Gigabit Ethernet, QoS, and Multimedia Applications.

Rivier College  
Course: CS575A, Advanced LANs  
Semester: Spring 2005  
Professor: Dr. Vladimir Riabov

Prepared by: Jeff Corbit  
Date: 04/28/2005
# Table of Contents:

1. Executive Summary ........................................................................................................ 3
2. Gigabit Ethernet Technology .......................................................................................... 4
   2.1. IEEE 802.3z and 802.3ab Protocol Architecture ....................................................... 4
       2.1.1. MAC Layer ........................................................................................................ 5
       2.1.2. Gigabit Interface Converter (GIC) ................................................................. 6
       2.1.3. Physical Medium ............................................................................................ 6
3. VoIP ..................................................................................................................................... 6
   3.1. VoIP Components ....................................................................................................... 7
   3.2. VoIP Network Requirements .................................................................................... 7
4. Ethernet Quality of Service (QoS) .................................................................................... 7
   4.1. Integrated Services ..................................................................................................... 8
       4.1.1. RSVP Protocol ................................................................................................. 9
   4.2. Differentiated Services ............................................................................................... 10
       4.2.1. Differentiated Services Components/Operation .............................................. 11
       4.2.2. Differentiated Services Topology ................................................................. 13
   4.3. Multi-Protocol Label Switching (MPLS) ................................................................... 14
       4.3.1. Implementation ............................................................................................... 14
       4.3.2. Traffic Engineering ......................................................................................... 15
5. Vendors ................................................................................................................................ 15
   5.1. Cisco Systems ............................................................................................................ 15
   5.2. Foundry Networks ...................................................................................................... 16
6. Market Analysis .................................................................................................................. 17
7. Conclusion .......................................................................................................................... 17
8. Glossary ............................................................................................................................... 17
9. References .......................................................................................................................... 19
1. Executive Summary

Since the 802.3z standard was approved by the IEEE in June of 1998, Gigabit Ethernet has done much to fulfill its promise with regard to providing an economical means of scaling Ethernet LANs and LAN backbones to keep pace with the ever growing demands placed by bandwidth intensive applications. Gigabit Ethernet’s backward compatibility with the huge installed base of 10/100 Mbps Ethernet, has ensured its place in campus networks as a means to seamlessly ease the congestion in backbone LANs, with little or none of the retraining or down time that accompanies switching to an alternative high speed technology.

However, despite the increased capacity delivered by Gigabit Ethernet, the focus of today’s network planners is on network convergence: delivering static application data (such as email, Word/Excel files, images, etc.) alongside real time multimedia applications such as Voice Over IP (VoIP) telephone service, and videoconferencing - all carried across the same network.

By their nature, real time applications such as VoIP and videoconferencing are very sensitive to latency, and subjecting the delivery of such data to connectionless protocols such as TCP/IP presents a significant challenge. For real time multimedia applications to function properly, some level of guaranteed Quality of Service (QoS) must be maintained.

Unlike ATM, Ethernet was not originally designed with Quality of Service (QoS) mechanisms in mind to guarantee specific bandwidth or latency rates to different types of network traffic. Rather, Ethernet dates back to very early in the evolution of LANs and the Internet, when simply negotiating common access to the shared medium LAN, and ensuring uninterrupted Internet service despite partial network outages were the primary goals.

Despite these challenges, Gigabit Ethernet still represents a viable platform on which to develop the next generation of ‘converged’ networks. The reasons for this have partly to do with other developments in network architecture, such the evolution and subsequent wide deployment of LAN switches. In addition, although QoS mechanisms are not integral in Ethernet’s design, there has been substantial development undertaken to implement QoS sophisticated schemes that dramatically improve Ethernet’s default ‘best effort’ approach to delivery. These include: RSVP protocol, Differentiated Services, and Multi-Protocol Label Switching (MPLS).

This paper briefly summarizes Gigabit Ethernet technology and VoIP as an example multimedia technology, and then goes into more detailed discussion of the nature of QoS required for converged networks, and how RSVP, Differentiated Services, and MPLS attempt to fulfill the necessary requirements.
2. **Gigabit Ethernet Technology**

Since it was first introduced, Ethernet has risen in prominence to become the most widely-spread LAN technology in use today. By some estimates, at least 85% of all network ports in the world are Ethernet. Gigabit Ethernet extends the data rate of its predecessor, 100 Mbps Ethernet (aka Fast Ethernet) by a factor of 10, while utilizing the same CSMA/CD protocol and frame format.

The initial version of Gigabit Ethernet ratified by the IEEE in June 1998 was 802.3z, which specifies 1000 Mbps rate transmission over fiber optic cable and ‘short-haul’ shielded copper cable. One year later, 802.3ab was officially adopted, which specifies transmission over Category 5 cable to achieve distances of up to 100 meters.

Gigabit Ethernet allows operation in both half-duplex mode (in hub configurations) and full duplex mode (used in LAN switch configurations).

### 2.1. IEEE 802.3z and 802.3ab Protocol Architecture

The protocol architecture used in Gigabit Ethernet is identical to 10/100 Mbps Ethernet from the LLC portion of the Data Link layer and above. However, in order to accomplish transmission over fiber optic cable, the 802.3z standard chose to adopt the medium access and encoding technologies already employed in the Fiber Channel protocol stack. Therefore, the medium access and physical layers actually implement Fiber Channel’s FC1 layer which encapsulates 8B/10B encoding (which helps balance the disparity of 1s and 0s that may be present in a transmission), and FC0 physical layer, which controls signaling over the physical medium. (See Figure 1 below).
In the case of 802.3ab, a separate encoding scheme (4D-PAM5) is substituted, with the physical layer signaling over Category-5 UTP.

2.1.1. MAC Layer

Gigabit Ethernet retains the same standard MAC frame format as its predecessors. However, a special technique is applied called Carrier Extension when Gigabit Ethernet operates in half-duplex mode and CSMA/CD protocol is required.

Carrier Extension appends special symbols to the end of any MAC frame that is less than 512 bytes in length, in order to ensure that the duration of the frame transmission time is greater than the time it takes the frame to propagate to the receiver station. This is necessary for the Collision Detection algorithm in CSMA/CD protocol to respond efficiently to collision of a transmitted frame at the receiver, and back-off a random time before re-transmitting. The only alternative to ensuring a frame transmission time greater than propagation time at 1Gbps speed would be to shorten the maximum cable length by a factor of 10 (to approx 20m), making it impractical for most uses.

In full-duplex, LAN switched configurations of Gigabit Ethernet, Carrier Extension is not necessary.
2.1.2. Gigabit Interface Converter (GIC)

The Gigabit Interface Converter operates just below the MAC layer, and allows for independent configuration of Gigabit Ethernet ports for short-wave laser (SX), long-wave laser (LX), or copper (CX) physical interfaces (Cisco Systems, Inc., 1989). This allows for additional flexibility in LAN switch configurations and network topologies.

2.1.3. Physical Medium

Between 802.3z and 802.3ab, there are four physical medium specifications for Gigabit Ethernet:

- **1000BaseSX**: short-wave laser (850 nm) over multimode fiber (62.5 or 50 µm), with transmission range up to 500 meters.
- **1000BaseLX**: long-wave laser (1300 nm) over multimode fiber (62.5 or 50 µm) or single mode fiber (10 µm), with transmission range up to 3 kilometers.
- **1000BaseCX**: short haul copper – shielded twisted pair with DB-9 connector. Transmission range limited to 25 meters.
- **1000Base-T**: 4 pairs of unshielded twisted pair. Transmission range up to 100 meters.

3. VoIP

Voice over Internet Protocol (VoIP) is prime example of a real-time media application for which the technology has recently evolved to the point where it is being more seriously considered for wide scale deployment by businesses and service providers that already have high speed network infrastructure (such as Gigabit Ethernet) in place.

It is evident to some organizations that VoIP represents an opportunity to gain efficiency by converging their voice networks onto their IP/Ethernet data network, thus eliminating the overhead of maintenance and support services for a separate circuit switched PBX network (Brunner, Ali, 2004).

VoIP telephony also offers attractive features not always found in traditional circuit switched phone networks, such as: encryption, ability to bypass long distance tolls, and computer integration.
3.1. **VoIP Components**

**Call Processing Server / IP PBX**
The call processing servers are responsible for controlling all VoIP connections for client end-devices.

**User End-Devices**
In most cases, an attached end-device is a desktop VoIP phone, though VoIP phones may also be software based in a computer integrated application.

**Media/VoIP Gateways**
Media gateways are responsible for analog-to-digital conversion of voice, creating IP packets, and interfacing with public switched telephone networks.

**IP Network**
For purposes of VoIP, the IP network replaces the circuit switching network of a PBX. VoIP packets are routed to destination devices (phones) which are assigned IP addresses, typically configured by a DHCP server.

3.2. **VoIP Network Requirements**

Aside from network bandwidth considerations (which varies substantially based on the theoretical number of calls that may be placed at the same time), the single most important requirement for VoIP is the existence of a QoS mechanism that can ensure VoIP packets receive priority over network resources such that packet loss, latency, and jitter are kept to minimum thresholds.

In order to maintain toll quality voice service, latency (end-to-end transmission delay of IP packets carrying voice data) should be less than 150 milliseconds (Brunner, Ali, 2004).

4. **Ethernet Quality of Service (QoS)**

Quality of Service (QoS) in the context of data networks (whether Ethernet, ATM, Frame Relay, etc.) refers to the ability to categorize different types of network traffic for the purpose of applying policy based prioritization of some traffic types over others. By the same token, QoS can also be seen as a way of allocating the limited physical resources of a network among competing types of traffic generated by different applications.

Although QoS was not engineered in to Ethernet at it’s inception, several efforts have been undertaken to develop QoS mechanisms that will provide Ethernet with a means of classifying and prioritizing different types of traffic.
4.1. **Integrated Services**

The Integrated Services working group was established by the IETF in the mid 1990s, and represents the first attempt to establish QoS capabilities that could work across the Internet. The Integrated Services model is centered around reserving network resources based on individual application traffic flows. For example, a single application running on a user’s desktop that wants to initiate a data transmission across the Internet must first request to reserve network resources along the entire path to the receiving station.

Two levels of service were offered under the Integrated Services model:

- **Guaranteed Service**: A packet forwarding service with a defined delay bound.
- **Controlled Load Service**: Standard ‘best effort’ as provided by TCP/IP protocol.

The main components of Integrated Services include:

- **Route selection**. Ideally, for a given traffic flow, the network would have the capability to examine each node along a path between the sender and receiver stations to determine if specific network resource (such as a router) meets the requested parameters for bandwidth, delay, etc. However, this proved to be an extremely difficult problem to solve for Integrated services, so route selection simply relies on the on the routing algorithm already implemented in router (such as OSPF).

- **Resource Reservation**. Integrated Services utilizes RSVP protocol, which was specifically designed by an IETF working group to implement resource reservation.

- **Admission Control**. In order for resource reservation to function properly, there needs to be a mechanism in place to continuously monitor traffic of the network, and block access of RSVP requests if sufficient network resources are not available to meet the request. An admission control agent monitors network traffic load based on measurement techniques such Exponential Averaging or Time Window, and either accepts or rejects reservation traffic after comparing with the resource parameters requested (x amount of bandwidth or delay).

- **Flow Identification**. Once an RSVP traffic flow is initiated and accepted onto the network, a router must be able to identify whether packets belong to an RSVP flow. Packet identification is accomplished by examining five fields in the IP packet header referred to as the ‘five-tuple’: Source ID Address, Destination ID Address, Protocol ID, Source Port, and Destination Port.
• **Packet Scheduling.** For packets identified on the network as RSVP, the final step in the mechanism is packet scheduling. At the router, scheduling is accomplished by having separate forwarding queues that map to the delay bounds and bandwidth requested by RSVP. Routers implement a number of different algorithms such as Weighted Fair Queuing to determine scheduling priority among the queues.

Ultimately the Integrated Services model met with only limited success due to a couple of significant pitfalls:

The first problem was that Integrated Services did not anticipate the explosive growth of the World Wide Web and the accompanying very large volume of short duration Web applications that generate transient data flows in the process of browsing. On a large LAN, the number of individual application traffic flows can easily number in the millions.

The second problem is that RSVP, the protocol established to accomplish resource reservation, requires a significantly large processing overhead, and would not scale to the size needed to handle all the potential traffic.

### 4.1.1. RSVP Protocol

Although individual traffic flows originate with a sending station, RSVP is a receiver oriented protocol in the respect that it is the receiving station that specifies the resource requirements (delay bound, bandwidth) to be reserved, and sends the instruction to reserve them.

The reservation state maintained in network resources (such as routers) for a particular traffic flow is referred to as a ‘soft state’ reservation. The reservation for a specific amount of resource is made based on parameters forwarded by RSVP, and a timer is set to hold the reservation for a default period. If no packets from the transmitting station are received (or cease), the timer expires and the reservation state is terminated.

Also, RSVP is designed such that it only reserves network resources in a single direction between the sending and receiving station. For applications that may transmit data in both directions between two stations, RSVP must explicitly reserve network resources in both directions.

**Operation**

RSVP is initiated with a PATH message sent from the source station to the destination station. The PATH message contains attributes of the reservation which is read by the routers along the transmission path. Each router along the path determines whether it has sufficient bandwidth to meet the request, and if so, the reservation is copied in a temporary state in the router.
If every router visited by the PATH message in transit to the destination agrees to a temporary reservation state, then the destination station returns a RSVP message back to the source. As the RSVP message transits all of the same routers on it’s ‘return’ path (specified from the PATH message), each of the router will update the temporary reservation to its reservation state table (Hartmann, 2004).

Once the RSVP message successfully reaches the sender (implying all reservations are set), the sending station may begin data.

4.2. Differentiated Services

Differentiated Services was developed in response to pressure from service providers who recognized that RSVP technology was not ready to scale to the level of providing efficient QoS across the WAN and Internet. A simpler, more scalable QoS architecture needed to be developed, and the IETF formed the Diffserv workgroup in 1997 to work on the standards for Differentiated Services.

In contrast to Integrated Services, the Differentiated Services model is oriented toward prioritizing and guiding aggregate traffic flows, rather than individual application flows. Differentiated Services uses a combination of edge policing, provisioning, and traffic prioritization to provide different levels of service to customers (Wang, 2001).
4.2.1. Differentiated Services Components/Operation

DSCP
The Differentiated Services (DS) architecture separates user network traffic into unique Forwarding Classes for prioritization. Encoding of forwarding class occurs at the IP packet level, and utilizes the existing 8 bit Type of Service (ToS) field in the IP packet header. The first 6 bits of the ToS field are overlaid with a code referred to as the Differentiated Services Code Point (DSCP). The 6 bit DSCP field allows for up to 64 classes of service to be defined, and also preserves backward compatibility with the 3 IP Precedence bits in the ToS field. The two remaining bits of the ToS field are unused (though reserved for explicit congestion notification).

![Diagram of DSCP onto IP ToS Header](Hartmann, 2004)

Traffic Classification/Admission Control
Admission control is performed at the edge of a DS network domain by means of traffic policing. The boundary nodes of the DS domain classify all incoming packets either according to Traffic Conditioning Agreements (TCA) with boundary domains, or by examining the same five-tuple fields of the IP header used to identify flows in Integrated
Gigabit Ethernet, QoS, and Multimedia Applications

Services. In addition, the aggregate traffic entering a DS domain must conform to a policy driven traffic profile. Non-conforming traffic is subject to further policing by a Traffic Conditioner.

Traffic Conditioner
A traffic conditioner has four components: Meter, Marker, Shaper, and Dropper. The Meter is the element that checks the incoming traffic flow to see if it matches the traffic profile for the DS domain. If it does, the traffic is admitted. Otherwise the incoming packets are subject to further conditioning. The actions that may be taken are the responsibility of the remaining three components. Depending on the policy, a Marker may simply mark (or re-mark) the incoming packet with an appropriate DSCP value. A Shaper delays incoming packets until the aggregate incoming flow complies with the traffic profile, and a Dropper simply drops packets that don’t meet the traffic profile.

Traffic Forwarding
The task of forwarding DSCP encoded packets is the responsibility of the interior nodes in a DS domain. At the individual packet level, packets are described as having a Per Hop Behavior (PHB) based directly on their DSCP value. Collectively, all packets that have the same DSCP value that are arriving at a particular node together are referred to as a Behavior Aggregate (BA).

Assured Forwarding Classes
All of the basic Differentiated Service components such as DSCPs, PHBs, and BAs describe how traffic is to be classified and grouped for prioritization, but none of them describe a particular prioritization scheme. Prioritization is finally defined in the concept of Assured Forwarding Classes.

The IETF has standardized two groupings of PHBs into Assured Forwarding Classes, called Assured Forwarding, and Expedited Forwarding.

- **Assured Forwarding** defines four traffic classes: AF1, AF2, AF3, and AF4. This represents a priority scheme in which AF1 is the lowest priority and AF4 the highest priority, with AF2 & AF3 as increments in between.

  Within each AF class, there are also three drop preferences; low(1), medium(2), and high(3), that define the order that packets are to be dropped from the forwarding node if congestion occurs. (AF1, high drop (3) packets are dropped first, and AF4, low drop (1) packets are dropped last).

  The AF* traffic classes defined by Assured Forwarding map back to DSCP values in the IP header (see figure 4 below).
Expedited Forwarding defines a low-loss, low latency forwarding class intended for delay sensitive real time traffic, such as voice and video. All packets classified with Expedited Forwarded use the reserved binary DSCP value 1011101. The mechanism used to ensure Expedited Forwarding class traffic has precedence over all other traffic classes is software priority queuing in the forwarding nodes.

4.2.2. Differentiated Services Topology

![Differentiated Services Topology Diagram](image)

Figure 5: Example, Differentiated Services Topology (Hartmann, 2004)
4.3. Multi-Protocol Label Switching (MPLS)

Multi-Protocol Label Switching (MPLS) is a packet forwarding technology that allows Ethernet based networks to behave more like a connection oriented service. In fact, one of the primary motivators in developing MPLS technology was to achieve a more efficient, scalable way to send IP packets over ATM networks than the existing LAN Emulation solution.

Arguably, MPLS itself does not represent a QoS technology, but its design has important implications for QoS by way of offering greater routing efficiencies and distribution of network traffic flows in the MAN and WAN backbone space.

IP packet forwarding, especially with QoS mechanisms such as Differentiated Services in place, means that routers must perform a complex set of tasks. With the ever growing quantity of traffic on the Internet, and the increase in the use of multimedia applications such as VoIP that are delay sensitive, routers on the backbone are taxed to keep up with the demand to forward traffic.

Label switching technology such as MPLS offers a way to help reduce the complexity of IP packet routing, though it comes at the cost of having to establish an end-to-end path before data can be forwarded (similar to RSVP, though MPLS handles aggregate traffic flows). Perhaps even more important is that MPLS offers a viable platform to apply traffic engineering on Ethernet based backbones, to more evenly distribute traffic for better resource utilization.

4.3.1. Implementation

MPLS employs forwarding of IP packets based on a short (4 byte) fixed length label in the packet header. The MPLS is embedded in between the Data Link Layer (layer 2) header, and the Network Layer (layer 3) header.

| Layer 2 Header | MPLS Label | Layer 3 Header |

Figure 6: Location of MPLS Header.

Similar in concept to the Integrated and Differentiated QoS models, MPLS operates in a domain of MPLS enabled devices. Packets entering an MPLS domain have labels inserted in front of the network layer header by a Label Switched Router (LSR). The label is actually referred to as a label ‘stack’, because depending on the MPLS routing mechanism used, the label may be appended with route information for the packet’s entire transit of the MPLS domain.

The key to MPLS packet forwarding is establishment of a Label Switched Path (LSP) through the network. When a packet enters an MPLS domain, the edge LSR processing the packet makes a route selection based on either standard IP hop-by-hop routing
methods (such as open shortest path), or by specifying the explicit route of the entire LSP before the packet is forwarded. In the case of explicit routing, the labels for each forwarding node along the path are stacked in the MPLS label.

In the case of hop-by-hop routing over MPLS, Label Distribution Protocol (LDP) is used by Label Switched Routers to maintain the label switching tables used to determine the next hop to be used for the packet LSP.

4.3.2. Traffic Engineering

One of the main strengths of MPLS is that it supports explicit routing of packets along a Label Switched Path. The option of explicit routing allows for traffic engineering of aggregate IP traffic flows. Traffic engineering refers to the process of balancing traffic across network resources to avoid congestion and achieve better network utilization.

When hop-by-hop routing by methods such as OSPF (open shortest path first) lead to congestion in a particular area of the network, explicit routing based on label switching technology offers a way to efficiently rout traffic away from congested areas. MPLS provides for additional protocols (CR-LDP and RSVP-TE) that build upon the advantages of explicit route traffic engineering.

5. Vendors

5.1. Cisco Systems

Cisco Systems, Inc. offers many network switch and routing products that support Gigabit Ethernet with QoS enabling technology. An example of Cisco’s product offerings is the 7300 Series edge routers which are MPLS enabled, and have built-in Gigabit Ethernet ports.
Foundry Networks also offers several Gigabit Ethernet capable routers for deployment in various contexts. An example of Foundry’s Internet/MAN routers with Gigabit Ethernet technology is the NetIron IMR 640:
6. Market Analysis

The success of Gigabit Ethernet in the market place is evident today. Businesses have been following deployment strategies suggested by equipment vendors to update their Backbone LANs with Gigabit capable switches and routers, and have either retrofitted critical servers with Gigabit network interface cards, or have opted to replace servers with Gigabit Ethernet ports built in. At this time, several PC vendors including Dell and IBM have started packaging Gigabit ports in all but their bottom of the line business desktops, either by including controllers directly on the motherboard, or with 10/100/1000 network interface cards.

VoIP has become a sought after technology, not only by businesses, but also by a growing number of residential broadband subscribers seeking to save money on long distance telephone charges as well.

7. Conclusion

Gigabit Ethernet provides network administrators with a well proven, reliable path to upgrade LANs to a high speed technology capable of supporting converged voice and data transmission on the same network. The backward compatibility with 10/100 Mbps Ethernet eliminates the need for retraining support personnel and minimizes the disruption of migrating to Gigabit speed components.

Although popular multimedia applications such as VoIP present significant Quality of Service challenges to implementation over Ethernet, advances in QoS mechanisms such as Differentiated Services, and technologies like MPLS that extend IP network capabilities, have helped close the gap to the point that VoIP is poised to grow dramatically in the coming years ahead.

8. Glossary

802.3ab: IEEE working group that defined the standard for transmission of Gigabit Ethernet over Category-5 copper cabling.

802.3z: IEEE working group that defined the standard for transmission of Gigabit Ethernet over fiber optic cable and shielded copper cable adopted from Fibre Channel medium interface standards.

ATM: Asynchronous Transfer Mode. Connection oriented network protocol widely deployed in network backbones that utilizes virtual circuit connections for reliable data transmission.
**CSMA/CD**: Carrier Sense Multiple Access/Collision Detection. Medium access protocol used by Ethernet that ‘listens’ to attached network cable for absence of transmissions by other stations before attempting transmit data over the network.

**Differentiated Services**: Policy based quality of service model developed by the IETF that prioritizes network traffic based on aggregate class of service.

**DSCP**: Differentiated Services Code Point. The encoding scheme employed by Differentiated Services which uses 6 bits of the IP Type of Service header to assign priority to IP traffic.

**Fast Ethernet**: Ethernet protocol designed to operate at 100 Mbps data rate.

**Gigabit Ethernet**: Ethernet protocol designed to operate at 100 Mbps data rate.

**IEEE**: Institute of Electrical and Electronics Engineers.

**IETF**: Internet Engineering Task Force. A collection of working groups developing standards for the Internet.

**Integrated Services**: Quality of Service model that seeks to explicitly reserve network resources as a method of prioritizing network traffic.

**MPLS**: Multi-Protocol Label Switching. A packet forwarding technology initially developed for integrating IP and ATM network traffic.

**OSPF**: Network routing protocol that forwards packets based on Open Shortest Path First algorithm.

**PBX**: Public Branch Exchange. A circuit switched telephone network widely installed in corporate/campus settings.

**QoS**: Quality of Service. Refers to the ability to differentiate network traffic.

**RSVP**: Resource Reservation Protocol. Protocol used reserve network resources within Integrated Services QoS model.

**VOIP**: Voice over Internet Protocol. A technology for transmitting encoded (digitized) voice traffic over data networks, as an alternative means to circuit switched networks for providing telephone service.
9. References


