Options</b> .....	<b>19</b>
Overlay Networks.....	20
<b>Conclusion</b> .....	<b>22</b>
<b>Checklist – Analyzing and Achieving QoS for Conferencing</b> .....	<b>23</b>
<b>Tips, Tricks &amp; Strategies</b> .....	<b>24</b>
<b>About Wainhouse Research</b> .....	<b>25</b>
About the Author .....	25
<b>About Network-i</b> .....	<b>26</b>

## List of Figures

Figure 1: Example of Network Contention.....	6
Figure 2: Typical Video-LAN Configuration.....	9
Figure 3: Typical LAN / WAN Network.....	10
Figure 4: WAN Bandwidth Options.....	11
Figure 5: Virtual Circuits in an ATM Transmission Link.....	14
Figure 6: Class of Service (CoS) Example.....	15
Figure 7: Impact on the IT Department.....	17
Figure 8: Applications on a Converged Network.....	19
Figure 9: Gradual Convergence via Overlay Networks.....	20

## Executive Summary

Today's conferencing applications are now IP friendly; they can run on either dedicated lines (like ISDN or telephone lines) or IP networks. However, as most network administrators know, conferencing applications can wreak havoc on unprepared corporate networks. The key to successfully deploying conferencing applications over corporate networks is the activation of Quality of Service (QoS).

QoS refers to a network's ability to reliably and consistently provide a certain level of throughput and performance. QoS for conferencing typically involves network availability, bandwidth, end-to-end delay, jitter, and packet loss. Simply stated, if your network doesn't conform to the minimum requirements in any of these areas, your conferences are doomed to fail.

So what's the hype behind QoS for conferencing? The fact is that not all organizations need to worry about QoS to enjoy IP based conferencing. Even though network providers and video conferencing system integrators shudder at the thought of conferencing without QoS, in reality the need for QoS depends upon your network configuration, the type of applications you run, and the expectations of your user community. In other words, organizations should do their homework before investing time and money in network upgrades. That's not to say that upgrades won't be required. However, responsible managers should understand their business requirements and all available options before signing on the dotted line of that new network contract.

QoS can be achieved in a variety of ways, including over-provisioning (deploying additional bandwidth), data prioritization, and the use of QoS-enabled overlay or converged networks. In most cases, deploying a less intelligent (non- or lower-QoS capable) network will result in higher IT and network management costs, and perhaps disappointed users. Therefore, network and conferencing managers should fight the temptation to deploy stripped-down networks as a means of cost-reduction. The money saved upfront will quickly be spent in back-office management.

From a 10,000 foot view, organizations have two main options for deploying QoS within their organizations; convergence or overlay. Convergence requires the use of QoS-capable WAN links throughout the organization. In many cases, this requires a fork-lift upgrade and migration of all network resources, which can place convergence out of reach of many cost-sensitive organizations. On the other hand, overlay networks allow a step-by-step migration from a non-QoS to a QoS network without the high cost and inherent risk of major network reconfigurations. In this way, overlay networks are a first step toward convergence.

This document is intended to provide network administrators, conferencing managers, and executives with a basic understanding of the need for QoS and the various options available.

## What Is QoS?

QoS, which stands for Quality of Service, is an overused (and often abused) term that describes the overall performance of a data network. Basically, QoS refers to how well and consistently a network can deliver predictable results. In other words, if a piece (or packet) of data enters a network in one location, just how well will that network be able to host that packet's travels from its point of origin to its final destination?

Just how do we measure QoS or network performance? For conferencing applications, we typically focus on five key performance parameters; availability (uptime), bandwidth, end-to-end delay, jitter, and packet loss.

*Quality of Service for conferencing typically involves availability, bandwidth, end-to-end delay, jitter, and packet loss.*

Availability refers to the percentage of time that the network is “up and running” and available to host data traffic. Measured in percentage of uptime, high quality network providers often provide three nines (99.9%), four nines (99.99%) or even five nines (99.999%) uptime guarantees. As a reference, a guarantee of “four nines of availability” indicates that the network will be down (or unavailable) for less than five seconds per month!

Bandwidth, also known as throughput, describes the capacity of an entire network or a portion of a network. Because the individual pieces of data that traverse a network are called bits, bandwidth is typically specified in bits per second (bps), kilobits per second (kbps), or megabits per second (mbps). For example, earlier modems provided a throughput of 300 bits per second. However, in today's conferencing networking environment, common network capacities include 384 kbps and 1.544 mbps, which can carry 384,000<sup>1</sup> and 1,544,000 bits simultaneously.

End-to-end delay, often called latency, refers to the average time it takes for a piece of data to travel successfully through the network from the point of origin to the destination. Note that one unavoidable source of latency is the transmission delay involved in the physical transmission of the piece of data. Since a piece of data cannot travel faster than the speed of light, the transmission delay for a bit to travel from New York to London, assuming a straight connection between those locations, would be 18.6 milliseconds<sup>2</sup> (or .019 seconds). Additional delays are introduced by routers and other processing equipment in the network.

Jitter is a measurement of the variation in the end-to-end delay between consecutive packets of information. In other words, if it takes packet “A” 31ms (milliseconds) and packet “B” 44 ms to traverse the network, the jitter will be 13ms. Network ratings often include a maximum (i.e. not to exceed) value, in milliseconds, for jitter.

Finally, packet loss is a measurement of the percentage of packets that will, for a variety of reasons, not reach their destination. Typically, packet loss is the result of congestion within the network when too many pieces of information arrive simultaneously at a network router or switch. In most cases, networks will simply re-transmit lost packets, but for real-time applications (like conferencing), this form of recovery is not acceptable.

Based on the above data, a typical network performance description might include 99.99% uptime, 100 mbps of bandwidth, with average latency of 34ms, maximum jitter of 12ms, and less than 1/100 of a percent of packet loss.

---

<sup>1</sup> A kilobyte (k) actually equals 1,024 (2<sup>10</sup>) bytes and a megabyte equals 1,024 kilobytes (or 2<sup>20</sup> bytes). However, when discussing network data rates and bandwidth, k is equal to 1,000 instead of 1,024.

<sup>2</sup> According to [www.javacommerce.com](http://www.javacommerce.com), the distance from NYC to London is 5,580 km (or 3,460 miles). Therefore, data traveling at the speed of light (299,792 km/hour or 186,282 miles/hour) would need .019 seconds to traverse that distance.

## Why Do I Need QoS for Conferencing?

The simple answer is that QoS helps us deal with contention between different users and applications for network resources. In this context, contention occurs when multiple applications simultaneously attempt to use the same network resources. Note that modern day networks are designed to operate well under a certain level of contention. However, when that contention reaches beyond an acceptable level, performance issues arise. Therefore QoS is needed to control the impact that applications can have on other applications.

### What generates contention?

There are two major factors that generate network contention; shared bandwidth and limited bandwidth.

- Shared bandwidth -- Contention is a byproduct of sharing network resources and bandwidth with other applications and users. Therefore, it is logical that in situations without shared bandwidth (e.g. dedicated lines that are used to support each application), one need not worry about contention. In this way, ISDN videoconferencing users and POTS (plain old telephone system) analog line audio conferencing users avoid contention by using dedicated lines for their applications.
- Limited bandwidth -- Since contention is caused by competition for existing network resources, it stands to reason that the greater your network resources, the less contention you will have. Taken to an extreme, one can all but eliminate the impact of contention just by providing adequate bandwidth. In fact, some organizations may find that over-provisioning<sup>3</sup> their network is a less expensive and better short-term option than enabling QoS. One must understand, however, that deploying additional bandwidth without QoS control will not provide performance guarantees. In addition, it stands to reason that using high bandwidth applications (like videoconferencing) will contribute to network contention.

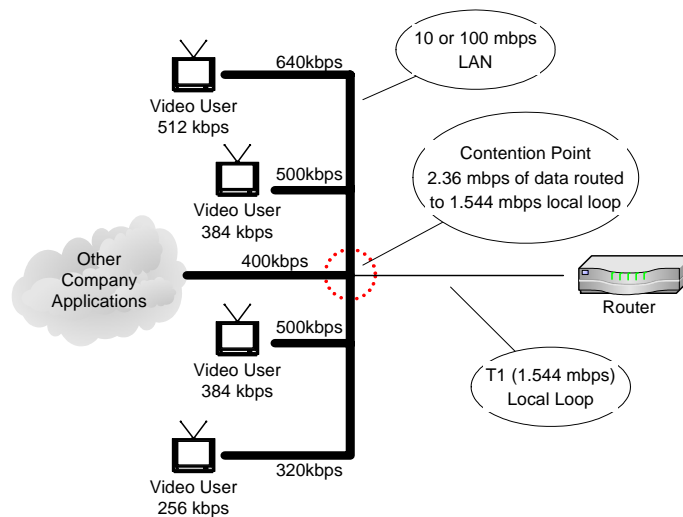


Figure 1: Example of Network Contention

<sup>3</sup> Over-provisioning involves deploying more bandwidth (capacity) than required by all applications in an effort to avoid problems caused by congestion and contention.

### *What Is the Impact of Contention?*

Contention causes resource allocation issues on the network. However, whether or not this causes a problem for your organization depends upon the answers to the following two questions:

**Question 1: How tolerant are your applications to poor QoS?** -- There are two kinds of applications; non-real time and real-time. Non-real time applications are those that do not require consistent and guaranteed throughput and network performance. In other words, these types of applications will operate properly even while facing QoS related problems like high levels of network contention, long delays, and even lost packets. Note that non-real time applications are not necessarily less important or critical for the users or the organization. However, in most cases, these applications are not “immediate” in their presentation of the requested data, and therefore can recover from network related glitches.

*Non-real time applications are not necessarily less critical than real-time applications... they are simply more tolerant to network performance issues.*

Email is an example of a critical, yet non-real time application. For example, when sending an email between Los Angeles to Tokyo, an end-to-end delay of a few seconds (or minutes) or a severe packet loss problem probably would go unnoticed by the email sender and recipient. Email programs are relatively tolerant of these types of QoS problems, and packets that are lost would simply be re-transmitted by the network.

However, real-time applications, like videoconferencing and audio conferencing, demand both high performance and consistent data throughput in order to operate effectively. In these situations, a delay of 200 ms (one-fifth of a second) or more may be noticeable and distracting to meeting participants. Furthermore, because the data must arrive in order and quickly, lost packets cannot be resent by the application. Therefore, the loss of even a few packets of data can cause significant audio and video artifacts, thereby interrupting the meeting in progress.

What does this all mean? Real-time conferencing applications require tighter control of network performance than non-real-time data applications like e-mail and web browsing.

*Real-time conferencing applications require tighter control of network performance than asynchronous data applications like e-mail and web browsing.*

**Question 2: How tolerant are your users to QoS problems?** -- As stated above, QoS problems may negatively impact conferences in progress. However, whether or not this is a problem worthy of remediation depends upon the expectations of your user community. If your users are willing to accept occasional (or even frequent) audio and video distortion during their meetings, QoS may not be necessary. However, most enterprise users expect solid and reliable performance during every meeting. Providing a consistent level of conferencing performance on a shared network requires either excessive amounts of bandwidth or the activation of QoS on the network.

*Providing a consistent level of conferencing performance on a shared network requires either excessive amounts of bandwidth or the activation of QoS on the network.*

## QoS - An End-To-End Concept

In order to enjoy quality of service performance during a conference, that quality of service must be in place from one end of the conference to the other. In this way, quality of service is an end-to-end concept, and the QoS provided by a network is only as good as the weakest link in the network chain. This can have a dramatic impact on the scope of work, cost, and reality of activating QoS on a company's network.

*Never forget that QoS is an end-to-end concept.*

When running time-sensitive applications on an IP network, one must realize that Ethernet networks were not designed to provide guaranteed levels of performance. Therefore, achieving quality of service requires that some combination of hardware, software, and/or intelligence be deployed throughout the following three parts of the network; the LAN, the "last mile," and the WAN.

### *QoS on the Local Area Network (LAN)*

Many people believe that quality of service is not an issue on the LAN due to the emergence of 10mbps, 100mbps, and Gigabit LANs. While true that most LANs do have adequate bandwidth to carry IP videoconferencing traffic, adequate bandwidth alone does not guarantee consistent conferencing performance.

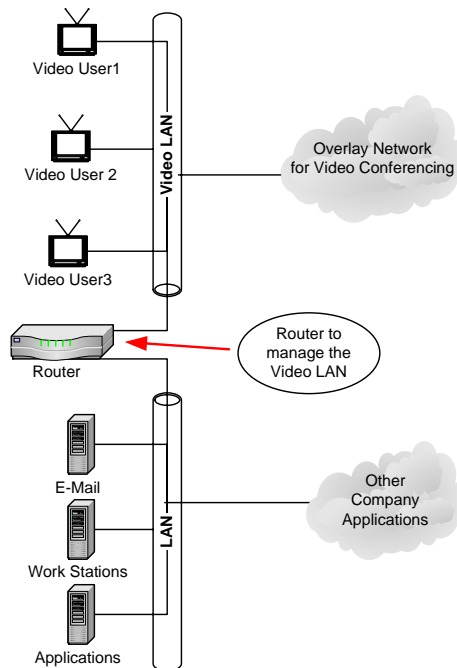
*Adequate bandwidth alone does not guarantee consistent conferencing performance.*

To understand this, we must consider that many networking companies recommend limiting LAN traffic to 30% or 40% of rated capacity. Therefore, a 100 mbps segment can actually handle only 30 mbps of traffic. Recognizing that a single LAN segment may actually serve 100 or more users, we can see how this segment could fill up very quickly. In fact, if even a few users activated bandwidth hungry applications (like FTP or even a web browser), conferences in progress may be impacted. There are two common solutions for this problem:

**Option 1 – Intelligent Subnetting** -- LAN contention can be managed by intelligently engineering and planning the network routing. For example, a company could choose to place all videoconferencing systems on only certain subnets to avoid co-mingling conferencing traffic with other LAN traffic. However, this is easier said than done and may involve a total re-design of the network and new IP address assignments. This is further complicated by the fact that in some cases a user's workstation is also their videoconferencing system. In the end, this is a costly, time consuming, and typically short-term solution.

**Option 2 – The V-LAN** -- To minimize the commingling of data and conferencing traffic on the LAN, many companies choose to create a totally separate overlay network, often called a Video-LAN. Note that due to the similar nomenclature, many people confuse Video-LANs with Virtual-LANs (or V-LANs). However, while a Virtual-LAN is a software-based grouping of workstations and servers, a Video-LAN involves a totally separate, parallel network and therefore provides total isolation of conferencing traffic from other LAN data. Since most organizations have a limited number of videoconferencing systems deployed in each facility, creating a Video-LAN is relatively inexpensive compared to the cost of deploying other QoS management schemes on the LAN.

The next diagram illustrates a typical V-LAN configuration, including the connection of the V-LAN to the LAN toward the edge of the network for control and management traffic.



**Figure 2: Typical Video-LAN Configuration**

### *QoS on the Last Mile*

QoS must also be maintained on the connections between the corporate LAN and the WAN. These connections, sometimes called the “local loop” or “last-mile” connection, typically consist of dedicated lines running from the end-user facility to the nearest point of presence (POP)<sup>4</sup> of the network provider.

The type of last mile connection utilized depends upon a number of factors including the required bandwidth and the types of lines available. However, typical connections include T1s, T3s, DSL, frame relay, and ATM. Each of these connections can provide varying amounts of bandwidth and QoS control, and therefore the last mile is another area of concern for network designers.

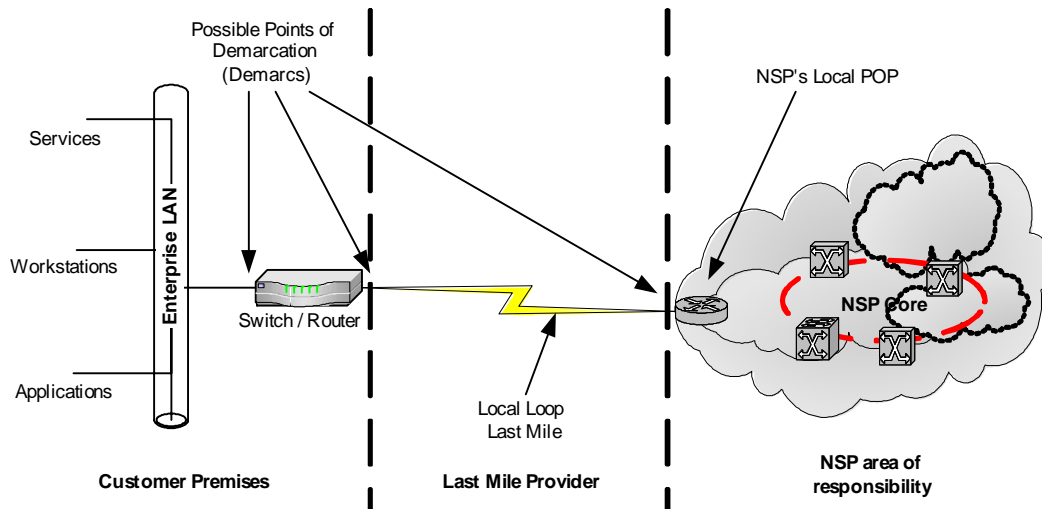
### *QoS on the Wide Area Network (WAN)*

For financial reasons, most companies purchase their WAN bandwidth from network service providers (NSPs) or telephone service providers (Telco’s). To do so, they contact the provider and request a network connection, of a certain capacity, and QoS level, between certain locations. As one might imagine, larger pipes, higher QoS levels, and longer distances increase the monthly price for those lines.

---

<sup>4</sup> According to <http://searchnetworking.techtarget.com>, a POP is an access point to the Internet. For our purposes, a POP is a space either owned or rented by a network carrier that provides access to their network backbone. For example, a network provider may have one or more New York and London POPs.

It should come as no surprise that NSPs can only monitor and manage performance of network segments under their control. Therefore, network providers will often provide QoS guarantees only for network traffic traveling between their network points of presence (or POPs). This part of the provider's network is called the network backbone or core. In some cases, the NSP's scope of responsibility, or demarcation, ends at the point where the traffic leaves their POP. Alternatively, if the NSP provides both the WAN and the last-mile, the NSP's scope of responsibility may include the last-mile connections and perhaps even the switch / router on the customer's premises. It is important for conferencing and network managers to clearly understand the NSP's scope of responsibility.



**Figure 3: Typical LAN / WAN Network**

### ***Internal vs. External Contention***

In the last section, we discussed the contention caused by applications competing for the same network resources. Since this traffic is all generated by the same company, this is called internal contention. However, companies that outsource their WAN bandwidth to network service providers may have to contend with traffic generated by other companies, called external contention. External contention represents not only a QoS issue, but also a potential security issue. Therefore, when companies contract for their WAN bandwidth, they must carefully consider their options and should seek SLAs (service level agreements) from vendors that would guarantee a minimal level of service.

***Internal contention is caused by applications within the same company competing for network resources. External contention is caused by competition for network resources from other companies.***

## WAN Bandwidth Options

Companies have a variety of options when seeking WAN bandwidth. Depending upon security requirements, QoS requirements, and budgetary restraints, bandwidth options range from dedicated (private) lines to the public internet.

### *Option 1 – Private / Leased Line WAN*

Some organizations opt to purchase or lease private lines between their major locations for intra-company communications. In effect, these companies are building their own dedicated network to host their internal traffic between offices. Although this option avoids external contention, QoS is still required to overcome internal contention for network resources.

### *Option 2 – Private VPN*

Some companies contract with a single network provider to provide a VPN<sup>5</sup>, or Virtual Private Network, between their facilities. In this case, all VPN traffic travels over a single provider's network (or backbone). This provides a number of benefits including lower cost and increased flexibility. Assuming the VPN supplier has properly configured their VPN service, VPN users should enjoy the same lack of external contention as dedicated bandwidth users. However, the concerns over internal contention still remain, and therefore QoS may still be required.

### *Option 3 – Public / Internet VPN*

Smaller companies may elect to connect their facilities using a VPN over the public Internet. Since the traffic is traveling over the public Internet, users are subject to both internal and external contention for network resources. In effect, this is a non-QoS connection.

The following table summarizes the most common types of WAN bandwidth and the potential for internal and external contention on those lines.

Type of WAN Connection	Internal Contention	External Contention	Cost
Private / Leased Line WAN	Yes	No (dedicated lines)	High
Private VPN	Yes	No (traffic separated by NSP)	Moderate
Public Internet VPN	Yes	Yes	Low

**Figure 4: WAN Bandwidth Options**

---

<sup>5</sup> A VPN, or virtual private network, is a semi-permanent and secure (encrypted) data connection between locations over a public, semi-public, or private network.

## Flavors of QoS

The previous sections described the meaning of QoS, situations under which QoS may be needed, and LAN/WAN options for companies needing QoS. In short, QoS is necessary on shared networks when multiple applications are competing for limited network resources. In most cases, once videoconferencing traffic runs over the organization's primary data network, some type of QoS is required. Organizations seeking to implement QoS have a variety of options at their disposal.

### *Option 1 - Best Effort (Lack of QoS)*

The first option is really the choice not to deploy quality of service at all. In effect, these users are choosing to accept the problems associated with running conferencing traffic over a non-QoS-enabled network. For example, organizations may deem that deploying QoS for a telecommuter's home network may be cost-prohibitive and unnecessary.

### *Option 2 – Intelligent Queuing*

As data traverses a network, it travels between devices called switches and routers. Similar to toll booths on an expressway, these network devices act as traffic cops directing traffic from one location to another. By controlling the priority these devices give to certain types of traffic, network designers can yield improved performance for certain applications.

When data arrives at a switch or router, it enters a queue (or a waiting line), and waits to be sent on its way. In addition, if too many pieces of data arrive at a single device simultaneously, the device may simply discard the newly arrived data to focus on processing the items already in the wait queues. Waiting in queues and overloading of switches and routers are common sources of latency (delay) and packet loss on data networks.

To improve performance, some network devices have multiple queues designed to process only certain types of traffic. Which queue a piece of network traffic reaches is determined by the value of certain bits in the packet's header<sup>6</sup>, which are set by the originating device or another device on the source network. Through this priority queuing methodology, time-sensitive voice and video data is processed more quickly than non-real time data, resulting in improved QoS performance for those applications.

The two most common forms of intelligent queuing are IP precedence and DiffServ. Both IP precedence and DiffServ utilize bits of the data packet's header for prioritization. IP precedence modifies bits 9 – 11 in the "header", or ID label, of the IP data packet and provides eight different classifications, ranging from a highest priority of seven to a lowest priority of zero. DiffServ uses bits 9-14 of the IP packet header<sup>7</sup> and provides up to 64 different classifications, called types of service (or TOS). Many videoconferencing endpoints, including current products from Polycom, TANDBERG, VCON, and Sony, have the ability to set the IP precedence and TOS bits

---

<sup>6</sup> According to [www.hyperdictionary.com](http://www.hyperdictionary.com), a header is "the portion of a packet, preceding the actual data, containing source and destination addresses, error checking and other fields."

<sup>7</sup> As of this writing, DiffServ actually reserves bits 9 – 16 of the IP packet header. However, bits 15 and 16 are currently unused by most networks. Therefore, DiffServ actually provides only 64 different classifications.

during IP video calls. For endpoints without this capability, external QoS management devices can be deployed to set these priority bits. However, in environments utilizing either IP precedence or DiffServ, QoS will not be achieved unless all switches and routers in the network have the ability to process these prioritization requests. Therefore, devices that can't process these priority tags must be upgraded. For this reason, the deployment of IP precedence and/or DiffServ can be expensive and time consuming for many organizations.

***In IP precedence and DiffServ environments, QoS will not be achieved unless all switches and routers in the data path can process these prioritization requests.***

It is important to recognize that while IP precedence and Diffserv prioritize data within the network, they do not guarantee the delivery of that data.

### ***Option 3 – MPLS (Multi-Protocol Label Switching)***

Technically speaking, MPLS is a packet switching protocol defining how information traverses a data network. However, for the purposes of this document, we will focus on MPLS' inherent QoS capabilities and the performance enhancements MPLS offers compared to other QoS methods.

In conventional networks, packet forwarding (routing) decisions are made by each router in the packet's path from source to destination. This requires each router to perform the following tasks:

- 1) Receive the packet in the main input queue
- 2) Open the packet header
- 3) Analyze the packet's priority bits – if QoS is enabled (IP precedence or DiffServ)
- 4) Move the packet to the appropriate priority queue – if QoS is enabled
- 5) Look up the routing information from a lookup table
- 6) Forward the packet to the next router

The above process, and especially the look up in step five, can cause delays, bottlenecks, and even lost packets depending upon the flow of traffic to this router.

In MPLS networks, the data packets are tagged (labeled) with more descriptive headers that include both priority and routing (destination) information. In effect, the packet's path through the network is pre-defined when it enters the network. Therefore, when the packet arrives at a router, the router can make forwarding decisions based on the contents of the packet header (or label) instead of a time-consuming look up.

As is the case with other QoS options, MPLS must be enabled on all routers and switches throughout the network in order to be effective. This may require hardware and/or software upgrades throughout the network. Once activated, MPLS provides a significant performance improvement and allows for effective convergence of voice, video, and data over a single data network. However, it is important to understand that MPLS, like the other queuing options, does not in itself guarantee packet delivery. Should network traffic exceed the capacity of the network routers, MPLS networks will experience delays and dropped packets. In effect, MPLS could be considered an intelligent form of best-effort and a solid mixture of performance and cost-effectiveness.

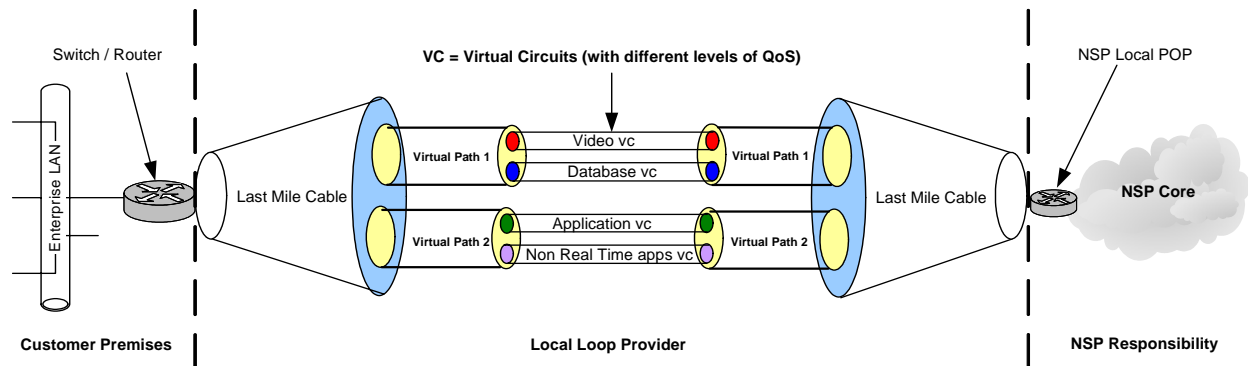
Due to the high cost, complexity, and the need to outfit the entire network, MPLS tends to be out of reach of most enterprise networks. However, a number of service providers have deployed MPLS over all or part of their network. By purchasing services from these companies, enterprise organizations can enjoy the benefits of MPLS on their WAN without the need to build an MPLS network.

## Option 4 - ATM

The QoS options described previously have involved the prioritization of data traffic within network switches and routers. In many situations, these “best effort” methodologies will yield the desired performance on the data network. However, when performance guarantees are required, ATM networks may be a better option.

ATM, which stands for asynchronous transfer mode, is a switching technology designed for high-performance multimedia networking. For our purposes, there are four important aspects of ATM as follows; virtual circuits, class of service, fixed-cell length, and hardware switching.

**Virtual Circuits** -- ATM allows the creation of virtual circuits (or VCs), which according to the International Engineering Consortium, is an end-to-end connection with defined end points and routes, but without any dedicated bandwidth.<sup>8</sup> Bandwidth for virtual circuits can be configured either as static (called a permanent virtual circuit or PVC) or dynamic (called a switched virtual circuit or SVC). The use of VCs means that data traveling across an ATM network rides on a virtual dedicated highway, totally segregated from other traffic. For this reason, ATM networks are able to provide performance guarantees and a level of QoS beyond that provided by best-effort, DiffServ, IP Precedence, and MPLS networks.



**Figure 5: Virtual Circuits in an ATM Transmission Link**

In this drawing, the video traffic is traveling along a virtual circuit (labeled “Video vc”) within virtual path 1 and provides a specific level of Quality of Service. The database traffic, application traffic, and non-real time application traffic are segregated in their own virtual circuits, and therefore the data will not commingle with other traffic.

***ATM networks are able to provide performance guarantees and a level of QoS beyond that provided by best-effort, DiffServ, IP Precedence, and MPLS networks.***

<sup>8</sup> Source: The International Engineering Consortium’s ATM Fundamentals Guide at [http://www.iec.org/online/tutorials/acrobat/atm\\_fund.pdf](http://www.iec.org/online/tutorials/acrobat/atm_fund.pdf)

**Class of Service (COS)** – A Class of Service is a set of QoS guidelines or policies that can be applied to specific network paths or virtual circuits. The following table shows examples of different classes of service:

Class	Availability	Delay / Latency (Round Trip)	Jitter	Packet Loss
1	99.999%	50ms	< 10 ms	< 0.1%
2	99.99%	60ms	< 15 ms	< 0.15%
3	99.9%	85ms	< 20 ms	< 0.7%
4	99.5%	110 ms	< 25 ms	< 1.0%

**Figure 6: Class of Service (CoS) Example**

In technical terms, class 1 is typically referred to as constant bit rate, or CBR. Class 2 and 3 are commonly referred to as variable bit rate – real time (VBR-rt) and variable bit rate – non-real-time (VBR-nrt) respectively. Finally, class 4 data is called UBR, or unspecified bit rate traffic.

As one might expect, cost is related to performance level, and therefore class 1 (CBR) bandwidth is more expensive than class 4 (UBR) bandwidth. The benefit of ATM is that it allows network managers to divide physical circuits into virtual circuits (see prior drawing) providing different classes of service. For example, a 1.544 mbps ATM circuit (delivered over a single network connection) could be parsed or divided into a 0.2 mbps (200k) Class 1 circuit for critical transaction data, a 1mbps Class 2 virtual circuit for video-conferencing traffic, and a 0.344 mbps Class 4 virtual circuit for accessing the Internet. By customizing the performance of these virtual circuits, NSPs can provide customers with a solution that precisely meets their requirements. In addition, since virtual circuits are defined in software, NSPs can modify the performance levels “on the fly” to meet changing customer requirements.

**Fixed-Cell Length --** On ATM networks, data for transmission is divided into 53 byte data packets (or ATM cells). This small and fixed cell size helps avoid delays caused by long data frames or packets on other network types.

**Hardware Switching --** ATM data headers are formatted to allow for high-speed, hardware-based switching. This allows all packet switching (and forwarding) to occur much more quickly than on routed networks (like those supporting DiffServ, IP Precedence, and MPLS). The result is that data packets traverse the ATM network much more quickly than on other network topologies.

The combination of the above four ATM characteristics makes ATM networks ideally suited to meet converged, high-performance, and multimedia networking requirements. Specifically, it allows network service providers (NSPs) to offer end-users a variety of managed, scaleable, and high-performance network services in a cost-effective manner. However, the high cost and complex management requirements make ATM better suited for WAN than LAN applications. Therefore, ATM is considered a WAN technology and is typically deployed by carriers and service providers.

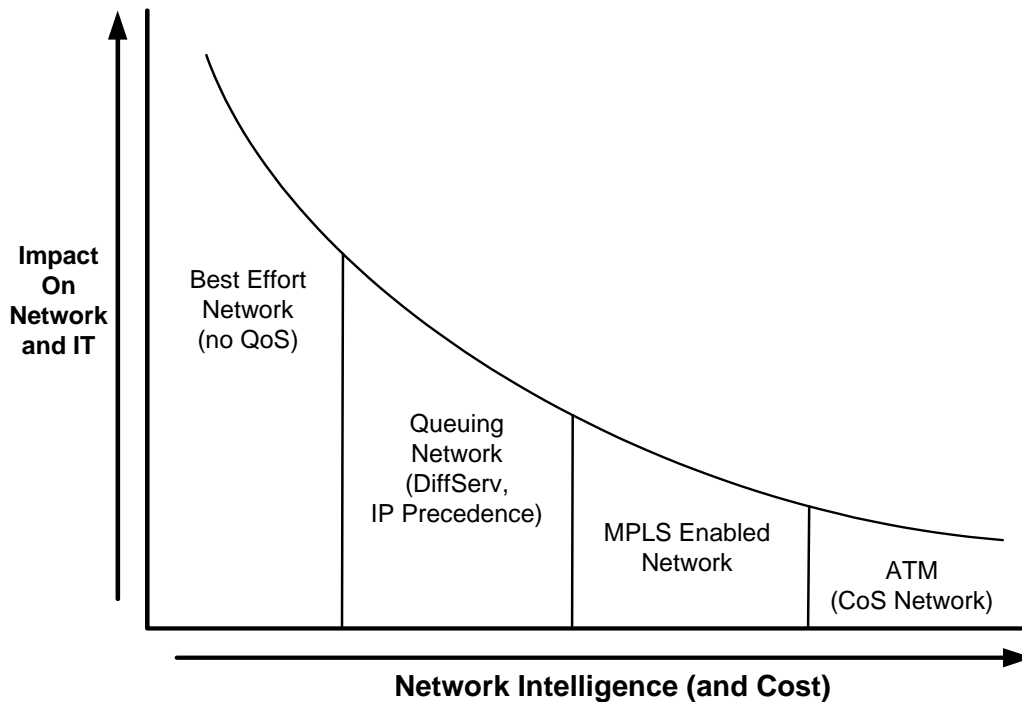
*Although relatively expensive, ATM networks are ideally suited to meet converged, high-performance, and multimedia networking requirements.*

### ***Option 5 – Non-Standards Based Solutions***

Although IP Precedence, DiffServ, MPLS, and ATM CoS are the most common “standards-based” forms of QoS delivered today, there are many other non-standards based solutions, each of which approach QoS problems from a different direction. For example, one company offers time-based packet sequencing as a means of scheduling packet delivery through IP networks. This requires the installation of specialized hardware throughout the network, but provides increased throughput and certain performance guarantees. Another company offers a QoS monitoring engine that analyzes network performance and re-routes traffic, in-real time, onto secondary (or overlay) networks to circumvent QoS issues. For example, using this technology a company could run its priority traffic over the Internet (using a secure, encrypted VPN) and divert it onto a more-expensive dedicated network when the Internet’s performance falls below a certain threshold. By using the Internet to carry much of their traffic, companies can significantly decrease their bandwidth requirements and networking costs.

## The Impact on the IT Department

No matter how the challenge is tackled, providing quality of service for conferencing applications will have an impact on an organization's IT or networking department. The magnitude of that impact, however, is proportional to the level of intelligence within the network and the management services included within the network offering as shown in the chart below:



**Figure 7: Impact on the IT Department**

As this chart illustrates, running conferencing applications on low-intelligence networks will increase the impact and burden on the IT / network department. For example, in non-QoS-enabled environments, the IT team will field various complaints ranging from inability to make connections to frozen video and choppy audio. In queuing and MPLS networks, these complaints will be less frequent, but not totally avoided.

***Running conferencing applications on low-intelligence networks will increase the impact and burden on the IT / network department.***

To minimize the impact on the IT department, companies should consider offloading the management of the conferencing network onto service providers. Although this sounds complicated, it simply requires the selection of a service provider that includes network monitoring and QoS guarantees as a part of their offering. In fact, many service providers bundle services like video endpoint management, scheduling software, automatic call launching, end-user training, and even online reporting packages with their network offering. Therefore, with a single PO, your company can enjoy the benefits of a high-performance network and other time-saving management tools. Although outsourced solutions like these may not be interesting to all organizations, due diligence requires that one consider the available options.

The take-away of this section is that saving money by running high-performance, real-time applications on non-QoS-enabled, limited intelligence networks will generate additional costs in terms of IT management and user dissatisfaction. Taken to an extreme, users will stop using these cost-saving and efficiency tools, like videoconferencing, because the service provided is not reliable or consistent. IT and conferencing managers should carefully consider these soft-costs when making decisions about network provisioning, topologies, and service providers.

## Real World Options

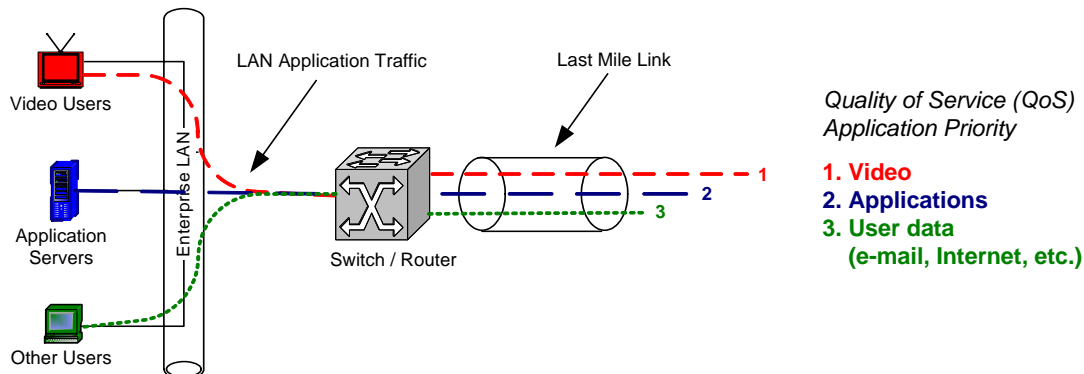
Most organizations have two options for deploying conferencing over QoS networks; convergence and overlay. Convergence networks host voice, video, and data over a single network. On the other hand, overlay solutions involve the deployment of separate networks for demanding applications.

### *Convergence*

At first glance, the concept of converging all types of data traffic onto a single network seems logical. Benefits include decreased cabling, economies of scale for management, and significant hardware cost savings. Furthermore, managing one network vendor is easier than managing several. Finally, purchasing a limited number of larger data pipes will decrease the cost per megabit of bandwidth.

However, we must consider that most data networks in use today are not QoS enabled. Therefore, chances are that the existing data network cannot host these demanding real-time applications. This means that a convergence initiative could require a total network migration from one data provider to another. This forklift upgrade would be expensive and time-consuming to activate, especially considering that only a limited number of applications and users would appreciate the benefits of the network enhancement. For these reasons, most companies have yet to activate convergence initiatives.

Nonetheless, for companies already using a QoS-enabled network, convergence may be cost-effective and relatively easy to achieve. The figure below shows the data flow for a typical converged network providing three different levels of QoS (video, application, and data).



**Figure 8: Applications on a Converged Network**

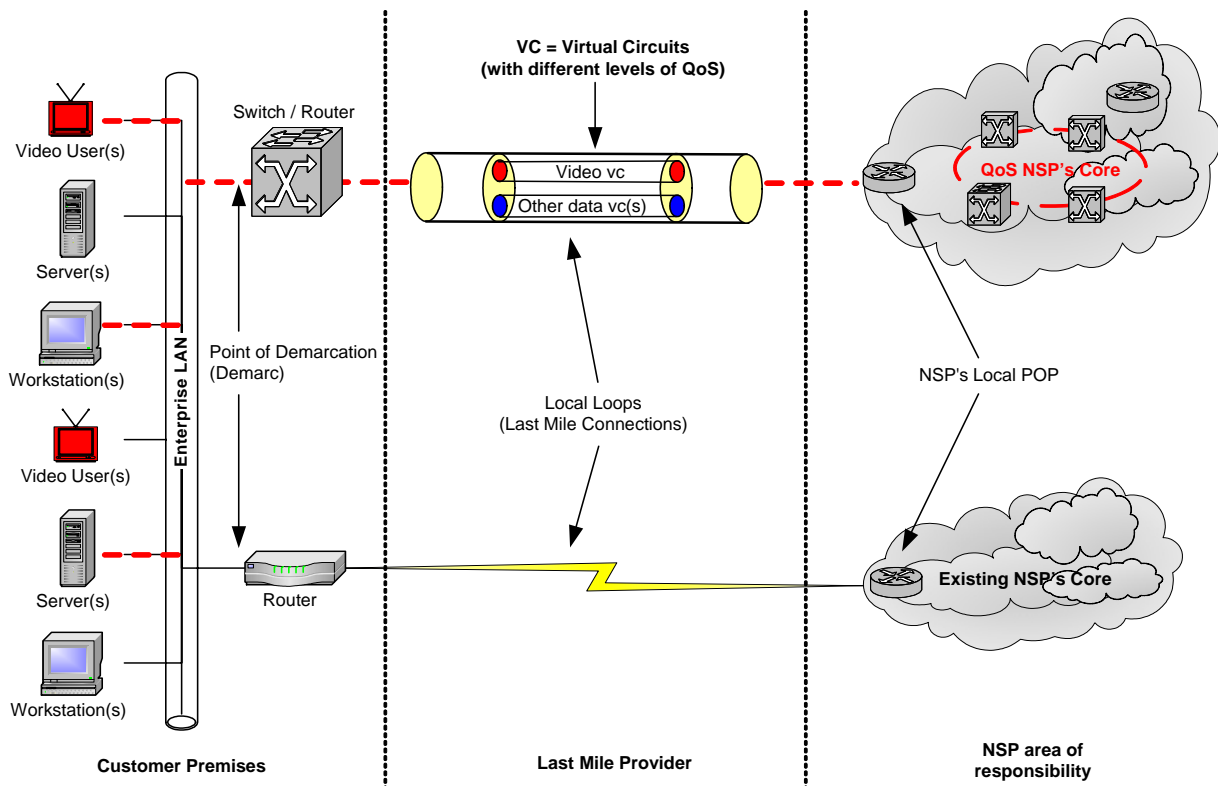
The key to finalizing your company's convergence plan requires an understanding of the QoS capabilities of your current network service provider(s).

## Overlay Networks

For many companies, the cost and time involved in convergence is prohibitive. Assuming performance improvements are necessary and cost-effectiveness is a priority, these companies should consider the deployment of an overlay network for their real-time conferencing applications. Overlay networks, often referred to as parallel networks, have several advantages including:

- Total isolation from the existing data network
- Strong performance and security
- Fast and easy deployment
- Relatively low cost (adding a few additional lines is less costly than a major network migration)

In addition, overlay network solutions allow companies to migrate specific types of traffic over to the overlay network without impacting other data applications. This gives network managers the ability to gradually migrate all traffic over to the new overlay network, thereby achieving convergence in a slow, controlled, and responsible manner. The diagram below demonstrates this staged approach to convergence:



**Figure 9: Gradual Convergence via Overlay Networks**

In the diagram above, this company has deployed a QoS overlay network (top half of the drawing) and is currently directing the traffic from some video users, servers, and workstations across the QoS network. Applications that don't require QoS continue to ride over the existing, non-QoS enabled, and less expensive WAN (lower half of the drawing).

Some IT managers will reject the idea of an overlay network because it highlights the fact that the existing data network is not equipped to handle conferencing traffic. The overlay network, however, allows managers to surgically attack the QoS problem by deploying only the bandwidth required to support real-time, multimedia applications. Therefore, overlay networks should be viewed as a first step to convergence, rather than an alternative to convergence.

*Overlay networks should be viewed as a first step to convergence,  
rather than as an alternative to convergence.*

## Conclusion

For many organizations, the only way to successfully host conferencing applications on the corporate LAN / WAN will be to activate quality of service on their network. This is a must to ensure that bandwidth hungry applications, like videoconferencing, do not step on other mission critical data applications.

As described earlier, QoS is an end-to-end concept. In other words, QoS must be in place on all network links in the data path in order to be effective. In effect, network performance is only as strong as the weakest link in the data chain. Therefore, the benefits of a QoS enabled WAN may be blocked by an inadequate local loop or enterprise LAN. This means that network administrators must view QoS from a bird's-eye view in order to plan a successful QoS project.

Depending upon network topology and cost limitations, companies can choose between a variety of QoS options including prioritization (DiffServ and IP Precedence), MPLS, and ATM / CoS networks. In most cases, higher performance networks are more expensive than those with lower levels of QoS. Therefore, organizations should carefully evaluate their requirements and plan their networks accordingly.

Although negatively viewed by many network administrators, overlay networks allow organizations to gradually deploy QoS on their networks without the need for major overhauls and forklift upgrades. By deploying a limited amount of QoS enabled bandwidth, administrators can selectively route demanding traffic over the QoS WAN while keeping lower priority traffic on the existing, less expensive network. Over time, additional traffic can be migrated over to the QoS WAN until such a time that all traffic is converged onto a single network. In this way, overlay networks can help companies bridge the gap between their current network and a single, high-performance, and cost-effective converged network.

The “hype” pervading the conferencing and networking industries is that the activation of QoS throughout the network is a pre-requisite for running real-time conferencing applications on a data network. This misconception has delayed, or even stalled, many potentially beneficial and cost-saving deployments of IP-based conferencing. The fact is that depending upon the current network architecture and user expectations, QoS enhancements may or may not be necessary for your organization. Therefore, instead of succumbing to the temptation to spend more than necessary on massive data pipes or forklift network upgrades, managers should carefully analyze the need for QoS within their organizations prior to signing on that dotted line.

## Checklist – Analyzing and Achieving QoS for Conferencing

The checklist below highlights some of the steps necessary to analyze and understand the level of QoS necessary for your organization's conferencing applications.

[ ]	Inventory	Generate a list of all existing equipment and systems, including video endpoints, gateways, gatekeepers, and bridges (MCUs).
[ ]	Analyze Usage	Gather any and all available information about conferencing system usage – by location, department, and person.
[ ]	Understand End-user Expectations	Meet with end-users to understand their requirements. Focus on applications, types of meetings, frequency of meetings, and performance expectations.
[ ]	Estimate Bandwidth Requirements	Based on the inventory, usage data, and end-user meetings, generate a bandwidth and QoS requirement chart between locations.
[ ]	Understand the LAN	Meet with IT and network managers to learn about the LAN capabilities throughout the organization. Explain concerns regarding high-bandwidth applications and discuss options including V-LANs and intelligent subnetting.
[ ]	Understand the WAN	Meet with IT and network managers to learn about the existing WAN in terms of bandwidth, current QoS levels, and QoS capabilities between locations.
[ ]	Generate Concern List	Based on BW requirements estimate and LAN / WAN meetings, generate a list of potential problem areas. Specifically highlight problem locations and users that may be impacted.
[ ]	Consider Convergence	With WAN data in hand, meet with IT managers and evaluate the possibility for a converged QoS WAN (may be impossible for cost / technology reasons).
[ ]	Meet with Overlay Vendors	Reach out to network service providers for pricing and availability on overlay networks for conferencing. Provide vendors with locations (including telephone exchanges), bandwidth, and QoS requirements.
[ ]	Generate Shortlist	Gather data from vendors and generate vendor shortlist.
[ ]	Calculate ROI	Calculate the costs and ROI (verses ISDN) for the various overlay options
[ ]	Analyze Cost vs. Benefit	Analyze the benefits offered by the overlay solutions compared to the total costs (including management). Consider SLAs offered by each vendor.
[ ]	Decide and Deploy	Assuming costs and ROI are acceptable, select and deploy overlay network. Meet regularly with IT team to continue gradual migration onto new converged network.

**Figure 10: Checklist for Videoconferencing Success**

## Tips, Tricks & Strategies

The following tips and tricks will help conferencing and network managers make wise decisions for their organizations.

- 1) Consider both hard and soft costs
- 2) Consider both hard and soft benefits
- 3) Keep things as simple as possible (avoid massive change and complexity)
- 4) Remember that the goal is to improve conferencing performance for your users  
(Focus on user needs, not your personal interest in new technologies)
- 5) Avoid the temptation to over-deploy, over-spend, or over-commit
- 6) Remember that this is not a competition (buy only the services you need)
- 7) Look inside before outside (figure out what you already have before reaching out for help)
- 8) Don't underestimate the high cost of IT management and user support
- 9) Take into account the short, medium, and long term goals of your organization
- 10) Analyze, plan, and deploy for both today and tomorrow
- 11) Deploy the bandwidth you need for the applications you need to support
- 12) Buy the best quality network services and high-performance bandwidth you can afford
- 13) Plan on staged deployments – by application, location, and department
- 14) Keep the users informed throughout the upgrade, migration, and improvement process
- 15) Develop and maintain partnerships with other involved departments such as IT, network management, network security, audio-visual, conferencing management, and facilities / corporate services.

## About Wainhouse Research

Wainhouse Research ([www.wainhouse.com](http://www.wainhouse.com)) is an independent market research firm that focuses on critical issues in rich media communications, videoconferencing, teleconferencing, and streaming media. The company conducts multi-client and custom research studies, consults with end users on key implementation issues, publishes white papers and market statistics, and delivers public and private seminars as well as speaker presentations at industry group meetings. Wainhouse Research publishes *Conferencing Markets & Strategies*, a three-volume study that details the current market trends and major vendor strategies in the multimedia networking infrastructure, endpoints, and services markets, as well as the segment report *Video Communications Management Systems*, the free newsletter, *The Wainhouse Research Bulletin*, and free e-zine, *ConferencingBuyer*. To learn more about conferencing, collaboration, and networking, please review other white papers and documents available from Wainhouse Research at [www.wainhouse.com](http://www.wainhouse.com).

### *About the Author*

**Ira M. Weinstein** is a Senior Analyst and Consultant at Wainhouse Research, and a 13 year veteran of the conferencing, collaboration and audio-visual industries. Prior to joining Wainhouse Research, Ira was the VP of Marketing and Business Development at IVCi, managed a technology consulting company, and ran the global conferencing department for a Fortune 50 investment bank. Ira's current focus includes IP video conferencing, network service providers, global management systems, scheduling and automation platforms, ROI and technology justification programs, and audio-visual integration. Mr. Weinstein holds a B.S. in Engineering from Lehigh University and is currently pursuing an MBA in Management and Marketing.. He can be reached at [iweinstein@wainhouse.com](mailto:iweinstein@wainhouse.com).

## About Network-i

This white paper was sponsored by Network-i ([www.network-i.net](http://www.network-i.net)). Headquartered out of London, England, since 1996 and an ISO 9001-2000 accredited Company, Network-i provides IT business solution services delivered over an end-to-end managed international network.

### Quality of Service

Network-i delivers IT business solutions with guaranteed service levels and a Quality of Service (QoS) that is hard to match in the industry.

### Network-i Managed Network Support

Networks have become the backbone of business; supporting mission critical systems and applications to deliver competitive advantage. The significance of the network has suddenly come to the forefront with organisations having to constantly assess their network architecture, security, reliability, performance and support. Backed by proven experience, Network-i has developed a comprehensive understanding of what it takes to provide a business-to-business communications backbone, built to deliver a comprehensive managed network service. For a single source, total turnkey solution tailored for your business requirements, Network-i provides the complete management solution that is 'supplier demarcation independent'.

### Network-i Managed Network Services

Network-i takes a highly proactive approach to ensure that the design of your network meets your business need, and that the operation of your network continually achieves predetermined service level agreements. Continuity of network service is a critical factor behind business productivity. Network-i can provide continuous monitoring of your network 24 hours per day, and provide status reports covering performance against agreed service levels, based on actual usage. Our detailed monitoring and reporting services ensure that bottlenecks or variances of traffic type against predicted levels can be spotted early and acted upon. Having such an early feedback mechanism provides essential information for organizations that are adopting more of an e-business approach to operations. As the pace of change increases, so does the need for network infrastructures to adapt to meet business application demands.

For enterprises that require a decision support network Network-i provides a 'state of the art,' executive level, 'always on' video conferencing service delivered via a dedicated video network along with the capability of delivering audio and video streaming service. For more information about our InVision services, please go to [www.network-i.net/is/](http://www.network-i.net/is/).

For a single source, total turnkey solution tailored to meet your specific business requirements, turn to Network-i for our experience, our expertise, and our 100% commitment to performance.