

# A Bandwidth-Broker Based Inter-domain SLA Negotiation

Haci A. Mantar<sup>1</sup>, Ibrahim T. Okumus<sup>2</sup>, Junseok Hwang<sup>3</sup>, and Steve Chapin<sup>4</sup>

<sup>1</sup> Department of Computer Engineering, Gebze Institute of Technology, Turkey

<sup>2</sup> Department of Computer Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY, USA

<sup>3</sup> Seoul National University, Seoul, Korea

<sup>4</sup> Department of Electronics and Computer Education, Mugla University, Turkey

**Abstract.** This work presents an Inter-Domain Bandwidth Broker (BB) based Service Level Agreements (SLAs) Negotiation Model for Differentiated Services (DiffServ) networks. A BB in each DiffServ domain handles SLAs on behalf of its domain by communicating with its neighboring peers. The proposed model uses a destination-based SLAs aggregation to increase signaling and state scalability, and it uses a BB-based inter-domain routing to increase resource utilization. The experimental results are provided to verify the achievements.

## 1 Introduction

The Differentiated Services (DiffServ) has become a key technology in achieving Quality of Services (QoS) in the Internet. DiffServ aggregates individual flows into different traffic classes at the edge of the network, and core routers within the domain forward each packet to its next hop according to the per-hop behavior (PHB) associated with the traffic class of the packet. Diffserv requires no per-flow admission control or signaling and, consequently, routers do not maintain any per-flow state and operation, and this greatly improves the scalability. However, due to the lack of admission control and signaling, DiffServ does not provide QoS guarantees to individual flows. Recently, many studies have focused on providing QoS guarantees in a domain. However, delivery of end-to-end QoS to support end-user applications requires the resource reservation in all the domains along the path. The resource reservation among different administrative domains is handled with service level agreements (SLAs) negotiation.

The Bandwidth Broker (BB) model [2] is a strong candidate for SLAs negotiation among DiffServ domains. As a central logical entity in each DiffServ domain, a BB is mainly responsible for the inter-domain SLA negotiation for its entire domain. The BB makes policy access, resource reservation and admission control decisions on behalf of its entire domain.

In [1], we present a Simple Inter-BB Signaling (SIBBS) protocol. A BB uses SIBBS to communicate with its peers to reserve resources for its inter-domain QoS traffic. SIBBS employs a core tunneling model, in which a pipe is pre-established between each possible source and destination domain and carries the

traffic of a particular class. Pipes are identified by a destination domain IP prefix and by DiffServ Code Point (DSCP). After a pipe is established, all individual reservations for particular destination and DiffServ class are multiplexed into the same pipe. The states of border routers and signaling messages exchanged between BBs are pipe-based. All the reservations in a pipe are considered as a single reservation from the intermediate BBs' point of view.

By aggregating individual reservations into an existing pipe, SIBBS significantly reduces the inter-BB signaling load and admission control time compared to per-flow/per-reservation schemes. Although SIBBS can be effectively used for small-scale networks such as VPN applications, it has a serious scalability problem when it is applied to large-scale networks such as the entire Internet. The number of pipes in core transit domains scale with  $O(n^2)$ , where  $n$  is the number of domains or networks in the Internet. When it is applied to the entire Internet, this number might be too high to be managed. Another problem is that the numbers of inter-BB signaling messages and states maintained by a BB are proportional to the number of pipes. This causes processing, signaling and state scalability problems.

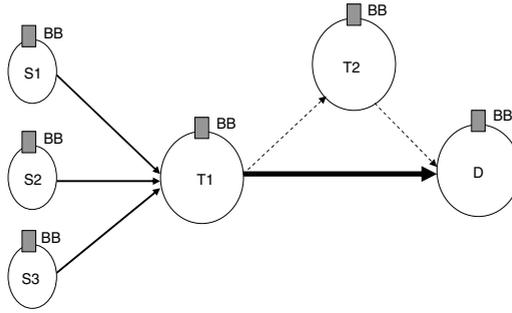
In this work we enhance SIBBS by adding a destination-based aggregation to make the inter-domain SLAs negotiation scalable and efficient. As the only entity responsible for QoS resource control in a domain, a BB uses this protocol to reserve/negotiate QoS resources with its peers located in the neighboring domains for its inter-domain QoS traffic.

## 2 Destination-Based Inter-domain SLAs Negotiation

The idea of destination-based aggregation is very simple. A provider domain's BB aggregates its customers' (upstream domains) pipe requests of the same destination region and QoS class, and makes a single reservation (setup a single pipe) with its downstream domain on behalf of all of its customers. We name this scheme enhanced SIBBS (eSIBBS).

Consider Figure 1, where  $S1, S2, S3$  represent source domains,  $T1, T2$  represent transit domains and  $D$  represents a destination domain. For simplicity, we assume that there are no end hosts located in transit domain and that requests are directed from  $S_i$  to  $D$ . We also assume that requests are based on bandwidth demands. As shown, source domains ( $S1, S2, S3$ ) establish a pipe to transit domain ( $T1$ ) for the traffic destined for  $D$ . The BB of  $T1$  aggregates its customers ( $S1, S2, S3$ ) requests and establishes a pipe to  $D$ . The resources of the pipe between  $T1$  and  $D$  is shared by all the customers of  $T1$ , and the resource negotiation is based on the pipe utilization. To reduce the numbers of signaling messages with  $D$ , the  $T1$  can make the reservation size more than the current demand (the sum of the demand from  $S1, S2$ , and  $S3$ ). The BB of  $T1$  can grant its customers requests as long as there is available resources in the pipe with  $D$ .

In the above aggregation scheme, we assume that the aggregated traffic follows a single path, determined by BGP-4, towards a destination (as if there is only a single provider for each destination). The resource provisioning and reservation is



**Fig. 1.** A simple multi-provider network

performed based only on the resource availability on that path. While this single path based (BGP-4 based) aggregation scheme increases scalability, it may not be sufficient for resource utilization efficiency.

Assume that only the path  $T1 - D$  has limited resources (Figure 1). When the traffic rate of the path  $T1 - D$  reaches its maximum capacity, the reservation requests coming from source domains will be rejected even if the alternative path,  $T1 - T2 - D$ , has available resources. This is because the traffic from  $T1$  to  $D$  is routed through the link  $T1 - D$ , which is the path given by BGP-4.

To alleviate this problem, we enhance the BB to play a role in inter-domain path selection. Among the possible candidate paths, provided by BGP-4, the BB selects a path that has available resources by signaling the candidate BBs. For example,  $T1$  has two alternative paths ( $T1 - T2 - D$  and  $T1 - D$ ) to send its traffic to  $D$ . When the BB of  $T1$  needs resource reservation to  $D$ , it can send a request to the BBs of  $D$  and  $T2$  asking for resources. Upon receiving the request, the BBs  $D$  and  $T2$  send reply back to the BB of  $T1$ . The BB of  $T1$  then chooses the path that has the requested resources. In case both of them have available resources, it can choose the least costly one.

Selecting a path based on the resource availability can significantly increase the resource utilization. But, sometimes it may not be possible to send the whole aggregated traffic through a single path. In this case, a domain can send the aggregated traffic over multiple paths. For example,  $T1$  can split and send its traffic over the paths  $T1 - D$  and  $T1 - T2 - D$ .

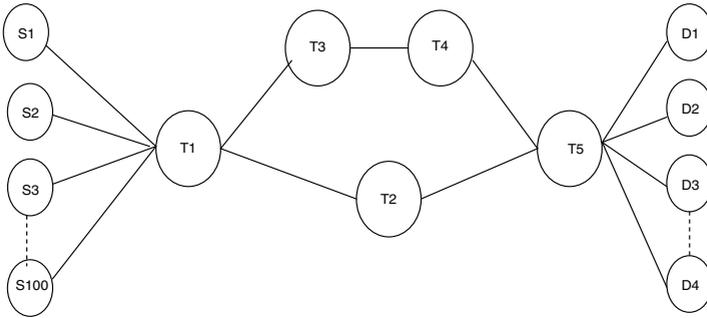
To split the aggregated traffic over multiple paths in a way that avoids out-of-order packet delivery, we use a hashing based scheme. The hashing can be done based on source and destination IP addresses, and possibly other fields of IP header. In hashing model, the traffic will be first distributed into  $N$  bins by using module- $N$  operation on the hash space. If the total traffic rate is  $X$  bps, each bin approximately receives the amount of  $X/N$  bps. The next step is to map bins to paths. The number of bins assigned to a path is determined based on its load portion.

With the hashing-based scheme, the number of states that border routers, which performs splitting, maintain is independent of the number of flows constituting the

aggregation. By receiving the load portion of each path from the BB, the routers just need to perform forwarding based on the result of module-N operation.

### 3 Evaluation Results

In this section, we present simulation results to verify that the proposed model is robust, scalable and efficient. Figure 2 illustrates the configuration of the test-bed, which consists of 100 source domains ( $S1, S2, \dots, S100$ ), 10 destination domains ( $D1, D2, \dots, D10$ ) and five transit domains ( $T1, T2, T3, T4, T5$ ). Each source domain and destination domain have 10 end hosts connected. Resource reservation requests are from source domains to the destination domains.



**Fig. 2.** BB Test-bed

We used real-time traffic traces collected from real networks for our experiments. We used traffic traces provided by CAIDA [7], which has advanced tools to collect and analyze data based on source and destination Autonomous Systems (AS) and traffic types for short time intervals. Source domains sent the traffic to the destination domain based on the traced data characteristics. In order to imitate the traced traffic characteristics, hosts located in source domains were configured to send UDP traffic with a variable average rate that reflects the traced data. The experiments are performed for virtual leased line (VLL) services [1][4]. The traffic rate of an individual request and a pipe is determined with the parameter-based scheme where each host requests reservation with the peak rate. All the experiments were run for 25-minute periods. The duration of a reservation was exponentially distributed with a mean of 1 minute. The reservation rates varied over the duration of the experiment, based on the rate profile originated from the traced data.

Table 1 shows the number of state messages that a BB (the BB of  $T1$  in this experiment) maintains in case of SIBBS and eSIBBS. The table shows that a BB can be a potential bottleneck in case of SIBBS when the number of source-destination pairs are large. On the other hand, this number is proportional to the number of destination domains in eSIBBS. The BB signaling scalability

**Table 1.** BB state scalability

# src.domains	SIBBS	eSIBBS
10	865	493
20	1877	497
30	2911	505
40	3818	507
50	4873	508
60	5756	509
70	6871	509
80	7953	509
90	8909	509
100	9929	509

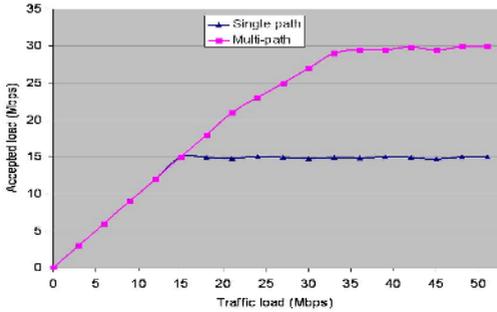
**Table 2.** BB signaling scalability (The # messages processed by the BB of  $T1$ )

# src domains	SIBBS	eSIBBS
10	1715	911
20	3707	915
30	5915	913
40	8567	921
50	9312	918
60	11471	917
70	13366	913
80	15213	915
90	17643	911
100	19789	909

**Table 3.** Border router state scalability

# src. domains	SIBBS	eSIBBS
10	187	93
20	191	97
30	285	105
40	391	108
50	487	109
60	579	109
70	678	109
80	785	109
90	892	109
100	991	109

results (the number of messages processed by a BB) are given in Table 2. Table 3 illustrates the comparison of SIBBS and eSIBBS in terms of border router state scalability. As depicted, when the number of sessions in each stub domain increases, the number of states in border routers increases. In case of eSIBBS,



**Fig. 3.** Accepted load in single and multi-path scheme

the number of states remain almost constant. (This measurement was done in the egress router of  $T1$ .)

In the last experiment (Figure 3), we evaluated the inter-domain multi-path scheme. We configured the paths  $T1 - T3 - T4 - T5 - Di$  and  $T1 - T2 - T5 - Di$  with 15Mbps and the rest of the links in the network with 30Mbps capacity. In single path case (when BGP-4 was used), the BB of  $T1$  sends all its traffic via  $T2$ . When that path reaches its capacity, the BB of  $T1$  rejects all the incoming requests. In multi-path case, the BB of  $T1$  accepts the requests up to 30Mbps. Because it sends through both providers ( $T2$  and  $T3$ ).

## 4 Conclusion and Future Work

This work shows the significant role a BB can play for service providers, and addresses many outstanding questions about the BB model. The experimental results show that an ISP can substantially improve its resource utilization, thereby increase its revenue, while requiring minimal changes in the underlying infrastructure. The scalability and utilization results provide the basic guidelines that an ISP should consider in defining SLAs.

In this work we have relied on VLL services (EF), however, it is obvious that there is a potential need for other services, so we will enhance SIBBS protocol to support other DiffServ classes (e.g., Assured Forwarding). In the future, we will also investigate the marketing aspect of SLAs among ISPs.

## References

1. QBone Signaling Design Team, “Simple Inter-domain Bandwidth Broker Signaling (SIBBS)”, <http://qbone.internet2.edu/bb/> work in progress.
2. K. Nichols, V. Jacobson, and L. Zhang. “A Two-bit Differentiated Services Architecture for the Internet” RFC 2638, July 1999.
3. S. Black et al., “An Architecture for Differentiated Services,” RFC2475, Dec. 1998.
4. V. Jacobson, K. Nichols, K. Poduri, “An Expedited Forwarding PHB,” RFC2598, June 1999. A

5. R. Braden, D. Clark, S. Shenker, "Integrated Services in the Internet Architecture: an Overview," RFC 1633, June 1994.
6. Y. Rekhter, T. Li, "A Border Gateway Protocol 4 ", RFC 1771, March 1995.
7. <http://www.caida.org/dynamic/analysis/workload/sdnap/>.
8. H. Mantar, J.Hwang, S. Chapin, I. Okumus "A Scalable Model for Inter-Bandwidth Broker Resource Reservation and Provisioning", IEEE Journal on Selected Areas in Communications (JSAC), Vol.22, No.10, December 2004
9. D. Di Sorte, G. Reali: "Pricing and Brokering Services over Interconnected IP Networks", Journal of Network and Computer Applications, 2005, pp. 249-283.