Providing QoS in MPLS-ATM Integrated Environment[1]

Sergio Sánchez-López, Xavier Masip-Bruin, Josep Solé-Pareta,
Jordi Domingo-Pascual

Departament d'Arquitectura de Computadors, Universitat Politècnica de Catalunya, Barcelona (Spain)
{sergio, xmasip, pareta, jordid} @ac.upc.es

Abstract. Multiprotocol Label Switching (MPLS) is an advanced forwarding scheme, which allows the network to achieve the benefits provided by Traffic Engineering (TE) techniques. The establishment of an end-to-end LSP between two IP/MPLS networks interconnected through an ATM backbone is still an open issue. This paper focuses in an MPLS-ATM environment, and addresses the problem of providing a fast LSP establishment, with certain QoS (Bandwidth guarantee), between two MPLS subnetworks interconnected through an ATM backbone. The Private Network to network Interface (PNNI) is used in ATM backbone as a routing and signaling protocol. In order to achieve the paper objectives, new PNNI elements are defined and evaluated.

1. Introduction

The current Internet popularization and growth imposes the necessity of optimizing the use of the network resources. Moreover, a redefinition of the Internet capabilities is being required due to the request of services with a certain Quality of Service (QoS) together with the progressive increase of traffic.

Best effort services, provided by conventional Internet, are not enough to absorb the QoS requested by new emerging applications. In this way, mechanisms such as the Resource Reservation Protocol (RSVP), differentiated services model, Multiprotocol Label Switching (MPLS) [1], Traffic Engineering (TE) and Constraint-based Routing (CR) have been defined by the Internet Engineering Task Force (IETF) in order to allow applications such as high-quality videoconference, real-time audio, etc. to guarantee their performance features.

Traffic Engineering pursues to accommodate the offered traffic in a network in order to optimize the resources utilization while avoiding the congestion. To achieve this, some tools such as Constraint-based Routing and MPLS have been defined. The main CR goal is to compute the optimum route between two nodes using QoS parameters. MPLS is an advanced forwarding scheme based on the allocation of a short fixed length label at the IP packets. MPLS sets up a Label Switched Path (LSP). A label Distribution Protocol (LDP) is used to distribute labels among the nodes in the path.

Nevertheless, data transport in Internet is currently supported by a heterogeneous set of network technologies, which have to coexist. One of these, the Asynchronous

[1] This work was supported by the CICYT (Spanish Ministry of Education) under contract TIC0572-c02-02 and by the Catalan Research Council (CIRIT) 2001-SGR00226

B. Stiller et al. (Eds.): QofIS/ICQT 2002, LNCS 2511, pp. 215-224, 2002.

Transfer Mode (ATM) technology is widely deployed in the backbone and in the Internet Service Provider environments. ATM can provide Traffic Engineering, QoS and fast data forwarding. The PNNI (Private Network to Network Interface) is a standard routing protocol over ATM defined by the ATM Forum [2]. Moreover, a PNNI Augmented Routing (PAR) and Proxy PAR was defined in order to allow non-ATM information to be distributed through an ATM backbone.
This paper analyses the case of two IP/MPLS networks interconnected through a backbone network, assuming MPLS as the mechanism to provide TE in the IP networks, and ATM technology in the backbone. In such a scenario, the interoperability between both MPLS and ATM to achieve MPLS connectivity through the ATM backbone is still an open issue. More specifically, the problem is how to set up an end-to-end LSP with QoS guaranties between two LSRs located in different MPLS domains. In this paper we address this problem. The remainder of this paper is organized as follows. In Subsections 1.1 and 1.2 a new mechanism to distribute MPLS label through an ATM network based on integrating MPLS and ATM is suggested. Moreover, these Subsections define the necessary mechanisms to set up an end-to-end LSP between two LSRs belonging to different domains. In Section 2 some proposals to provide QoS in these LSPs are suggested. In Section 3 a performance evaluation is carried out. Finally, Section 4 concludes the paper.

1.1. MPLS and ATM Integration

The existing solutions described in [3,4] require either having an IP/MPLS Router on top of the ATM switches, or establishing a tunnel across the ATM backbone. 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 Augmented Routing (PAR)
As it was indicated in [5], PAR is a common work performed by ATM Forum and the IETF. PAR is based on the PNNI [2], which is a routing and signaling protocol used in ATM networks and defined by the Forum ATM. PAR allows non ATM information to be transported through the ATM network. This information is transparent to non PAR-capable ATM switches. Our objective is to use the PAR mechanisms in order to transport the MPLS information between MPLS subnetworks connected through an ATM backbone.
PAR uses specific PNNI Topology State Elements (PTSEs) in order to transport the non-ATM information. PTSEs are encapsulated within PNNI Topology State Packet (PTSP), which is flooded throughout the ATM backbone. PAR uses the PNNI flooding mechanisms. So far, PAR defines specifics PAR Information Groups (PAR IG) in order to describe IPv4 and VPN information. We consider that the ATM Border Switches (BSs) are PAR capable and the external devices (IP Routers, LSRs, etc.) are directly connected to BS. An External device has to register its information in BS in order to be distributed to other BSs. Moreover, each external device has to obtain the information from the BS, to which is attached. The Proxy PAR [5] is the protocol defined in order to allow external devices to register and obtain information to and from the BSs. A Proxy PAR works as client and server mode. While the client mode is a simple procedure installed on the external device, the server functions are performed on the BS.

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. This new PAR IG allows MPLS labels to be distributed through ATM backbone. Concerning the BR architecture, we suggest to use a Proxy PAR Capable Label Switching Router (PPAR-LSR) consisting of the following elements: A 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. An ATM Switch as a Proxy PAR server. And a forwarding table in order to establish a relation between MPLS labels an ATM outgoing interfaces.
A more detailed description of this solution can be found in [6].

1.2. LSP Establishment in MPLS/ATM Environments

Let us consider the scenario depicted in Fig. 1. Once we have a solution to transport MPLS information from a MPLS subnetwork (ND1, Network Domain 1) to other MPLS subnetwork (ND3, Network Domain 3) through an ATM backbone (ND2, Network Domain 2), the next step is to set up an end-to-end LSP between two LSRs, each one in a different MPLS subnetwork. The solution suggested in Subsection 2.1 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. The problem appears when there are more than one LSR until to reach the destination LSR, i.e. the destination LSR belongs to a MPLS subnetwork (e.g. LSRy in ND3). In this situation is needed to use a LDP in ND3 in order to set up the LSP until the destination LSR. In order to solve this problem, we suggest proceeding as follow [7]:

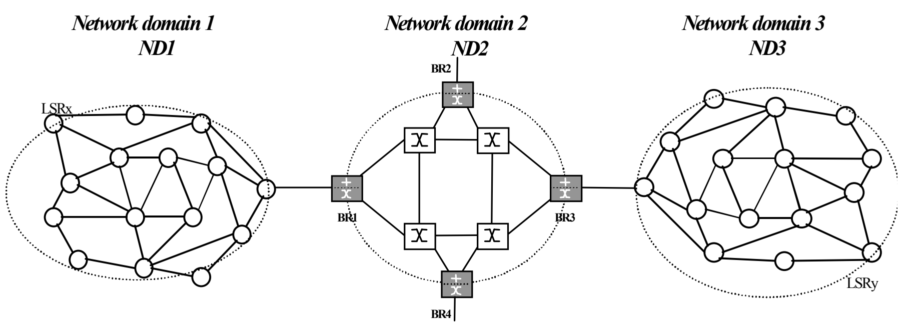

Fig. 1. Network Topology scenario

The Proxy PAR client on each BR registers the MPLS protocol along with labels, and all the address prefixes, which can be reached. Every server bundles its state information in PAR PTSEs, which are encapsulated within a PTSP, which is sent to a neighboring peer.

Using the received PAR MPLS devices Definition IG, every server side generates an MPLS topology database. Each client side will use the query protocol to obtain information about registered services by other clients.

LSRx decides sets up a connection with LSRy. Therefore, an LSP has to be created between both nodes. An RSVP-TE is used as LDP to distribute labels along the path.

We assume that an Ordered LSP Control performs the establishment of the LSP. Because of the implemented functions in the BRs both MPLS networks acts 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 BR1 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 reach ingress BR1 in ND2, RSVP Resv message is returned to the source LSRx with the requested label. Simultaneously, PNNI signaling protocol sets up a VC from the ingress BR1 to the egress BR3. A UNI 4.0 signaling subset is used by the PNNI as signaling mechanism. Therefore, a call set up is triggered in the ingress BR1 when the RSVP path message reaches the ingress BR1. A SETUP message is sent to the egress BR3 in order to establish the connection. Just as was mentioned in Subsection 2.1, 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 [7] 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 (0x60), which corresponding to MPLS. An Identifier Type with a value (0x02) assigned to Resource. In this case, the identifier corresponds to the MPLS VCID [3]. However, in our case the VCID is not necessary because we have used the PAR mechanisms, explained in Subsection 2.1, in order to distribute MPLS labels. Thus, we proposed replacing the VCID field with the destination LSR IP address. In this way, when the SETUP message reaches the egress BR3, it has enough 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. At the moment that the path is set up, the ND2 behavior is as if it was a unique LSR, called Virtual LSR (LSRV). This node is the last node along the path in ND1, while in ND3 it is responsible of setting up the path towards a destination node LSRy, carrying out the source node functions

This method, in addition to provide an end to end LSP establishment, optimizes the time of setup in comparison with the methods presented in [3,4].

2. Providing QoS

The conventional way to compute routes in the MPLS domain is using a dynamic routing protocol, such as Shortest Path First (SPF). Since SPF does not take in consideration any QoS attributes (e.g. available bandwidth), LSPs computed by this protocol could lead the network to congestion situations. Although through longer paths (more hops counts), using TE based routing algorithms congestion could be avoided and network resources optimized. One tool to achieve this is the Constrain-based Routing (CR). The main goals of CR are both selecting routes with certain QoS requirements and increasing the network utilization. This Section is devoted to apply the CR concept to our proposal in order to reduce the congestion effects in the different MPLS domains composing our scenario.

Let us to consider the scenario shown in Fig. 1, and assume that the ATM backbone is able to set up a VC between the ingress and the egress BR 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 resource information are considered perfectly updated in all the LSRs.

According to this, the mechanism to be used for setting up an LSP with Bw guarantees works as follows: 1) LSRx requests an LSP in order to set up a connection from LSRx to LSRy. 2) The required amount of Bw is requested in this LSP. Therefore, in order to compute the optimal route, the LSRx (step 3) computes the optimal route applying the Widest-Shortest Path (WSP) routing algorithm. The WSP calculates the path with the minimum hop count (shortest) among all the possible paths, and in the case that there is more than one with the same minimum hop count, it selects the path with maximum Bw available. 4) Once the WSP algorithm has selected the path, the LDP triggers a mechanism in order to distribute labels. Thus, an RSVP Path message is sent to LSRy. 5) When RSVP Path message reaches the ATM ingress BR1, it finds a label associated to the destination IP address (in this case, a subnetwork IP address). Then, an RSVP Resv message is returned to LSRx with the label and a relation <label in, label out> is set up in each LSR of the path in order to establish the LSP. 6) Simultaneously, the ingress BR1 triggers a SETUP message with the new GIT proposed in Subsection 2.2 with a Generic Identifier Element containing the destination IP address. Therefore, when the SETUP message reaches the egress BR3, the LSR side of the BR3 can use the destination address in order to compute the LSP. Now, assuming that within the ATM network a VC linking BR1 and BR3 with the amount of Bw requested by LSRx could be set up, the routing algorithm existing in the ND3 (step 7) calculates a LSP in this domain to reach the final destination. 8) Once this path is computed, a RSVP Path is sent to LSRy. The setup LSP will be finished when the RSVP Resv will return from LSRy to the egress BR3. 9) Simultaneously, the egress BR returns a CONNECT message to the ingress BR1 in order to establish the corresponding ATM VC.

At the point of the process, where the SETUP message reaches the egress BR3 (step 6) the routing algorithm used to calculate the path within the remote domain (ND3) cannot be a CR algorithm because the SETUP message do not contain any QoS parameters. So then routing algorithm, such as the SPF, should be used there. As a consequence, the LSP set up in ND3 could not offer the Bw requested by LSRx and in case of congestion in that LSP the loss of information would be unavoidable (no QoS is providing).

Obviously the ideal case is using the same constrain-based routing algorithm (for example WSP) in ND1 than in ND3. Therefore, the egress BR has to know, in addition to the destination IP address, the Bw requested by the LSRx. For fixing this issue, we suggest adding a new Identifier Type in the Generic Identifier Element proposed in Subsection 2.2. This can be easily done because an Identifier related standard/application may have multiples Identifier Type and the number of identifiers is limited by the GIT maximum length, which is 133 octets (see Subsection 4.1 of [4]).). We suggest to define the following identifier: Identifier Type = Resource, Identifier Length = 3 octets and Identifier Value = requested BW in Mbps. The format of the new GIT is shown in Fig. 2.

Information Element Identifier= Generic identifier transportIE (0x7F)			1
Ext	Coding standard	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
@IP MPLS node destination (4 octets)			11
Identifier type=Resource(0x02)			12
Identifier length=3 octets(0x03)			13
Requested BW (3 octets)			16

Fig. 2. Generic Identifier with BW

In this way, when the RSVP Path message reaches the ingress BR1, the destination IP address and the requested Bw by LSRx are both transferred to the ingress BR1. These values are carried in the GIT and they are transported by the ATM SETUP message. Now, the egress BR3 has a value for a QoS parameter, which allows it to compute the path using a CR (as for example WSP) algorithm. Now we can come out with to different situations, namely finding a route with the requested Bw, or not finding it. As a consequence the last step of the algorithm described above, (step 9) has to be reformulated as follows: 9) If within ND3 a LDP with the requested Bw could be set up, the egress BR3 would return a CONNECT message to the Ingress BR1 in order to establish the corresponding ATM VC. If that LSP could not be found, the egress BR3 would return a RELEASE message to the ingress BR1, the VC would not be created, and a RSVP ResvTear message would be sent from the ingress BR1 to the LSRx in order to reject the connection request.

2.1. An Optimized Solution

The solution proposed so far still has a drawback, which is the possibility of wasting a lot of time. This happens when, after establishing the first segment of the LSP (the one of the subnetwork 1) and the VC through the ATM backbone, the end-to-end path cannot be completed due the impossibility of finding a LSP with the required Bw in the second subnetwork. In order to overcome this drawback, we suggest an optimized solution consisting of define a mechanism able to distribute information on the whole network topology and available resource between all the MPLS subnetworks interconnected through the ATM backbone. In this way, the subnetwork 1 would know in advance the possibility to establish an end-to-end path. Nevertheless, since Network Operators commonly do not like to share such an amount of information with others Network Operators, this solution have no chance to be applied when the different MPLS subnetwork belong to different administrative domains.

Let us consider again the scenario shown in Fig. 1 and assume that the egress BR3 has a Traffic Engineering Database (TED) where the ND3 topology and available resource information is kept. 15 nodes and 27 links compose the MPLS subnetwork topology. We only consider the link available bandwidth as available resource. Therefore, the TED could be 27x2 matrix where a metric (available Bw) is assigned to a link (couple of nodes). A node is represented as a natural number between 0 and 14. Moreover, the considered range is between 0 and 2.5Gbps where the metric is indi-

cated in Mbps. Thus, the couple link-metric could be represented by 4 bytes and, consequently, the TED size is 108 bytes. We assume that the relation between the natural number assigned to a node and the node IP address is known by the BRs. Therefore, the node IP address is not included in the TED.
In order to transport the TED through the ATM backbone, we suggest a new PAR PTSE named PAR MPLS TED Service Definition IG. The element format is shown in Fig. 3.

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)**

Fig. 3. PAR MPLS TED Services definition IG

Now 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 TED. Then Proxy PAR server floods the PTSEs throughout the ATM backbone. Once the flooded information reaches all the BRs, each one uses the Proxy PAR client to obtain that information. Considering the Fig. 1 as scenario, ND1 distributes the ND3 TED using its own flooding mechanisms.
From here on, LSRx computes the end to end path using both ND1 and ND3 TED. The CR algorithm uses ND1 TED to compute the path from LSRx to ingress BR1 and it uses the ND3 TED to compute the path from Egress BR3 to LSRy. If there are both routes then a RSVP Path messages is sent to ND3. RSVP path message contains the Explicit Routing Object (ERO) in order to contain the routes. The ND3 route is distinguished from ND1 route by the ND3 subnetwork IP address.
Next, the ingress BR1 triggers a SETUP message in order to set up a VC. We can observe that now it is not necessary to include the required bandwidth in the SETUP message because the ND3 path has been computed using the required BW. Therefore, the GIT has only to include the computed path. A new GIT format is suggested and it is shown in Fig. 4.
Finally, when SETUP message reaches the egress BR, the route included in the messages is used by the RSVP Path message in order to set up the LSP.
We have considered that the TEDs are always updated. Therefore, the path has always the Bw specified in the TEDs. Nevertheless, this is not always the case. It may be possible that the TED would not be updated, for example due to fast changes in the link states. The impact of possible network information inaccuracy in the mechanisms proposed in this paper will be a subject for further studies. Preliminary we have sug-

gested in [8] a new QoS routing algorithm in order to reduce LSP blocking probability and the routing inaccuracy without increasing the routing control information overhead and avoiding rerouting.

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
@IP MPLS node destination (4 octets)			11
Identifier type=Resource(0x02)			12
Identifier length=5 octets(0x05)			13
PATH			18

Fig. 4. Generic Identifier with PATH

3. Performance Evaluation

The topology used to evaluate the routing algorithms behavior is shown in Fig.1, where two different link capacities are used. Links represented by a light line are set to 12Mbps, and links represented by a dark line are set to 48Mbps. The incoming requests arrive following a Poisson distribution and the requested bandwidth is uniformly distributed between two ranges, (16-64kbps) to model audio traffic and (1-5Mbps) to model video traffic. The holding time is randomly distributed with a mean of 120 sec. Finally, the existent Topology and available resources database in each border router is assumed perfectly updated.

Results shown in Fig. 5 have been obtained using an extension of the ns2 simulator named MPLS_ns2 developed by the authors of this paper. In this way, in Fig. 5 the LSP Blocking ratio as a function of the traffic load is analyzed. This value is defined according to expression (1):

$$LSP\_Blocking\_Ratio = \frac{\sum_{i \in rej\_LSP} LSP_i}{\sum_{i \in tot\_LSP} LSP_i} \qquad \textbf{(1)}$$

where rej_LSP are the set of blocked demands and tot_LSP are the set of total requested LSPs.

Two different routing algorithms are applied in order to demonstrate the QoS parameters influence in the number of rejected LSPs. Thus, the SPF and the WSP behavior are shown for audio and video traffic.

The WSP selects the route in accordance with the residual bandwidth. The SPF does not take into consideration any QoS parameters to select the path. Thus, SPF algorithm does not block any LSP and all the traffic runs by the same shortest path. This situation produces congestion when the Bw of the traffic is higher that the Bw of the path. Therefore, all the LSPs requested from this moment will lose information. In our simulations we consider that the LSPs computed by the SPF, which lose information due to the congestion, are as blocked LSPs. Fig. 1 exhibits the goodness of the WSP in front of the SPF algorithm. We can see that the blocking is lower when a QoS parameter is included in the path selection process, that is, when the WSP is used.

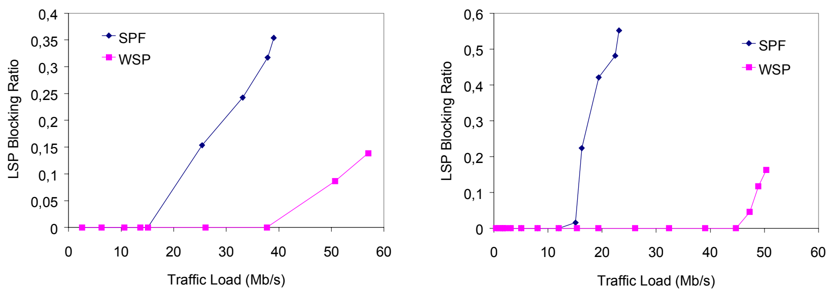

Fig. 5. LSP Blocking Ratio for audio and video traffic load

Once the solution proposed in this paper is implemented in the topology shown in Fig.1, we evaluate the following situation. The first segment of the LSP (the one of the ND 1) and the VC through the ATM backbone are established. Then, the end-to-end path cannot be completed due the impossibility of finding a LSP with the required Bw in the network domain 3, ND3. The results of several simulations are presented in Fig. 6. Three simulation groups have been performed. Each group corresponds to the average of five simulations. Each simulation differs from other in which is the destination node.

In this evaluation, rej_LSP are the set of blocked demand due the impossibility of finding a LSP with the required Bw in the network domain 3, ND3 and tot_LSP are the set of total requested LSPs.

The optimized solution presented in Subsection 3.1 avoids this situation because the source node in ND1 has a TED with the topology and available Bw of the ND1 and the ND3. In this way, the route computed by the routing algorithm will be always established avoiding the LSP blocking.

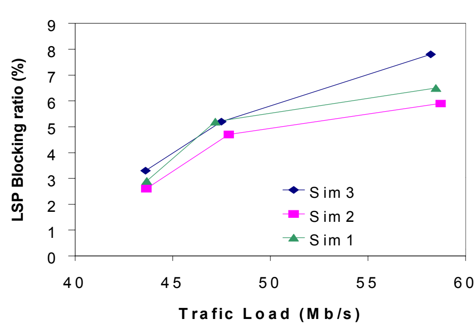

Fig. 6. LSP Blocking Ratio due the impossibility of finding a LSP with the required Bw in the ND3

4. Conclusions

Different approaches have been considered in order to achieve the interoperability between MPLS and ATM technologies. Firstly, the distribution of MPLS information through an ATM backbone has been solved either using ATM-LSRs in the ATM network or tunneling a LDP through an ATM 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 signaling protocol through an ATM cloud. In this paper a more appropriate and general solution is proposed. This solution is based on the MPLS and ATM integration using the PNNI Augmented Routing. Therefore, a new Border Router architecture, a new PAR PTSEs and a new Generic Information Element have been defined.

Secondly, once the MPLS information has been distributed throughout the BRs, a new mechanism in order to set up an end to end LSP between two different MPLS domains, which are connected via an ATM backbone is suggested.

Finally, a method in order to apply the Constraint-based Routing concept to this scenario is proposed. As a consequence, we have achieved to set up an end to end LSP with the required QoS.

References

1. E.C. Rosen. A. Viswanathan and R. Callon, Multiprotocol Label Switching Architecture, IETF RFC 3031, July 2000
2. ATM Forum, Private Network-Network Interface Specification Version 1.0, af-pnni-0055.000, March 1996.
3. K. Nagami, Y.Katsube, N. Demizu, H. Esaki and P. Doolan, VCID Notification over ATM link for LDP, IETF RFC 3038, Jan.2001.
4. 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, Jan.2001.
5. ATM Forum, PNNI Augmented routing (PAR), Version 1.0, af-ra-0104.000, January 1999.
6. S. Sánchez-López, X. Masip-Bruin, J. Domingo-Pascual, J. Solé-Pareta, A Solution for Integrating MPLS over ATM, in Proc 15th International Symposium on Computer and Information Sciences (ISCIS2000), Oct. 2000, pp.255-303.
7. 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
8. X. Masip-Bruin, S. Sánchez-López, J. Domingo-Pascual, J. Solé-Pareta, Reducing LSP Blocking and Routing Inaccuracy by Using the BYPASS Based Routing Mechanism, Internal Report, UPC-DAC-2001-41, December 2001.