



QoS

Supporting Technologies
September 2006



INTRODUCTION:

As introduced in IP Infusion's Metro Ethernet white paper, Quality of Service is one of the key characteristics in the Next Generation Network (NGN) to support triple play services. We all know service providers are always in search for ways to generate new revenues, entrepreneurs are always in look for new business models, consumers are always interested in new services to ease their day to day activities, and new generation always seeks for new entertainment for fun. Because of all these, real time applications like streaming video, Voice over IP (VoIP), IPTV, e-commerce, on-line gaming, video-on-demand, chat room, instant messaging... have taken on a higher demand in recent years than expected. With the success of the Internet and its ubiquitous reach, it has become an obvious choice for service providers to use the IP network to provide new services. As a result, the convergence of data, voice and video started in the 90s has since then been the key evolution in the technology world. The merge of the data network and the voice network has post new challenges and requirements. New technologies are being defined to meet these challenges and requirement.

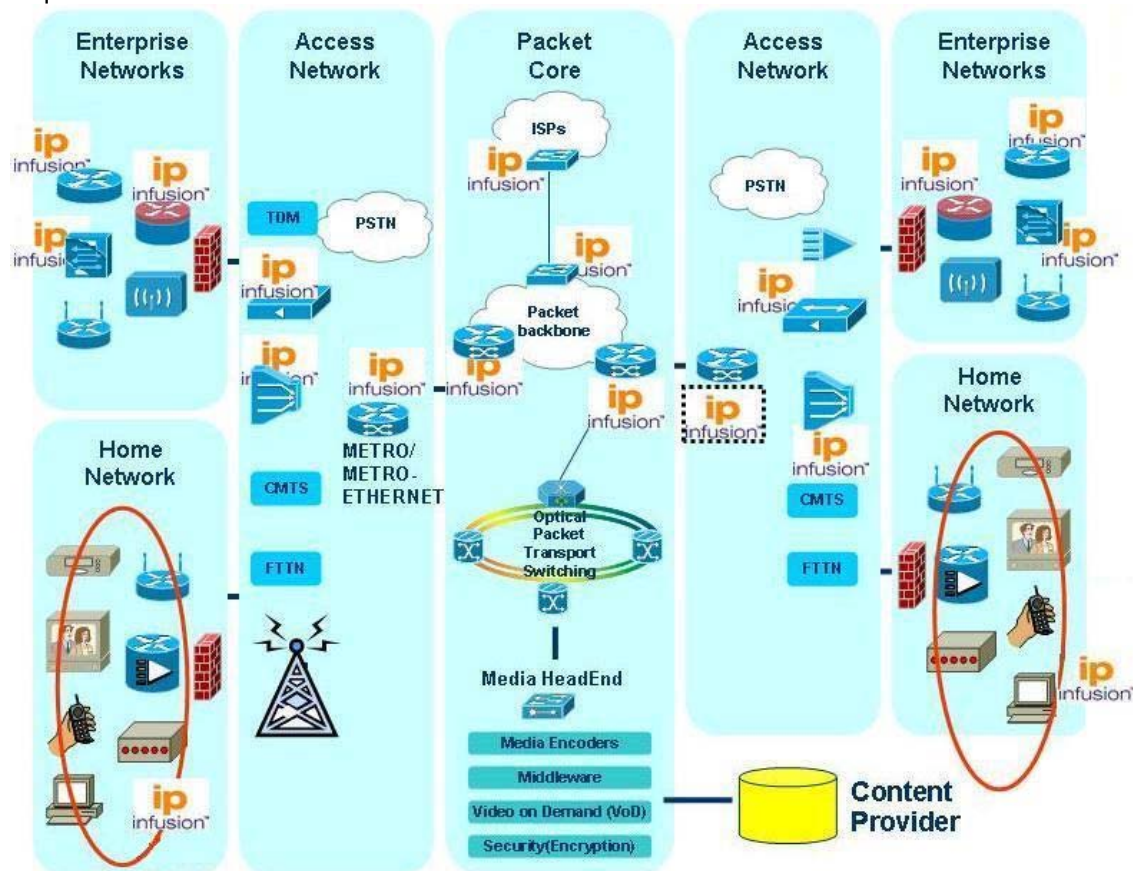


Figure 1 - QoS Everywhere

In order for service providers to bill their customers accordingly, they need to be able to deliver corresponding level of services based on users' subscription (premium, gold, silver...) and the nature of the services (data, voice, video...). Networks then need to be able to differentiate traffic based on corresponding criteria. This new requirement in the network makes Quality of Service (QoS) an essential element in the Next Generation Network (NGN). In this paper we are going to examine different technologies used for supporting QoS. When the Internet Protocol (IP) was first defined back in the late 70's, it was target to carry data and thus QoS was not being considered at the time. Packets were delivered on best effort and packet delay and loss were more tolerable. Today, as new services (like on-line gaming, video-on-demand ...) roll out, the nature of traffic in the IP network has changed. Network needs to consider how to deliver timing and loss sensitive traffic, like streaming multimedia and Voice over IP (VoIP), in a high quality and non-disruptive fashion. Voice and video traffic need to be distinguished from data traffic and be treated accordingly to guarantee the level of services it requires.

The fundamental nature of QoS is the ability to distinguish a traffic type and provide it with resources it deserves. Resources can be reserved in a static way but more often need to be allocated dynamically to maximize the usage of resources and to accommodate the ever changing traffic flows. In the following sections, we will look at how IEEE 802.1p/Q, Integrated Services (IntServ), Differentiated Services (DiffServ) and Multi Protocol Label Switching (MPLS) address QoS.

IEEE 802.1P/Q IN SUPPORT OF QOS

IEEE standards 802.1p and 802.1Q are protocols defined to support QoS at the Media Access Control (MAC) level. 802.1p enables traffic prioritization and dynamic multicast filtering. 802.1Q specifies the VLAN tag, which is composed of a 12-bit VLAN ID and a 3-bit Prioritization field. Within the 3 prioritization bits, 8 levels of priority can be defined. Level seven, being the highest level, is usually used for network-critical traffic such as routing update information. Levels six and five are usually used for delay-sensitive traffic, such as interactive video and voice. Levels four to one are usually used for different types of traffic from streaming multimedia to critical-data traffic. Zero is for best effort data. Based on the value of the p-bit, switches can then decide the treatment of the packets. For example, packets with level seven will be placed in highest priority queue.

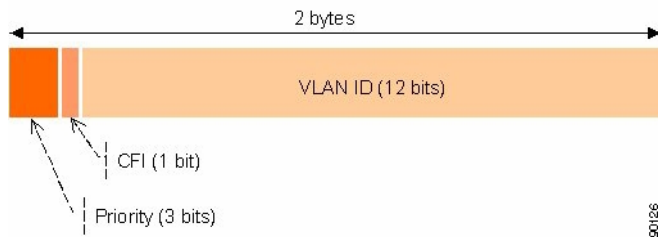


Figure 2 - 802.1Q header

For service providers who offer Layer 2 services, they can use the IEEE 802.1p/Q standard to provide QoS. VLAN IDs will be prepended onto all incoming frames from customer circuits. By doing this, the service provider can support multiple customers using the same circuit, but still maintain a separation between them. Each customer's traffic is identified by a different

VLAN tag. And with the priority field, service providers can offer different classes of service to their customers.

[Note: IEEE 802.1ad defines stacking of VLAN IDs (Q-in-Q), which enables a hierarchical VLAN structure. This structure allows service providers to use the outer VLAN tag to segregate between services, and subscribers to use the inner VLAN tag for internal usage within their own organization. More detail of this standard will be discussed in later section.]

SYSTEM OVERVIEW

Before discussing the following three technologies, let's have a quick review of a typical system architecture. At a high level in any networking system three functional planes can be defined, namely the control plane, the data plane and the management plane. The control plane is where the signaling and control protocols reside; it has the intelligence to make decisions on how the traffic should be forwarded. This forwarding information is then passed to the data plane, which is where hardware implements the logic for fast data forwarding. The management plane is where the intelligence for managing and monitoring the network and the system resides.

INTSERV IN SUPPORT OF QOS

In realizing the lack of QoS support in traditional IP network, IntServ was defined with the goal of enabling end-to-end QoS. The IntServ architecture assumes some explicit setup mechanism is used to convey information to the network elements such that the network can provide requested services to the "flows" that require them. A flow is a unidirectional data stream between two applications represented by the 5-tuple: source IP address, source port number, destination IP address, destination port number and the transport protocol. IntServ reserves resources on a per-flow basis. In order to support IntServ, signaling protocols are used in the control plane for reserving resources required by a particular flow along the data stream path.

ReSource reservation Protocol (RSVP) is the signaling protocol most commonly used to support IntServ. Two types of service can be requested via RSVP, namely the Guaranteed Services (GS) and Controlled Load (CL). GS targets real-time applications with strict bandwidth and latency requirements. CL targets traditional applications with performance of a lightly loaded best-effort IP network.

RSVP is used by a Host application to request the network resources it needs for a given flow. Two message types are defined by RSVP, the PATH message and the RESV message. PATH messages are sent periodically from the sender. A PATH message contains the Sender Template (source address, source port and data format), the Sender Tspec (description of the characteristic of the traffic) and the previous hop address (to be used for by the RESV message). The PATH message propagates along the route as determined by some underlying routing mechanism to the receiver.

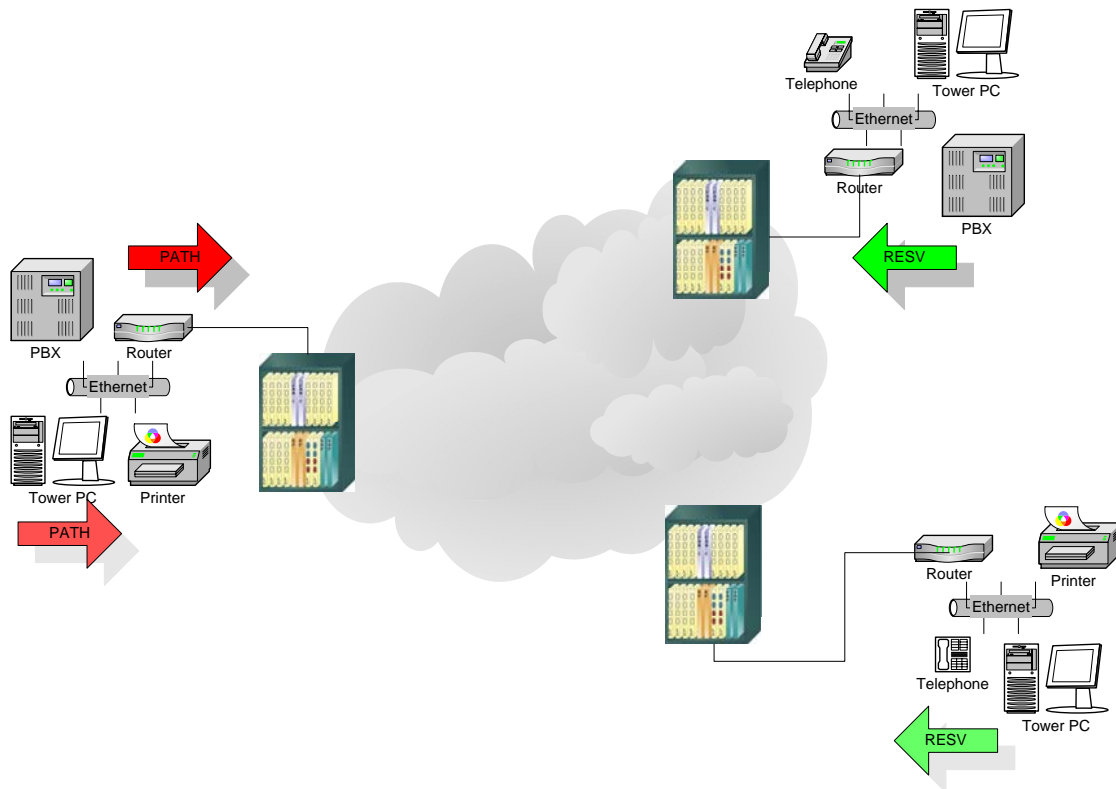


Figure 3 - RSVP

Upon receiving the PATH message, the receiver responds with the RESV (resource reservation) message. A RESV message contains the “flowspec”, which defines the desired QoS, and the “filterspec”, which defines the flow (the set data packets) to receive the QoS defined by the flowspec. The RESV message must travel the exact “reverse” route as the PATH message, by using the previous hop information in the PATH message. The RSVP process must pass both the admission control and the policy control on each and every node on the path for a RSVP session to be successfully established. Each router on this path must create and maintain the “reservation state”. These reservation states are soft states, refresh messages need to be sent periodically to maintain the reserved path. Since RSVP requires all nodes to maintain reservation states for each flow, it is not scalable for large network. One implementation to overcome this disadvantage is to define a hierarchical RSVP network as defined in RFC3175, where the core network resources are reserved for aggregated flows when the edge network resources are reserved on a per-flow basis, thus reducing the network states to be maintained.

DIFFSERV IN SUPPORT OF QOS

DiffServ was designed with the goal to address the scalability issue IntServ has and target to support large networks. Instead of dealing with micro-flows like IntServ, DiffServ deals with aggregated data flows, called classes. DiffServ uses the TOS field in the IP header and

renames it Differentiated Service field (DS). A DiffServ Code Point (DSCP) is a 6-bit long field embedded in the DS field and is used for classifying packets. DiffServ, therefore, can support up to 64 different classes of service. Packets (even from different applications) which have the same codepoint value and travel to the same direction belong to the same Behavior Aggregate (BA). At any given node, the same treatment (for example: scheduling, queuing, policing, shaping) will be used for those packets of the same BA. This characteristic of the DiffServ network is called Per Hop Behavior (PHB). Three kinds of PHB are defined: Expedited Forwarding PHB (has one class defined), Class Selector PHB (has eight classes defined) and Assured Forwarding PHB (has four classes with priority high, medium and low in each class defined).

Expedited Forwarding PHB is for guaranteed service applications like VoIP and video streaming, which require low-latency, low-loss and low-jitter. It is similar to RSVP in the sense that resource reservation is made prior to the packets being sent.

The Class Selector PHB has a format of xxx000, where xxx is of different value depends on the service required. Routers already supporting QoS and using the 3-bit precedence field in the TOS field can use the same classification and forwarding scheme to map to the class selector codepoint. This makes the DiffServ-compliant nodes be backward compatible with the IP-precedence aware nodes.

The Assured Forwarding PHB has four classes with three priority level defined in each class. With this structure, the service provider can establish different SLA levels, billing schemes and penalty policies for exceeding allowable limits.

Following is the table of the recommended codepoint values of the different PHBs. An application can also define its own PHB.

<u>Expedited forwarding</u>	<u>Class selector</u>	<u>Assured forwarding</u>
101110	Class 0 - 000000	Assured forwarding, Class 1, Low - 001010
	Class 1 - 001000	Assured forwarding, Class 1, Medium - 001100
	Class 2 - 010000	Assured forwarding, Class 1, High- 001110
	Class 3 - 011000	Assured forwarding, Class 2, Low - 010010
	Class 4 - 100000	Assured forwarding, Class 2, Medium - 010100
	Class 5 - 101000	Assured forwarding, Class 2, High - 010110
	Class 6 - 110000	Assured forwarding, Class 3, Low - 011010
	Class 7 - 111000	Assured forwarding, Class 3, Medium - 011100
		Assured forwarding, Class 3, High - 011110
		Assured forwarding, Class 4, Low - 100010
		Assured forwarding, Class 4, Medium - 100100
		Assured forwarding, Class 4, High - 100110

Table 1 - PHB Code Points

One of the drawbacks to DiffServ is provisioning; setting up the different classes throughout the network is time-consuming and requires network experts. Tools like Remote Monitoring (RMON) can be used to ease the provisioning operation.

MPLS IN SUPPORT OF QOS

MPLS was originally designed to use labels for fast forwarding. This technology has further evolved to support QoS in the IP network and is now the most desired mechanism for a service provider in offering various service level agreements (SLAs) to the subscriber. Two types of routers are defined in an MPLS network, namely the Label Edge Router (LER) and the Label Switch Router (LSR). LERs reside on the edge of the MPLS network while LSRs reside in the core of the network. When a packet enters an MPLS network, a label will be assigned to it based on the packet's Forwarding Equivalency Class (FEC). A FEC could be the source IP address, the destination IP address or any information that matches the policies used to place the packet in a particular service class. Labels are of local significance and will be swapped at each hop along the way. This label is inserted in the layer 2 header for fast forwarding. As the packet traverses the network, the label will be used as an index into a policy table to determine how this packet should be treated. Upon exiting the MPLS network, the label will be stripped off and the packet is forwarded to the destination address.

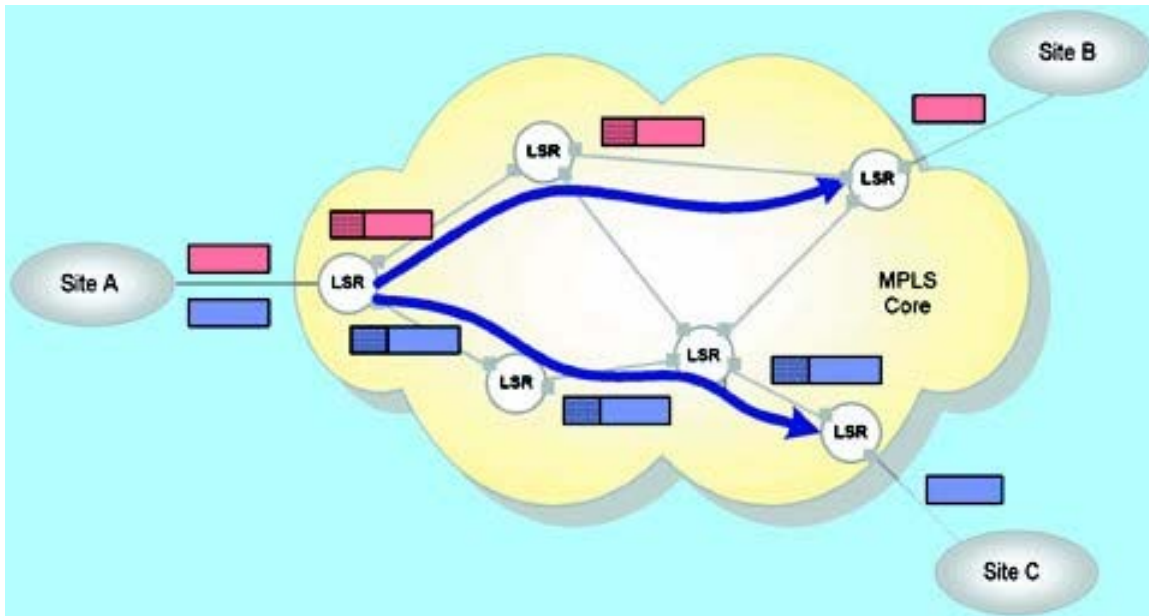


Figure 4 - MPLS network

With this methodology, the forwarding intelligence is only required at the edge (LERs) to bind the FEC to the label. LSRs only need to perform label look up and forward the packets accordingly, and in conjunction with high speed hardware could tremendously improve the network performance.

The paths the MPLS packets traverse are called Label Switching Path (LSP). MPLS uses an underlying routing mechanism (static route or dynamic routes set up by routing protocols) to: identify the individual hops in the LSPs and; uses signaling protocols like Label Distribution Protocol (LDP) or ReSource reserVation Protocol-Traffic Engineering (RSVP-TE) to distribute the labels along the path.

LDP was defined specifically for MPLS while RSVP-TE enhanced the existing RSVP standard to meet the MPLS requirement. In both cases, the MPLS signaling protocol works with the underlying routing mechanism (for route information) to establish an LSP based on a FEC. Two mechanisms for distributing the labels are defined: downstream on demand and unsolicited downstream.

In the first case, a LSR will explicitly ask its down stream next hop for a label to bind to a certain FEC. In the second case, LSR distributes bindings to LSRs that have not explicitly asked for it. Once the labels are distributed, the data plane will have enough information to forward packets of the same FEC with the same treatment. Typically, an LER will have the following information in its forwarding table: FEC, outgoing Label, and outgoing interface. An LSR will have the following information in its forwarding table: incoming label, incoming interface, outgoing label, and outgoing interface.

Reliability and scalability are keys in providing QoS. MPLS uses Graceful Restart (GR) and Fast ReRoute (FRR) to address reliability and uses label stacking to address scalability. GR defines the mechanism to preserve labels and forwarding information, with the end goal resuming traffic flow in the shortest possible time when a link or node fails. With label stacking, multiple labels can be inserted in the packet header. The idea of label stacking is to enable a hierarchical MPLS network structure; this is extremely useful when one wants to overlay one network on the other or tunnel one type of traffic through another. [For example, label stacking is used in support of L3-VPN where the inner label carries the BGP information and the outer label carries the MPLS information.]

Though MPLS itself is well defined for supporting QoS, DiffServ is widely used in supporting QoS, so there is a need to support DiffServ over MPLS. MPLS uses its 3-bit Experimental field (EXP field) to map to the DSCP value for support up to eight classes in a network. For networks requiring up to 64 classes (the max that DiffServ can support), MPLS use the label itself to map the DSCP value.

COMPLEMENTARY STANDARDS IN SUPPORT OF QOS

IEEE 802.1ad, Provider Bridges, is defined by IEEE to define an architecture and associated bridge protocols used to provide separate instances of MAC services to multiple independent users of a Bridged Local Area Network. This is done in a manner that does not require cooperation among the users, and requires a minimum of cooperation between the users and the provider of the MAC services. With 802.1ad, each Ethernet frame can be encapsulated with a Service VLAN tag (outer tag) and a client VLAN tag (inner tag). With this ability to stack the VLAN ID, it is possible to preserve the customer VLAN structure and therefore enable customer separation and traffic differentiation through the service provider network. The VLAN tag can accommodate only up to 4096 VLANs and the scaling issue is being addressed by 802.1ah.

802.1ah, Provider Backbone Bridges, is also defined by IEEE to specify means for interconnecting Provider Bridged Networks. It defines an architecture and bridge protocols compatible and interoperable with Provider Bridged Network protocols and equipment allowing interconnection of multiple Provider Bridged Networks, to allow scaling to at least 2^{20} Service VLANs. 802.1ah encapsulates an end-user's Ethernet frame inside a service provider MAC header and hence this standard is also referred to as MAC-in-MAC. The work for 802.1ah is an ongoing effort at this time. But Service Providers are very excited about the scalability that will be made possible by 802.1ah.

Remote Monitoring (RMON), mentioned in previous section, was defined by the Internet Engineering Task Force (IETF). RMON was originally standardized by RFC 1271 in November 1991, updated by RFC 1757 in February 1995, and again updated by RFC2819 in May 2000. A set of MIBs is defined in the RFC to enable RMON-complined network management device to remotely monitor and manage network devices. It also enables these management devices to communicate with each other to exchange network information.

IP INFUSION IN SUPPORT OF QOS

Equipment manufacturers always face the decision of which standards to adapt. Market requirement, network characters, capital expenditure, existing infrastructure, traffic condition are some of the key factors that drive the decision. More than often one or more standards are required to support a certain feature. QoS is a good example of such scenario. IP Infusion is committed to stay on top of the latest standard and develops QoS solutions to meet all possible customers requirement. In our product portfolio, we have support of all the technologies/standards mentioned in this paper, including IEEE 802.1p/Q, Integrated Services (IntServ), Differentiated Services (DiffServ), Multi-Protocol Label Switching (MPLS), 802.1ad, 802.1ah and RMON among the others. In addition, we provide solution target applications like VPN, VPLS, Traffic Engineering... We also provide integrated solutions, which integrate our software onto QoS capable hardware to shorten customer's time to market.

Another key factor to equipment manufacturers is an open architecture which allows them to easily upgrade and add features. IP Infusion defines the ZebOS architecture to enable additional features to be easily added on. This architecture also provides the framework that allows IP Infusion software to run on any operating systems and any hardware platform. Please visit our website, www.ipinfusion.com, for an architectural review, product information and other related white papers.



© Copyright 2006 IP Infusion Inc. All Rights Reserved.

ZebOS and IP Infusion are registered trademarks and the ipinfusion logo is a trademark of IP Infusion Inc. All other brands or product names are trademarks or registered trademarks of their respective holders. All specifications within this document are subject to change without notice. Contact Sales for current feature availability.