2001

Inter-Domain LSP Setup Using Bandwidth Management Points

Ibrahim Taner Okumus
Syracuse University

Junseok Hwang
Syracuse University

Haci Ali Mantar
Syracuse University

Steve J. Chapin
Syracuse University, chapin@syr.edu

Inter-Domain LSP Setup Using Bandwidth Management Points

Ibrahim Taner Okumus,Junseok Hwang, Haci Ali Mantar,Steve J. Chapin,
Syracuse University

Abstract— **Bandwidth Management Points (BMP) are a necessity to manage the intra and inter-domain resources in the Internet. In this paper we propose a way to setup inter-domain Label Switched Path (LSP) with the help of a BMP in a MultiProtocol Label Switching (MPLS) over Diffserv network. We use extended Simple Inter-domain Bandwidth Broker Signalling Protocol (SIBBS) to distribute the labels inter-domain. We also use a BMP to interact with the MPLS to setup the intra-domain LSP and to provision the intra-domain traffic. With the help of a BMP, we show how end-to-end Quality of Service (QoS) can be achieved.**

I. Introduction

Recent developments in the Internet technology opened new horizons for the future of the Internet. Long awaited QoS-enabled technologies are on its way. Initial QoS research in internetworking focused on Resource Reservation Protocol (RSVP) and Intserv. It was seen that scalability problems prevent these technologies from broadly answering the needs of the current Internet. This led to the development of Differentiated Services (Diffserv), which eschews per-flow QoS reservation in favor of simplicity in routing through the network core. Diffserv is not as strong as RSVP-Intserv in terms of QoS but it is more scalable and simple to implement in a domain. Another technology is MultiProtocol Label Switching (MPLS). MPLS does not have strong QoS features but it is very useful for Traffic Engineering (TE) purposes and brings the fast forwarding to the backbones by deploying IP Switching.

To have robust QoS features and control over the domain, we need to implement MPLS technology over Diffserv domain. Using MPLS inside a domain is well defined and MPLS technology is being studied intensively [1], [2], [3], [4]. But there was not any focus on the use of MPLS technology inter-domain . If we want to have end-to-end QoS in an MPLS network we need to solve the inter-domain Label Switched Path (LSP) setup problem. In this paper we will answer these issues on Diffserv over MPLS networks by using BMPs[5], [6].

There are different approaches for interdomain label distribution. One of the approach is to use Border Gateway Protocol (BGP) to distribute the label between peer edge LSRs [7]. In this approach authors propose to piggyback the label information with the BGP update message. This is done by using the BGP-4 Multiprotocol Extensions attribute [9]. Another document gives a framework for interarea LSP setup [10]. The focus of this document is on setting up LSPs across Interior Gateway Protocol (IGP) areas. IGP areas are the areas separated to make the autonomous system more manageable. Inter-area approach tries to set up an LSP across areas without prior knowledge of the resources available across the area boundaries.

We are proposing a different approach to setup inter-domain LSP. In order to connect LSPs in different domains we are proposing to use a specialized Bandwidth Broker agent, called Bandwidth Management Point (BMP). BMPs acting as Bandwidth Broker agents in every autonomous system (AS) manage the resources in the domain and communicate with peer BMPs to manage the inter-domain resources. General structure of the BMP will be in conformance with the QBone Bandwidth Broker Architecture. We will also use the SIBBS inter-domain signalling protocol that is developed by QBone Signalling Design Team[11].

Section 2 describes the system architecture. Section 3 gives a detailed explanation of intra-domain LSP set up and our approach to establish inter-domain LSP and inter-domain label distribution. Section 4 contains concluding remarks.

II. Diffserv Over Mpls With BMP Architecture

Differentiated Service Network Model: Differentiated Service (Diffserv) is the IP QoS network service architecture proposed in RFC 2475 [12]. In DiffServ, a 6-bit field known as the Differentiated Service Code Point (DSCP) field of the TOS (Type Of Service) is proposed for use to specify QoS service semantics. This approach realizes scalability by aggregating traffic classification states into a finite number of classes which can be defined from the DSCP field. The assigned class of the traffic at the edge of the network will define the per-domain-behavior (PDB) and associated per-hop-behavior (PHB), which in turn will define how to treat each packet for each router on the route. Standardized classes are: default forwarding (DF), expedited forwarding (EF)[13], and assured forwarding (AF)[14]. The DF PHB is the most common best-effort forwarding behavior which is available in the current Internet. The AF PHB group provides delivery of IP packets in four independent forwarding classes and assigns three different levels of drop preferences in each forwarding class. The EF PHB can be used for the services requiring a low loss, low latency, low jitter, and assured bandwidth within DiffServ network.

The key elements of the proposed DiffServ network architecture include DiffServ boundary nodes, DiffServ interior nodes, and bandwidth management points (BMP) [6],[8].

Bandwidth Management Point System Architecture: A Bandwidth Management Point (BMP) is the key element of our system architecture. Following are the building blocks of a Bandwidth Management Point. The BMP allocates and controls the bandwidth share between different interconnecting networks and serving DSCP service classes.

- *The Bandwidth Management Point* (BMP) calculates current service demands, available network resources, and their values. Using such measures, the BMP makes a decision or sets a policy for admission control, network resource provisioning, SLS (Service Level Specification) configuration, and bandwidth exchange. Based on the decision, the BMP can configure inter- and intra-domain paths with the interaction of the routing mechanism.

- *The Bandwidth Measurement Base* (BMB) is a MIB (Management Information Base) which measures the traffic load and demand statistics of the BMP domain. A Management Creation Point (MCP) relates BMB with the SLAs (Service Level Agreements) of interconnections.
- *Interior-BMP* (I-BMP) is a processor which provides a mechanism for managing the network resources within a BMP domain. Network provisioning among different service classes and different users within a BMP domain using decision information is one of the key tasks of the I-BMP. Processing local user requests for end-to-end services, and computing and allocating network resources while satisfying user requirements at the lowest cost possible are the tasks of the I-BMP.
- *Exterior-BMP* (E-BMP) is a processor which provides a mechanism to manage the interconnection SLS with other BMP domains and interconnection points. Therefore, it is concerned with resource allocation and provisioning at the network boundary among multiple domains. This process should solve a complex set of brokerage problems among the alternative options. SLS requests, configurations, and updates with multiple interconnection points should be handled by this E-BMP.
- *Service-Provisioning-Points* (S-P-P) are network nodes and interfaces which the I-BMP and E-BMP can configure using their own management information and decisions. The ingress and egress border routers of its own network or outsourced network capacity from other interconnecting network domains could be such provisioning points. These configurations include the capacity constraint and queue depth adaptation parameters for each class.
- *SLS-Enforcement-Points* (S-E-P) are network nodes and interfaces involved in inter-domain resource allocation (admission) based on the SLS. The I-BMP and E-BMP configure these points and have them perform the admission tasks (where SLS is enforced).
- *User-BMP-Interconnection-Points* (UBIP) are points at which user interconnection-access networks (such as those in the IntServ) subscribe the aggregation network with BMP (such as those in DiffServs using BMP). Individual service flows are generated and terminated by the end hosts connected to the user network. User signaling such as RSVP will communicate with S-E-P or directly with BMP for interconnection subscription management. Large user networks may employ the BMPs for resource management and aggregation purposes, especially those involved in multiple subscriptions to multiple networks.

MultiProtocol Label Switching: MPLS is a technology that combines the best features of the layer 2 switching and layer 3 routing to increase the performance and efficiency of the network (see [15]). An MPLS capable router is known as Label Switching Router (LSR).

MPLS uses labels to switch packets. Labels are short fixed length identifiers that are used during forwarding process. Every packet is assigned a label and the LSR decides the forwarding path looking at the packet's label. Labels are distributed via Label Distribution Protocol (LDP). Labels are associated with forwarding equivalency classes (FEC) and are local to that particular link. Peer LSRs distribute labels to each other. LSRs decide the next hop by looking up the *(label, incoming port)* pair in the forwarding table, and extracting a corresponding *(label, outgoing port)* pair. The path that LSP takes can be either decided by using conventional routing algorithms or can be defined explicitly. One of the strongest feature of MPLS is its traffic engineering capability. With this capability, domain administrators can easily manage the resources inside the domain by setting up various LSPs and can offer quality of service to the customers.

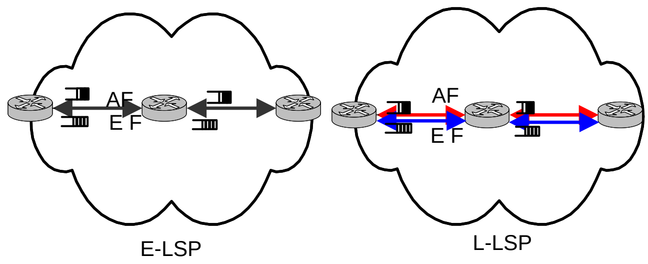

Fig. 1. E-LSP: Packets use same LSP. LSR puts them in different queues. L-LSP: Packets use different LSPs and different queues.

Explicit LSPs are also known as LSP tunnels [16]. Label distribution protocol, such as CR-LDP, RSVP-TE, associates QoS features with the tunnel.

III. LSP Tunnel Setup In Diffserv Domain

In our model an end-to-end QoS path is a concatenation of separate LSPs. Setting up a label switched path (LSP) tunnel in a domain is well defined. An LSP is a path defined by two end points, the LSRs in between, and the labels associated with that path. In order to support traffic engineering in a domain, conventional routing protocols are extended [17], [18]. These extended protocols have ability to carry QoS constraints and determine a path that conforms with these constraints. An Ingress LSR (entry point to an MPLS network) uses this information to determine the path that LSP should take. Once the path is decided, the ingress LSR initiates label distribution if there isn't already one setup for the same destination. All the LSRs on the path assigns a label for the LSP. Once the labels are assigned, LSP setup is complete and LSP is ready to use.

In our architecture, BMPs are the entities that are responsible for the resource management within a domain (and between domains through cooperation with other BMPs). When a reservation request comes to the ingress LSR it notifies the BMP. In our model, LSP should be associated with a Diffserv QoS classification.

A. Label - PHB Scheduling Class (PSC) Match

The basic LSP types for Diffserv domain are suggested in [19]. There are two different types of LSPs that can be used in a Diffserv domain. These LSPs are EXP-Inferred-PSC LSPs (E-LSP) and Label-Only-Inferred-PSC LSPs. An MPLS shim header contains a 3 bit EXP field, reserved for experimental use. In a QoS context this field is used to determine the QoS features that the label should be treated with. Within a domain, an administrator can use combination of these methods to meet the QoS demands. E-LSP can support up to eight Behavior Aggregates (BA). Each BA can span multiple Ordered Aggregate (OA). This means that one LSP can support multiple different Per Hop Behaviors. In this case determination of the PHB to be applied depends on the EXP field of the label.

Using L-LSP, separate LSPs can be established for a single *(FEC,OA)* pair[19]. PSC information is signalled during LSP establishment. This means a specific label is bound to that LSP and each LSR on the path knows exactly what kind of treatment that LSP should get. In this case label has the information of what PSC that LSP has and EXP field is used to determine the drop precedence. Every LSR keeps the DSCP-LSP mappings. Ingress LSR looks at the DSCP of the packet and puts it into the LSP that has been associated with that specific QoS level. Intermediate LSRs check the label to decide the QoS treatment that a packet gets. Egress LSR pops the label and forwards the packet with its original DSCP.

B. Simple Interdomain Bandwidth Broker Signalling

For the interaction between BMPs and domains we used Simple Interdomain Bandwidth Broker Signalling Protocol (SIBBS). SIBBS is under development by QBone Signalling Design Team [20]. SIBBS only specifies inter-domain signalling protocol. The Bandwidth Management Point in a domain can receive Resource Allocation Requests (RAR) from three different sources. One is the host in the domain that BMP controls, and the other is the peer BMP, and the last is a third-party agent acting on behalf of a host or application. The BMP responds with a Resource Allocation Answer (RAA) to the request. The request may have certain side effects also, such as altering the router configurations at the access, at the inter-domain borders, and/or internally within the domain, and possibly generating additional RAR messages requesting downstream resources[20]. For security reasons every BMP authenticates the messages it receives from other BMPs and signs the messages it sends to other BMPs, which was suggested as an important issue in [21].Fig. 2 is the simple pseudo-code for the extended SIBBS.

C. Intra-Domain LSP Setup In A Diffserv Environment Using BMP

Following is an example scenario of setting up an LSP in a Diffserv domain with the help of a BMP (Fig. 3). Suppose the host at AS1 wants a QoS path with the server at AS4.The host generates a reservation request to BMP1. BMP1 first have to check whether the requesting entity have right to ask for the request. BMP1 also checks whether the request conforms with the SLA between the requester and the service provider. BMP1 verifies that there is enough resources to support the requested traffic. Traffic Engineering extended routing protocol gives the possible paths and the egress points to the destination considering given QoS constraints. BMP1 is the responsible entity to decide which path to use and which egress point to use. This information is used to setup an LSP between the host and the egress point. LSP is setup by ingress router by distributing labels for the specified flow along the predetermined path. When the LSP setup is complete, BMP1 sends back a resource allocation answer to the host.

```
Originating BMP:
get RAR;
IF (authentication and Resources and SLA conformation
   and Policy conformation)
THEN egress-router = egress router;
   Path=(Org-router,...,egress-router);
      IF( Label-Insert )THEN
      Label = Request-Label;
      ELSE;
   RAR = RAR-swap ( BMP-ID, BMP-Signature);
   Forward RAR; Wait RAA;
ELSE Return RAA-Insert( Reason Code);

Transit BMP:
get RAR;
IF (authentication and Resources and SLA conformation
   and Policy conformation)
THEN egress-router = egress router;
   Path=(ingress-router,.., egress-router);
      IF( Label-Insert ) THEN
      Label = Request-Label;
   ELSE;
   RAR = RAR-Swap ( BMP-ID, BMP-Signature);
   Forward RAR;Wait RAA;
ELSE Return RAA-Insert( Reason Code);

Destination Domain:
IF (authentication and Resources and SLA conformation
   and Policy conformation)
THEN egress-router = dest-router;
   Path=(ingress-router,.., dest-router);
      IF( Label-Insert ) THEN
      Label = Request-Label;
   ELSE;
   Forward RAR; // to the end-system;
ELSE Return RAA-Insert( Reason Code);
RAA Processing:

IF( RAA) THEN
   IF(Label-flag) THEN
      RAA = RAA-Insert( Label);
   ELSE;
   RAA = RAA-Insert( BMP-ID, BMP-Signature);
   Allocate Resources();
ELSE
   return
   RAA-Insert(BMP-ID,BMP-Sign,Reason Code);
```

Fig. 2. Modified SIBBS

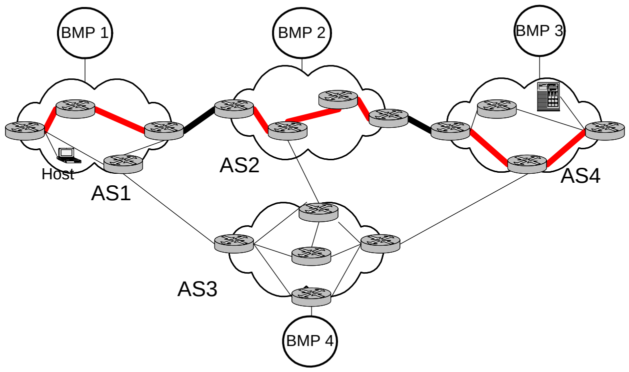

Fig. 3. End to End LSP Setup

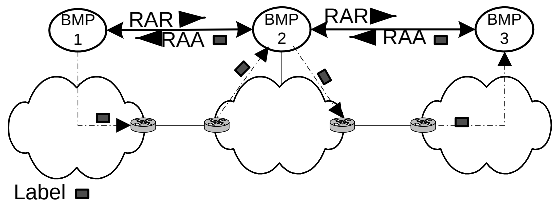

Fig. 4. BMP Signalling and Label distribution

D. Inter-Domain LSP Setup In A Diffserv Environment Using BMP

We use an extended QBone SIBBS signalling protocol for our model. RARs travel downstream and RAAs travel upstream. In MPLS LSP setup, labels also distributed upstream. There is a logical match between these two mechanisms. If we think of ASs as nodes, labels should be distributed from downstream ASs to upstream ASs. We propose to use the extended SIBBS for the inter-domain label distribution (Fig.4). As an extension, we propose to add optional *label-insert* flags and *label* object to the protocol.

There are two different cases that require LSP setup process to run on the inter-domain. In the first case, two domains connect for the first time and set up the initial LSPs. In the second case, domains decide whether they accept a new flow into the tunnel and if necessary increase the tunnel capacity. Second case is directly related to the dynamic provisioning and in this case it is not necessary to request a label from the downstream AS [22]. Here, we are considering the initial LSP setup case.

In order for the BMP1 to send the Resource Allocation Request to the downstream BMP, BMP1 has to verify that LSP setup is possible in its domain. Then BMP1 sends the RAR to BMP2. BMP1 sets Label Request Flag in the RAR if there are no established LSP tunnel from egress to the ingress router or if the request can not be aggregated with other flows.Transit BMP gets the RAR, performs intra-domain LSP setup procedures. If LSP setup fails the BMP sends back a negative RAA with a reason code. If result is positive, then it sends the RAR to the next domain.

The destination BMP gets the RAR, checks whether it is possible to reach the destination, and set up an LSP that supports the requested QoS. If the outcome is negative then the BMP sends back an RAA with a reason code. In case of a positive outcome, it forwards the request to the destination host. The destination host performs the routine checks specified in the SIBBS and if the outcome is negative, it sends back a negative RAA to the BMP. If the outcome is positive, it sends back a positive RAA to the BMP. If the label request flag is set, the BMP asks the specified ingress router to assign a label for that flow, puts that label into the RAA and sets Label Insert Flag, and then sends RAA back to the upstream BMP.

A transit BMP gets the RAA. If Label Insert Flag is set, the BMP extracts the label from the RAA and sends the label to the assigned egress. The BMP asks for a new label from the ingress router, then inserts that label in place of the extracted label. The transit BMP does not need to modify the Label Insert Flag. Transit BMP sends RAA back to the upstream BMP.

The origin BMP (BMP1) gets the RAA. If the Label Insert Flag is set, extracts the label information from the RAA, and informs egress router about the label. BMP1 modifies the RAA which includes taking out the label from the RAA and setting Label Insert Flag to 0 and forwards the RAA to the requesting end-system.

When the host gets a positive RAA, this means that all the ASs on the path to the destination setup LSPs to support the requested QoS parameters. All those intra-domain LSPs connected to each other with inter-domain LSPs. As a result of this, the host has an end-to-end QoS path to the destination.

IV. Conclusion

One of the advantages of using MPLS and Diffserv together is the scalability of the approach. Diffserv has a potential 64 different classes. Currently only 14 of them are defined. This means that this is the maximum number of behaviors we must support between domains. Since all the flows with the same QoS class gets the same treatment we can easily aggregate them at the egress router. Another advantage of using MPLS is that it is easy to identify a flow by looking at the label. At the egress point, the LSR pushes one label for the flow and one for the tunnel. When the flow comes out of the tunnel at the ingress of the other domain, the ingress LSR pops the first label and swaps the second label. Another advantage of using MPLS with Diffserv is that we eliminate IP lookups at border routers. We are planning to use our BMP architecture to test the under-development SIBBS protocol and show the benefits of using BMP in Diffserv, MPLS and MPLS over Diffserv environments.

We are currently working on the implementation of the mechanisms (BMP, SIBBS, etc.) to simulate our architecture and approach.There are still many open issues in this area. To match the QoS features of LSPs between peer ASs we need a mechanism to negotiate the QoS parameters [23]. Also another open issue we are planning to solve in future works is interdomain traffic engineering in MPLS and Diffserv networks.

By using MPLS-Diffserv-BMP altogether the following problems are solved:

- Inter-domain LSP setup
- Inter-domain flow management for Diffserv
- Fast forwarding at border routers
- Inter-Domain TE

References

[1] Anoop Ghanwani, Bilel Jamoussi,Don Fedyk,Peter Ashwood-Smith,Li Li, Nancy Feldman,"Traffic Engineering Standards in IP Networks Using MPLS" , IEEE Communications, Dec 1999.

[2] Yoshihiro Ohba,"Issues on Loop Prevention in MPLS Networks",IEEE Communications, Dec 1999.

[3] Thomas M. Chen,Tae H. Oh,"Reliable Services in MPLS", IEEE Communications, Dec 1999.

[4] Daniel O. Awduche,"MPLS and Traffic Engineering in IP Networks",IEEE Communications, Dec 1999.

[5] K. Nichols,V. Jacobson,L. Zhang, "A Two-bit Differentiated Services Architecture for the Internet",Internet draft,Nov 1997.

[6] Junseok Hwang,"A Market-Based Model for the Bandwidth Management of Intserv-Diffserv QoS Interconnection: A Network Economic Approach",Ph.D. Thesis,University of Pitssburg,2000.

[7] Yakov Rekhter,Eric C. Rosen, "Carrying Label Informartion in BGP-4", Internet draft, January 2001, work in progress.

[8] Junseok Hwang, Hak-Jin Kim, Martin Weiss,"Interprovider Differentiated Service Interconnection Management Models in the Internet Bandwidth Commodity Markets", Telematics and Informatics, Special Issues on Electronic Commerce,2001.

[9] T. Bates,R. Chandra,D. Katz, Y. Rekhter,"Multi-Protocol Extensions for BGP-4", RFC2283.

[10] S. Venkatachalam,S. Dharanikota, "A Framework for the LSP Setup Across IGP Areas for MPLS Traffic Engineering", Internet Draft, November 2000

[11] QBone Signalling Design Team, http://qbone.internet2.edu/bb/index.shtml.

[12] S. Blake et al. "An Architecture for Differentiated Services", RFC2475, December 1998.

[13] V. Jacobson,K. Nichols,K. Poduri, "An Expedited Forwarding PHB", RFC 2598, June 1999.

[14] J. Heinanen,F. Baker,W. Weiss,J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, June 1999.

[15] E. Rosen,A. Viswanathan,R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[16] George Swallow,"MPLS Advantages for Traffic Engineering",IEEE Communications Magazine, Dec 1999.

[17] Tony Li,Henk Smit, "IS-IS extensions for Traffic Engineering", Internet Draft, September 2000, work in progress.

[18] Dave Katz,Derek Young,Kireeti Kompella,"Traffic Engineering Extensions to OSPF", Internet Draft, work in progress.

[19] Francois Le Faucheur et al. "MPLS Support of Differentiated Services", Internet Draft, February 2001, work in progress.

[20] P. Chimento et al. QBone Signalling Design Team, "QBone Bandwidth Broker Architecture", Work in progress http://sss.advanced.org/bb/bboutline2.html.

[21] Junseok Hwang, Joon-soo Park,"Agent-based Secure Bandwidth Transaction Service Mechanisms using RBAC(Role-Based Access Control) Models",International Conference on Telecommunications System, March 2001.

[22] Haci Ali Mantar, Junseok Hwang, Ibrahim T. Okumus,Steve J. Chapin,"Inter-Domain Resouce Reservation via Third Party Agent",to appear on Fifth World Multi-Conference on Systemics, Cybernetics and Informatics 2001,Jun 2001.

[23] Junseok Hwang, "Cross-Network Open Provisioning Intelligent Network (COPIN) for Bandwidth Transaction Services in the Next Generation Internet",IEEE Intelligent Network Workshop 2001 (IN2001), May 2001.