

# A SCALABLE AND EFFICIENT INTER-DOMAIN QoS ROUTING ARCHITECTURE FOR DIFFSERV NETWORKS<sup>1</sup>

Haci A. Mantar<sup>\*</sup>, Junseok Hwang<sup>+</sup>, Steve J. Chapin<sup>\*</sup>, Ibrahim Okumus<sup>\*</sup>

<sup>\*</sup>*L. C. Smith College of Engineering and Computer Science*

<sup>+</sup>*School of Information Studies, Syracuse University*

*{hamantar, jshwang, iokumus}@syr.edu, chapin@ecs.syr.edu*

**Abstract:** With Bandwidth Broker (BB) support in each domain, Differentiated Services (Diffserv) is seen as a key technology for achieving QoS guarantees in a scalable, efficient, and deployable manner in the Internet. In this paper we present a Route Server (RS) architecture that is compatible with the BB model for inter-domain QoS routing. The RS is responsible for determining QoS routes on behalf of all the routers and for exchanging routing information with its neighboring peers. It optimizes network resources by taking the intra-domain resource utilization state into account for selecting a route. It also achieves scalability by pre-computing a limited number of paths for each destination region and mapping all the packets to one of these paths regardless of their sources.

**Key words:** Bandwidth Broker (BB), Inter-domain QoS Routing, Route Server.

## 1. INTRODUCTION

Providing end-to-end QoS in a multi-domain environment is a very complex problem. Each domain is under different administration so that QoS *policies* and *services* in one domain might be significantly different from those in other domains [1,4,5]. A domain administration has control only of its own domain resources and discloses very restricted information about its internal network to others (e.g., competitors).

Recently, many studies [1,3-9] have addressed the important role of *central servers* for controlling access and managing resources in a domain. A well-recognized central server is a Bandwidth Broker (BB) [1]. The idea behind the BB model is to provide Intserv-type end-to-end quantitative QoS guarantees in Diffserv-enabled networks without per-flow state in the network core. With such a scheme, control functionalities such as policy control, admission control, and resource reservation are decoupled from routers into the BB and therefore the BB makes policy access and

---

<sup>1</sup>This work is supported by the National Science Foundation under grant ANI-0123939

increases the scalability of scalable as well as the likelihood of the deployment of QoS because of the minimum changes required in network infrastructure, and simplification of accounting and billing. Similarly, the IETF proposed COPS, where a Policy Decision Point performs tasks similar to those of a BB. In this sense, the BB-supported Diffserv model is realized to be a key technology by many researchers [1,3-9] for making end-to-end QoS possible across multi-domains networks.

Having said all this, the BB-supported Diffserv model still has the following open issues: 1) how to get dynamic link states without signaling with routers; 2) how to assure quantitative QoS guarantees with no reservation in core routers; 3) how to get QoS and cost information about the networks beyond its domain (e.g., which provider it should choose for border-crossing traffic); 4) how to perform service mapping in the inter-connection points; 5) how to manage its domain resources efficiently; 6) how to communicate and reserve resources with neighboring BB for border-crossing traffic; 7) how to achieve inter-domain scalability (signaling between BB and state in border routers). In our previous work [3,6,7], we presented solutions for problems 6 and 7. The concept of this paper is a base for the solution of problems 1-5.

We propose a route server (RS) architecture that works in conjunction with a BB reservation model (independently BGP-4) and makes scalable and efficient QoS routing possible in a multi-domain, multi-policy, multi-technology environment with restricted routing information. In particular, we define two *central* routing entities called Exterior Route Server (eRS) and interior Route Server (iRS) as main components of a BB (Figure 1). The eRS is mainly responsible for routing information distribution between domains and path computation and selection at a high level. The iRS maintains intra-domain QoS information and presents to the eRS in a simplified fashion. At run time the BB uses both iRS and eRS databases to perform admission control without invoking routing protocol and interaction nodes along the path.

## 2. DIFFSERV ROUTE SERVER ARCHITECTURE

### 2.1 Assumptions and Definitions

As shown in Figure 1, our BB architecture comprises three main components: the Exterior Route Server (eRS), the Interior Route Server (iRS), and the Reservation Module (RM). Before going into more detail on these components, we make the following definitions.

**Service level agreement (SLA):** We assume that domains have *bilateral* SLAs with their adjacent neighbors, which can act as both customer and provider.

**Destination region:** Due to signaling and state scalability problems with per-flow routing schemes, a destination region (*destID*) in RS represents a domain, a network, or a set of points identified by CIDR (IPv4/X where  $X < 32$ ).

**Point-to-Destination Services (*dest\_serv*)** specifies the upper bounds of QoS constraints (such as delay, loss ratio) that identifiable traffic will be experienced from ingress point of the provider domain to the final destination region, regardless of the network utilization. Each domain assumed to have limited numbers of pre-defined (see next section) *dest\_serv* for each destination region and all incoming traffic is mapped to one of these *dest\_serv*.

**Edge-to-Edge QoS Class** defines upper bounds of QoS constraints that identifiable traffic will experience from the ingress nodes to the egress nodes of a domain,

regardless of network utilization. Conceptually, this is very similar to Per-domain behavior (PDB) defined in [2]. (From now on, the terms “edge-to-edge QoS” and “PDB” will be used interchangeably.) Following [2], we assume each domain has limited number of PDBs.

**Inter-domain routing information** is solely maintained by the eRS of each domain, independently of BGP-4. As shown in Figure 1, each eRS communicates and exchanges routing information with its peers via the same TCP session that is already established between BBs for resource reservation (see SIBBS [3] for detail). A routing information message basically includes *destID*, and *dest\_serv*,  $\langle destID, dest\_serv \rangle$ . For a *destID*, there might be multiple paths, each of which represents a different *dest\_serv*. A *dest\_serv* includes the service ID (*servID*) and the *cost* of the service in addition to QoS parameters.

**Maintenance of intra-domain network state:** The iRS acts as a link state protocol peer with the nodes in the network and gets link state data as part of the protocol operation. As shown in Figure 1, each node sends its state directly to the iRS instead of flooding it to all the nodes in the domain.

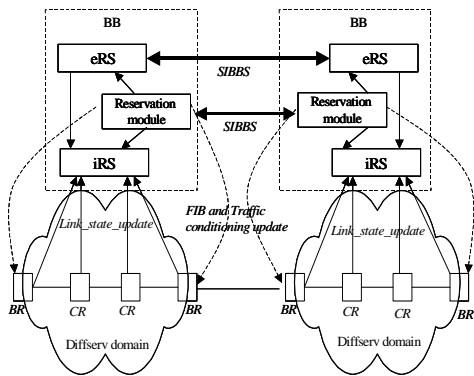


Fig.1. Diffserv QoS Architecture

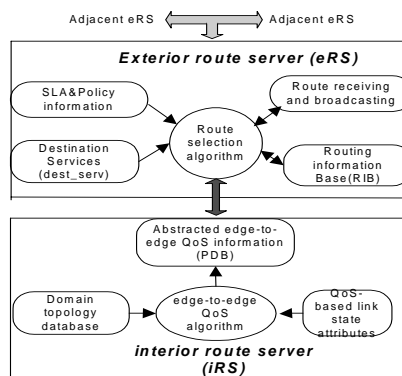


Fig.2: Functional decomposition of eRS

## 2.2 Overview

Based on the above assumptions and definitions, the RS model can be briefly described as follows: the iRS dynamically receives link utilization state from each node, computes the cost and available bandwidth (BW) of pre-established paths between each ingress-egress pair and then maintains this information in the abstracted edge-to-edge database (Figure 2). The eRS performs inter-domain route selection for each  $\langle destID, dest\_serv \rangle$  among the routes received from its peers by taking intra-domain network utilization into account. Note that the available BW in a route is not directly taken into account in route selection and it is not included in route advertisement. It is explicitly not the task of an eRS to locate routes that are guaranteed to have resources available at the time of route advertisement. This is motivated by the fact that even if available BW is included in route advertisements, it is not guaranteed to have available BW at the time of reservation because of dynamic network conditions.

Note that neither the iRS nor eRS algorithms affect existing reservations. They can be considered as modules that work in the background and maintain the databases as

shown in Figure 2. In other words, they are not triggered by individual reservation requests. When a BB needs to make a reservation, it checks the abstracted edge-to-edge QoS database maintained by the iRS for the intra-domain admissibility test and RIB maintained by the eRS for selecting the next BB.

### 2.3 Edge-to-Edge QoS Routing Model

**Quantitative Per Hop Behavior (PHB):** We assume that each PHB is assigned to a certain share of link capacity and that each PHB can use only its share, no preemption. The surplus capacity of a PHB can only be used by best-effort traffic. With the exception of EF [1], all PHBs in Diffserv context define relative QoS assurance. To have *quantitative* QoS guarantees in a scalable manner, we associate an upper delay bound  $d$ , loss ratio bound  $l$ , and *cost*,  $\langle d, l, cost \rangle$  to each PHB (e.g.  $d < 3ms, l < 10^{-2}$ ). It is assumed that  $d$  and  $l$  are pre-determined at the stage of network dimensioning (configuration) [2,3,9], which is done over long time intervals (e.g., days, weeks). The cost is dynamically changed with link utilization.

**PDB (edge-to-edge QoS) and Path:** Similarly to PHB, PDBs are determined at the stage of network dimensioning [2,3,9]. A PDB is expressed by the sum of PHB constraints of all the nodes along the path. In offline, the iRS establishes a path between each ingress-egress pair, one for each PDB. Here, the path can be an MPLS tunnel or any other existing scheme. A path, in turn, a PDB is associated with delay bound  $D$ , loss ratio bound  $L$ , and cost value  $COST$ ,  $\langle D, L, COST \rangle$ .

Because of the scalability problem, it is clearly recommended that both PHB and PDB QoS constraints, for guaranteed services, be independent of network utilization [2]. To achieve this in a scalable manner, we introduce a *self-adaptive per-node provisioning* algorithm. The iRS installs PHB parameters ( $d, l$ ) in each node. The algorithm, in each node, then dynamically monitors each PHB queue size and adjusts the scheduler rate and the buffer size for each PHB to meet pre-defined PHB constraints with respect to the traffic rate. It also calculates the current available BW and *cost* and sends them to the iRS. So, there is no need for a signaling scheme that configures each node along the path when a new reservation is granted. Due to communication overhead and the scalability concerns, we use a threshold-based triggering [4] of link state updates, where a new update is sent when the change in cost since the last update exceeds some predetermined threshold. By having the cost and available BW of each PHB in each node, the iRS dynamically computes the available BW and cost of each path (PDB) and stores this information in the edge-to-edge QoS database. Here, the cost is used by the eRS for route selection and the available BW is used by the BB for resource reservation. Although an iRS has detailed knowledge of its domain, it represents the domain in an abstract fashion to an eRS. The eRS sees the domain as if it consists only of border routers.

### 2.4 eRS Route Selection Algorithm

The eRS route selection algorithm is based on three main components, PDBs and their cost, destination services (*dest\_serv*), and the route received from peers. As mentioned before, from the eRS point of view, PDBs and *dest\_serv* can be considered as static components (modified at the stage of network dimensioning). The cost of PDBs and the route received from peers are dynamically changed with the network conditions. An eRS may have multiple peers from just a few up to hundreds, and each

of them can be potential customers and providers. Upon receiving a route from its peer(s), it performs the following algorithm:

- Discard all the routes that failed in policy and SLA checks.
- Add PDBs to each *dest\_serv(n)* received from peers, and compare with *dest\_serv(I)*. All the results that satisfy *dest\_serv(I)*'s requirements are selected as candidate routes.
- Store all candidate routes in RIB in the order of increasing cost, and send least costly one to peers.
- Repeat steps 2-3 for each *dest\_serv(n)* (1,2...n).

## **2.5 Resource Reservation**

By having the up-to-date intra-domain network state (available BW in each path) in edge-to-edge database and inter-domain routes in eRS's RIB, the resource reservation is very simple and straightforward. Upon receiving a resource reservation request (RAA), the BB first queries the eRS's RIB for possible next BB, which supports the given QoS requirements, and the associated egress router. It then queries the edge-to-edge QoS database for intra-domain resource availability. If both of the queries succeed, it sends RAR to next BB (see 3, 6 and 7 for details). It does not have to check individual links in its domain. Also, since the route in RIB has the cost associated with it, it chooses the least costly one among multiple possible routes.

## **3. CONCLUSION**

In this paper we have presented a Route Server (RS) architecture that is compatible with the BB model for QoS routing. We addressed the role of RS in solving complex inter-domain routing in a scalable and efficient fashion. In particular, we first showed that how a Diffserv domain can provide quantitative QoS guarantees across its cloud and how a dynamic network state can be maintained in a scalable way. We also presented an inter-domain scheme that provides QoS route information to a possible destination region by taking inter-domain policies into account. Finally, we briefly describe the effectiveness of RS in reservation state. In our future work we integrate RS with our existing inter-domain BB implementation.

### **References:**

- [1] K. Nichols et al. "A Two-bit Differentiated Services Architecture for the Internet" rfc2638
- [2] K.Nichols, B. Carpenter "Definition of Differentiated Services Per Domain Behaviors" RFC 3086
- [3] QBone Signaling Design Team, "Simple Inter-domain Bandwidth Broker Signaling (SIBBS)", <http://qbone.internet2.edu/bb/>
- [4] G. Apostolopoulos et al. "Server Based QoS Routing", Globecom99, Rio de Janeiro, Brazil
- [5] P. Aukia et al. "RATES: A server for MPLS Traffic Engineering", IEEE network magazine, 2000
- [6] H. Mantar, et al. "Inter-domain Resource Reservation via Third-Party Agent", SC 2001
- [7] I. Okumus, J.Hwang, H. Mantar, S.Chapin, "Inter-domain LSP Setup Using Bandwidth Management Points", Globalcom 2001
- [8] P. Auki et al. "RATES: A server for MPLS Traffic Engineering", IEEE network magazine, 2000
- [9] P. Trimintzios et al. "An Architecture for Providing QoS in Differentiated Services" IM2001