A Framework for End-to-End Quality of Service Provisioning Models in IP based Networks

Kasigwa J., Baryamureeba V., Williams D.
jkasigwa@ics.mak.ac.ug, barya@ics.mak.ac.ug, d.williams@lsbu.ac.uk
Institute of Computer Science, Makerere University - Uganda
www.makerere.ac.ug/ics (2004)

Abstract: Providing end-to-end Quality of Service (QoS) to diverse classes of traffic in IP (Internet Protocol) based networks has become a great challenge for most network researchers. IP based networks are expected to provide end-to-end QoS guarantees to connections in the form of bounds on delays, jitter and losses. The proposal known as Differentiated Services (DiffServ) is a very promising approach for implementing QoS in IP-based networks and is being discussed and developed. The challenge has been to deliver an integrated end-to-end QoS framework in IP based networks across multiple network domains. The key contribution of this paper is the integrated novel framework for end-to-end QoS provisioning models and admission control in IP based networks. We have shown that, when a mapping function is applied at the boundary of the IntServ and DiffServ domains, end-to-end QoS can be achieved.

Keywords: Admission control, End-to-end QoS, DiffServ, Mapping function

I. INTRODUCTION

End-to-end QoS provisioning is very important for real-time connections in which losses are irrecoverable and delays cause interruptions in service. QoS is defined as the specification of level of performance or behavior from a network that is typically guaranteed to an application or user. The specific parameters which define QoS vary depending on the application and user requirements. QoS is mainly measured in terms of delay, jitter, and packet loss. However, they are not reliable and very difficult to measure precisely [1]. To achieve end-to-end QoS, the network can be divided into three parts: QoS definition, Access network QoS, and Backbone network QoS [3].

(i) *QoS Definition*: Customers, service providers and the network itself must have a common vocabulary that defines QoS, the expected QoS and the provided QoS, and must adhere to a means of moderating QoS levels across the network.
(ii) *Access network QoS*: Bandwidth management and admission control are critical at the point where customer traffic enters the network. End-to-end QoS begins at the customer premises and must be granular enough to differentiate the service requirements of multiple traffic classes.
(iii) *Backbone network QoS*: The network backbone administers QoS differently compared to the access network. Instead of individual traffic streams, backbone network QoS works on aggregates of customer traffic. To support end-to-end service guarantees, the backbone network must be able to differentiate traffic classes to satisfy the overall service levels that have been promised to customers.

One of the challenges faced by the internet service providers (ISPs) is to come up with an integrated end-to-end QoS framework, where IP-based network traffic receives end-to-end QoS support. Another critical problem in QoS research is the lack of common definition for functionalities, delivery methods, and implementation [2]. Specifically, the problem is to provide guaranteed end-to-end delay bound, delay-jitter bound, and packet losses to a real-time traffic (such as voice) from source to destination. The standards pertaining to IP based networks do not specify any mechanism for this problem, and the implementations, if any, are vendor specific [12], [14].

A considerable amount of research has been dedicated to the study of end-to-end QoS, though there are results to calculate bounds on end-to-end delays for certain schemes, there are hardly any mechanisms to provide the required end-to-end QoS bounds [3],13]. It is also important to note that due to the heterogeneous composition of the network, switches and routers could be running different scheduling and admission control algorithms. In this case the calculation of end-to-end delay bounds becomes exceedingly difficult and often inaccurate.

Overview: The rest of this paper is organized as follows: Section II discusses brief overview of previous work; Section III presents our proposed novel framework for end-to-end QoS provisioning models and the mapping function; Section IV concludes the paper and proposes some future enhancements to the presented framework.

II. STATE OF ART

The Internet Engineering Task Force (IETF) has proposed many service models and mechanisms to meet the demand for QoS in IP based networks. Notably among them are the Integrated Services (IntServ) [6], Differentiated Services (DiffServ) [5], Resource Reservation Protocol (RSVP) [7], and Multi-protocol Label Switching (MPLS) [23]. All of these methods mainly deal with the technical solutions for classifying traffic and managing congestions. But from deployment and trial experiences, it is quite clear that there are frustrations and difficulties to implement these end-to-end QoS techniques [4]. The current best-effort service model in IP-based networks does not guarantee timely delivery of packets and is subject to variable queuing delays and congestion loses [9].

The IntServ architecture intends to provide end-to-end bandwidth reservations by maintaining per-flow state information along the path from the sender to the receiver. The IntServ model is characterized by resource reservation [6]. For real time applications, before data is transmitted, the applications must first set up paths and reserve resources. IntServe suggested a fundamental change to the Best Effort design of the internet, by proposing the maintenance of per-flow state in the routers. This obviously was not scalable in IP based networks, the number of entries in the routing table were very many. IntServ also suggested an explicit out-of-band signaling protocol for the reservation of resources [RSVP]. This scalability problem resulted in the DiffServ approach where QoS is not achieved by resource reservations for individual flows, but by assigning packets to certain service classes.

The DiffServ model provides a scalable means of service differentiation in IP based networks [5]. No per-flow state needs to be maintained in the routers, neither is there an explicit connection setup phase. The DiffServ architecture is composed of a number of small functional units implemented in the network nodes, which includes the definition of a set of Per-Hop Behaviors (PHBs), packet classification and traffic conditioning functions like metering, marking, shaping and policing. The DiffServ model is scalable as compared to the IntServ model. At the network level, the DiffServ architecture allows for service differentiation. Bernet et al. [3], describes a framework combining the IntServ and DiffServ ideas in order to provide end-to-end QoS delivery to applications. The IntServ reservations are bridge-spanned over DiffServ regions. A resource provisioning in the DiffServ regions realizes the resource allocation for the IntServ flows. These resources may be statically or dynamically allocated.

In MPLS model [10], the path forwarding and traffic management state are established for traffic streams on each hop along a network path. Traffic aggregates of varying granularity are associated with a Label Switched Path (LSP) at an ingress node, and packets within each label switched path are marked with a forwarding label that is used to lookup the next-hop node, the per-hop forwarding behavior, and the replacement label at each hop.

III. THE PROPOSED FRAMEWORK FOR END-TO-END QOS PROVISIONING MODELS

The proposed framework for end-to-end QoS provisioning models is dynamic and interactively adjusts to the varying QoS needs across multiple domains. The dynamic nature of the proposed solution enables QoS-provisioning to meet the mechanisms and challenges of achieving end-to-end QoS and maintaining it. The proposed solution is built on ideas of Markus, et al [8], on achieving end-to-end QoS in the next generation networks.

The significance of this work is that, end-to-end QoS in IP-based networks can be achieved if the DiffServ backbone network is used to connect IntServ access networks. In our context, we advocate for the DiffServ model in the backbone network, which lays a valuable foundation to achieve end-to-end QoS. The proposed framework is integrated through five models across multiple domains.

A.. THE ORGANIZATIONAL MODEL

End-to-end QoS provisioning in IP based networks is required both in the access network and the backbone network. Each component is provisioned separately and then a mapping function is applied to provide the complete end-to-end QoS path. This allows an end-to-end path to be segmented into; the access network; the backbone network and the inter-domain QoS negotiation.

i. The Access network

The basic framework of IntServ is implemented by four components: the signaling protocol (e.g. RSVP), the admission control routine, the classifier and the packet scheduler. In this model, applications must set up paths and reserve resources before transmitting their data. Network elements will apply admission control to those requests. In addition, traffic control mechanisms on the network element are configured to ensure that each admitted flow receives the service requested in strict isolation from other traffic. When a router receives a packet, the classifier will perform multi-field classification and put the packet in a specific queue. The packet scheduler will then schedule the packet according to its QoS requirements.

ii. The Backbone network

To combine the advantages of DiffServ (good scalability in the backbone) and IntServ (per flow QoS) guarantee), a mapping from IntServ traffic classes to DiffServ classes has to be performed. The mapping function is used to assign an appropriate DSCP code to packets arriving from a flow specified in the IntServ domain. This is to ensure that the appropriate QoS can be achieved for IntServ flows when running over a DiffServ domain.
To achieve the above goal, we introduce a mapping function at the Access Network Gateway (ANG) between the DiffServ and the IntServ domain as shown in Figure 1 below. Every packet in the flow from an IntServ domain has a *flow ID* indicated in the *flow-id* field in the IP header. The *flow ID,* attributed with the mapping function is used to determine which flow the packet belongs to.

Packets specified by *the function* parameters in IntServ domain are first mapped to the corresponding PHBs in the DiffServ domain by appropriately assigning a DSCP according to the mapping function. The packets are then routed in the DiffServ domain where they receive treatment based on their DSCP code. The packets are grouped into behavior aggregates (BAs) in the DiffServ domain.

Table 1: An example of the mapping function between IntServ and DiffServ domains

Traffic flow	**Flow ID**	**PHB**	**DSCP**
r = 0.7 Mb, b = 5000 bytes	0	EF	101110
r = 0.7 Mb, b = 5000 bytes	1	EF	101110
r = 0.5 Mb, b = 8000 bytes	2	AF11	001010
r = 0.5 Mb, b = 8000 bytes	3	AF11	001010
r = 0.5 Mb, b = 8000 bytes	4	AF11	001010

Table 1 shows an example mapping function which has been used to classify different traffic flows. A flow in IntServ domain specified by r = 0.7 mb, b = 5000 bytes and flow ID = 0 is mapped to EF PHB (with corresponding DSCP 101110) in DiffServ domain, where r means token bucket rate and b means token bucket depth.

The sender initially specifies its requested service by specifying traffic flows. Traffic flows are differentiated by the flow ID. It is also possible that different flows can be mapped to the same PHB in DiffServ domain. Without making any significant changes to the IntServ or DiffServ infrastructure and without any additional protocols or signaling, it is possible to provide end-to-end QoS to IntServ application when IntServ runs over a DiffServ network.

iii. Inter-domain QoS negotiation

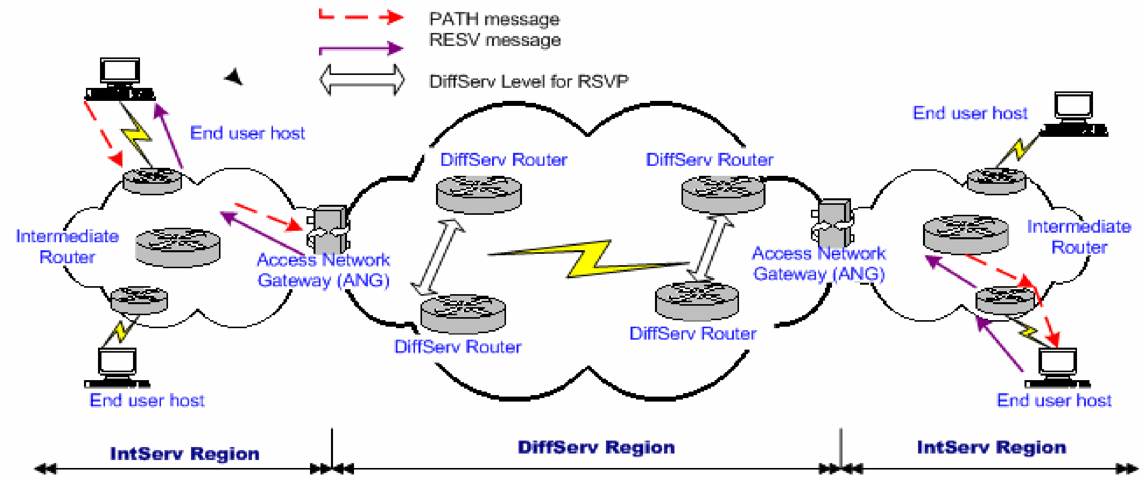

Fig1.The proposed end-to-end QoS framework

To negotiate end-to-end QoS specification across domains, the service classification and admission control between ANGs is needed. The initial signaling process to obtain end-to-end quantitative QoS starts when the end user host generates an RSVP PATH message. The generated PATH message is forwarded to the ANG along the access routers and intermediate routers. Then ANG sends the received PATH message towards the DiffServ region. To forward the PATH message, ANG maps the message to the DiffServ service level. The service mapping is possible by defining a new class or utilizing the existing classes. After the mapping function procedure, the PATH message is routed based on PHB. When the PATH message gets to the destination IntServ region, the message is processed according to the standard RSVP processing rules. When the PATH message reaches the destination end user host, the host generates an RSVP RESV message. And the RESV message is routed to the source end user host along the reverse path.

The requests may be rejected at any node in the IntServ region according to IntServ admission control. Also, ANG triggers the admission control when it receives the RESV message. The ANG compares the resources requested in the RESV message to the corresponding DiffServ service level. If there are enough resources in the DiffServ region and the request fits in the customer’s SLS (Service Level Specification), the request is granted. If not, the RESV message is not forwarded and the appropriate RSVP error message is sent back to the receiver. Figure 1 shows the initial QoS negotiation procedure in the proposed framework.

B. THE INFORMATIONAL MODEL

The informational model handles service level agreements (SLAs) between the service providers and customers. SLA is a contract between a service user and a service provider outlining the desirable QoS, contracted QoS, offered QoS, and expected QoS. To control conformity to a SLA, the network QoS should be evaluated in terms of generic criteria i.e. availability, reliability, delay, jitter and bandwidth capacity that apply to all the required QoS levels [11]. The quantitative aspects can then be measured through specific parameters of each visibility level. The resulting end-to-end QoS is then the aggregation flow of intermediate QoSs.

C. THE ARCHITECTURAL MODEL

The architectural model performs an end-to-end QoS delivery across multiple domains. In backbone routers, a behavior aggregate classifier is needed to recognize packets for a required class of service (CoS). Depending on the CoSs pre-defined in the domain; buffer acceptance and admission control algorithms have to be chosen accordingly. In boundary routers, classification has to be SLA based. Classifiers can be multi-field to restitute flow identification from multiple header fields. Markers are also needed to set the DiffServ field to the CoS resulting from classification. Within egress routers, markers act as a mapping function that sets the DiffServ field to a CoS defined in the next domain. Meters will compare traffic profiles to inter-domain contracts. Nonconforming packets that cannot be shaped will be dropped by policers at ingress routers, or delayed by shapers at egress nodes. Traffic control mechanisms act on packets and enable very short reaction delays.

D. THE FUNCTIONAL MODEL

In the functional model, QoS provisioning must be dynamically performed at inter-domain nodes to adapt to local QoS and offer the required end-to-end QoS. This adaptation is needed to make up for the QoS deficiency or to reduce an eventual surplus in the QoS offered. End-to-end QoS support can be achieved through an integration of inter-domain and intra-domain resource allocations. The end user negotiates a SLA with its domain and the resulting agreement includes among others two important parts: the user expected end-to-end QoS expressed in QoS metrics as per the customer SLA; and the traffic profile to which customer promises or network should conform to, expressed in a meaningful way to the service provider. The domain operates on predefined QoS strategies and protocols use the internal QoS support based on the negotiation and the mapping functions. In the negotiation function, before a host can send its traffic, it has to request a QoS from its domain QoS broker [15].

E. THE ADMISSION CONTROL MODEL

In order to provide end-to-end QoS in IP based networks, a dynamic admission control scheme is definitely needed to work in conjunction with the mapping function. In this section, we present a simple measurement-based admission control (MBAC) algorithm for DiffServ backbone networks. The admission control decisions are based on bandwidth reservations and periodically measured & exponentially averaged link loads. If any link load on the ANG is over the applicable threshold, access is denied. Link loads are periodically sent to bandwidth broker (BB) of the routing domain, which makes the admission control decisions.

The Admission Algorithm:

Let be μ be the link bandwidth, r^{α} the amount of bandwidth requested by flow α, and $\hat{u}$ be the measured load of current traffic. Then the measured sum algorithm accepts α into the network only if the following condition is satisfied:

$$\hat{u} + r^{\alpha} < c\,\mu,$$

where $0 < c < 1$;

c is a user-specified utilization target. When the bandwidth utilization approaches 100%, the variation of packet delay grows large and the algorithm will make wrong admission decisions as a result. Therefore a utilization target is set and the network is kept operating under the desired utilization.

IV. CONCLUSION

In this paper, we addressed the end-to-end QoS provisioning framework in diverse network situations. It utilizes the scalable feature in DiffServ model and the per-flow resource management in IntServ model. The framework is able to meet the end-to-end QoS requirements across multiple domains by use of a mapping function at the ANG.

The most important entities in our framework are Access Network Gateways (ANG). Since all nodes except ANG keep the standard DiffServ and IntServ/RSVP mechanisms, the framework is so flexible and scalable. This framework can be applied as the end-to-end QoS model in the next generation IP-based networks. The combination of admission control and provisioning models is key to end-to-end QoS solution. It can be implemented without massive change to existing architectures, and can satisfy the needs of business customers of IP-based networks.

For the future work, enabling end-to-end QoS may call for a huge variety of mechanisms that intervene at various times and execution levels. Each of these architectural components performs scheduling and admission control algorithms whose selection is dictated by the QoS strategy.

REFERENCES

[1] M. Albrecht, M. Kster, (2003). End-to-end QoS Management for Delay-sensitive Scalable., The Proceedings of the 25 Annual Conf. on Local Computer Networks (LCN'03), Tampa, FL.

[2] V. Fineberg (2002). A Practical Architecture for Implementing End-to-End QoS in an IP Network,. IEEE Communications Magazine, January 2002, pp.122-130.

[3] Y. Bernet, R. Yavatkar, F. Baker, L. Zhang, et al. (2000). A Framework For Integrated Services Operation Over Diffserv Networks. IETF, Internet Draft ¡draft-ietf-issll-diffserv-rsvp-05.txt.

[4] C. Samir, B. Jongbok, (2002). Quest for the end-to-end network QoS. *2002 . Eighth Americas Conference on Information Systems, pp 1919-1925.*

[5] S. Blake, D. Black, M. Carlson, (1998). Architecture for Differentiated services. IETF RFC 2475.

[6] R. Braden, , S. Shenker, D. Clark, (1994). Integrated Services in the Internet Architecture: An Overview, RFC 1633. Available: http://www.ietf.org/rfc/rfc1633.txt?number=1633

[7] R. Braden, L. Zhang, et al.(1997). Resource ReSerVation Protocol (RSVP), version 1 Functional Specification, RFC 2205. Available: http://www.ietf.org/rfc/rfc2205.txt?number=2205

[8] A. Markus, K. Michael, M. Peter, F. Matthias (2000); End-to-end QoS Management for Delay-sensitive Scalable Multimedia Streams over DiffServ. Published in the Proceedings of the 25th Annual Conference on Local Computer Networks (LCN'00), Tampa, FL, USA.

[9] D. Clark, and W. Fang, (1998). Explicit allocation of best-effort packet delivery service. IEEE/ACM Transactions on Networking, 6(4):362-373.

[10] E. Rosen, A. Viswanathan, R. Callon 2001). Multi-protocol Label Switching (MPLS) Architecture, RFC 3031. Available: http://www.ietf.org/rfc/rfc3031.txt?number=3031

[11] N. Saad, N. Simoni, J.F. Pernet, (1998). ”Tactical Network Management for Preserving the QoS”, published IEEE Milcom'98.

[12] I. Stoica, and H. Zhang, (1999). ”Providing guaranteed services without per-flow management,” In: Proc. ACM SIGCOMM '99, Boston, MA, pp. 81-94.

[13] O. Schelén, “Quality of Service Agents in the Internet”, Ph.D. thesis, Division of Computer Communication, Dept. of Computer Science and Electrical Engineering, Luleå University of Technology, 1998. http://www.cdt.luth.se/~olov/publications/thesis.pdf

[14] R. Gurin, , V. Peris, (2003). ”Quality of Service in Packet Networks: Basic Mechanisms and Directions”, Computer Networks, 31:169-189.

[15] M. Gnter, T. Braun, (1999). Evaluation of Bandwidth Broker Signaling. Proc. of ICNP'99, IEEE Computer. Society. Sec 3.4.1