Available online at www.sciencedirect.com

ScienceDirect

Computer Networks 51 (2007) 835–852

Computer Networks

www.elsevier.com/locate/comnet

# Fast setup of end-to-end paths for bandwidth constrained applications in an IP/MPLS–ATM integrated environment

Sergio Sánchez-Lopez *, Xavier Masip-Bruin, Josep Sole-Pareta, Jordi Domingo-Pascual

*Advanced Broadband Communications Laboratory (LCABA), Universitat Politècnica de Catalunya,*
*Av. Victor Balaguer s/n, 08800 Vilanova i la Geltru-Barcelona, Catalunya, Spain*

Received 8 July 2005; received in revised form 7 January 2006; accepted 12 May 2006
Available online 24 July 2006

Responsible Editor: J.C. de Oliveira

---

## Abstract

Transport networks are currently being moved towards a model of high performance Internet Protocol/Multiprotocol Label Switching (IP/MPLS) routers interconnected through intelligent core networks. Currently, Asynchronous Transfer Mode (ATM) technology has been widely deployed in several network backbones along with the Private Network-to-Network Interface (PNNI) protocols as the control plane. In order to cope with the increasing Internet traffic demands in the current context, fast setup of end-to-end paths with the required Quality of Service (QoS) is necessary.

This paper analyzes the case of two IP/MPLS networks interconnected through an ATM core network, assuming MPLS as the mechanism to provide Traffic Engineering in the IP networks, and a PNNI-based control plane in the core network. This paper aims to define a mechanism needed to set up a fast end-to-end QoS Label Switched Path (LSP) between two Label Switched Routers (LSRs) belonging to different IP/MPLS domains. First, the fast end-to-end setup is achieved by modifying the network backbone control plane. Second, two different aggregation schemes are proposed to summarize the QoS network state information to be transported through the ATM core network. Therefore, both the efficient aggregation schemes and the fast mechanism allow source routing to set up a path faster than the existing methods and to reduce the blocking probability using a summary of the available resource information.

*Keywords:* IP/MPLS; ATM; PNNI

---

## 1. Introduction

The exponential growth of real-time multimedia applications, such as Voice over IP (VoIP), video-conference, Video on Demand (VoD), tele-engineering, etc., is directing the network evolution towards transport infrastructures enabling the provisioning of connections with certain performance guarantees, such as Quality of Service (QoS) requirements (for instance, high bandwidth, low end-to-end delay, low delay jitter or minimum losses). Therefore, the real-time applications require the utilization of

---

* Corresponding author. Tel.: +34 93 896 77 67; fax: +34 93 896 77 00.

*E-mail address:* sergio@ac.upc.edu (S. Sánchez-Lopez).

1389-1286/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2006.05.011

real-time channels, which must be set up with specific traffic characteristics and QoS requirements. The required time to set up an end-to-end real-time channel is one of the fundamental metrics to be taken into consideration in the real-time context and interactive multimedia applications.

In the Internet, these transport networks are moving toward a model of high performance IP/MPLS routers interconnected through intelligent backbones, which directly provide an infrastructure of new IP services, compatible with the already existent IP services. An intelligent backbone supports emerging requirements such as dynamic and rapid provisioning of connections, automatic topology discovery, reactive traffic engineering and a fast restoration. A key issue in order to achieve these functionalities is the definition of a control plane, which is responsible for the routing and signalling processes. This control plane must be independent of the subnetwork control planes connected through the backbone.

Initially, ATM networks were expected to replace the current router-based Internet. Although this did not happen, ATM switches are widely used as the backbone technology in many core networks. The control plane functionalities of such ATM core networks are implemented by the Private Network-Network interface (PNNI) [1] proposed by the ATM Forum, which consists of a routing protocol and a signalling protocol. A typical scenario in the current Internet is the case of two IP/MPLS subnetworks interconnected through a PNNI ATM backbone. In such a scenario, the interoperability between both MPLS and ATM technologies in order to achieve MPLS connectivity across the ATM backbone is still an open issue. More specifically, the problem is how to set up a fast end-to-end Label Switched Path (LSP) with QoS guaranties between two Label Switched Routers (LSRs) located in different MPLS domains.

Different approaches have been considered in order to achieve the interoperability between MPLS and ATM technologies. Thus, the distribution of MPLS information through an ATM backbone has been solved by either using ATM-LSRs in the ATM network or tunnelling a LDP through an ATM Virtual Path (VP). The main drawbacks of these solutions are respectively the addition of an IP/MPLS router over each ATM switch and the encapsulation and transport of a signalling protocol through an ATM cloud. In this paper a more appropriate and general solution is proposed. There are two main basics of this mechanism. The former, this solution is based on two previous published works [12,13], that is the MPLS and ATM integration using the PNNI Routing and a new mechanism to set up a fast end-to-end LSP between two different MPLS domains, which are connected via an ATM backbone. The latter, we extend and improve the mechanism by including a new method to apply the Constraint-based Routing concept to this scenario. Consequently, the overall mechanism proposed in this paper allows an end-to-end LSP from a source node to a destination node belonging to two different MPLS networks to be fast established through an ATM-based backbone with the required QoS.

The remainder of this paper is organized as follows: Section 2 presents the different mechanisms existing to distribute MPLS information through an ATM network. Section 3 describes a new mechanism to set up a fast end-to-end LSP. Section 4 is devoted to applying the Constraint-Routing concept to our fast setup mechanism proposal to reduce the congestion effects in the different IP/MPLS domains composing our scenario. Section 5 carries out a performance evaluation of the proposed mechanisms and finally Section 6 concludes the paper.

## 2. MPLS over ATM: state of the art

In this section the already existing solutions (mostly specified in RFCs of the IETF) to establish LSPs through ATM networks are discussed. There are indeed three different approaches to distribute MPLS information through an ATM network, namely to distribute MPLS Labels between ATM LSRs, tunnelling through ATM and communicating the Virtual Connection Identifier (VCID) within ATM signalling messages. In the next subsections, we describe the performance and the drawbacks of these proposals.

### *2.1. Labels distribution between ATM-LSRs*

In [2] a solution for distributing MPLS labels through an ATM network is presented. This solution considers that all the network nodes are ATM-Label Switching Routers (ATM-LSRs). In this way, ATM routing and signalling algorithms are not necessary. ATM-LSR implements the label switching control and forwarding components specified in [3]. Moreover, it is composed by Label

Switched Controlled ATM (LC-ATM) interfaces in order to forward cells between them, using the Virtual Circuit Identifier (VCI) or Virtual Path Identifier (VPI) field to carry MPLS labels. An ATM-LSR is defined with the restrictions imposed by the standard ATM. Therefore, the ATM-LSR behaviour is affected by the characteristics of the ATM switches. Since the label swapping function is performed in the fields VCI or VPI of the cell header, the size and placement of the labels are fixed by the cell header format. A different situation is produced when there are non-LSR-capable ATM switches in the ATM network. In this case, the connection between two ATM-LSRs can be set up via either ATM PVC (Permanent Virtual Circuit) or ATM SVC (Switched Virtual Circuit). When the PVC is set up, MPLS cannot use the VPI field because it is used to establish the path. Hence, in this case labels must be transported on the VCI field. This implies a size restriction of the MPLS label. On the other hand, if the path is a SVC then both VPI and VCI field are used in the connection establishment. Therefore, MPLS can use neither VPI field nor VCI field in order to transport labels.

### 2.2. Tunnelling through ATM

In order to solve the drawback of the previous solution, where the connection between two ATM-LSRs is an ATM SVC, document [4] describes a method to transport a label identifier through ATM switches. This identifier is named Virtual Connection Identifier (VCID) and its value is the same in both edges of the path. In this way, an LSP can be identified with an ATM VC. The In-band notification procedure is a method in order to perform the VCID notification. In this process the messages are sent in a VC, which is set up for the requested session. A control message, used as LDP, is sent across the VC. An example is shown in Fig. 1. A VC is established between ATM-LSR1 and ATM-LSR2. ATM-LSR1 selects a value for the VCID. It sends a VCID PROPOSE message with VCID value and an Identifier (ID) message. Moreover, a relation between the outgoing label (VCI/VPI) and the VCID is carried out by the ATM-LSR1. Then, ATM-LSR2 receives the message and performs an association between the VCID received and the incoming label (VCI/VPI). After that, ATM-LSR2 sends an ACK message with the same VCID value and the message ID. ATM-LSR1 compares the received values with its registered values (VCID and message ID). If the values are the same then ATM-LSR1 sends an LDP REQUEST message with the ID message. When the ATM-LSR2 receives this message sends an LDP MAPPING message with the VCID value in its label TLV. The VCID can be communicated through RSVP as a label value. The main drawback of this solution is that the RSVP messages have to be tunnelled across the ATM network, therefore, increasing the setup time as we demonstrate in Section 5.

### 2.3. Communicating the VCID within ATM signalling messages

Other mechanisms have been defined in the literature in order to establish a VCID end-to-end in a path. We will point out the mechanism described in [5], which introduces a method to transport the VCID in an ATM signalling message. This method differentiates two kinds of sessions, Long-live and

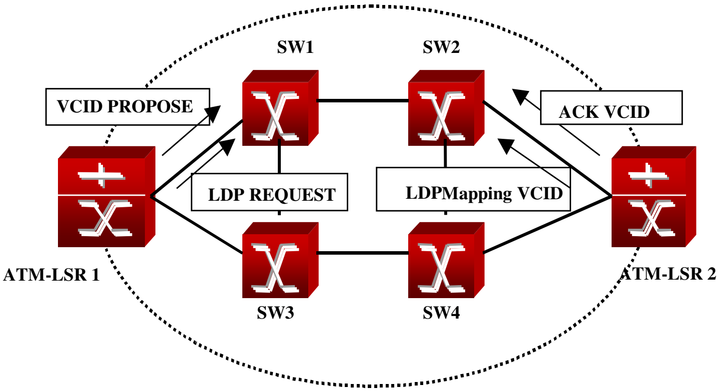


Fig. 1. VCID notification procedure.

| | | | |
|---|---|---|---|
| Information Element Identifier= Generic identifier transportIE (0x7F) | | | 1 |
| Ext | Coding | IE instruction field | 2 |
| Length of contents of information element | | | 3-4 |
| Identifier related standard/aplication = MPLS (0x06) | | | 5 |
| Identifier type=Resource(0x02) | | | 6 |
| Identifier length=4 octets(0x04) | | | 7 |
| **VCID**<br>**(4 octets)** | | | 11 |

Fig. 2. VCID format.

QoS-sensitive session. On the first, the call setup is started when a Long-live session is detected. On the second, the call setup is started by a signalling protocol (e.g., Reservation Protocol, RSVP). When RSVP requests a connection to the ATM network, an ATM signalling process is triggered. This procedure is based on B-ISDN signalling where a session identifier and an element for transporting session information are both included. This element is named Generic Identifier Element (GIT) and allows user identifiers to be transported end-to-end in ATM networks. GIT is an optional information element in both Q.2931 and Q.2971 UNI signalling protocol. The signalling messages (as SETUP, RELEASE, etc.) are transferred between end users through the ATM network. Each one has to contain three GITs, which are transparent to the ATM network. The format of a GIT is shown in Fig. 2.

The existing solutions described above require either all the nodes in the network being ATM-LSRs, or establishing a tunnel across the ATM backbone. Certainly, even though these approaches provide an end-to-end QoS path the setup time is too high for real-time connections. Thus, we aim at proposing a mechanism to provide fast end-to-end QoS constrained paths in our integrated network scenario. As stated before, our proposal contains two main basics, a signalling protocol to establish fast LSPs and providing QoS capabilities to such LSPs. Both issues are carefully described in the next two sections.

## 3. PNNI adaptation for LSP setup in MPLS–ATM integrated environments

A more appropriate and general solution would be that integrating both MPLS and ATM technologies. In this section, we deal with such a solution based on the PNNI routing protocol in order to achieve a QoS path setup faster than the existing solutions and, therefore, more appropriate for the real-time path setup.

### 3.1. A brief review of the PNNI routing protocol

PNNI has the becoming potential for efficiently supporting an intelligent backbone control plane because it includes many interesting features (Table 1).

PNNI consists of two separable parts, a signalling protocol and a routing protocol. PNNI signalling is based on a subset of the UNI 4.0 (itself based on Q.2931) [6] and supports the connection setup and release procedures. PNNI routing is based on:

- *Hierarchical structure*. The lowest level is organized in groups of peer physical nodes (i.e., they have similar features). A Logical Group Node (LGN) in the upper level represents each group, and a particular node, called “Peer Group

Table 1
PNNI features

- High routing scalability with multiple (up to 104) levels of hierarchy
- Security features at the UNI
- Support for modification of call parameters after call establishment
- Unique connection identifiers (e.g., to be used for accounting purposes)
- Support for OAM information elements
- Traffic management functions. Unidirectional and bi-directional support of Permanent Virtual Connection (PVC), Switched Virtual Connection (SVC), and Soft-Permanent Virtual Connection (SPVC)
- Automatic topological and resource discovery
- Support for Closed User Groups (CUG) and VPN
- Support for separation between signalling and transport networks
- Support for multicasting
- Support for connection tracing functions
- Resilience functions in case of failures (pre-planned or on-demand). “Slow” re-rerouting for optimization purposes (make before brake)
- Call Admission Control (CAC)
- Constraint-based source routing using cost, capacity, link constraints, propagation delay constraints, etc.
- Support for protection and restoration mechanisms and crankback capabilities

Leader" (PGL), summarizes topology information within the peer group and performs the functions needed to represent the subnet at the upper level. The hierarchical structure is completed by creating higher levels until the whole network is included in a single highest level.

- *Exchange of information*. Once a hierarchical structure is established, each node exchanges HELLO messages with its neighbours in order to determine the local state information. This information includes a node Identifier, a group Identifier, and the status of its links to the neighbours. Each node bundles its state in a PNNI Topology State Element (PTSE). PTSEs are encapsulated within a PNNI Topology State Packet (PTSP), which is flooded throughout the network. Then, a topology database, consisting of a collection of all the PTSEs received at nodes, provides the information required in order to compute a route from any source to any destination node belonging to the network.
- *Flooding mechanism*. The flooding is the advertising mechanism used in PNNI. It is based on the hop-by-hop propagation of PTSEs throughout the peer groups. This mechanism has to assure that all the nodes belonging to the same peer group have a similar topology database.
- *Distribution of non-ATM information across the ATM network*. PNNI Augmented Routing (PAR) [7,8] is a common work performed by the ATM Forum and the IETF [9,10]. PAR is defined as an extension of the PNNI routing to allow the distribution of information concerning the services provided by the network clients through the ATM network. Client network nodes have to discover and register the different services offered by all the devices interconnected through the ATM network. In order to achieve this, a proxy PAR is defined [9]. Proxy PAR works in client mode on the IP routers and interacts with those ATM switches having the Proxy PAR server mode installed. One of the main advantages is that due to the simplicity of the client implementation, it can be immediately incorporated into the existing devices.
- *Path selection*. Routing algorithms selecting paths according to the required QoS still remain as an open issue since the PNNI protocol only provides a routing framework and so does not standardize a method to find the most appropriate path that satisfy the requested QoS.

### 3.2. MPLS–ATM integration based on PNNI

One of the objectives of this paper is to use the PAR mechanisms in order to transport the MPLS information between MPLS subnetworks connected through an ATM backbone. PAR defines specifics PAR Information Groups (PAR IG) in order to describe IPv4 and VPN information. We consider that the ATM Border Switches (BSs) (i.e., edge nodes in the ATM cloud) are PAR capable and that the external devices (IP Routers, LSRs, etc.) are directly connected to the BSs. An External device has to register its information in a BS so that this information can be distributed to other BSs. Moreover, each external device has to obtain the information from the BS, to which is attached. In order to use the PAR mechanisms to transport MPLS information, a new PAR IG has to be defined and the architecture of the Border Router (BR) has to be modified. According to this, we suggest defining the so-called PAR MPLS services Information Group as shown in Fig. 3. This new PAR IG allows

| C | IG Name | Nested in |
|---|---|---|
| 768 | PAR Service IG | PTSE (64) |
| 776 | PAR VPN ID IG | PAR Service IG (768) |
| 784 | PAR IPv4 Service Definition IG | PAR VPN ID IG (776) / PAR Service IG (768) |
| **792** | **PAR MPLS Services Definition IG** | **PAR Services IG (768)** |
| 800 | PAR IPv4 OSPF Service Definition IG | PAR IPv4 Service Definition IG (784) |
| 801 | PAR IPv4 MOSPF Service Definition IG | PAR IPv4 Service Definition IG (784) |
| 802 | PAR IPv4 BGP4 Service Definition IG | PAR IPv4 Service Definition IG (784) |
| 803 | PAR IPv4 DNS Service Definition IG | PAR IPv4 Service Definition IG (784) |
| 804 | PAR IPv4 PIM-SM Service Definition IG | PAR IPv4 Service Definition IG (784) |

Fig. 3. PAR MPLS services IG format.

MPLS labels to be distributed through the ATM backbone. Concerning the BR architecture, we suggest using a Proxy PAR Capable Label Switching Router (PPAR-LSR) consisting of the following elements (see Fig. 4): an LSR with routing and specific MPLS functions, a Proxy PAR client, added to the LSR in order to register and obtain information of the Proxy PAR server and finally, an ATM Switch as a Proxy PAR server. A forwarding table will establish a relation between MPLS labels an ATM outgoing interfaces.

A more detailed description of this solution can be found in [11].

### 3.3. A fast LSP setup

Let us consider the scenario depicted in Fig. 5. Once we have a solution to transport MPLS information from a MPLS subnet (ND1, Network Domain 1) to other MPLS subnet (ND3) through an ATM backbone (ND2), the next step is to set up a fast end-to-end LSP between two LSRs, each one in a different MPLS subnet. The solution suggested above has the advantage that allows a path to be set up between an LSR of the ND1 and an LSR directly connected to a BR of ND2 (e.g., LSR9 in Fig. 5). The problem appears when there are more than one intermediate LSRs, i.e., the destination LSR belongs to another MPLS subnet (e.g., LSR8 in ND3). In this situation is needed to use a Label Distribution Protocol (LDP) in ND3 in order to set up the LSP to the destination LSR. In order to solve this problem, we suggest proceeding as follows [12].

The Proxy PAR client on each BR registers the MPLS protocol along with labels, and all the address prefixes belonging to the subnetworks directly connected to the ATM backbone, which

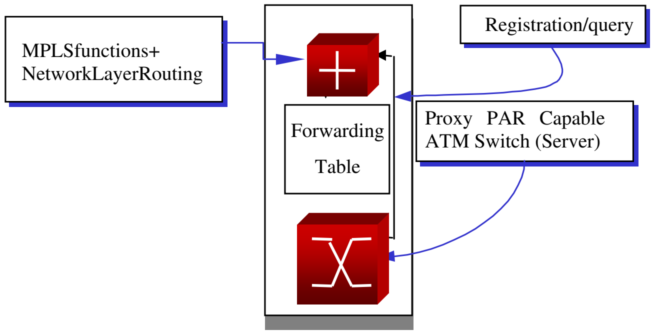


Fig. 4. Border router architecture.

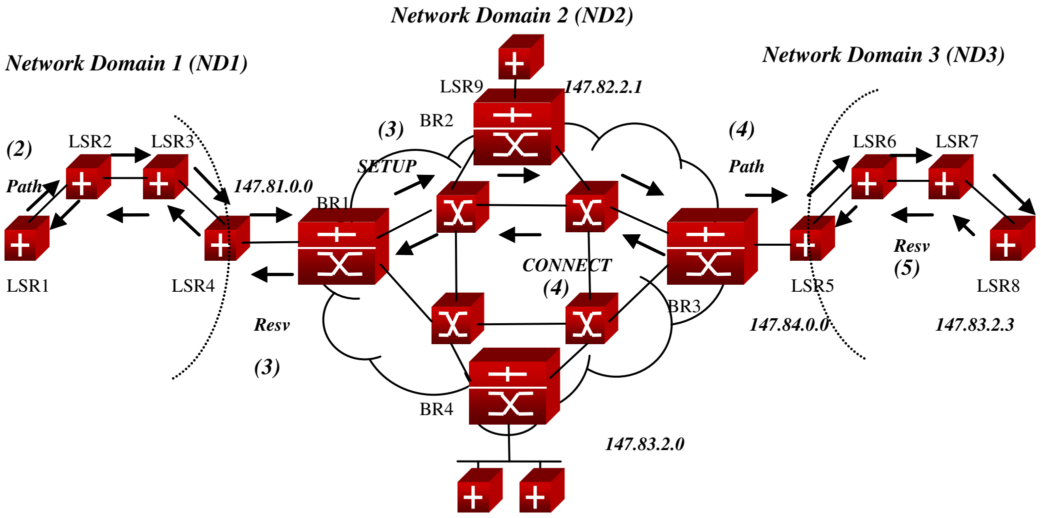


Fig. 5. Network topology scenario.

can be reached. Every server bundles its state information in PAR PTSEs, which are encapsulated within a PTSP, which is sent to a neighbouring peer. Using the received PAR MPLS devices Definition IG, every server side generates an MPLS topology database as shown in Fig. 6. Each client side will use the query protocol to obtain information about registered services by other clients.

Just to illustrate the performance of the proposed mechanism let suppose that LSR1 decides to set up a connection with LSR8. Therefore, an LSP has to be created between both nodes. We assume that RSVP-TE is used as LDP to distribute labels along the path. We assume that an Ordered LSP Control (that is, the set up of an LSP starts at an egress LSR and is forwarded in the upstream direction) performs the establishment of the LSP. Because of the implemented functions in the BRs both MPLS networks act as if the BRs were an end node. As a consequence, after a *Path* message with a label request object has been sent along the path, the ingress BR returns a *Resv* message to the source MPLS node as if it were the end node along the path. Moreover, the egress BR triggers a *Path* message to the destination MPLS node, as if it were the source node.

When the RSVP *Path* message reaches the ingress BR in ND2 (***2***), the RSVP *Resv* message is returned to the source LSR with the requested label (***3***). Simultaneously, the PNNI signalling protocol sets up a VC from the ingress BR1 to the egress BR3. A UNI 4.0 signalling subset is used by the PNNI as signalling mechanism. Therefore, a call setup is triggered in the ingress BR when the RSVP *Path* message reaches the ingress BR1. A *SETUP* message is sent to the egress BR in order to establish the connection. Just as was mentioned in Section 2, the *SETUP* message can contain some Generic Identifier Elements (GITs). They are transparently transported through the ATM backbone. Therefore, they can be used to carry non-ATM information such as LDP information. In this way, we suggested in [12] using a new GIT in order to start a LDP in ND3. We proposed adding a new GIT in the *SETUP* message with the following characteristics: an Identifier related standard/application field with a value (0 × 60), which corresponds to MPLS; and an Identifier Type with a value (0 × 02) assigned to Resource. In this case, the identifier corresponds to the MPLS VCID presented in Section 2. However, in our scenario the VCID is not necessary because we have used the PAR mechanisms, explained in Section 2, in order to distribute MPLS labels. Thus, we proposed replacing the VCID field with the destination LSR IP address (Fig. 7). In this way, when the *SETUP* message reaches the egress BR3, it has enough

BR1

| @IP Dest | @ATM | Label |
|---|---|---|
| | | |
| 147.82.2.1 | @BR2 | 0.50 |
| 147.84.0.0 | @BR3 | 0.40 |
| 147.83.2.0 | @BR4 | 0.30 |

BR3

| @IP Dest | @ATM | Label |
|---|---|---|
| | | |
| 147.82.2.1 | @BR2 | 0.50 |
| 147.81.0.0 | @BR1 | 0.20 |
| 147.83.0.0 | @BR4 | 0.30 |

BR2

| @IP Dest | @ATM | Label |
|---|---|---|
| | | |
| 147.81.0.0 | @BR1 | 0.20 |
| 147.84.0.0 | @BR3 | 0.40 |
| 147.83.2.0 | @BR4 | 0.30 |

BR4

| @IP Dest | @ATM | Label |
|---|---|---|
| | | |
| 147.82.2.1 | @BR2 | 0.50 |
| 147.84.0.0 | @BR3 | 0.40 |
| 147.81.0.0 | @BR1 | 0.20 |

Fig. 6. BRs topology database.

| | | | |
|---|---|---|---|
| Information Element Identifier = Generic identifier transport IE (0x7F) | | | 1 |
| Ext | Coding standard | IE instruction field | 2 |
| Length of contents of information | | | 3-4 |
| Identifier related standard / aplication = MPLS (0x06) | | | 5 |
| Identifier type = Resource(0x02) | | | 6 |
| Identifier length = 4 octets(0x04) | | | 7 |
| **@IP MPLS node destination (4 octets)** | | | 11 |

Fig. 7. New generic identifier element.

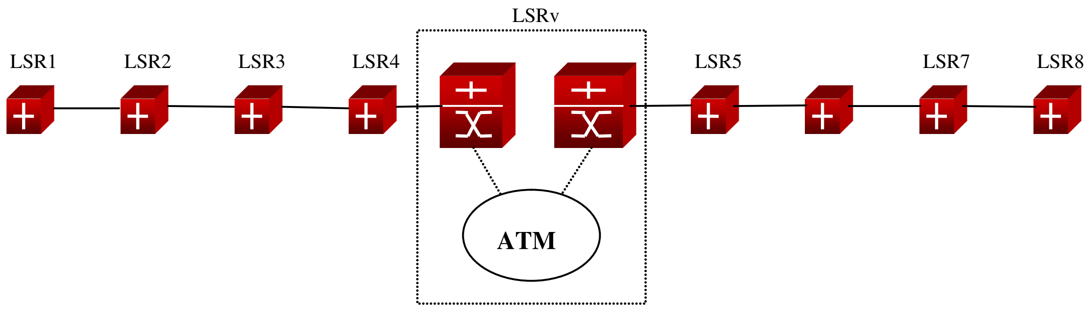


Fig. 8. Network topology with the ATM backbone as an MPLS node.

information about the protocol used in the connection (MPLS), type identifier (Resource) and destination IP address. With this information, the egress BR3 can start an LSP establishment in ND3 (*4*) and (*5*). Once this process has finished, the network can be modelled according to the topology shown in Fig. 8. As we can see, at the moment that the path is set up, the ND2 behaviour is as if it was a unique LSR, called Virtual LSR (LSR$_V$). This node is the last node along the path in ND1, while in ND3 it is responsible for setting up the path towards a destination node LSR8, carrying out the source node functions.

This method, in addition to provide an end-to-end LSP establishment, optimizes the setup time in comparison with the methods presented in Section 2. In order to prove this, a mathematical model was proposed in [12]. Section 5 presents the numerical results.

## 4. Providing QoS for the fast LSP setup

Traditionally, routes in an IP/MPLS domain are computed using a dynamic routing protocol, such as a Shortest Path First (SPF). Since SPF does not take into consideration any QoS attributes, LSPs computed by this protocol could lead the network to congestion situations. Using Traffic Engineering (TE)-based routing algorithms, congestion could be reduced and network resource utilization optimized. One tool to achieve this is Constraint-based Routing (CR). The main goals of CR are both selecting routes with certain QoS requirements and increasing network utilization. This section is devoted to applying the CR concept to our fast setup mechanism proposal to reduce the congestion effects in the different IP/MPLS domains composing our scenario.

Consider the scenario shown in Fig. 9, assuming that the ATM backbone is able to set up a VC

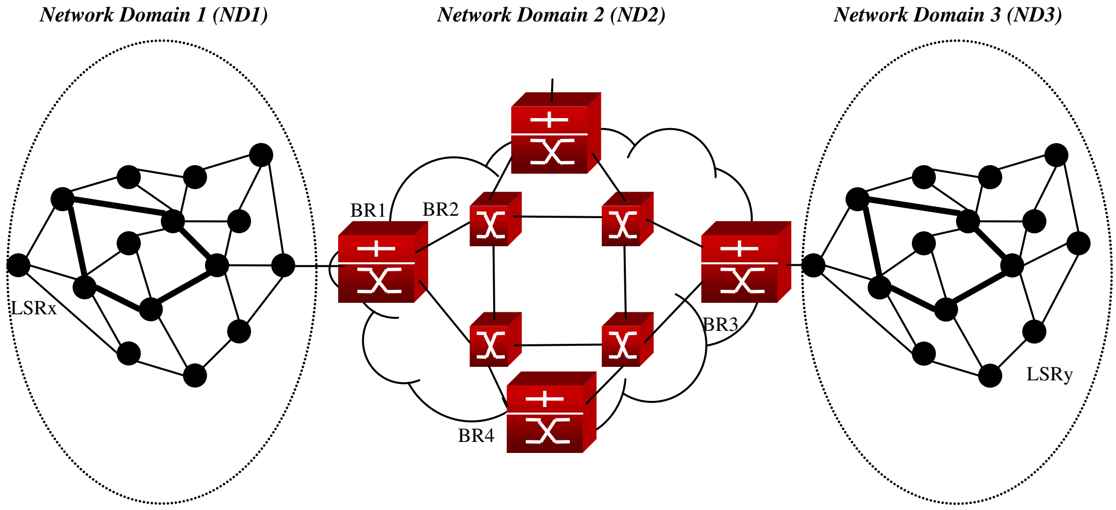


Fig. 9. Network topology for simulation.

between the ingress and egress BRs with the QoS values requested by the LDP of ND1. For simplicity a single QoS parameter, the requested Bandwidth (Bw), is used and the topology and available resources information are considered perfectly updated in all the LSRs.

Considering these factors a new method to provide QoS to our fast LSP setup mechanism is proposed [13]. The establishment of LSPs with Bw guarantees works as follows:

An incoming LSP demand requests LSR$x$ for a connection from LSR$x$ to LSR$y$ with specific Bw guarantees.

LSR$x$ computes the optimal route applying a Constraint-based Routing algorithm using the bandwidth requested by the incoming LSP demand.

Once the CR algorithm has computed the path, the RSVP triggers a mechanism to distribute labels. Thus, an RSVP *Path* message is sent to LSR$y$.

When the RSVP *Path* message reaches the ATM ingress BR1, it finds a label associated to the destination IP address (in this case, a network IP address). Then, an RSVP *Resv* message is returned to LSR$x$ and a relation ⟨*label in*, *label out*⟩ is set up in each LSR along the path to establish the LSP. Simultaneously, the ingress BR1 triggers a SETUP message with the new GIT proposed in Section 3 with a Generic Identifier Element containing the destination IP address.

Once the SETUP message reaches the egress BR3, BR3 sends a RSVP *Path*. Message to ND3. It is assumed that within the ATM network a VC with the amount of Bw requested by LSR$x$ can be set up between BR1 and BR3. Simultaneously, the egress BR3 returns a CONNECT message to the ingress BR1 to establish the corresponding ATM VC.

BR3 sends an RSVP *Path* message to LSR$y$ and the ND3 ingress node computes the LSP in ND3. The LSP setup will be finished when the RSVP *Resv* returns from LSR$y$ to the egress BR3.

At the point in the process where the ND3 ingress node computes the path, it occurs that the routing algorithm used to calculate the path to LSR$y$ cannot be a CR algorithm, because the received RSVP *Path* message from BR3 does not contain any QoS parameters. Accordingly, a certain routing algorithm such as SPF should be used there. Consequently, the LSP setup in ND3 cannot guarantee the Bw requested by LSR$x$, and in case of congestion in that LSP the loss of information would be unavoidable (no QoS is provided).

The ideal case uses the same Constraint-based Routing algorithm in ND1 as in ND3. For this to be possible, the egress ND3 ingress node has to know, in addition to the destination IP address, the Bw requested by LSR$x$. To resolve this issue, this paper proposes adding a new Identifier Type in the Generic Identifier Element proposed in Section 3. This can easily be done because an Identifier Related Standard/Application may have multiple Identifier Types, and the GIT maximum length (133 octets [10]) limits the number of identifiers. The following identifier definition is proposed: Identifier Type = Resource, Identifier Length = 3 octets and Identifier Value = requested BW in Mbps. The format of the new GIT is shown in Fig. 10.

In this way, when the RSVP *Path* message reaches the ingress BR1, both the destination IP address and the Bw requested by LSR$x$ are transferred to the ingress BR1. These values are carried in the GIT and they are transported by the ATM SETUP message. When this message reaches BR3, BR3 sends a RSVP *Path* messages to ND3 with the requested Bw. Now, the ND3 ingress node has a value for a QoS parameter to compute the path using a CR algorithm.

Extending this concept, new additive QoS measures could be considered. Consequently, the Generic Identifier Transport (GIT) element will increment its size, for example, three octets for each QoS measures (delay, jitter, etc.) and the Constraint-based Routing algorithm will have more than one parameter to select the best route.

The solution proposed above has a minor drawback. When the end-to-end path cannot be completed due to the impossibility of finding an LSP with the required Bw in ND3 (i.e., the LSP is blocked), then paths already set up in ND1 and ND2 have to be torn down. In this situation, we have to take into consideration the cost of tearing

| | | | |
|---|---|---|---|
| Information Element Identifier = Generic Identifier transportIE (0x7F) | | | 1 |
| Ext | Coding standard | IE Instructions field | 2 |
| Length of contents of information element | | | 3-4 |
| Identifier related standard/application = MPLS (0x06) | | | 5 |
| Identifier type = Resource (0x02) | | | 6 |
| Identifier length = 4 octets (0x04) | | | 7 |
| **IP/MPLS destination node address (4 octets)** | | | 11 |
| Identifier type = Resource (0x02) | | | 12 |
| Identifier length = 4 octets (0x03) | | | 13 |
| **Requested Bw (3 octets)** | | | 16 |

Fig. 10. Generic identifier with BW.

down the path in both domains. Therefore, in Section 5 we will evaluate the LSP blocking ratio in ND3 to determine the effects produced by this drawback in our mechanism.

### 4.1. QoS mechanism based on an aggregated traffic engineering database (ATED)

A possible solution to overcome this drawback could be to define a mechanism able to distribute a summary of topology and available resource information of the entire network between all the MPLS networks interconnected through the ATM backbone. This information should be transported from ATM egress node (BR3) to ATM ingress node (BR1). Therefore, the BR1 knows the ND3 network status before sending back the *Resv* message to source node in ND1. If a route with the required QoS does not exist in ND3 the path in ND1 is not established. One option could be to transport a Traffic Engineering Database (TED) containing a summary of topology and available resource information about MPLS network in ND3. The TED depends on the number of ND3 links and each link has associated both the amount of bandwidth not yet reserved at each of eight MPLS priority levels and a pair of node Identifiers. In this way, the overhead introduced in the ATM backbone due to the TED size depends on both the MPLS network size in ND3 and the mechanism to summarize information.

We propose using the aggregation concept, which is used on hierarchical networks (e.g., ATM PNNI networks [1]). It consists in reducing the amount of topology and available resource information in each Peer Group that composing the network. Taking into consideration this concept, we may reduce the amount of topology information in an IP/MPLS network (i.e., ND3) using an aggregation scheme.

Several aggregation schemes have been proposed, analyzed and compared [15–17]. We propose two new aggregation schemes, namely Full Mesh Aggregation (FMA) Scheme and Asymmetric Simple Aggregation (ASA) Scheme. This adaptation consists in using the MPLS TE parameters to generate an Aggregated Traffic Engineering Database (ATED). We only consider the unreserved Bandwidth as a QoS parameter in order to simplify the aggregation scheme definition.

The Full Mesh Aggregation Scheme works as follows:

Select all border nodes of the network. Consider a border node as ingress and egress node (Fig. 11(1)).

Compute all the routes between an ingress node and an egress node (Fig. 11(2)).

Select the path that provides the best QoS (Fig. 11(3)).

Repeat steps 2 and 3 for each pair of ingress/egress nodes.

The aggregated unreserved bandwidth for each MPLS priority level computed by the FMA scheme is defined as follows:

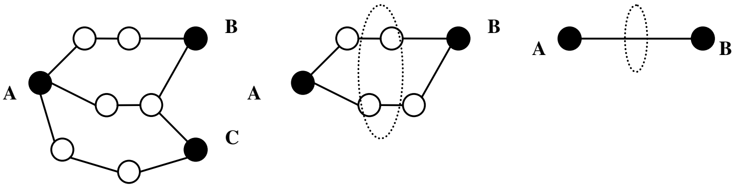
$$\underset{k=1\ldots8}{B_k^{ij}} = \max_{\forall R_{ij},i\neq j}\left\{\min_{l\in R_{ij}}[B_k(l)]\right\},$$

where

- $B_k^{ij}$ is the aggregated unreserved bandwidth for the $k$ priority level allocated to a pair of border nodes (ingress node $i$ and egress node $j$).
- $i$, $j$ are an ingress and an egress node of the network.
- $R_{ij}$ is a path between $i$ and $j$.
- $l$ is a link of the path $R_{ij}$.
- $B_k(l)$ is the unreserved bandwidth for the $k$ priority level in the link $l$, which belongs to the path $R_{i,j}$. How to define $B_k(l)$ is an open research issue. We assume that it is a known value.

The Asymmetric Simple Aggregation Scheme works as follows:

Select all border nodes of the network. Consider a border node as ingress and egress node (Fig. 12(1)).



(1) (2) (3)

Fig. 11. Steps for an FMA scheme.

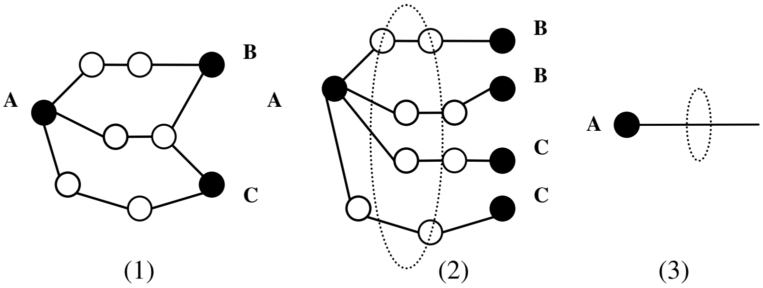


Fig. 12. Steps for an ASA scheme.

Select an ingress node and compute all the routes between that ingress node and all the egress nodes (Fig. 12(2)).

Select the path that provides the best QoS. (Fig. 12(3)).

Repeat steps 2 and 3 for each ingress node.

The aggregate unreserved bandwidth for each priority level computed by the ASA scheme is defined as follows:

$$\underset{k=1\ldots 8}{B_k^i} = \max_{S_{ij},\forall j\neq i}\left\{\min_{l\in S_{ij}}[B_k(l)]\right\},$$

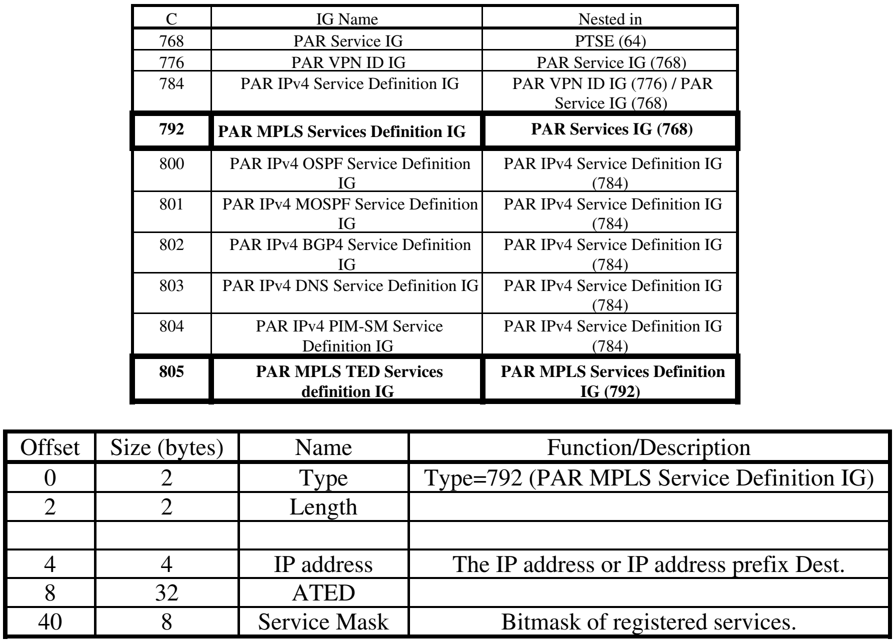
where

- $B_k^i$ is the aggregated unreserved bandwidth for the $k$ priority level allocated to the ingress node $i$.
- $i$,$j$ are an ingress and an egress node of the network.
- $S_{ij}$ is a path between $i$ and any $j$.
- $l$ is a link of the path $S_{ij}$.
- $B_k(l)$ is the unreserved bandwidth for the $k$ priority level in the link $l$, which belongs the path $R_{ij}$. We also assume that it is a known value.

In order to transport an ATED through the ATM backbone, a new PAR PTSE named PAR MPLS ATED Service Definition IG is defined. The element format is shown in Fig. 13. Thus, the process to establish the LSP end-to-end is as follows:

Each BR uses the Proxy PAR client to register the MPLS information and the ATED associated with the IP/MPLS Network connected to the BR.

The Proxy PAR server floods PTSEs throughout the ATM backbone. Once the flooded information reaches all the BRs, each one uses the Proxy PAR client to obtain that information.



| C | IG Name | Nested in |
|---|---|---|
| 768 | PAR Service IG | PTSE (64) |
| 776 | PAR VPN ID IG | PAR Service IG (768) |
| 784 | PAR IPv4 Service Definition IG | PAR VPN ID IG (776) / PAR Service IG (768) |
| **792** | **PAR MPLS Services Definition IG** | **PAR Services IG (768)** |
| 800 | PAR IPv4 OSPF Service Definition IG | PAR IPv4 Service Definition IG (784) |
| 801 | PAR IPv4 MOSPF Service Definition IG | PAR IPv4 Service Definition IG (784) |
| 802 | PAR IPv4 BGP4 Service Definition IG | PAR IPv4 Service Definition IG (784) |
| 803 | PAR IPv4 DNS Service Definition IG | PAR IPv4 Service Definition IG (784) |
| 804 | PAR IPv4 PIM-SM Service Definition IG | PAR IPv4 Service Definition IG (784) |
| **805** | **PAR MPLS TED Services definition IG** | **PAR MPLS Services Definition IG (792)** |

| Offset | Size (bytes) | Name | Function/Description |
|---|---|---|---|
| 0 | 2 | Type | Type=792 (PAR MPLS Service Definition IG) |
| 2 | 2 | Length | |
| | | | |
| 4 | 4 | IP address | The IP address or IP address prefix Dest. |
| 8 | 32 | ATED | |
| 40 | 8 | Service Mask | Bitmask of registered services. |

Fig. 13. PAR MPLS ATED services definition IG.

LSR$x$ computes the path using the ND1 ATED and sends an RSVP *Path* message to ND3. The RSVP *path* message contains the Explicit Routing Object (ERO), which registers the route.

The RSVP *Path* message reaches BR1, which checks the ATED to verify whether a path with the required Bw exists on ND3. If there is no available path with the required Bw, then the connection will be refused. Otherwise, the setup process follows.

The ingress BR1 triggers a SETUP message to set up a VC and sends an RSVP *Resv* message to LSR$x$ with the required label.

When the SETUP message reaches the egress BR3, BR3 sends an RSVP Path message to ND3. ND3 ingress node computes last part of the end-to-end path in ND3. The LSP setup will be finished when the RSVP *Resv* returns from LSR$y$ to the egress BR3. Simultaneously, the egress BR3 returns a CONNECT message to the ingress BR1 to establish the corresponding ATM VC.

Since Network Operators commonly do not like sharing such an amount of information with other Network Operators, this solution may be applied when the MPLS networks belong to different administrative domains because the ATED information is a summary of the complete topology and resource information of ND3. It is assumed that the ATEDs are always updated. This means that the ND3 ATED has to be flooded throughout the ATM backbone every time the update process modifies the ATED. This increases overhead in the backbone, and it could be a problem in large networks. Consequently, a trade-off between the ATED accuracy and overhead in the ATM backbone has to be considered. The impact of possible network information inaccuracy in the mechanisms proposed in this paper will be a subject for further studies. Finally, we propose using a QoS routing algorithm, proposed in [14] to reduce the LSP blocking probability resulting from selecting the path under inaccurate routing information.

## 5. Performance evaluation and numerical results

On the one hand, the effectiveness of the fast LSP setup proposal is demonstrated by comparing it with other LSP establishment mechanisms. Firstly, in order to compute the time needed by the LSP establishment, mathematical models of our proposal and the existent mechanisms are presented. These models are calculated under certain specific working conditions and scenarios, which are described in Section 3.

On the other hand, in order to evaluate the effects over the routing algorithm by applying the solutions presented in Section 4 about providing QoS to the end-to-end LSP, the topology depicted in Fig. 9 is used. Moreover, numerical results are obtained using an extension of the ns2 simulator named MPLS_ns2 developed by the authors of this paper.

### 5.1. The fast LSP setup solution

Let us consider the following notation and the topology shown in Fig. 5:

| | |
|---|---|
| $T_{\mathrm{LSP}}$ | total time needed to establish the end-to-end LSP |
| $N_1$ | number of nodes along the path in the ND1 |
| $N_2$ | number of switches crossed by the MPLS traffic in ND2 |
| $N_3$ | number of nodes along the path in ND3 |
| $\mathrm{BW}_1$ | link bandwidth in ND1 |
| $\mathrm{BW}_2$ | VCC bandwidth in ATM domain |
| $\mathrm{BW}_3$ | link bandwidth ND3 |
| $\mathrm{BW}_L$ | link bandwidth in ATM domain |
| $t_R$ | delay in LSR |
| $t_S$ | delay in switch |
| $S_{\mathrm{PACK}}$ | IP packet size |
| $S_C$ | ATM cell size |
| $S_{\mathrm{PATH}}$ | *Path* message size |
| $S_{\mathrm{RESV}}$ | *Resv* message size |
| $t_{\mathrm{PACK}}$ | time needed by a IP packet in order to reach the egress BR from the source node |
| $t_{\mathrm{PATH1}}$ | time needed by a *Path* message in order to reach the ingress BR |
| $t_{\mathrm{RESV1}}$ | time needed by a *Resv* message in order to reach the source node from the ingress BR |
| $t_{\mathrm{SET}}$ | time needed by a *SETUP* message in order to reach the egress BR from the ingress BR |
| $t_{\mathrm{CON}}$ | time needed by a *CONNECT* message in order to reach the ingress BR from the egress BR |
| $t_{\mathrm{PATH3}}$ | time needed by a *Path* message in order to reach the destination node from the egress BR |
| $t_{\mathrm{RESV3}}$ | time needed by a *Resv* message in order to reach the egress BR from the destination node |

According to the notation shown above, the time needed to setup an LSP between two MPLS nodes situated in two different domains connected through an ATM backbone, is represented by

$$T_{\mathrm{LSP}} = t_{\mathrm{PATH1}} + \max[t_{\mathrm{RESV1}}, t_{\mathrm{SET}} + \max(t_{\mathrm{PATH3}} + t_{\mathrm{RESV3}}, t_{\mathrm{CON}})]. \tag{1}$$

Eq. (1) depends on the number of nodes existing on each domain along the path. In order to simplify the expressions the following considerations have been taken into account: equal delay time in all the LSRs existing in the MPLS domains, equal ATM signalling message size for all the ATM messages and there are only two messages in the VC setup. Applying this we have that

$$T_{\mathrm{LSP}} = \left(\frac{S_{\mathrm{PATH}}}{\mathrm{BW}_1} + t_R\right)N_1 + \max\left\{\begin{array}{l}\left(\frac{S_{\mathrm{RESV}}}{\mathrm{BW}_1} + t_R\right)N_1, \\ \left(\frac{S_C}{\mathrm{BW}_L} + t_S\right)N_2 + \max\left[\begin{array}{l}\left(\frac{S_{\mathrm{PATH}}}{\mathrm{BW}_3} + t_R\right)N_3 + \left(\frac{S_{\mathrm{RESV}}}{\mathrm{BW}_3} + t_R\right)N_3 \\ \left(\frac{S_C}{\mathrm{BW}_L} + t_S\right)N_2\end{array}\right]\end{array}\right\}. \tag{2}$$

In order to compare with other LSP establishment mechanisms, we can analyze the behaviour of different topologies under both these different mechanisms and the mechanism suggested in this paper. In order to perform this, it is necessary to bear in mind that different cases are possible:

- The ATM network is made of ATM LSRs such as is described in Section 2. In this case, we can consider that all the nodes are MPLS capable nodes and the RSVP is used as LDP. We will simulate its behaviour as if the network was three MPLS domains. Eq. (2) is

$$\begin{aligned} T_{\mathrm{LSP}} &= t_{\mathrm{PATH}} + t_{\mathrm{RESV}} \\ &= \left(\frac{S_{\mathrm{PATH}}}{\mathrm{BW}_1} + t_R\right)N_1 + \left(\frac{S_{\mathrm{PATH}}}{\mathrm{BW}_2} + t_S\right)N_2 \\ &\quad + \left(\frac{S_{\mathrm{PATH}}}{\mathrm{BW}_1} + t_R\right)N_3 + \left(\frac{S_{\mathrm{RESV}}}{\mathrm{BW}_1} + t_R\right)N_1 \\ &\quad + \left(\frac{S_{\mathrm{RESV}}}{\mathrm{BW}_2} + t_S\right)N_2 + \left(\frac{S_{\mathrm{RESV}}}{\mathrm{BW}_3} + t_R\right)N_3. \end{aligned} \tag{3}$$

- The ATM network does not implement the PNNI so that the BRs do not have topology information about the other BRs. One method to set up the LSP is proposed in Section 3. In this case (2) is

$$\begin{aligned} T_{\mathrm{LSP}} &= t_{\mathrm{PATH1}} + t_{\mathrm{TUNNEL-PATH2}} + t_{\mathrm{PATH3}} + t_{\mathrm{RESV3}} \\ &\quad + t_{\mathrm{TUNNEL-RESV2}} + t_{\mathrm{RESV1}} \\ &= \left(\frac{S_{\mathrm{PATH}}}{\mathrm{BW}_1} + t_R\right)N_1 + \frac{S_{\mathrm{PATH}}}{\mathrm{BW}_2} + \left(\frac{S_{\mathrm{PATH}}}{\mathrm{BW}_3} + t_R\right)N_3 \\ &\quad + \left(\frac{S_{\mathrm{RESV}}}{\mathrm{BW}_3} + t_R\right)N_3 + \frac{S_{\mathrm{RESV}}}{\mathrm{BW}_2} + \left(\frac{S_{\mathrm{RESV}}}{\mathrm{BW}_1} + t_R\right)N_1. \end{aligned} \tag{4}$$

As we know, the path is simultaneously established in ND1, ND2 and ND3. Because of this, the source node could send traffic before the LSP setup process in the ATM backbone or in the ND3 had been finished. In order to test the effects produced by this bug, an analysis has been performed. The different cases are:

The LSP is completely established when the source node starts sending traffic. The condition is determined by the $t_{\mathrm{RESV1}}$ according to

$$t_{\mathrm{RESV1}} \geqslant t_{\mathrm{SET}} + \max(t_{\mathrm{PATH3}} + t_{\mathrm{RESV3}}, t_{\mathrm{CON}}). \tag{5}$$

And if the expression is

$$t_{\mathrm{RESV1}} < t_{\mathrm{SET}} + \max(t_{\mathrm{PATH3}} + t_{\mathrm{RESV3}}, t_{\mathrm{CON}}). \tag{6}$$

Two possible cases appear, always considering that the backbone ATM VC is established:

The first case can be represented by

$$t_{\mathrm{RESV1}} + t_{\mathrm{PACK}} \geqslant t_{\mathrm{PATH3}} + t_{\mathrm{RESV3}}. \tag{7}$$

In this case, as in the first case, the path is completely established before the traffic flows along the path.

The second case can be represented by

$$t_{\mathrm{RESV1}} + t_{\mathrm{PACK}} < t_{\mathrm{PATH3}} + t_{\mathrm{RESV3}}. \tag{8}$$

In this case, depending on the first packet size of the MPLS traffic, it is possible that the path set up in the ND3 is still in progress when the first packet flows from the source node to the ingress BR. The condition that the first packet size must fulfil so that when this packet reaches the ingress BR the end-to-end LSP will be completely established is

$$t_{\mathrm{PACK}} = t_{\mathrm{PATH3}} + t_{\mathrm{RESV3}} - t_{\mathrm{RESV1}}. \tag{9}$$

According to (9), the first packet size in order to have the end-to-end LSP completely established is

$$\begin{aligned} S_{\mathrm{PACK}} = \frac{\mathrm{BW}_1\mathrm{BW}_2}{N_1\mathrm{BW}_2 + \mathrm{BW}_1} \times \Bigg[&\left(\frac{S_{\mathrm{PATH}} + S_{\mathrm{RESV}}}{\mathrm{BW}_3} + 2t_R\right)N_3 \\ &- \left(\frac{S_{\mathrm{RESV}}}{\mathrm{BW}_1} + 2t_R\right)N_1 - t_S N_2\Bigg]. \end{aligned} \tag{10}$$

Once the analytical expressions have been obtained, we are going to compute the numerical results to compare the different methods explained above. In order to perform the graphic representation, the following values will be constants: $\mathrm{BW}_1 = BW_2 = 2$ Mbps, $\mathrm{BW}_L = 155$ Mbps, $t_R = 71$ µs, $t_S = 10$ µs, $S_C = 53$ bytes, $S_{\mathrm{PATH}} = 112$ bytes and $S_{\mathrm{RESV}} = 120$ bytes. The rest of the parameters will be modified to obtain a meaningful set of results from the suggested method.

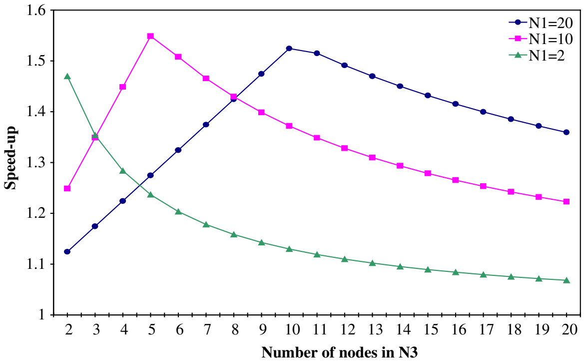


Fig. 14. Speed-up LSR.

Firstly, we compute the speed-up, i.e., how many times faster our proposal is compared to the other mechanism in setting up the end-to-end LSP. Therefore, using (2) and (3), we obtain Fig. 14. Similarly, using (2) and (4) we obtain the result shown in Fig. 15. In both cases, the $\mathrm{BW}_2$ is 34 Mbps. Moreover, in order to reduce the complexity we consider that the VC between BR1 and BR3 is already set up in the second case. We can observe that the speed-up enhances when the N3 is increased to a maximum point where N1 is the double of N3. From here on, when N3 increases its value, the speed-up decreases until reaching a permanent value over 1. Therefore, in this way, we have achieved a method to allow the LSP to be set up faster than the other methods.

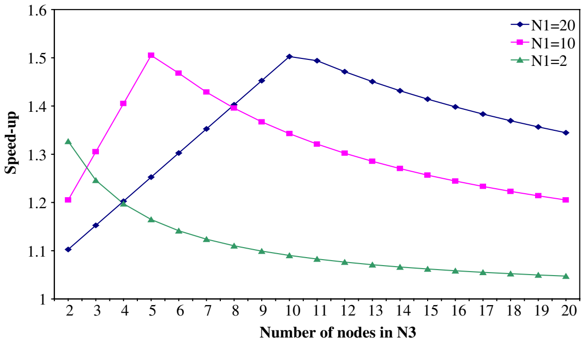


Fig. 15. Speed-up RSVP TUNNEL.

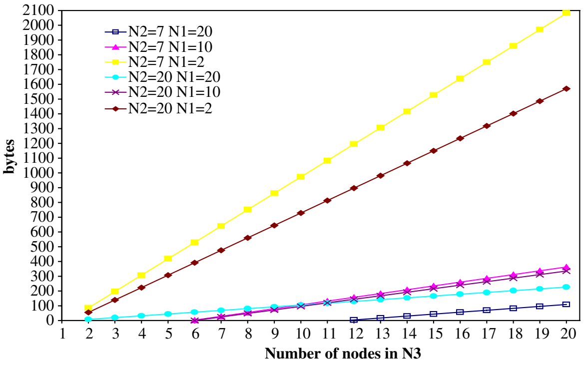


Fig. 16. Packet size for $\mathrm{BW}_2 = 34$ Mbps.

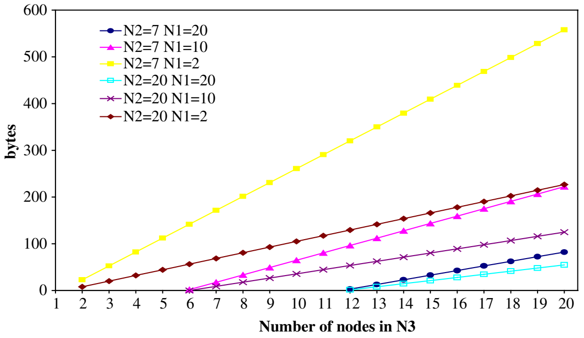


Fig. 17. Packet size for $\mathrm{BW}_2 = 2$ Mbps.

Now we are going to compute the IP packet size in order to fulfil the condition in (7). Using (10) with $\mathrm{BW}_2 = 34$ Mbps and $BW_2 = 2$ Mbps we obtain the results shown in Figs. 16 and 17, respectively. We observe that the worst case is produced when $\mathrm{BW}_2$ is 34 Mbps and N1 is seven times higher than N3 where the size packet is over 1500 bytes. It could be possible that the size of the packets were lower than 1500 bytes. In this case the IP packets will arrive before the LSP is set up, so that the egress BR will place the packets in a queue until the *Resv* message from destination node is received.

### 5.2. Numerical results and evaluation of the proposal for providing QoS

An LSP is blocked when the route computed in the source node by a well-known CR algorithm, the WSP (Widest Shortest Path) cannot be set up because a link of the path does not fulfill the QoS requirements. Traditional mechanisms to set up an LSP solve this problem by triggering either a crankback process (e.g., in ATM networks) or a time-out (e.g., in IP/MPLS networks), which avoid setting up an LSP without the QoS requirements requested by the connection. On the other hand, in the mechanism proposed in this paper the LSP blocking problem must be taken into especial consideration. This mechanism is based on a parallel process used to establish an end-to-end path in order to reduce the setup time. This parallel process consists in establishing a part of the total path in each network domain at the same time. In this situation, a part of the total path could be established in ND1 but ND3 could not establish the last part of the path because does not fulfill the QoS requirements. As a consequence, a Teardown process must be triggered form destination network domain to source network domain to avoid sending traffic from the source node. Depending on the network size (number of nodes) the teardown message could not arrive on time. This situation must be avoided as far a possible. Specifically, the LSP blocking problem that occurs in the scenario is depicted in Fig. 5 as follows.

Consider that the first segment of the LSP (in ND1) and the VC through the ATM backbone has been established. Whereas, last part of the LSP in ND3 cannot be established because an LSP with the required Bw does not exist. Then the end-to-end path cannot be completed due to the impossibility of finding an LSP with the required Bw in network ND3. Thus, a teardown process must be triggered to tear down the established path in ND1 and ND2. The cost of this process will depend on the LSP Blocking Ratio produced in ND3.

The solution presented in Section 4 reduces the LSP Blocking Ratio, since BR1 checks the ND3 ATED to know whether a path with the required Bw exists in ND3. If the path exists in ND3, the process to set up the total path follows as mentioned above. Otherwise, the path is not established in ND1. Because of applying this solution, the number of teardown actions in ND1 is also reduced. On the other hand, it must be taken into account that transporting the ND3 ATED throughout the ATM backbone produces signalling information overhead. This overhead depends on the ATED size and the update mechanism.

Consider again the scenario shown in Fig. 9 and assume that the egress BR3 has a Traffic Engineering Database (no aggregated TED) where the ND3 topology and available resource information is kept. There are 15 nodes, 27 links and 5 border nodes (ingress and egress) in each IP/MPLS network topology. The TED is assumed to have been made using "extended link attributes", so its size depends on the number of ND3 links. Each link has associated the amount of bandwidth not yet reserved at each of eight priority levels (32 bytes) and a number of node IDs (8 bytes). Therefore, in our scenario the TED size is over 1 kbyte. On the other hand, using an aggregation scheme the size of the TED is as follows. The Aggregated TED (ATED) based on the FMA scheme depends on the number of routes between all the border nodes. Therefore, the ND3 ATED contains information about the four routes existing between the ingress node and the rest of the border nodes. Thus, the ATED size is 192 bytes. The ATED based on ASA scheme depend on the number of border nodes. In particular, the ND3 ATED, which will be transported across the ATM backbone, contains the information related with the ingress node that is directly connected to the backbone. Therefore, the ATED size is 32 bytes.

Fig. 9 shows the topology used to simulate the behaviour of the aggregation schemes. Two different link capacities are used. Links represented by a light line are set to 622 Mbps, and links represented by a dark line are set to 2.5 Gbps. The incoming requests arrive following a Poisson distribution and the requested bandwidth is uniformly distributed between two ranges: range 1 is between 0.5 Mbps and 1 Mbps; range 2 is between 1 Mbps and 5 Mbps. The holding time is randomly distributed, with an average of 120 s. Finally, the existing topology and available resources database in each border router is assumed to be perfectly up-to-date. The experiments have been repeated 10 times with a 95% confidence interval.

In general, the FMA scheme performs an ATED larger than the ASA scheme. However, the information of the FMA ATED is more accurate that the information contained in the ASA ATED. Fig. 18 compares the ATED size (number of entries on the database) for both aggregation schemes and

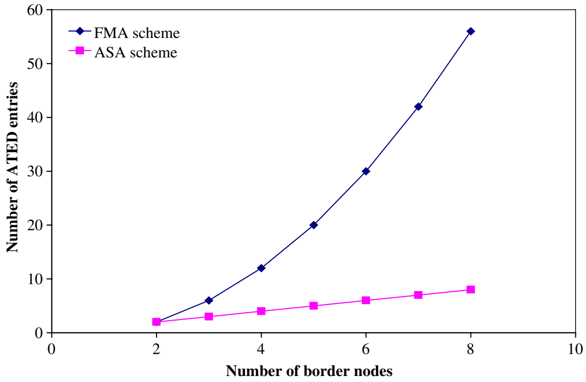


Fig. 18. Comparison between FMA and ASA ATED size.

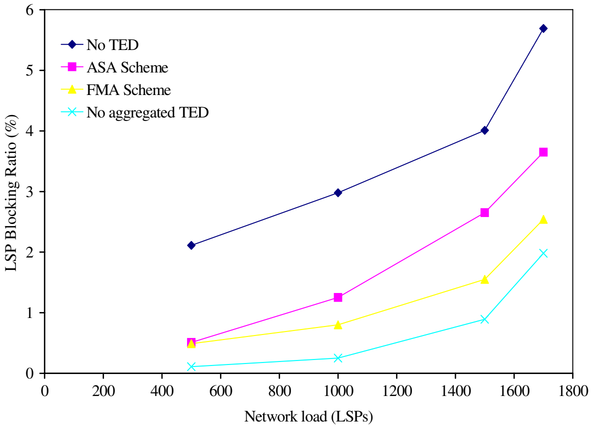


Fig. 19. LSP blocking ratio due to the ATED inaccuracy.

Fig. 19 shows the inaccuracy effects, that is, the LSP blocking ratio due to the ATED inaccuracy. The worst case is when a TED is not transported from ND3 to ND1. The best case occurs when an aggregation scheme is not applied to the TED. In this situation, BR1 knows the complete topology and available resources information of ND3 and may decide to establish the total path. Between both cases are the aggregation options. We observe that the ASA scheme reduces the LSP Blocking Ratio more than the FMA. As mentioned above, it is due to the larger inaccuracy introduced by the ASA scheme. Therefore, according to the results presented in Fig. 19, we can conclude that using the aggregation schemes the LSP blocking can be reduced around 34% for the ASA scheme and around a 61% for the FMA scheme.

In addition, each time that the ATED is updated in ND3, it must be transported across the ATM backbone using the flooding mechanism. This process could produce a huge signalling overhead in the ATM backbone depending on the update frequency. Hence, this solution has to consider a trade-off between ATED accuracy and the overhead produced by flooding the ATED through the ATM backbone. Reducing the update frequency of the ATED across the ATM backbone is proposed according to the overhead produced.

Once the different aggregation schemes have been analyzed, we propose using the FMA scheme because produces less LSP Blocking Ratio. Moreover, Using a QoS routing mechanism to reduce the LSP blocking probability and routing inaccuracy effects is recommended as well. How this mechanism works is explained in [14].

## 6. Conclusions

In this paper, the most relevant documents related to set up an end-to-end path have been analyzed and, consequently, a new method to establish a fast path setup has been suggested. This method is based on two essential concepts: the overlay model and a parallelism technique. The first is based on the current Internet structure, i.e., client subnetworks interconnected through backbones, which are managed and controlled by network operators. The subnetwork clients do not obtain topology information from the backbone and the backbone does no know the topology and available resources information of the subnetworks connected to it. In this way, each subnetwork and backbone has to use its own routing and signalling mechanism in order to establish a path. Moreover, a frequent situation is that the backbone uses a different network technology to the used in the subnetwork clients (e.g., IP/MPLS in the subnetwork client and ATM in the backbone). The second concept is based on the path establishment by segments, i.e., each subnetwork and the backbone setups a part of the total path and each part is approximately established at the same time. Note that the existing methods cannot achieve the same kind of parallelism because they do not have available resource information from the destination domain. In this way, the proposals of this paper achieves that subnetworks belonging to the same administrative domain and connected through a backbone, which has a different network technology, can setup a fast end-to-end path with the required QoS. Therefore, the distribution of MPLS information through an ATM backbone has been solved both using the PNNI Augmented Routing and aggregating routing information according to two new aggregation schemes.

In conclusion, a new Border Router architecture, a new PAR PTSEs and a new Generic Information Element have been defined. Once the MPLS information has been distributed throughout the BRs, a new mechanism to set up an end to end LSP between two different MPLS domains, which are connected via an ATM backbone is suggested. Finally, this paper proposes a method to apply the Constraint-based Routing concept, based on including two new aggregation schemes (ASA and FMA) to summarize available resource information, to interconnect IP/MPLS networks through an ATM backbone. Consequently, we have achieved to set up a fast end-to-end LSP with the required QoS.

## Acknowledgements

This work was partially funded by the Spanish Ministry of Education (MEC) under the CATARO project (TEC2005-08051-C03-01) and the Catalan Research Council (CIRIT) under contract 2005 SGR00481 and the European Commission through Network of Excellence E-NEXT under contract FP6-506869.

## References

[1] ATM Forum, af-pnni-0055.002, Private Network to Network Interface Specification Version 1.1, April 2002.
[2] B. Davie et al., MPLS using LDP and ATM VC Switching, IETF RFC 3035, January 2001.
[3] E.C. Rosen, A. Viswanathan, R. Callon, Multiprotocol Label Switching Architecture, IETF RFC 3031, July 2000.
[4] K. Nagami, Y. Katsube, N. Demizu, H. Esaki, P. Doolan, VCID notification over ATM link for LDP, IETF RFC 3038, January 2001.
[5] M. Suzuki, The assignment of the Information Field and Protocol Identifier in the Q.2941 Generic Identifier and Q.2957 User-to-user Signaling for the Internet protocol, IETF RFC 3033, January 2001.
[6] ITU-T Recommendation Q.2931 B-ISDN-DSS2-UNI Layer 3 Specification for basic call/connection control, February 1995.
[7] ATM Forum, PNNI augmented routing (PAR), Version 1.0, af-ra-0104.000, January 1999.
[8] R. Haas, P. Droz, D. Bauer, PNNI augmented routing (PAR) and proxy-PAR, Computer Networks 34 (2000) 399–418.
[9] P. Droz, T. Przygienda, Proxy-PAR, IETF RFC 2843, May 2000.
[10] T. Przygienda, P. Droz, R. Haas, OSPF over ATM and Proxy-PAR, IETF RFC 2844, May.
[11] S. Sánchez-López, X. Masip-Bruin, J. Domingo-Pascual, J. Solé-Pareta, A solution for integrating MPLS over ATM, in: Proceedings of the 15th International Symposium on Computer and Information Sciences (ISCIS2000), October 2000, pp. 255–303.
[12] S. Sánchez-López, X. Masip-Bruin, J. Domingo-Pascual, J. Solé-Pareta, J. López-Mellado, A path establishment approach in an MPLS–ATM integrated environment, IEEE GlobeCom, November 2001.
[13] S. Sánchez-López, X. Masip-Bruin, J. Solé-Pareta, J. Domingo-Pascual, Providing QoS in an MPLS–ATM integrated environment, QoFIS 2002, October 2002.
[14] X. Masip-Bruin, S. Sánchez-López, J. Solé-Pareta, J. Domingo-Pascual, QoS routing algorithms under inaccurate routing information for bandwidth constrained applications, IEEE ICC, May 2003.
[15] A. Iwata et al., QoS aggregation algorithm in hierarchical ATM networks, IEEE ICC 1998.
[16] B. Awerbuch et al., Routing through network with hierarchical topology aggregation, in: IEEE Symposium on Computer and Communication, 1998.
[17] Ben-Jye et al., Hierarchical QoS routing in ATM Networks based on MDP cost function, IEEE ICON 2000.

**Sergio Sánchez-Lopez** received his BS degree in telecommunication engineering in 1989 from Polytechnic University of Catalonia, his MS degree in electrical engineering in 1996 from University of Barcelona and he got his PhD in Telecommunications Engineering in 2003 from Polytechnic University of Catalonia, Spain. In 1989 he joined the Computer Architecture Department of UPC. Since 1992 he has been an Associate Professor with this department. He is member of the Advanced Broadband Communications Centre of UPC and his publications include several book chapters, two papers in relevant international journals and more than 20 papers in international refereed conferences. His current research interests are in Broadband Internet and high Speed and Optical Networks, with emphasis on traffic engineering, routing algorithms, Control Plane and Multi-Domain Resilience. He has participated in the LION project and currently he is participating in NOBEL Integrated Project and in e-Photon/ONe Network of Excellence of the European VI Framework Program. He is member of the program committee of ICTON'05 and he is also the Local Organization Vice-Chair for INFOCOM'06.

**Xavier Masip-Bruin** received the MS and PhD degrees from the Technical University of Catalonia, Barcelona, Catalonia, Spain, both in Telecommunications Engineering, in 1997 and 2003 respectively. He is currently an Associate Professor of Computer Science with the Technical University of Catalonia. His current research interests lies in the field of Broadband Communications, QoS management and provision and traffic engineering. His recent work has focused on QoS provisioning both in IP/MPLS and Optical Transport Networks. His publications include around 40 papers in national and international refereed journals and conferences. Since 2000 he has participated in many research projects: in IST projects E-NEXT, NOBEL and EuQoS; and in Spanish research projects SABA, SABA2, SAM and TRIPODE.

**Josep Sole-Pareta** was awarded his Master's degree in Telecommunication Engineering in 1984, and his PhD in Computer Science in 1991, both from the *Universitat Politècnica de Catalunya* (UPC). In 1984 he joined the Computer Architecture Department of UPC. Since 1992 he has been an Associate Professor with this department. He is co-founder and member of the Advanced Broadband Communications Centre of UPC (http://www.ccaba.upc.es). His current research interests are in Broadband Internet and High Speed and Optical Networks with emphasis on traffic engineering, traffic characterisation, traffic management, MAC protocols and QoS provisioning. He has participated in many ACTS IST European project devoted to

Optical Networking, such as SONATA, LION, DAVID and The action COST 266. Within the VI Framework Program he is participating in NOBEL (IP-project), in e-Photon/One (Network of Excellence) and in the Action COST 291.

**Jordi Domingo-Pascual** is Professor of Computer Science and Communications at the Universitat Politècnica de Catalunya (UPC) in Barcelona. There, he received the Engineering Degree in Telecommunication (1982) and the PhD Degree in Computer Science (1987). Since 1983 he is lecturer at the Computer Architecture Department. His research topics are Broadband Communications and Applications. Since 1988 he has participated in RACE projects (Technology for ATD and EXPLOIT), in several Spanish Broadband projects (PLANBA: AFTER, TR1 and IRMEM), ACTS projects (INFOWIN, MICC, IMMP) and IST projects (LONG, E-NET, EuQoS, E-NEXT). Since 1995 he is researcher at the Advanced Broadband Communications Center of the University (CCABA). Currently is in charge of the i2CAT next generation Internet infrastructure project (GigaCAT).