Using Programmable Interfaces for Networks as the Key to Provide Quality of Service in the Internet

Jairo A. Gutierrez
University of Auckland, New Zealand

j.gutierrez@auckland.ac.nz

Terence Wee
National Computer Systems Pte Ltd, Singapore

ylwee@mbox.ncs.com.sg

Abstract

The growth of the Internet with its increasing levels of traffic, more ambitious applications and the convergence of communication technologies results in poor quality of service. New technologies such as those discussed in this paper have been developed to aid in introducing quality of service to the Internet. However, those efforts may seem to be in vain due to the heterogeneous nature of IP-based networks. This paper suggests an integrated approach to provide end-to-end quality of service on the Internet based on the use of programmable interfaces. The need for dynamic modification of policies and configurations within IP routers and switches is already beginning to emerge and that functionality will help fulfill customers' demands for differentiated services.

Keywords: Programmable networks, Quality of Service, Programming Interfaces, Multiple Protocol Label Switching (MPLS), Differentiated Services, Asynchronous Transfer Mode (ATM)

Introduction

The objective of enabling the development of higher-level multimedia services with guaranteed quality of service (QoS) on networks has prompted developments that attempt to accommodate these new application requirements. Several architectures have been proposed, and a common basic functionality is emerging. Any new architecture that intends to satisfy the ever-growing need for bandwidth in the Internet while providing support for QoS guarantees, needs to concern itself with the following aspects (Zhang L., Deering S., Estrin D., Shenker S., Zappala D., 1993; Biswas J. et al, 1998):

- Flow management: identifying the traffic characteristics of a flow so that the network can specify the quality of service to be delivered to that flow

- Compatibility with a wide range of routing protocols (Callon, R. et al, 1997)

- Resource reservation

- Admission control

- Packet scheduling: including packet filtering and classification

These aspects clearly call for routers with powerful features that can be easily configured and modified in order to support customers' demands for differentiated services.

Background

The Internet is a phenomenon that has grown at an exponential rate, causing several new technologies to emerge to cope with the number of users of this global communications network. TCP/IP was the protocol of choice because of its simplicity and ability to find a route from a particular source to a particular destination. However, as the network grows, there have been greater demands for additional services and real-time applications to work over the Internet. Because TCP/IP is primarily a best-effort protocol, it is not able to provide the QoS that is required by real-time applications and their users. To complement this deficiency,

other protocols such as Integrated Services (IntServ) and Resource Reservation Protocol (RSVP), Differentiated Services (DiffServ) and Multi-Protocol Label Switching (MPLS) have been developed as well as the standards promoted by the IEEE P1520 Group.

The Integrated Services (IntServ) model is based on reservations-based traffic engineering assumptions. It reserves resources explicitly using a dynamic signalling protocol (RSVP) and employs admission control, packet classification, and intelligent scheduling to achieve a desired QoS (Trillium Digital Systems Inc., 1998). The Integrated Services model has two services categories, and they are Guaranteed Delay and Controlled Load services.

RSVP is a resource reservation set-up protocol designed for an integrated service Internet. It is used by a host to request specific qualities of service from the network for particular application data streams or flows. RSVP is also used to establish and maintain state information in all nodes along a flow so as to provide the requested service. However, RSVP itself is not a routing protocol, but instead it is considered a signalling protocol similar to that used in ATM networks.

The Differentiated Services (DiffServ) model is based on reservation-less traffic engineering assumptions. It classifies packets into a small number of service types and uses priority mechanisms to provide adequate QoS to the traffic. No explicit resource reservation or admission control is employed, although network nodes do have to use intelligent queuing mechanisms to differentiate traffic (Trillium Digital Systems Inc., 1998).

MPLS stands for Multi-Protocol Label Switching, and it is a protocol that assigns a particular FEC (Forwarding Equivalence Class) to a particular packet as it enters the network. The FEC to which the packet is assigned is encoded as a short fixed length value known as a "label". Once a packet is assigned to a FEC, no further header analysis is done by subsequent MPLS capable routers and all forwarding is driven by the labels (Rosen E. C., Viswanathan A., Callon A., 1999).

In the IEEE P1520 standardization project (Working Group on IEEE P1520, 2000) it is the objective to allow routers and switches to be programmable by third parties that will have access to the scheduling mechanisms and policies of the router, thereby allowing a greater degree of flexibility and freedom for service provisioning. In line with this the IP Sub Working Group (SWG) of P1520 has written a white paper and an architectural framework document in (Working Group on IEEE P1520, 2000) that depicts the

positioning of various building blocks for programmability of the network. This paper is a continuation of the above effort, in order to realize the building blocks that support services on the Internet. The fundamental building blocks consist of a set of virtualized resources of the network elements (routers and switches) structures in a fashion that make these resources available to multiple services on the network and accessible through open and well defined interfaces for easy programming and fast and efficient access. This paper explores the use of these virtualized resources to support the functionality of Multi-Protocol Label Switching (MPLS). MPLS (Callon, R. et al, 1997) is taken merely as a case in point, to illustrate the principle of structuring the resources in a function independent fashion and then applying functions or algorithms upon these resources in a disciplined manner. This paradigm represents a shift from the traditional approach of each function, protocol or algorithm being developed in an independent fashion, thereby leading to an immense number of interactions between functions that must be considered individually.

Multi-Protocol Label Switching

MPLS, used with or without RSVP (Zhang L. et al, 1993; Baker F., Krawczyk J., Sastry A., 1998), fits within the framework of Integrated Services (Braden R., Clark D., Shenker S., 1994). Before suggesting detailed interfaces for IP routers or switches it is necessary to understand the basic operations of this proposed standard. The roots of MPLS originated from technologies such as IP switching, developed by Ipsilon Networks Inc. (Sunnyvale, California), and tag switching, developed by Cisco Systems Inc. (San Jose, California). Thus, Multi-Protocol Label Switching is the use of fixed length labels to decide packet handling (Xiao X., and Ni L.M., 1999). An MPLS router, called the label-switch router (LSR), examines only the label in forwarding the packet. MPLS also needs a protocol to distribute labels to set-up label-switched paths (LSPs). The protocol used to distribute labels is known as the label distribution protocol (LDP). Whether a generic LDP should be created or RSVP should be extended for this purpose is an issue that is yet to be decided (Xiao X., and Ni L.M., 1999). MPLS can also be piggybacked by routing protocols. A LSP is similar to an ATM virtual circuit (VC) and is unidirectional from the sender to the receiver. MPLS LSRs use the protocol to negotiate the semantics of how each packet with a particular label from its peer is to be handled. Therefore, when a packet enters an MPLS domain, it is classified and routed at the ingress LSR. MPLS headers are then inserted to the packet. When a LSR receives a labelled packet, it will use the label as the index to look up the forwarding table. This is faster than the processes of parsing the routing table in search of the

longest match done in IP routing. The packet is processed as specified by the forwarding table entry. The outgoing label then replaces the incoming label, and the packet is switched to the next LSR. Inside an MPLS domain, packet forwarding, classification and QoS service is determined by the labels and the COS (Class of Service) fields. This makes core LSRs simple. Before a packet leaves a MPLS domain, its MPLS label is removed (Xiao X., and Ni L.M., 1999). As far as the original packet is concerned, the routers carrying it through the MPLS network appear as a single hop (Stephenson A., 1998).

Despite its advantages MPLS does however have one major drawback as the protocol to implement QoS in the Internet. The architecture and protocols defined by MPLS require a much more extensive change to conventional IP networks than the other protocols discussed so far. This need for major upgrades will certainly retard the widespread implementation of MPLS, and will perhaps limit the use of LSRs to core routers in the Internet.

The overall description suggested by the Cisco White Paper: "Tag Switching: Uniting Routing and Switching for Scalable, High-Performance Services" [retrieved on September 22, 1998 from the web site http://www.cisco.com/warp/public/732/tag/ tagsw_wp.htm] can be augmented to fully describe an MPLS internetwork. The main components are:

- Edge Label Switching Routers: Located at the boundaries of the network, edge LSRs apply labels to packets.

- Core Label Switching Routers: Core LSRs switch labeled packets or cells based on the labels.

- Label Distribution Protocol (LDP): Used with standard network-layer routing protocols to distribute label information between devices in a Label Switched network.

The basic processing within a MPLS internetwork can be summarized as follows:

1. All LSRs use standard routing protocols to identify routes through the network.

2. Core LSRs use the tables generated by the standard routing protocols to assign and distribute label information via LDP. LSRs receive LDP information and build forwarding tables that make use of the labels.

3. When an edge LSR receives a packet for forwarding across the network, it:

- Analyzes the network-layer header

- Performs applicable network-layer services (such as security, bandwidth management, etc)

- Selects a route for the packet from its routing tables

- Applies a label and forwards the packet to the next-hop label switch

4. The Core LSR receives the labeled packet and switches the packet based solely on the label, without reanalyzing the network-layer header.

5. The packet reaches the edge LSR at the egress point of the network, where the label is stripped off and the packet delivered.

Because MPLS decouples the label distribution mechanisms from the data flows, it supports several physical and link-layer technologies. For ATM, the label is placed in the ATM cell header in the VPI/VCI field. In a LAN, the label is placed after the MAC header. Standard routers can act as core LSRs by running Label Switching software (adding the capability to switch labeled packets based on the label values), and by supporting LDP.

An Integrated Approach to end-to-end Quality of Service on the Internet

The Internet QoS protocols (Integrated Services, Differentiated Services, RSVP, and MPLS) use different technologies to achieve their goal. Whether the protocols become widely adopted by hardware and software vendors largely depends on several factors, some of the most important ones are the scalability of the protocol as well as the granularity of the service classes offered.

In general, Integrated Services (Braden R., Clark D., Shenker S., 1994) is more appropriate for networks with fewer flows, such as those found in LANs and corporate networks. These networks are usually considered to be at the boundaries of the Internet. Differentiated Services (Yoram B. et al, 1999) on the other hand is more appropriate for networks found further within the Internet. For example, ISP networks should implement Differentiated Services because the amount of flows that they would have to manage would be significantly larger than those found in corporate networks. Also, by the very nature of the protocol, ISPs would be able to provide differentiated services to their clients while still able to manage QoS efficiently. MPLS, because of its architecture, is more appropriate for
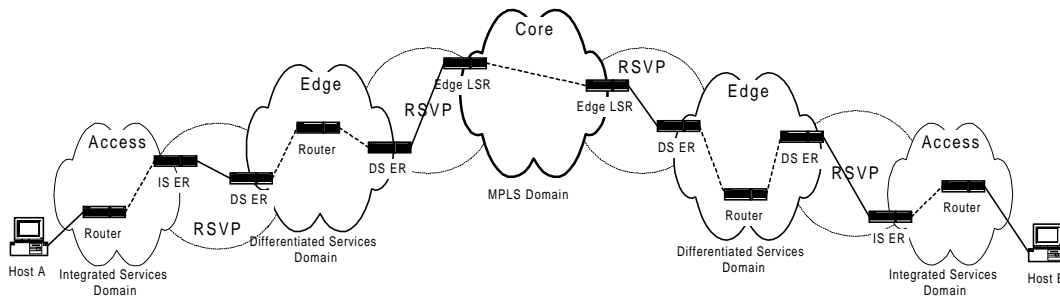
Figure 1 - An Integrated Approach to QoS

routers and networks deep within the Internet, where the number of flows to handle is much greater than those found in Differentiated Services networks. With such great number of flows to manage, the fine granularity of service classes as well as the label switching technology will significantly provide QoS.

As for RSVP, it is the signalling protocol that has been mentioned by each protocol's working group to be a possible candidate for their signalling needs. It is already known that RSVP works very closely with the Integrated Services protocol to deliver end-to-end QoS. In fact, it is considered part of the Integrated Services architecture. It has been suggested that RSVP be adopted in Differentiated Service networks (Yoram B. et al, 1999) so as to provide end-to-end QoS over a Differentiated Service domain. RSVP has also been suggested as a candidate to be used with MPLS (Guerin R., Gan D.H., Kamat S., Li T., Rosen E., 1997) to set up Label Switched Path tunnels. In fact, there is also a MPLS and RSVP Working Group in the IETF, and this working group suggests the use of RSVP for Traffic Engineering in MPLS networks (Awduche D., Malcolm J., Agogbua J., O'Dell M., McManus J., 1999).

TCP/IP, though lacking in QoS characteristics, is required as the common protocol used by the approaches discussed above. The QoS features of ATM on the other hand give us an idea of the characteristics that each of these approaches is trying to achieve. With this in mind, the following diagram illustrates a framework of protocols and their possible mapping with each other to provide end-to-end QoS.

From Figure 1, a data packet originating from an Integrated Services domain, requiring a certain level of QoS, may traverse through several other networks before reaching its destination. These networks might be Differentiated Services networks or MPLS networks. However, if there is a service mapping between the different networks, end-to-end QoS can be achieved by the packet when it arrives at its destination.

Therefore, it is relatively important for each of the protocols discussed so far to have a common understanding of each other's commands and parameters. Only then will QoS agreements in a network be communicated to neighbouring networks, and thus assuring that packets in a particular flow will receive the same if not similar treatment, in terms or delays, delivery rates or latency, across the different networks.

Many current real-time technologies will benefit from such "integration" between the different protocols. The following is a discussion of some of those technologies, namely Internet telephony and programmable networks.

Internet telephony is all about carrying voice over the Internet or corporate Intranet instead of the circuit-switched network. This has been a recent buzzword in the communications industry because of its ability to save significant amounts from organization's monthly phone bills. Internet telephony can be done either using the Public Services Telephone Network (PSTN) and going through a H.323 gateway, or using a computer with Voice Over IP (VOIP) software. There may be several combinations of user types. For instance, a caller can be calling through a H.323 gateway, while the receiver is on a computer with VOIP software, such as Microsoft NetMeeting or IDT Net2Phone. Significant savings can be achieved because calls between a caller and a H.323 Gateway are seen as a local call by the telecommunications company. And if the destination of the call is transnational, the connection is digitised and carried as Internet traffic from caller to receiver. As far as the telecommunications company is concerned, the call is a local call and therefore charged accordingly. The "digitised" portion of the call is charged as Internet traffic, which is significantly cheaper than international toll calls, resulting in huge savings for the corporation.

However, packets carrying Internet telephony data are delay-sensitive and therefore require a certain amount of service quality assurance, which until now only ATM net-

works are able to provide. The Internet carries all types of data, both delay-sensitive and data-sensitive. In the above discussion, the service mappings between the protocols would allow a certain QoS provided in one network to be carried over to other networks. Also, by having the different routers across different networks understand RSVP, vital resources for delivering delay-sensitive packets such as those generated by Internet telephony, can be reserved, providing end-to-end QoS. Also, Integrated Services do not scale well when there are a large number of flows to maintain. By providing the necessary mappings across Integrated Services, Differentiated Services and MPLS, the necessary differentiation as well as aggregation of traffic can take place, resulting in QoS for real-time communications.

Another benefit of "integrating" the protocols is the use of Application Programming Interfaces for Networks (Working Group on IEEE P1520, 2000). Programmable networks involves the ability to use the distributed object computing paradigm, and represent a telecommunications network as abstractions of objects, each with its own states and software interfaces by which that state can be manipulated in order to achieve a global objective. As multimedia communications become increasingly prevalent, it becomes important to deal with special conditions of multimedia traffic and provide for these special conditions with the architecture of an IP router/switch and its management and control. There is an opportunity to reconcile the perspectives of the computing and communication communities in new network architectures that support service creation, QoS control, and the joint allocation of computer and communication resources. It is thus implied that programmable networks were developed to capitalise on this opportunity. The trust of the IP community is towards supporting a very few differentiated services, with different treatment for aggregation of traffic in each service class. One such research direction is to use MPLS in a trunk dedicated to a particular service class and RSVP for setting up and maintaining these trunks. The P1520 group foresees the need for dynamic modifications of policies and configuration within IP routers in the near future. Applications of P1520 could include adjustment of resource allocation per differentiated service class, traffic grooming by source or destination address or customisation of per-flow processing. By providing service mappings between the four protocols discussed in this paper, it would be easier to implement programmable interfaces such as P1520 over a heterogeneous network such as the Internet.

Conclusions

The main difference between the Internet community's approach and the P1520 approach is that the P1520 group attempts to standardize programming interfaces rather than specific algorithms or protocol semantics. Thus, for example in the area of MPLS it is desirable to come up with a set of minimal specifications of interfaces in an Interface Definition Language (IDL), that captures the basic programmability requirements of IP Routers and Switches from the perspective of MPLS related algorithms. Thus whether the algorithms are doing traffic engineering, label distribution protocol or RSVP; they would all use the same common set of interface definitions. In a manner similar to what has been done for ATM networks, we propose that the initial set of definitions targeted for study should be centered around the namespace and capacity region (QoS related) resources. This allows service mappings between protocols, and those mappings are the building blocks of the framework proposed in Figure 1. The ultimate goal being the provision of integrated end-to-end Quality of Service in today's Internet.

References

Awduche D., Malcolm J., Agogbua J., O'Dell M., McManus J. (1999, June). Requirements for Traffic Engineering over MPLS. MPLS Working Group. *Internet Draft <draft-ietf-mpls-traffic-eng-01.txt>*

Baker F., Krawczyk J., Sastry A. (1998, April). RSVP Management Information Base. *Internet Draft <draft-ietf-rsvp-mib-v2-00.txt>*

Biswas, J., Lazar, A.A., Huard J.F., Lim, K., Mahjoub, S., Pau, L.-F., Suzuki. M., Torstensson, S., Wang W., and Weinstein, S. (1998, October). The IEEE P1520 Standards Initiative for Programmable Network Interfaces. *IEEE Communications Magazine*, 64-70.

Braden R., Clark D., Shenker S. (1994, June). Integrated Services Architecture, *RFC 1633*

Callon, R., Doolan, P., Feldman, N., Fredette, A., Swallow, G., Viswanathan, A. (1997, November). A Framework for Multiprotocol Label Switching. Network Working Group. *Internet Draft <draft-ietf-mpls-framework-02.txt>*

Guerin R., Gan D.H., Kamat S., Li T., Rosen E. (1997, November). Setting up Reservations on Explicit Paths using RSVP. *Internet Draft <draft-guerin-expl-path-rsvp-01.txt>*

Rosen E. C., Viswanathan A., Callon A. (1999, April). Multiprotocol Label Switching Architecture. Network Working Group. *Internet Draft <draft-ietf-mpls-arch-05.txt>*

Stephenson A. (1998, November). Diffserv and MPLS: A quality choice. *Data Communications*, 73-77.

Programmable Interfaces for Networks

Trillium Digital Systems Inc. (1998, April). Trillium IP Quality of Service White Paper. http://www.trillium.com/whats-new/wp_ipqos.html

Working Group on IEEE P1520 (2000, January). Application Programming Interfaces for Networks (DRAFT WHITE PAPER), http://www.ieee-pin.org/

Yoram B., Binder J., Blake S., Carlson M., Carpenter B.E. , Srinivasan K., Davies E., Ohlman B., Verma D., Wang Z., Weiss W. (1999, February). A Framework for Differentiated Services. Differentiated Services Working Group. *Internet Draft <draft-ietf-diffserv-framework-02.txt>*

Xiao X., and Ni L.M. (1999, March/April). Internet QoS: A Big Picture, *IEEE Network*, 8-18.

Zhang L., Deering S., Estrin D., Shenker S., Zappala D. (1993, September). RSVP: A New Resource ReSerVation Protocol. *IEEE Network*, 8-18.

Biography

Jairo is a Senior Lecturer in Information Systems at the University of Auckland. Previously he worked as an R&D Manager, Systems Integration Consultant, and Information Systems Manager. He also conducted seminars on LAN/WAN technologies. He teaches data communications and computer networking. His current research topics are in network management systems, programmable networks, and high-speed computer networking.

Terence was a student at the University of Auckland, where he was tutoring in System Analysis and Design. Currently, he is working as a consultant at National Computer Systems Pte Ltd, Singapore, dealing with Internet technologies in education and promoting e-learning.