Testbed Implementation of Control Plane Extensions for Inter-Carrier GMPLS LSP Provisioning

Aruna Prem Bianzino, Jean-Louis Rougier, Stefano Secci
Telecom ParisTech, Paris, France
{bianzino,rougier,secci}@telecom-paristech.fr

Ramon Casellas, Ricardo Martinez, Raül Muñoz
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Castelldefels, Barcelona, Spain
{ramon.casellas, ricardo.martinez, raul.munoz}@cttc.es

Nabil Bachir Djarallah, Richard Douville, Hélia Pouyllau
Alcatel-Lucent Bell Labs, Paris, France.
{nabil_bachir.djarallah, richard.douville, helia.pouyllau}@alcatel-lucent.com

***Abstract*—This paper presents a testbed implementation of an inter-carrier GMPLS (Generalized Multi Protocol Label Switching) service architecture recently proposed. This architecture couples the Path Computation Element (PCE)-based control plane with a service plane managing discovery, composition and activation functions of inter-carrier service elements. The testbed implements the required PCE Communication Protocol (PCEP) and Resource Reservation Protocol with Traffic Engineering (RSVP-TE) extensions, together with service request filtering operations performed with a policy based architecture[1].**

I. Introduction

A dynamic routing architecture suitable for inter-carrier, connection-oriented, service provisioning has not been implemented yet, mainly because of privacy, billing and monitoring issues. However some important steps in this direction are being made. The Internet Engineering Task Force (IETF) has defined an extension to the GMPLS technology, called inter-Autonomous System (AS) GMPLS, which enables the establishment of inter-carrier, explicitly routed connections with stringent quality of service (QoS) and availability constraints [1]. Recently, the authors of [2] have proposed additional extensions to the GMPLS technology in multi-AS environment, in order to enable automatic provisioning of inter domain TE services. The idea is to introduce a distributed inter-carrier *service plane*, coupled with a PCE-based control plane, through which carriers interact by discovering carriers' service elements, by composing the service elements into a multi-domain service, by instantiating and enabling the service, and finally by triggering management and network plane operations to finally establish and maintain the connection. In this framework, routing is source-based at the AS-level and distributed at the router-level [3]. Some form of cooperation among carriers is needed to over-ride privacy, billing and monitoring issues and for managing service-related data. Hence, we believe that the proposed architecture can be very interesting within a carrier alliance for instance.

[1]Work partially funded by the INCAS S.JRA of the EU IST Euro-NF Network of Excellence, the CELTIC TIGER2 and ANR ACTRICE projects.

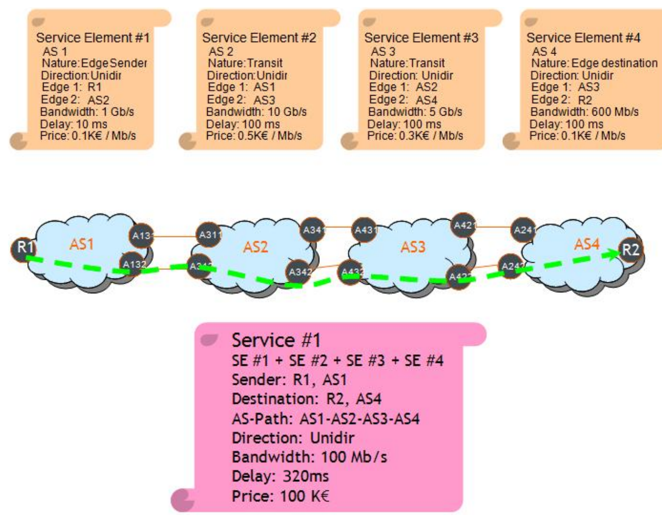

Fig. 1. Inter-carrier network service as composition of service elements [2]

In this paper, we describe the inter-AS GMPLS testbed implementation conducted at the CTTC facility in order to validate of the aforementioned architecture.

Section II resumes the inter-AS GMPLS architecture and the extension propositions. Section III provides an overview of the reference topology. Section IV details the testbed implementation steps, describing path computation and path signaling issues, and related interworking functional aspects. Finally, Section V concludes the paper. To simplify, in the following the terms carrier, AS or domain are used interchangeably.

II. The inter-AS GMPLS architecture

The GMPLS architecture allows establishing Label Switched Paths (LSPs) within carrier boundaries. The GMPLS protocol family intrinsically includes TE features, enabling to route LSPs explicitly taking TE constraints into account. Further extensions, detailed below, support the configuration of inter-AS LSPs [4].

A. Inter-AS LSP signaling

The RSVP-TE signaling protocol [5] is used to establish GMPLS LSPs. The inter-AS LSP signaling can be done in three different ways:

- LSP Nesting: An intra-domain LSP is used between domain border routers to transport inter-domain LSPs sharing a common intra-domain subpath.
- Contiguous LSP: A single end-to-end LSP is signaled across the domains. There is a single signaling session between the head-end router and the tail-end one.
- LSP Stitching: In this mode, the local intra-domain LSPs are signaled separately, and then stitched at the boundaries to form a single inter-domain LSP.

In our testbed we implemented the contiguous LSP mode.

B. Inter-AS LSP computation

An LSP is to be signaled over a pre-computed (router-level) path. A head-end router has full topology visibility within its domain boundaries, hence, can only compute an end-to-end intra-domain path, but not an end-to-end inter-domain one. Two methods can be adopted for the inter-AS path computation:

- The per-domain path computation method. The source or ingress router determines the next domain and the ingress router in the next domain, and computes the corresponding subpath. Then the path computation is moved to the ingress router of the next domain (by the signaling protocol), and so on up to the tail-end router. This simple method does not allow computing a shortest inter-domain path and can lead to several crankbacks that might affect the stability of the control plane.
- The cooperative PCE-based path computation method. It takes as input the AS chain - i.e., the succession of ASs to be crossed - and relies on computation entities present in each AS, the PCEs, to collaboratively compute an inter-AS shortest path along the given AS chain.

As highlighted in [2], the cooperative PCE-based method shall be preferred to allow a composed end-to-end service billing. In the PCE-based architecture [6], the PCEs serve requests sent by Path Computation Clients (PCCs) - i.e., routers or switches - using information in the local TE database. A PCE can query the PCEs of other domains to collaborate in this computation, acting in turn as a PCC; a PCE communication protocol (PCEP [7]) was defined to relay these request and answer messages. In the inter-domain path computation context, the Backward Recursive Path Computation (BRPC) [8] seems to be the procedure that meets best the operator and the supplier requirements in terms of complexity and network information hiding. It consists of computing iteratively, at each PCE of the AS chain and starting from the tail-end AS, an inverse tree of constrained shortest paths, with one branch for each ingress AS Border Router (ASBR) - and toward the destination. The tree is sent back to the previous AS, which does the same, and so forth up to the source AS. Obviously, at least one PCE is required in each domain. No TE information exchange is required between PCEs.

C. Service plane related extensions

Hence, the IETF developed solutions for inter-AS LSP set-up. However, some missing points are needed for the deployment of inter-carrier GMPLS network services. First, for the PCE-based architecture, the standardization does not indicate how the input AS chain is calculated. Then, being the set-up of an inter-carrier tunnel subject to strong business, security, and confidentiality aspects, a trusted multi-carrier service architecture would be needed to ensure billing, and to manage routing and signaling requests at provider boundaries. Such procedures are beyond the scope of the IETF, but have been defined within the ACTRICEproject. Douville, Le Roux, Rougier and Secci in [2] introduce the notion of a inter-carrier service plane and structure the lifecycle of an inter-carrier GMPLS LSP service with seven functional steps. The service plane assembles carriers (alliance flavor) interested in settling inter-AS network services; it manages inter-carrier service elements through which each carrier announces its service offer in terms of Service Level Specifications (SLSs) and potentially according to an adopted business model of monetary costs. The authors indicate the IPsphere Forum framework as a potential framework within which implementing such a service plane. It is worth briefly resuming the proposed seven functional steps in [2], highlighting those directly implemented in our testbed:

1) *Service Discovery*: The inventory of all the service elements offered by the carriers of the alliance is acquired.
2) *Service Elements Composition*: The service plane is asked for a constrained shortest AS path (for point-to-point tunnels) or AS tree (for multipoint tunnels). An ad-hoc routing algorithm has been defined in [3]. It consists of a composition of service elements - following the idea described in [9] - at the source AS. An example of service element composition - for a LSP from node R1 to node R2 - is depicted in Fig. 1.
3) *Service Instantiation*: The point availability of the service elements composing the AS chain is verified. A Service Identifier (SID) is generated to identify the service during its lifecycle, and distributed among the involved ASs through the service layer. Every involved AS sends back a message to grant/refuse the availability of the required service element, and to possibly negotiate some SLSs or (if allowed) the cost of the service.
4) *Service Activation*: This step consists of triggering the service establishment: an activation message is distributed within the service plane to all the ASs, including also the SID. Then, the service plane sends to the management plane the filtering policy associated to the SID, useful to filter future inter-AS PCEP and RSVP-TE messages. If this is successful, the management plane configures the head-end router in the network plane, establishing the inter-AS LSP, passing the SID, the AS chain, and the TE parameters.
5) *Path Calculation*: This step consists of computing the inter-AS path at the network plane. Since it is active part of the testbed implementation, this step is detailed

in the following.

6) *Service Signaling*: At this step the inter-AS LSP is contiguously signaled at the network plane. This step is detailed in the following too.
7) *Service Maintenance*: After the inter-AS LSP establishment, it may fail or be closed. A particular protection strategy may be provided in case of failure. If a failure cannot be recovered, a status message is sent to the service plane and the source AS is notified and may proceed with a new service request.

In the rest of the paper, we focus on the network plane (i.e. essentially the last steps) and its interaction with the management and the service planes. We detail how these steps have been implemented in our GMPLS testbed.

III. The reference topology

We chose the testbed scenario so as to include four ASs and many alternative AS paths between them (meaning with AS paths or AS chains the list of ASs to be crossed to connect a source to a destination router or AS). Moreover, we chose to use different parallel inter-AS links between AS pairs in order to arise possible flaws of the proposed solution.

Following these guidelines, the chosen reference topology is shown in Fig. 2. The PCE deployment model is based on collocating a PCE in one of the nodes at each AS.

The testbed has been built upon the existing CTTC ADRENALINE+ (All-optical Dynamic REliable Network hAndLINg IP/Ethernet Gigabit traffic) testbed, presented in [10] and [11].

Fig. 3. The ADRENALINE+ testbed

On this existing testbed, there was no Exterior Gateway Protocol (EGP) implementation available since it has originally been conceived for intra AS uses only. Hence, in order to implement a multiple AS scenario and inter-AS links, we decided to insert static routing information on the ASBR in order to guarantee inter-AS reachability.

IV. Implementation steps

The testbed implementation plan has been divided into steps, each one facing a logically separated problem. In summary, the main issues have been:

- The extension of the starting RSVP-TE implementation for the multi-AS environment;
- The inclusion of the SID object into RSVP-TE, and its management;
- The extension of the existing implementation of BRPC from the intra-AS inter-area to the inter-AS scope;
- The inclusion of the SID object into PCEP, and its management;
- The PCEP - RSVP-TE interworking aspects;
- The implementation of a separate policy management module to implement the policy manager functionalities.

In the following we detail each implementation step that was needed to test the theoretical architecture and to solve the above mentioned issues.

A. RSVP-TE extension to the inter-AS scope

Our first focus has been on issues related to inter-AS path signaling itself. As a first step, we thus deployed a scenario using per-domain path computation (see section II), instead of the cooperative PCE-based approach. In the per-domain path computation, we need to cope with the inter domain visibility issue - that is, any node can only count on the local knowledge of its routing domain, given by OSPF-TE. The initial Explicit Routing Object (ERO) used by the inter-AS RSVP-TE Path message only contains the strict list of unnumbered interface ID subobjects (nodeID, interfaceID) down to proper egress ASBR and then the final destination nodeID as a “loose” subobject. Once the RSVP-TE message reaches the ASBR of the domain, the ERO (containing at this point only the destination as “loose”) has to be expanded once in order to include the next hop as “strict” (nodeID of the ingress ASBR of the next AS). At the ingress ASBR of the following AS, the ERO has to be expanded again to include the strict list of subobjects down to the proper egress ASBR of the actual AS (or to the final destination in the case of destination AS), and so on up to the destination. This iterative procedure is depicted in Fig. 4 for a LSP-TE signaling with RSVP-TE from node N1 to node N4. Every node receiving an RSVP-TE PATH message erases itself from the contained ERO, and then looks for the next strict hop (if present), or expands the ERO with the above explained procedure, in the case of an ASBR. This issue was not considered on the existing testbed implementation of RSVP-TE since it was designed for multi-area, intra-AS scenarios. Indeed, in that case it was possible to rely on the area border routers belonging to multiple areas and thus having visibility on the topology of each area belonging to; such nodes are thus able to expand the ERO of the received packets that have to cross an area boundary.

B. RSVP-TE extension with the Service ID object

As stated in [2], it is needed that RSVP-TE transport a Service ID object in order to apply policy management during the LSP establishment. Such an object was initially proposed in [12] as an “order ID” to be carried by the PCEP. Then in [2] it has been redefined as “Service ID (SID)” to be

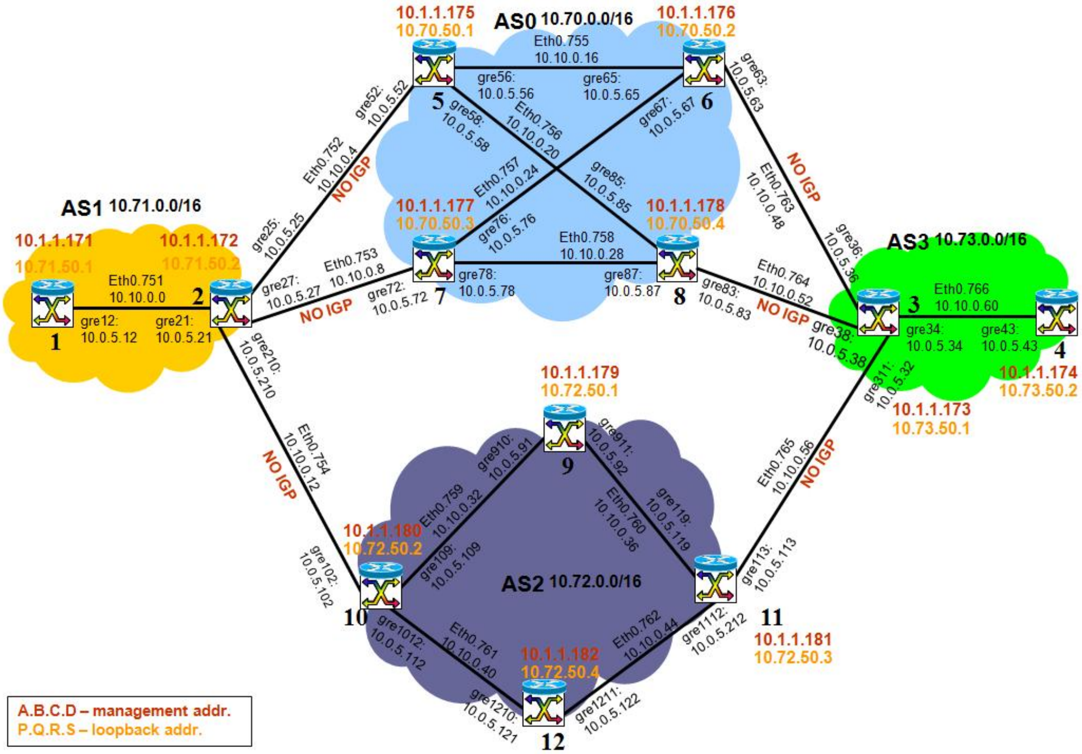

Fig. 2. The reference topology

carried by PCEP and RSVP-TE. It is an integer number that every RSVP-TE PATH message has to carry and every ingress ASBR has to check when receiving an inter-AS RSVP-TE PATH message, in order to retrieve the associated context and to validate/reject the received path establishment request: it allows discriminating between authorized inter-AS RSVP-TE PATH messages and unauthorized attempts, guaranteeing that only the first ones are processed, while the others are rejected.

The SID values are local to each AS, and are first exchanged during the aforementioned service instantiation phase. Hence, a multi-AS service is identified by many local SIDs, one per AS[2]. To model the processing of SIDs by a given entity, a simple rule based on a function which is applied to the Label Switched Path Identifier (LSPID) and, on the basis of the destination node, can compute the SID to be used, and to be coherently checked at the reception. In Fig. 5 an example depicts the evolution of SID value for an LSP request from node N1 to node N4 and a LSPID value of 7.

C. BRPC algorithm extension to the inter-AS scope

The BRPC path computation procedure was implemented in the ADRENALINE testbed. However, it was designed for the multi area, intra-AS scenario case only, with a PCE in each area. In this environment, border routers belong to multiple areas and have interfaces on each one while, in the multiple AS environment, the ASBRs are part of a single domain with the

[2]In [2] a single SID is depicted for the sake of clarity.

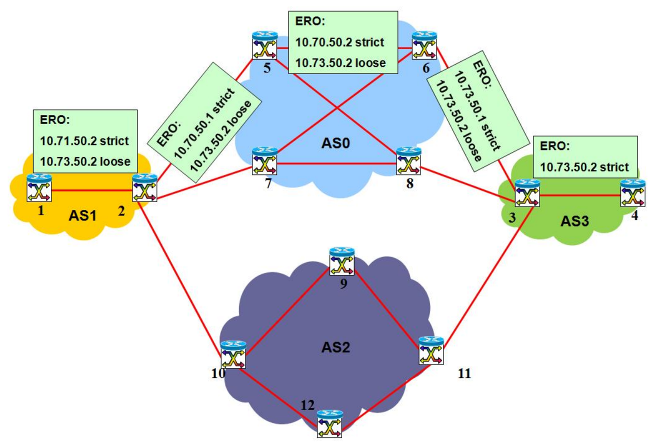

Fig. 4. Example of ERO expansion for a path from N1 to N4

inter-AS links not belonging to any domain, neither announced by any IGP instance.

In order to overcome this issue, when an upstream PCE receives a tree of paths routed to the destination from the following downstream PCE, every branch starting at a given downstream ASBR is mapped to the proper inter-AS link and then to corresponding egress ASBR of the upstream domain. To do so, we use the aforementioned static information regarding the inter-AS links in order to extend and prune the BRPC tree branches in a first step. In this way, the tree of

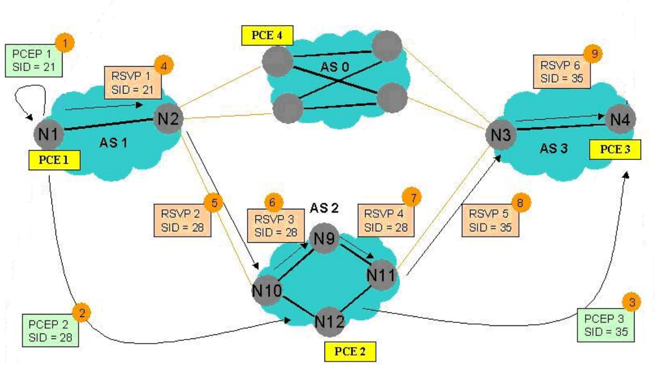

Fig. 5. SID lifecycle in RSVP-TE and PCEP

paths can be treated by the current PCE exactly as if we were in a multi-area context, reusing the same implementation of the BRPC algorithm. This approach supports the presence of multiple parallel links between AS pairs: if an incoming path can be mapped to more than one of the ASBRs in the upstream domain an extended path to each of the ASBRs including its (nodeID, interfaceID) is created. All the created paths are then added to the solution set, which is sorted using the total path metric so only the best is kept, while the others are pruned consistently with the BRPC algorithm. Fig. 6 depicts an example of the above described procedure for a path computation from node N1 to node N4 with an AS chain equal to (AS1, AS0, AS3).

Another issue to be taken into account is the possibility of receiving a path computation request from a node to itself, situation that triggers an error and a corner-case that did not appear in the multi-area BRPC testbed implementation since, if the destination is an area border router, it is seen as the "last" node of the upstream area. However, in a multi-AS environment, if the destination of the path to be computed is an ingress ASBR, the destination PCE is asked a path computation from the ASBR to the ASBR itself. This particular case has been taken into account and has been overcome with proper check and exception additions.

D. PCEP extension with the SID object

According to the proposed architecture, and in order to be able to apply policy management during the path computation, every PCEP PCReq message has to carry a SID object and every PCE has to check it in order to validate/reject a received path computation request: it allows discriminating between authorized PCReq messages (to be processed) and unauthorized attempts (to be discarded). The SID computation is performed similarly than for the inter-AS RSVP-TE extension (Fig. 5). The PCEP implementation has been extended to include the object with proprietary Ctype objects. Fig. 7 shows the SID object as seen in a capture of a PCEP PCReq by the PCE1 for the PCE2, concerning a path computation request for a path going from node N1 to node N4. The capture has been obtained using a version of the Wireshark tool properly extended in order to be able to correctly decode the new object.

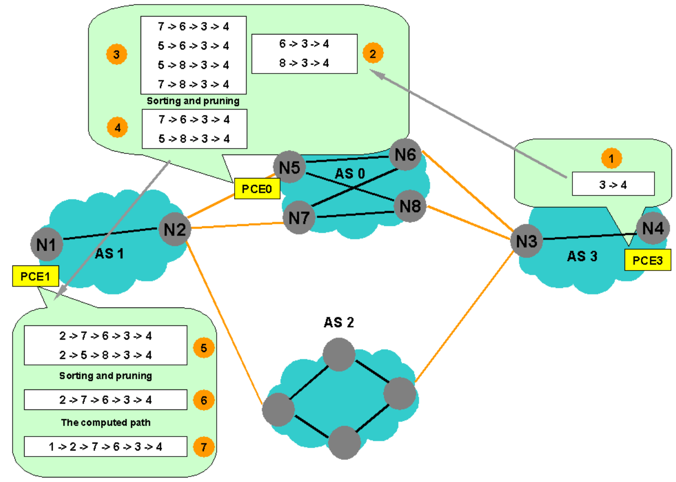

Fig. 6. Received ERO mapping procedure

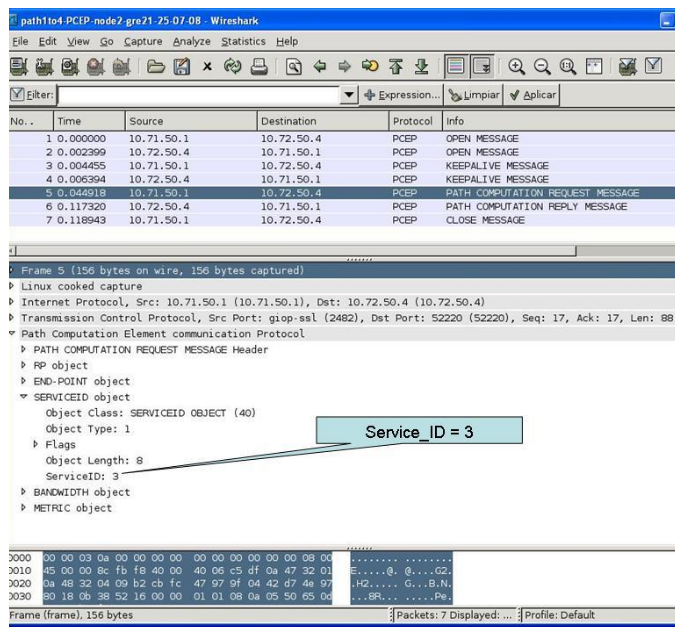

Fig. 7. Wireshark capture of a PCEP packet with the SID object

As highlighted by the zoom, the SID object contains the value assumed by the SID for the particular request.

E. PCEP and RSVP-TE interworking

When a node requests an LSP, it queries the local PCE for the ERO to be used in the RSVP-TE path signaling procedure. The ERO is computed by the PCEs with the above mentioned procedure and returned to the requesting node. This allows forcing policies on the followed path differently from the case of simple OSPF-TE use. Moreover, this allows a global path computation on a specific AS chain given by the service plane. Hence, now the optimal "strict" ERO (i.e., the complete list of sets (nodeID, interfaceID) representing the crossed nodes in the computed constrained path from the source toward the destination) is returned by the PCE and used by RSVP-TE to reserve a path and establish the LSP.

F. Policy management module

Meeting the requirements highlighted in [2], now being formalized in [13], the policy management functionalities have to be implemented in a module, named Policy Decision Point (PDP). This module, running on an external node, has to reply to the following types of request:

- SID request by PCC: a PCC requesting path computation asks to the PDP for the SID to be used in the request.
- SID and TE parameters check request: a PCE required for a path computation asks to the PDP if the parameters included in the request are coherent with the negotiated ones. This check includes a control on the LSP ID[3]; it must belong to a negotiated interval, i.e., the TE parameters (in our testbed we used just the bandwidth) must correspond to the negotiated ones for the LSP ID; moreover, the SID on the request must follow the right computation rules.
- SID request for RSVP-TE PATH message: a node sending a RSVP-TE PATH message asks to the PDP for the SID to be used in the PATH message.
- SID check for RSVP-TE PATH message: a node receiving a RSVP-TE PATH message asks to the PDP if the SID used in the PATH message is consistent with the negotiated one. This check includes the control on the LSPID already mentioned above.

An ad-hoc protocol has been designed to handle the communications between the PDP and the different Policy Enforcement Points (PEP).

V. Perspectives

Currently, the MPLS-TE and GMPLS technologies are deployed mainly within carrier boundaries to support real-time and interactive services. The extension of these services in the inter-carrier scope requires supporting inter-AS QoS guarantees between carriers. The IETF has worked on extending existing protocols and architectures required to set-up inter-domain connections. The ACTRICE project has defined the blocks to be added in order to automate inter-AS services [2]. However, a working implementation of such an extended architecture was missing to validate the architecture functioning.

This paper reports an inter-AS GMPLS testbed implementation for the multi-carrier service plane architecture proposition of the ACTRICE project. It is devoted to the automatic provisioning of inter-AS GMPLS services. This experimentation extended the existing CTTC ADRENALINE+ testbed to the inter-carrier scope, integrating the service architecture-related data (AS chain, "Service IDentifier" object and inter-AS service TE metrics) and the required protocol extensions (RSVP-TE and PCEP extensions to carry the SID object). The main contributions are the integration, the testing and the validation of the Service IDentifier object in the PCEP and RSVP-TE, and the related management by the PDP entity.

[3]in our testbed implementation, LSP ID and Tunnel ID take the same value, even if normally can take different values.

We aim at overcoming the following issues in further work:

- The current BRPC implementation suffers from a non-confidentiality issue when an explicit list of crossed equipments is shared with other domains. A possible solution is described in [14], where it is proposed to substitute the explicit list by another more abstract one, anyway carrying all the needed information.
- The current experimentation relies on service-related data coming from the service plane. One step further is to directly integrate the service plane in the experimentation platform in order to test and validate the gain provided by direct and automated interaction between the service plane with the existing network (data and control) and the Management planes.
- We took into account wavelength continuity and required bandwidth constraints while minimizing TE metric path costs. At a next step multiple TE metrics (e.g. the delay, the jitter, etc.) could be taken into account when computing and signaling an inter-domain path.
- Other service-related constraints, such as the reliability, should be implemented in future work so as to compute and signal diverse inter-AS LSPs.

References

[1] A. Farrel, J.-P. Vasseur, A. Ayyangar, "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", RFC 4726, Nov. 2006.

[2] R. Douville, J.-L. Le Roux, J.-L. Rougier, S. Secci, "A service plane over the PCE architecture for automatic multidomain connection-oriented services", *IEEE Communications Magazine*, vol. 46, no. 6, Jun. 2008.

[3] S. Secci, J.-L. Rougier, A. Pattavina, "AS Tree Selection for Inter-Domain Multipoint MPLS Tunnels", in *Proc. of 2008 IEEE Int. Conference on Communications (ICC 2008)*, pp. 5863-5868, May 2008, Beijing, China.

[4] A. Farrel, A. Ayyangar, JP. Vasseur, "Inter-Domain MPLS and GMPLS Traffic Engineering - Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 5151, Feb, 2008.

[5] L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, Jan. 2003.

[6] A. Farrel, A. Vasseur, J. Ash, "A path computation element (PCE) based architecture", RFC 4655, Aug. 2006.

[7] J.-P. Vasseur, J.-L. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, Mar.2009.

[8] J.-P. Vasseur, R. Zhang, N. Bitar, J.-L. Le Roux, "A backward recursive PCE-based computation (BRPC) procedure to compute shortest constrained inter domain traffic engineering label switched paths", draft-ietf-pce-brpc-09, wip, April 2008.

[9] J. Xiao, R. Boutaba, "QoS-aware service composition and adaptation in autonomic communication", *IEEE Journal on Selected Areas in Communications*, Vol. 23, pp: 2344-2360 (2005).

[10] R. Martinez, R. Munoz, M. Requena, J. Sorribes, J. Comellas, G. Junyent, "ADRENALINE Testbed: architecture and implementation of GMPLS-based network resource manager and routing controller", in Proc. of *TRIDENTCOM 2006*, Barcelona (Spain). Mar. 1-3, 2006.

[11] R. Munoz, C. Pinart, R. Martinez, J. Sorribes, G. Junyent, A. Amrani, "The ADRENALINE Testbed: Integrating GMPLS, XML and SNMP in transparent DWDM networks", *IEEE Communications Magazine*, Vol. 43, no. 8, pp: 40-48, August 2005.

[12] J.-L. Le Roux, R. Jacob, R. Douville, "Carrying a Contract Identifier in the PCE communication Protocol (PCEP)", draft-leroux-pce-contract-id-01.txt, Mar. 2007

[13] J.-P. Vasseur, J.-L. Le Roux, "Policy-Enabled Path Computation Framework", RFC 5394, Oct. 2008.

[14] R. Bradford, J.-P. Vasseur, A. Farrel, "Preserving topology confidentiality in inter domain path computation using a key based mechanism", draft-ietf-pce-path-key-06, wip, Mar. 2009.