A Managed Bandwidth Service for IP Networks

C. Bouras C. Chantzi V. Kapoulas A. Sevasti

Computer Engineering and Informatics Dept., University of Patras, 26500 Rion, Greece, and Research Academic Computer Technology Institute, Riga Feraiou 61, 26221 Patras, Greece.
E-mail : bouras@cti.gr

Abstract

The work presented here attempts to face the problem of bandwidth management in IP networks, especially in cases where the available resources have to be shared among many users with demanding and conflicting needs. We have designed a Managed Bandwidth Service (MBS) that allows users to make bandwidth reservations throughout the topology of the network infrastructure to which they have access based on the establishment of Virtual Private Networks (VPNs) using the Multiprotocol Label Switching (MPLS) technology. In this paper we present the MBS developed along with our on-going and future work on the subject.

Keywords: Managed Bandwidth Service, Multiprotocol Label Switching, Virtual Private Networks, Traffic Engineering, Quality of Service

1. Introduction

The past several years have witnessed dramatic growth in demand for bandwidth and IP services. Service providers (SPs) face the continuing challenge of how best to design a networking infrastructure that is flexible enough to address the ever growing customer demands for increased bandwidth and expanding network access. Customers are clamoring for streamlined access methods that allow quick provisioning of bandwidth and services. SPs require tools to control and manage their networks in a variety of ways without overhauling their existing infrastructure, while at the same time fulfilling the customer demands.

Therefore, in the last few years, great attention has been given to the design and implementation of mechanisms for network bandwidth management. In cases there appears lack of resources, the service provided does not meet the desired, promised or expected quality levels resulting, thus, in underlying networks exhibiting denial of service, long delays and facing several bandwidth availability problems. It is becoming increasingly vital to ensure high-quality network services for bandwidth critical applications which require support for the transmission of voice, video and other kinds of multimedia data while at the same time, lack or ineffective allocation of resources results in limitations to the number of users that can be simultaneously served, regardless of whether they are running bandwidth critical applications over the underlying network or not.

All the existing Managed Bandwidth Services that have been developed in order to deal with the above problem are implemented using ATM technology. Until recently, ATM switches enjoyed a capacity and interface throughput advantage over routers, leading to widespread adoption of ATM switches at the core of service provider networks. Some of the largest IP networks are based on a Layer 2 switched core. This design – called the overlay model – allows the provider to achieve virtual connectivity across physical backbone links - beneficial in terms of flexibility and traffic engineering. However, like the Frame Relay switches before them, capacity-constrained ATM switches are being steadily removed from the core of most IP networks and replaced with core IP routers. The reason is speed. In this case an end-to-end MBS service using ATM technology cannot be supported in many locations. However, there is still a need for provisioning end-to-end quality of service guarantees over the new network infrastructures because the basic requirement of users is to carry IP traffic with a bandwidth guarantee.

The rest of the paper is organized as follows: In section 2 describes our MBS from an operational point of view, while section 3 deals with the implementation issues of the service. Finally, we describe our on-going and future work on the service, as well as our conclusions with respect to the usability and usefulness of the MBS, in section 4 and 5 respectively.

2. Description and Operation of the MBS Service

The implementation of the MBS service will be based on the establishment of VPNs between end users across the MPLS network. It will allow the definition of MPLS VPNs for connecting the participant sites and providing them with the appropriate network resources taking into account the requirements in bandwidth, duration of the established VPNs, traffic profile and a complete set

of network parameters. The MPLS VPNs will be customisable with management capabilities and performance properties comparable to a dedicated physical network.

Bandwidth reservation will be implemented via the creation of virtual private networks with MPLS techniques between two end users of the network.

More specifically, the user of the network will be able to make a request for the establishment of an MPLS VPN between his site and another site connected to a network node. The service must provide the user with a user-friendly graphical interface for making requests and receiving the system's responses, and the manager with a simple interface for responding to user's requests, presenting and monitoring the network status, and configuring and managing MPLS VPN properties such as the virtual topology, bandwidth requirements of virtual links and VPN membership information.

Before making his request, the user will have to be authenticated to use the service via an appropriate authentication mechanism that must be provided by the service within the user's interface. This way, any non-authenticated users will be prevented from using the service. Hence, in order to make use of the service, the user will have to make a request via the appropriate interface defining the desired features of the requested VPN including the sites participating in the VPN (members of a VPN are described by the member end host's IP addresses and/or the member subnets' network prefixes), bandwidth requirements, and period of time that the VPN will be available to the user. Once the request is submitted, this information will be sent to the network manager who will establish the specified VPN.

After specifying the VPN description, the service manager will submit the request of setting up this VPN. Subsequently, the service application will send the appropriate setup messages to a software application that will run on the router that connects the backbone network to the site of the user that has requested the VPN establishment. The job of this application is to act as a proxy for the setup messages between the application of the MBS service and the routers in the network where the MBS will be deployed. This enables the service application to be executed remotely from anywhere in the Internet.

Continuously, each VPN request will be processed in network routers and if the network can bear the load of the new request for the requested time period, the network will determine the physical path that the user's traffic must follow, will reserve the respective network resources via the signalling mechanism (RSVP-TE) and will establish the VPN through the appropriate procedures. Once the process is terminated, a notification message will be sent to the user and manager of the service. Then, the manager will have to update the service database in order to reflect the current state of the network.

3. Implementation issues

3.1. Admission Control and Path Selection

As it has been mentioned before, the request data concerning the establishment of the VPN will be sent to the ingress router that connects the user's equipment to the backbone network. Then, it will be determined whether the network can support the specified request taking into account the requested features. Continuously, the path that the user's traffic must traverse will be selected and established using a signalling mechanism, which will make the bandwidth reservations on every link of the path.

All these procedures require detailed knowledge about the network topology as well as dynamic information about network loading. Thus, a primary requirement is the support of a framework for information distribution. This component can easily be implemented by defining relatively simple extensions to the IGP so that link attributes are included as part of each router's link-state advertisement. IS-IS extensions can be supported by the definition of new Type Length Values (TLVs), while OSPF extensions can be implemented with Opaque LSAs (a new class of Link-State Advertisements – LSAs). Each LSR (Label Switched Router) will maintain network link attributes and topology information in a specialized Traffic Engineering Database (TED). The TED will be used exclusively for calculating explicit paths for the placement of LSPs across the physical topology. A separate database will be maintained so that the subsequent traffic engineering computation is independent of the IGP and the IGP's link-state database (Figure 1).

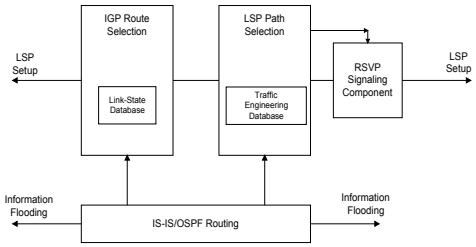

Figure 1. LSR Block Diagram [3]

After network link attributes and topology information are flooded by the IGP and placed in the TED, each ingress LSR uses the TED to calculate the necessary path. The path for each LSP will be represented by an explicit route, which is defined as a preconfigured sequence of LSRs that should be part of the physical path of the LSP.

The ingress LSR determines the physical path for each LSP by applying a Constrained Shortest Path First (CSPF) algorithm to the information in the TED. CSPF is a shortest-path-first algorithm that has been modified to take into account specific

restrictions when calculating the shortest path across the network.

As CSPF considers each candidate node and link for a new LSP, it either accepts or rejects a specific path component based on resource availability. The output of the CSPF calculation is an explicit route consisting of a sequence of LSR addresses that provides the shortest path through the network that meets the constraints. This explicit route is then passed to the signalling component, which establishes forwarding state in the LSRs along the LSP. The CSPF algorithm is repeated for each LSP that the ingress LSR is required to generate.

3.2. LSP Establishment

After the path calculation, the respective LSP must be established. In order for an LSP to be setup, labels are negotiated and distributed through signalling messages that LSRs use. Since the overhead of manually configuring an LSP is very high, most service providers will want to automate the process by using a signalling protocol. The signalling protocol will distribute labels and establishes LSP forwarding state in the network nodes selected by the path calculation process that was mentioned above.

The signalling protocol that establishes LSP state across the network plays an important role in automating the whole process. In the design of the MBS service, the signalling component that is proposed to use is RSVP-TE (RSVP extensions for traffic engineering). It provides the mechanism to setup an explicitly routed LSP that differs from the normal path calculated by the IGP and performs downstream label allocation, distribution, and binding on demand among LSRs in the path, thus establishing path state in network nodes. RSVP-TE allows a service provider to dynamically establish LSPs across their network, making the network more responsive to changing conditions while saving time and reducing operating expenses. In addition, it is ideal for use as a signalling protocol to establish LSPs because its soft state can reliably establish and maintain LSPs in an MPLS environment, it allows network resources to be explicitly reserved and allocated to a given LSP, and supports the establishment of explicitly routed LSPs that provide load-balancing capabilities equivalent to those currently provided by ATM and Frame Relay.

Hence, the network will perform dynamic online path calculation for the LSP. The service manager will configure the constraints for each LSP and then the network itself will determine the path that best meets these constraints. Continuously, the forwarding state will be installed across the network using the signalling capabilities of RSVP-TE [4], and mainly the following new objects:

- The *Explicit Route Object* allows an RSVP PATH message to traverse an explicit sequence of LSRs that is independent of conventional shortest-path IP routing.
- The *Label Request Object* permits the RSVP PATH message to request that intermediate LSRs provide a label binding for the LSP that it is establishing.
- The *Label Object* allows RSVP to support the distribution of labels without having to change its existing mechanisms. Because the RSVP RESV message follows the reverse path of the RSVP PATH message, the Label Object supports the distribution of labels from downstream nodes to upstream nodes.

Once the LSP is setup, the desired requested bandwidth will then be available end-to-end on the explicit route for the user's traffic. However, it might be necessary to reroute the LSP because of the failure of a link or router along the LSP's path that generally requires that the LSP be rerouted.

It is extremely important that the flow of traffic is not disrupted when an LSP is rerouted. A smooth transition requires support for a concept called *make before break* – the new LSP must be established and the traffic transferred to it before the old LSP is torn down. One of the benefits of RSVP signalling is that it provides an elegant solution to this challenging problem.

3.3. MPLS VPN Establishment

The previously mentioned LSP must be configured on a per-VPN basis. This LSP concerns a bandwidth reservation that the user requests and it is used to forward user's data through the established VPN. In fact, this LSP is a private LSP and its establishment is of high importance in the MPLS VPN architecture. Hence, after the establishment of the appropriate LSP, a VPN has to be established between the users that to communicate.

During the VPN setup the most important configuration step is to provide edge routers with VPN membership information and the globally unique VPN identifier that was chosen by the service application. This information is needed so edge routers can inject packets appropriately into the VPN. The second VPN-specific step is to establish security associations between the edge routers. The security associations are used to provide authentication and encryption of data that travels over the VPN.

Using the membership information, the ingress edge router can correctly identify packets that belong to a VPN. It then injects the packet in the appropriate LSP and tags the packet with the globally unique VPN id, which is added to the encapsulating header at the ingress edge router to differentiate between packets so as to enable per-VPN forwarding and resource management.

In the MPLS VPN architecture [1], [5] the customer sites are linked with the provider network

via CE-routers. The CE-routers are connected to the PE-routers (Provider Edge routers), which serve as the edge devices of the provider network. The core devices in the provider network (P-routers) provide the transport across the provider backbone (Figure 2).

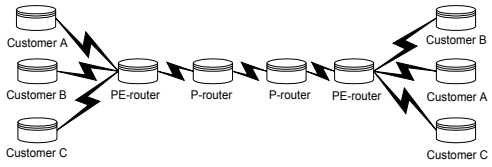

Figure 2. An MPLS VPN network

The design decisions that must be made in order for the network to be able to support the creation of MPLS VPNs concern the selection of the routing protocols via which the MPLS VPN is established. These protocols include the routing protocol between the PE-routers, the routing protocol between the PE- and CE-routers, and the backbone IGP.

In order to enable data transfer between sites attached to different PE-routers, the relevant routing information needs to be exchanged between PE-routers. Therefore, it is needed a routing protocol that will transport all customer routes across the provider network while maintaining the independency of individual customer address spaces. The best approach is to run a single routing protocol between PE-routers that will exchange customer routes without the involvement of the P-routers. This solution is scalable as the number of routing protocols running between PE-routers does not increase with increasing number of customers and P-routers do not carry customer routes. The protocol that is most suitable for the exchange of customer routes between PE-routers is Border Gateway Protocol (BGP).

However, with the deployment of a single routing protocol (BGP) exchanging all customer routes between PE-routers, an important issue arises – *how can BGP propagate several identical prefixes, belonging to different customers, between PE-routers*. The answer to this question is the expansion of customer IP prefixes with a unique prefix (called Route Distinguisher) that will make them unique (called VPN addresses) even if they were previously overlapping. Therefore, BGP between PE-routers must support exchange of traditional IP prefixes as well as VPN prefixes. Consequently, Multi-Protocol BGP (MP-BGP) is used in the MPLS VPN backbone to carry VPN routes between PE-routers.

In addition, CE-routers need to communicate with PE-routers to which they are connected. The routing protocols that are supported for PE-CE routing information exchange include static routing (recommended for simple VPN sites), RIP (for sites where there are more subnets per site or where the service provider does not manage the CE), BGP and OSPF (should only be used for extremely large VPN customers).

In conclusion, the MPLS VPN architecture will use BGP in two different ways:

- Multiprotocol BGP can be used between the PE-routers for the propagation of VPN routes across the MPLS VPN backbone.
- BGP can be used as the PE-CE routing protocol to exchange VPN routes between the PE- and CE-routers.

3.4. Packet Forwarding

Provided all the necessary setup procedures are completed, the network will start to forward user's packets. The packet-forwarding component in the backbone network is MPLS. MPLS is responsible for directing a flow of IP packets along a predetermined path across the network, i.e. across the LSP. When the ingress LSR receives an IP packet, it adds an MPLS header to the packet and forwards it to the next LSR in the LSP. The labelled packet is forwarded along the LSP by each LSR until it reaches the egress LSR where the MPLS header is removed and the packet is forwarded based on the IP destination address.

The packet-forwarding process at each LSR is based on the concept of label swapping. Each MPLS packet carries a 4-byte encapsulation header that contains a 20-bit fixed-length label field. When a packet containing a label arrives at an LSR, the LSR examines the label and uses it as an index into its MPLS forwarding table. Each entry in the forwarding table contains an interface-inbound label pair that is mapped to a set of forwarding information that is applied to all packets arriving on the specific interface with the same inbound label.

When an IP packet traverses the MPLS backbone network it is processed at three points in the network: as the packet arrives at the ingress LSR of the MPLS backbone, as it is forwarded by each LSR along the LSP, and as it reaches the egress LSR of the MPLS backbone.

At the ingress of the backbone network, the IP header is examined by the ingress LSR. Based on this analysis, the packet is classified, assigned a label, and forwarded toward the next hop in the LSP. Thus, the packets belonging to the same VPN will be forwarded across the same LSP.

Once the packet begins to traverse the LSP, each LSR uses the label to make the forwarding decision. The incoming interface and label are used as lookup keys into the MPLS forwarding table. The old label is replaced with a new label, and the packet is forwarded to the next hop along the LSP. This process is repeated at each LSR in the LSP until the packet reaches the egress LSR.

When the packet arrives at the egress LSR, the label is removed and the packet exits the MPLS backbone. The packet is then forwarded based on the

destination IP address contained in the packet's original IP header according to the destination of the path.

More specifically, after the MPLS VPN setup packets will be forwarded as follows.

Once the users' routes are distributed across the MPLS backbone, all routers are ready to begin the packet forwarding. The customer traffic between CE-routers and PE-routers is always sent as pure IP packets, satisfying the requirement that the CE-routers run standard IP software and are not MPLS VPN-aware.

In a very simplistic approach to packet forwarding across MPLS VPN backbone, the PE-routers might just forward IP packets received from the CE-routers towards other PE-routers. However, this approach would fail as the P-routers have no knowledge of the customer routes and therefore cannot forward customer IP packets. A better approach is to use MPLS LSPs between the PE-routers and a label to determine the proper LSP to use.

Customer VPN packets will be forwarded across the MPLS VPN backbone encapsulated in a MPLS label stack composed of two labels:

- The top label in the stack is a label assigned by the RSVP-TE toward the egress PE-router
- The second label in the stack is the VPN label assigned by the egress PE-router and propagated to other PE-routers via the MP-BGP

The propagation of the VPN label is performed via the following steps:

1. Egress PE-routers assign a label to every VPN route received from attached CE-routers. This label is then used as the second label in the MPLS label stack by the ingress PE-routers when labeling VPN packets.
2. VPN labels assigned by the egress PE-routers are advertised to all other PE-routers together with their VPN prefix in MP-BGP updates.
3. The ingress PE-router has two labels associated with a remote VPN route – a label for BGP next hop assigned by the next-hop P-router via RSVP-TE as well as the label assigned by remote PE-router and propagated via MP-BGP update.

Hence, RSVP-TE propagates the label at the top of the stack (which is responsible for packet transfer in the backbone) and MP-BGP distributes the second-level label among the PE-routers. This label is recognized only by the egress PE-router that has originated it and would not be understood by any other router.

4. Future work

In the future, the service is going to be implemented based on the proposed design. By implementing the proposed mechanisms we intend to create a flexible and efficient managed bandwidth service that will provide the members of the network with the necessary capacity, according to their needs.

In addition, it is suggested that the service includes a charging scheme, in order to increase the network's capabilities and, thus, to provide the users with a better service. For example, the users could be charged according to the usage of the network (usage-based charging). The users that request for more bandwidth will be charged according to the requested bandwidth.

5. Conclusions

Nowadays, Internet Service Providers are constantly facing the challenge of managing their networks to support extremely rapid growth rates while also maintaining a reliable infrastructure for mission-critical applications. Multiprotocol Label Switching (MPLS) has emerged as the enabling technology for the new public network because it supports a number of applications such as traffic engineering and virtual private networks that can be used to provide users with services of guaranteed bandwidth. MPLS also extends the abilities of traditional routing with constrained based routes that can be used for introducing QoS service and better load balancing.

References

[1] "Introduction to Cisco MPLS VPN Technology", Cisco VPN Solutions Center: MPLS Solution Provisioning and Operations Guide, Chapter 1 http://www.cisco.com/univercd/cc/td/doc/product/rtrmgmt/vpnsc/mpls/1_2/prov_gd/vpn_ug1.pdf

[2] Blake S. et al., "An Architecture for Differentiated Services", RFC 2475, December 1998

[3] Chuck Semeria, "Traffic Engineering for the New Public Network", White Paper, Juniper Networks, 2000 http://www.urec.cnrs.fr/hd/MPLS/JUNIPER/Traffic_engineering_publis_net.pdf

[4] Chuck Semeria, "RSVP Signalling Extensions for MPLS Traffic Engineering", White Paper, Juniper Networks, 2000 http://www.juniper.net/techcenter/techpapers/200006.pdf

[5] E. Rosen, Y. Rekhter, "BGP/MPLS VPNs", RFC 2547, March 1999

[6] Eric C. Rosen et al., "BGP/MPLS VPNs", Internet draft (draft-ietf-ppvpn-rfc2547bis-00.txt), July 2001

[7] "MPLS – An Introduction to Multiprotocol Label Switching", White Paper, Nortel Networks http://www.nortelnetworks.com/corporate/technology/mpls/collateral/55053.25-04-01.pdf

[8] "Multiprotocol Label Switching Overview", Cisco IOS Switching Services Configuration Guide, XC-57-74 http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fswtch_c/swprt3/xcftagov.pdf