On the Challenges of Establishing Disjoint QoS IP/MPLS Paths Across Multiple Domains

Marcelo Yannuzzi, Xavi Masip-Bruin, Sergio Sánchez and Jordi Domingo-Pascual, Technical University of Catalonia

Ariel Orda, Technion, Israel Institute of Technology

Alex Sprintson, Texas A&M University

This work was funded by the Spanish Ministry of Education under contract TEC2005-07520-C03-02.

ABSTRACT

MPLS is being actively adopted as the core switching infrastructure at the intradomain level. This trend is mainly attributable to the undeniable potential of MPLS in terms of Virtual Private Networks (VPNs) management, traffic engineering (TE), QoS delivery, path protection, and fast recovery from network failures. However, little progress has been made to attain the expected extension of MPLS label-switched paths (LSPs) across domain boundaries. Among the problems that remain unsolved is how to efficiently find and establish primary and protection interdomain LSPs for mission-critical services subject to QoS constraints. This article explores the major limitations hindering the deployment of these kinds of LSPs across multiple domains, in the context of the current interdomain network model. We describe the critical problems faced by the research community, and present our vision on how to rationally overcome some of the problems exposed. Our perspective is that we should be prepared for rather coarse-grained solutions as long as we need to coexist with the current interdomain network model.

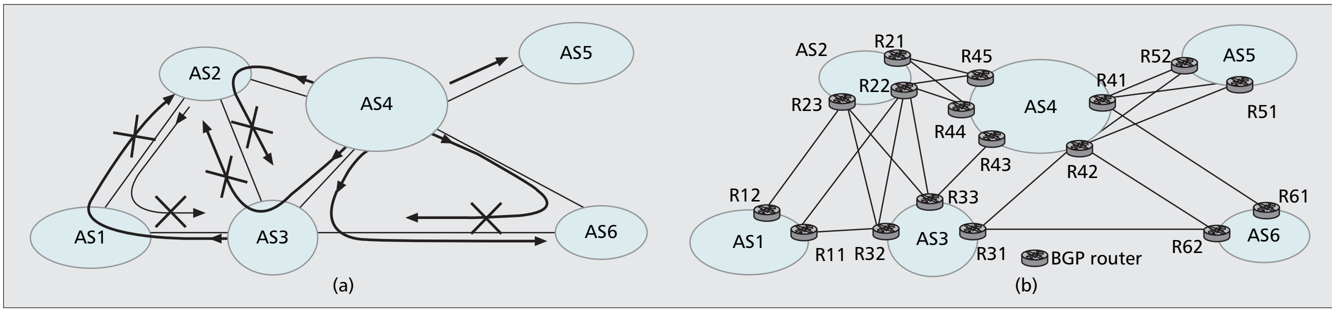
INTRODUCTION

Many research efforts have been and are still devoted to improve different facets of MPLS in the context of a single domain. At present, a significant part of these efforts are expected to move into the interdomain area. The reason for this is two-fold. On the one hand, customers are requiring from their Internet service providers (ISPs) the capability to extend their MPLS-based Layer 2 and Layer 3 Virtual Private Network (VPN) services across domains. Such services typically support some mission-critical applications and IP telephony, demanding hard QoS guarantees and fast restoration capabilities from the network. On the other hand, ISPs are eager to offer these services, so the research community is facing the challenge of devising the most suitable way of extending the reach of QoS-constrained MPLS Label Switched Paths (LSPs) beyond domain boundaries.

A recently chartered Working Group (WG) by the IETF has started to address the issue. Their first contribution is the introduction of a new network component inside each domain called the path computation element (PCE) [1]. The WG is expected to draft solutions and provide guidelines for a wide range of unsolved problems, including:

- The extension of MPLS Traffic Engineering (MPLS-TE) capabilities across domains
- The design of novel communication protocols to handle requests for the computation of paths subject to multiple constraints within and between domains [2]
- The definition of the extensions needed for some of the existing routing and signaling protocols

From this range of open problems, in this article we focus on exploring the major limitations hindering the deployment of primary and protection interdomain LSPs for mission-critical services subject to given QoS constraints. Our interest is in advance path protection strategies (i.e., backup paths need to be established jointly with the primary LSPs). The rationale for this approach is that in many practical settings it might not be possible to restore all QoS protected paths after a failure. This typically depends on the type of failure, and the amount of traffic that needs to be restored. Furthermore, restoring interdomain QoS LSPs after a failure might take an unacceptably long time for a number of mission-critical applications. Thus, for this kind of applications, switching promptly from a primary



■ **Figure 1.** *a) The export policies and the power-law relationship of the Internet topology make AS-graphs inadequate to find disjoint paths across domains; b) the same topology but at a border router level.*

to a backup path in the event of a failure can be guaranteed by provisioning two disjoint QoS paths between the source and destination nodes.

The subject of this article is to explore the challenges in doing so at the interdomain level. As we show, the problem of finding two disjoint QoS paths in the context of the current interdomain network model yields solutions that are far from optimal. We hope that some of the discussions presented in this article will encourage researchers to explore novel Internet models that support the establishment of optimal or near-optimal disjoint QoS paths.

The rest of this article is organized as follows. First, we analyze the main limitations imposed by the current interdomain network model, and then we describe the features of the PCE-based proposal coming from the IETF. Finally, we expose the major challenges to be faced, and present our vision on how to overcome some of the problems exposed.

LIMITATIONS IMPOSED BY THE CURRENT INTERDOMAIN NETWORK MODEL

The current interdomain network model introduces a series of limitations that hinder the computation and establishment of high-quality disjoint LSPs across domains. These limitations can be grouped into three categories:

- Lack of a model for traffic engineering (TE) information exchange between domains
- Policy-based routing
- Scarce path diversity

LACK OF A MODEL FOR TE INFORMATION EXCHANGE BETWEEN DOMAINS

At present, the information exchange between domains at the control plane level is conveyed by the interdomain routing protocol, that is, the Border Gateway Protocol (BGP). Although BGP supports the distribution of some limited TE information (e.g., with the BGP communities attribute [3]), in practice, BGP only advertises reachability information between domains. BGP routers never exchange network "state" information, such as path bandwidth utilization or path delays, which are essential for TE purposes. Furthermore, BGP routers are completely unaware of the topology of the Internet. A BGP router handles destination prefixes, and the next-hop to reach each destination. This approach has been proven to supply a scalable interdomain control plane. Unfortunately, it hinders the deployment of TE mechanisms capable of coping with the existing QoS and resilience demands at a multidomain level. Overall, at present there is neither a model nor a valuable mechanism for distributing TE information (or TE demands) among domains.

POLICY-BASED ROUTING

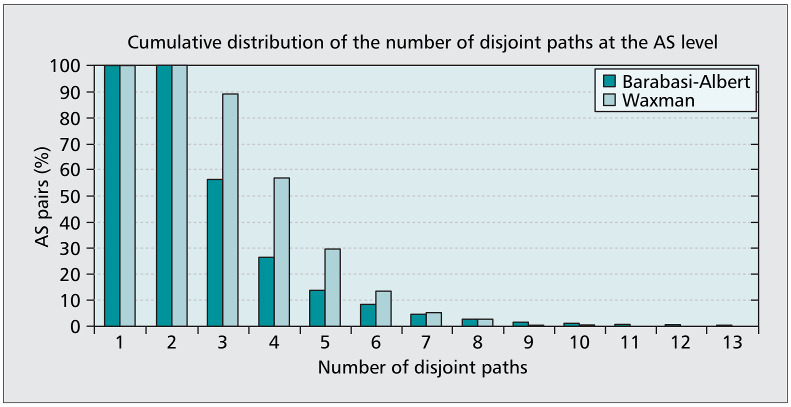
There are two types of business relationships between domains or autonomous systems (ASs), namely, customer-provider and peer-to-peer, which correspond to the two different traffic exchange agreements between neighboring domains. These relationships imply the following export policies of the ASs.

Customer-Provider Advertisements — Each AS advertises to its providers its own IP prefixes and those learned from its customers, but never those learned from its peers or from other providers. In addition, each AS advertises to its customers all the reachable IP prefixes it knows.

Peer-to-Peer Advertisements — Each AS advertises to its peers its own IP prefixes and those learned from its customers, but never those learned from its providers or other peers.

Figure 1a illustrates the effect of the export policies. The figure shows six interconnected ASs. Let us suppose that AS1 is a customer of AS2 and AS3, which are in turn peers of AS4. Let us also suppose that AS2 and AS3 are peers. In addition, AS5 is a customer of AS4, and AS4 are providers of AS6. The arrows in the figure represent the flow of BGP advertisements for the set of prefixes owned by AS4, according to the export policies. At a pure AS-graph level, AS3 has four possible paths to reach AS4 (i.e., one through AS1, one through AS2, one through AS6, and the one directly linked to AS4). However, the export policies determine that the path directly connecting AS3 and AS4 is actually the only one available for AS3.

The overall effect of the export policies is two-fold. First, interdomain routes cannot be inferred from the topology. These set of rules turn interdomain routing into being policy-driven rather than topology-driven or network-state driven, so finding disjoint paths across domains,



■ **Figure 2.** *Number of disjoint paths in a power-law topology (Barabasi-Albert) and a non-power-law topology (Waxman).*

at least at present, is strictly limited by these rules. Second, the algorithms for finding optimal disjoint QoS paths typically rely on a directed graph abstracting the network topology. However, in [4] we show that a multidomain network cannot be abstracted as a directed graph in the presence of the export policies. Thus, intradomain algorithms such as the ones proposed in [5] cannot be simply extended for AS-diverse routing.

Scarce Path Diversity

In addition to the reduction in the candidate paths due to the export policies, other factors contribute to the problem of the scarce path diversity between nodes located in distant ASs. The power-law relationship of the Internet topology, which was first reported in [6], is one of the main contributors to the problem. It reveals the hierarchical nature of the Internet and exposes the issue that only a very few highly connected transit ASs keep the Internet as a whole. At present, only around 20 of these large transit ASs exist, which means that, at the AS-level, the core of the Internet is very small. It also means that the ASs located at the edge of the Internet tend to connect to this highly connected group of ASs, which translates into very few AS-paths between distant ASs.

Another main contributor to the scarcity of paths is BGP. BGP introduces two major limitations. First, while a BGP routing table typically contains more than one candidate route toward a destination prefix, BGP routers allocate only one route (the best route) in the forwarding table. BGP routers typically select the shortest AS-path as the best route. This route is the one they use to forward packets and the only one advertised to other BGP peers. This reduces the number of routes handled by upstream domains, supplying a scalable routing approach, but it also drastically reduces the path availability information flowing upstream.

The second limitation introduced by BGP in terms of path diversity is that, for the sake of scalability, BGP handles and advertises highly aggregated information. To be precise, the reachability information advertised by BGP routers only contains AS-path information, that is, a set of destination prefixes and the list of AS hops that need to be traversed to reach those destinations. Such a list of AS hops offers highly aggregated information by completely hiding the internal structure of the ASs. The advantage of this lack of internal visibility is that it makes BGP highly scalable. A disadvantage is that although several disjoint paths might be available along an AS-path, they cannot be determined. For example, Fig. 1b discloses the internal structure of the ASs in Fig. 1a. For the sake of simplicity, we have only depicted the border routers. Without loss of generality, we assume high path diversity between the nodes inside the ASs. Figure 1 shows that, at the AS-graph level, there are no disjoint paths between AS1 and AS5 (all available paths traverse AS4). Yet, at the router-level, there are in effect two disjoint paths between the nodes in AS1 and AS5. In order to assess how some of the above limitations affect the number of disjoint paths between domains, we have conducted two different experiments.

Experiment 1 — The goal of the first experiment is to study how the power-law relationships of the Internet topology contribute to the scarceness of link-disjoint paths at the AS-level. We compared the number of disjoint paths using ten AS-level topologies generated by means of the BRITE topology generator [7]. We used two different models for generating the test topologies: a Waxman model and a Barabasi–Albert model. The Waxman model uses a probability function for interconnecting nodes based on the distance that separates them on the plane [8]. In this model, the node degrees are uniformly distributed, and hence they do not follow a power-law. The Barabasi–Albert model establishes links based on the preferential attachment principle [9]. This model follows a power-law. We used the default parameters provided by BRITE (e.g., $\alpha = 0.15$ and $\beta = 0.2$ for the Waxman model).

All topologies have the same number of ASs and links, namely, 100 ASs and 400 links. For each topology, we computed the number of link-disjoint paths between each pair of ASs. The average number of disjoint paths for topologies belonging to these models is depicted in Fig. 2. The figure compares the percentage of AS pairs that have at least n disjoint paths, for each $n \geq 1$. Our results show that for small values of n, the power-law topology has a smaller number of disjoint paths. For example, in the Waxman model, 57 percent of the AS pairs have at least four disjoint paths, compared with just 26 percent for the Barabasi–Albert model. This means that almost three-quarters of the AS pairs have less than four disjoint paths in a power-law topology and this is just from the topology perspective. Additional reductions need to be considered after introducing BGP and the export policies.

The tail of the distribution shows that only a small number of ASs have a large number of disjoint paths between them. This small group of ASs represents the highly connected core of the Internet, which is almost a full-mesh. Unfortunately, most of the candidate disjoint paths between the ASs in the core are unavailable in practice, due to the export policies between domains (recall that a provider does not supply transit for the packets exchanged between its peers).

Experiment 2 — Our goal in this experiment is to study the effect of topology aggregation on the maximal number of node-disjoint paths at the router-level. With topology aggregation, the ASs do not reveal the details of their internal structure, but rather supply an aggregated representation to the outside world. Using BRITE, we constructed several hierarchical network topologies that include 20 ASs, 200 nodes, and 874 links. The topologies are constructed using a top-down approach. A set of ASs is generated first, according to the Barabasi–Albert model. Next, for each AS in the AS-level topology, BRITE generates a router-level topology using the Waxman model. Once again, we used the default parameters provided by BRITE for both models. Next, we constructed a corresponding aggregated topology by using a virtual node model. In such a model, each AS is substituted by a single node, while two parallel links between the same pair of ASs are substituted by a single link. We then computed the maximum number of node-disjoint paths between each pair of routers for all topologies. We used node-disjoint paths because the aggregated topology does not provide information about the availability of link-disjoint paths that run through each AS. Next, we compared the percentage of router pairs that have at least n disjoint paths for each $n \geq 1$ with and without aggregation. The experiment results for a typical topology are depicted in Fig. 3. The results show that the number of disjoint paths in the aggregated topology can be up to 30 percent less than that in the original topology. Further reductions need to be considered after introducing the BGP decision process and the export policies in the aggregated topology (see [10] for simulation results about the scarcity of paths due to BGP).

Overall, the power-law relationship among ASs, together with the limitations imposed by BGP aggregation and the export policies, make AS-graphs inadequate to find disjoint paths across domains. Given that the AS-paths are the only information available in practice that can be used for interdomain TE purposes, the provisioning of disjoint LSPs with QoS constraints is simply unfeasible in the framework of the current interdomain network model.

PCE-based Approach

The limitations exposed above have motivated the creation of the PCE WG within the IETF. The aim of this initiative is to standardize a PCE-based model to distribute the computation of TE LSPs among different areas of a single domain or within a small group of domains. This model is not considered to be applicable to the entire Internet, and this stems from the fact that there is no such demand at the moment. Most of the ongoing work at the IETF is still focused on interarea (single domain) issues. Even though the interdomain case has begun to be analyzed, the discussions are in an early stage. This section provides an overview of the key aspects of this model, and succinctly explores its possibilities in terms of provisioning primary and backup QoS LSPs across domains. Besides the recent standardization of the architecture [1], all the work in the WG is in the draft stage. Many issues remain open, so from the alternatives that are being discussed, we present the one that we consider supplies the most practical approach.

This approach proposes a decoupled architecture, in which path computation tasks are performed by a device that is detached from the head-end MPLS Label Switching Router (LSR). Such a device is referred to as the PCE. Each domain may allocate one or more PCEs, depending on its size. For instance, large transit domains can be split into several areas, and use one PCE to handle the path computations within each area. For the distributed computation of interarea LSPs, a communication protocol is used between the PCEs of the involved areas [2]. Actually, the same model applies at the interdomain level, so the set up of LSPs spanning multiple domains involves at least one PCE per domain [1].

Each PCE is capable of computing primary and backup QoS paths within a domain or an area of a domain. To accomplish this task, the network state information of the domain (area) is gathered into a Traffic Engineering Database (TED). The TED is fed by the intradomain routing protocols (e.g., OSPF-TE or IS-IS-TE) and "raw" BGP information, that is, by the set of BGP routes that are available before BGP chooses the best route. This increases the number of candidate paths inside the TED. The PCE uses the information contained in the local TED to find primary and backup QoS paths by means of heuristics especially designed to tackle the intractability of the path computation problem [5]. By detaching the path computation tasks from the routers, dedicated PCEs can relief the LSRs from intensive computations such as finding disjoint QoS paths.

The WG has already drafted the first version of the communication protocol between the LSRs and the PCEs as well as between cooperating PCEs [2]. In [2] the LSRs are termed path computation clients (PCCs). The protocol specifies both the PCC-PCE communication, and the PCE-PCE communication for the distributed computation of LSPs. The PCC-PCE part of the protocol supports path requests subject to multi-

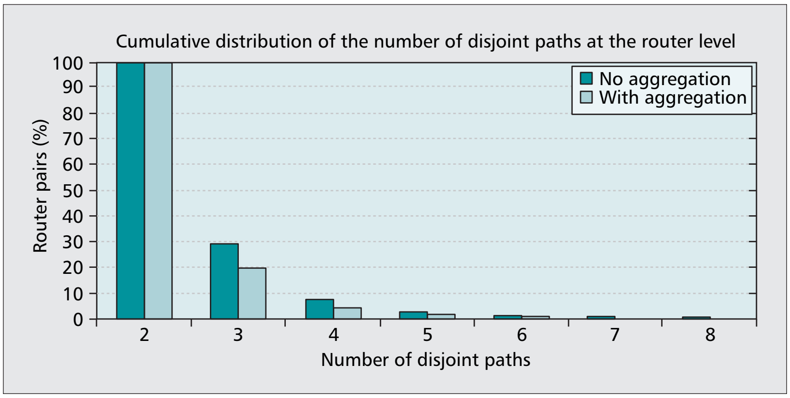
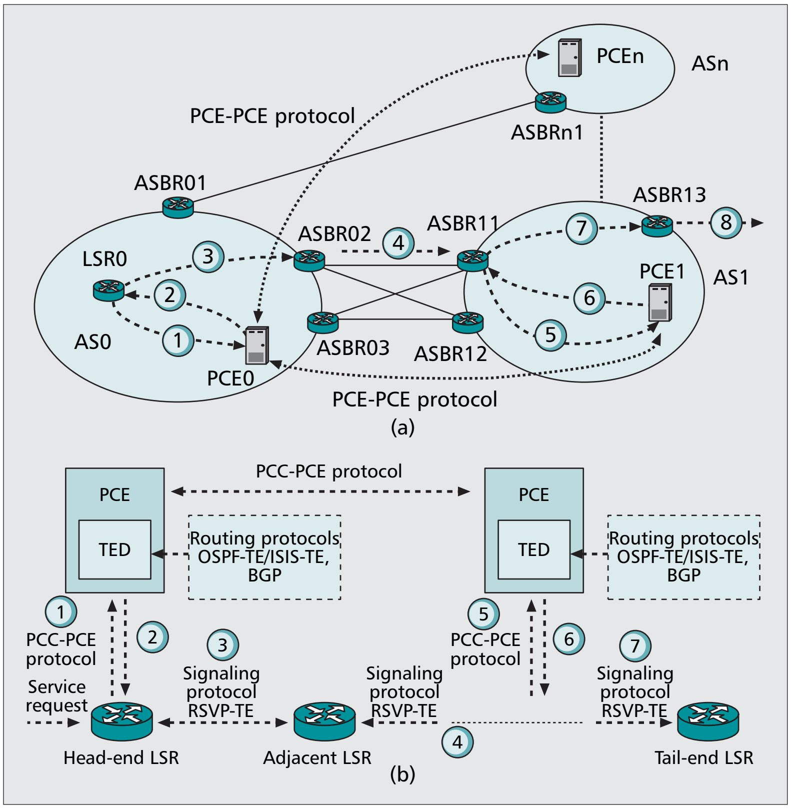


Figure 3. *Number of disjoint paths with and without aggregation.*



■ **Figure 4.** *a) Per-domain LSP computation based on ERO expansion; b) Request/Response messages and protocols involved.*

ple QoS constraints; it is able to return multiple (disjoint) paths, and takes into consideration features such as security and policies. Accordingly, some of the limitations exposed in the previous section are partially addressed by means of this approach.

Figure 4a illustrates the PCE-based architecture. The LSR0 in AS0 is the head-end of a requested LSP toward a destination node located in a distant AS (not depicted in the figure). When LSR0 receives the LSP request, the following sequence of actions occurs:

1 LSR0 requests PCE0 to compute the path.

2 PCE0 queries the TED in AS0 and computes the segment of the interdomain LSP up to the next-hop (NH) AS border router (ASBR). If more than one candidate path exists, the heuristic algorithm in PCE0 selects the "best" segment towards the destination (we discuss this selection process in the next section). Suppose that PCE0 selects ASBR11, so that it responds LSR0 with a set of strict hops toward this node. Notice that the NH ASBR denotes the ingress ASBR of the downstream domain, so that the NH ASBR and the PCE computing the local segment of the path belong to different domains.

3–4 These steps represent the signaling messages, that is, the resource reservations and explicit path routing performed by a protocol like RSVP-TE.

Once the signaling messages reach ASBR11, the same process occurs inside AS1, which is represented as the actions from 5 to 8, and this process is repeated on a per-domain basis until the destination AS is reached.

Figure 4b shows a more detailed description of the sequence of actions and the role of the different protocols involved in the set up of an interdomain LSP. The distributed path computation approach explained above is referred to as Explicit Route Object (ERO) expansion [11]. The name comes from the RSVP-TE ERO, which allows signaling a mix of strict and loose hops to be used in the path. A hop may be even an "abstract" node such as an entire AS. Abstract and loose hops are expanded inside each transit domain to a set of strict hops between the ingress ASBR and the NH ASBR.

This approach has two practical advantages. First, it supplies a scalable path computation scheme, since the responsibility and "visibility" of each PCE ends up in the corresponding NH ASBRs. Second, it supplies an appealing approach to ISPs, since it leverages confidentiality by hiding the internal network topology of downstream domains. The approach is simple, since each PCE computes a piece of the LSP based on its knowledge of the state of resources within its AS, and the reachability information obtained from BGP. Unfortunately, the major drawback of computing paths by segments is that the resulting paths are likely to be far from optimal. For instance, it is a well-known fact that high-quality paths are frequently uncorrelated with the routing choices made by BGP.

The issue that remains wide open is how to exploit the PCE-based model to compute *high-quality* primary and backup LSPs across a small group of domains in a viable way, that is, without adversely affecting the scalability and the confidentiality features of the above approach. In the sequel, we explore the key challenges raised by this issue.

Major Challenges to be Faced

Splitting the computation of primary and backup interdomain LSPs introduces a number of problems that need to be addressed in order to avoid coarse-grained solutions. We examine the key challenges to be faced and present our perspective on the road to solve them.

TE Information Exchange Model Among Domains

With the current PCE model each segment of an interdomain LSP is derived from a very limited visibility of the state and topology of the network. As a result, no guarantee exists that the optimal QoS path (e.g., the shortest path) will be found. In fact, once a PCE has chosen the NH domain and established its segment of the LSP, there is no guarantee that a viable QoS path will be discovered through the NH domain. This is because no information apart from IP reachability is exchanged between domains. When this occurs, the NH ASBR signals back an error message indicating that its domain is unable to set up the next segment of the LSP (such an error can occur either while computing the path segment or while signaling its establishment along the domain). When the PCE receives this error message, it iteratively tries other downstream domains until it succeeds or rejects the path request. This trial and error provisioning and signaling process is referred to as *crankback*.

An alternative approach is to work toward a TE information exchange model between domains. This model could be supported by the PCE-based architecture; thus, it could be applied to a small group of neighboring domains. In this framework, domains become capable of exchanging some highly aggregated topology and state information, which can be used to compute "entire" LSPs from the source PCE. In order to preserve the confidentiality of ISPs and also keep the model scalable, domains never advertise their internal structure, but rather supply an aggregated representation (AR) to the outside world. Thus, a key aspect is to find an adequate AR that captures the available path diversity of QoS paths across a small group of domains. Certainly, a trade-off exists between the optimality of the resulting QoS paths and the size of the AR. Two different ARs based on the advertisement of the available disjoint paths between the border nodes of a domain can be found in [4] and [12].

The advantage of the AR is that it facilitates the computation of entire (primary and backup) *shortest* paths directly from the source PCE [4]. Since the source PCE only knows an AR of the whole network, the resulting paths are still a mix of strict and loose hops. The list of strict hops could include the source node, the list of border nodes to be traversed across the different domains, and the destination node. Thus, approaches like this still need to rely on the ERO expansion, but with the advantage of increasing the number of strict hops conveyed in the signaling messages.

Issues such as how TE information is to be distributed and updated need to be carefully investigated. One possibility could be to use the PCE protocol. However, [2] only proposes a request/response protocol for the computation of paths, making it inappropriate for this purpose. Sound alternatives are to propose extensions to the current specification of the PCE protocol, or to develop a new one facilitating the advertisement of TE information among a small group of PCEs.

Routing Decision

Once each PCE knows an AR of the multidomain network, a routing decision is required in order select the paths for the PCC requests. For example, in [12] the disjoint paths computed by a source PCE are both routed along the same chain of domains, since it is assumed that the AS-level path is known in advance (e.g., is precomputed by BGP). The AR in this case is basically an abstraction of an AS-path. This routing approach has two major weaknesses. First, high-quality paths (e.g., the shortest paths between the source and destination nodes) are not guaranteed to be discovered, since they might not belong to the precomputed AS-path. Second, when several disjoint LSPs need to be established following the same AS-path, the utilization of network resources at the interdomain level could be quite inefficient. Instead, if a source PCE is not constrained to route the LSPs along a given AS-path, then the shortest paths can be found and a more efficient use of the overall network resources can be achieved [4].

An alternative routing scheme is to avoid handling an AR of the small-sized multihomed network, but refine the selection of the NH domain and then repeatedly solve this problem on a per-domain basis. Two heuristics in this direction have been recently proposed in [13], but the focus there is on the selection of a single path. The resulting paths under these routing schemes are expected to be of higher quality than those that can be obtained with the current PCE-based approach. Still, these routing schemes cannot guarantee to find optimal QoS paths (e.g., the shortest paths) across domains. Another alternative that can be used as an interim solution (e.g., before the deployment of the PCEs) was proposed in [14]. This proposal exploits the multiconnectivity between peering ASs in order to find disjoint LSPs along a chain of domains.

Overall, the key issue is that the resulting paths from the current PCE routing scheme (Fig. 4) are far from optimal, so alternative routing strategies like [4, 12, 13] deserve to be investigated.

Strategy for the Computation of Restoration Paths

The issue that arises is whether to compute the primary and restoration paths at the same time, or one after the other. The latter case is subject to the well-known *trap topology problem* [12], hence network resources can be consumed more efficiently when both paths are computed simultaneously. Accordingly, the heuristic algorithm controlling the decisions made by a PCE should be able to compute disjoint paths at the same time.

Fast Restoration After A Failure

With local restoration, each AS can potentially protect its corresponding segment of a path. However, fully relying on this approach means that each domain needs to trust the restoration decisions made by downstream domains, which might not be acceptable for some ISPs as well as for some mission-critical applications. Indeed, after a distant failure in an interdomain path, the source node has neither guarantee that the path will be restored nor that the restored one will actually comply with the QoS constraints. Thus, precomputed restoration paths with prompt failure detection and fast restoration from the source LSR become necessary in some cases. From our perspective, some applications can be protected by means of local protection, (i.e., on a per-segment basis), while others will need novel mechanisms at the application level to promptly detect a failure and switch to a backup path.

With local restorations, each AS can potentially protect its corresponding segment of a path. However, fully relying on this approach means that each domain needs to trust the restoration decisions made by downstream domains, which might not be acceptable for some ISPs.

Conclusions

The PCE-based model facilitates the provisioning of primary and backup QoS LSPs across domains. The current proposals for finding such paths are based on a coarse selection of the paths by the source domain, and then rely on the

It is difficult to predict the requirements in the longer term, so it seems sound to deeply analyze the possibilities that the PCE-based model offers before pushing for the standardization of coarse solutions.

ERO expansion technique within the subsequent domains traversed. The strengths of this approach are its scalability and the preservation of the confidentiality of ISPs networks. The main weakness is that the resulting paths are far from optimal.

An important subject is how to exploit the PCE-based model to compute high-quality primary and backup LSPs across a small group of domains in a viable way. From our perspective, approaches tending to endow this model with the capability of aggregating and distributing enriched TE information, allowing a source PCE to compute entire near-optimal LSPs, are worthy of being explored. From a practical viewpoint, especially for the short-term requirements driving the extension of TE LSPs across domains, coarse-grained solutions might be enough. However, it is difficult to predict the requirements in the longer term, so it seems sound to deeply analyze the possibilities that the PCE-based model offers before pushing for the standardization of coarse solutions.

REFERENCES

[1] A. Farrel, J. P. Vasseur, and J. Ash, "A Path Computation Element (PCE)-Based Architecture," IETF RFC 4655, Aug. 2006.

[2] JP. Vasseur *et al.*, "Path Computation Element (PCE) Communication Protocol (PCEP) — Version 1, " Internet draft, draft-ietf-pce-pcep-02.txt, June 2006, work in progress.

[3] B. Quoitin *et al.*, "Interdomain Traffic Engineering with BGP," *IEEE Commun. Mag.*, May 2003.

[4] A. Sprintson *et al.*, "Reliable Routing with QoS Guarantees for Multi-Domain IP/MPLS Networks," Res. rep. RR-CBA-2006-4, http://elio.ac.upc.edu/~dacsecre/reports/2006/31/RR-2DPs_Interdomain.pdf, Dept. of Comp. Architecture, Tech. Univ. Catalonia, Spain, Sept. 2006.

[5] A. Orda and A. Sprintson, "Efficient Algorithms for Computing Disjoint QoS Paths," *Proc. IEEE INFOCOM*, Hong Kong, Mar. 2004.

[6] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On Power-Law Relationships of the Internet Topology," *Proc. SIGCOMM*, Sept. 1999.

[7] A. Medina *et al.*, "BRITE: An Approach to Universal Topology Generation," *Proc. MASCOTS*, Aug. 2001.

[8] B. Waxman. "Routing of Multipoint Connections," *IEEE JSAC*, Dec. 1988.

[9] A. L. Barabasi and R. Albert, "Emergence of Scaling in Random Networks," *Science J.*, vol. 286, 1999, pp. 509–12.

[10] C. de Launois, "Unleashing Traffic Engineering for IPv6 Multihomed Sites," Ph.D. thesis, Univ. Catholique de Louvain, Sept. 2005.

[11] J P. Vasseur, A. Ayyangar, and R. Zhang, "A Per-domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)," Internet draft, draft-ietf-ccamp-inter-domain-pd-path-comp-03.txt, Aug. 2006, work in progress.

[12] F. Ricciato, U. Monaco, and D. Ali, "Distributed Schemes for Diverse Path Computation in Multi-Domain MPLS Networks," *IEEE Commun. Mag.*, vol. 43, no. 6, June 2005.

[13] C. Pelsser and O. Bonaventure, "Path Selection Techniques to Establish Constrained Interdomain MPLS LSPs," *Networking 2006*, Coimbra, Portugal, May 2006.

[14] R. Romeral and D. Larrabeiti, "Combining Border Router Policies for Disjoint LSP Computation," *Proc. V Wksp. G/MPLS Networks*, Gerona, Spain, Mar. 2006.

BIOGRAPHIES

MARCELO YANNUZZI (yannuzzi@ac.upc.edu) received a degree in electrical engineering from the University of the Republic (UdelaR), Uruguay, in 2001, and a D.E.A. from the Department of Computer Architecture, Technical University of Catalonia (UPC), Spain, in 2005, where he holds a teaching and research position. He was formerly with the Uruguayan telco and with the Physics Department at UdelaR. His research interests are in the areas of interdomain routing, traffic engineering, and network stability in both IP and optical networks.

XAVI MASIP-BRUIN (xmasip@ac.upc.edu) received M.S. and Ph.D. degrees from UPC, both in telecommunications engineering, in 1997 and 2003, respectively. He is currently an associate professor of computer science at UPC. His current research interests lie in broadband communications, QoS management and provision, and traffic engineering. His publications include around 50 papers in national and international refereed journals and conferences. Since 2000 he has participated in many research projects: IST projects E-NEXT, NOBEL, and EuQoS; and Spanish research projects SABA, SABA2, SAM, and TRIPODE.

SERGIO SANCHEZ LOPEZ (sergio@ac.upc.edu) got his Ph.D. in telecommunications engineering in 2003 from the Department of Computer Architecture, UPC. Since 1992 he has been an associate professor with this department. His publications include several book chapters, papers in relevant international journals and more than 50 papers in international refereed conferences. His current research interests are in broadband Internet and high-speed and optical networks, with emphasis on traffic engineering and control plane.

JORDI DOMINGO-PASCUAL (jordid@ac.upc.edu) is a full professor of computer science and communications at UPC in Barcelona. He is co-founder and researcher of the Advanced Broadband Communications Center of the university (CCABA). His research topics are Broadband Communications and Applications, IP/ATM integration, QoS management and provision, traffic engineering, IP traffic analysis and characterization, and QoS measurements.

ARIEL ORDA [S'84, M'92, SM'97, F'06] (ariel@ee.technion.ac.il) received B.Sc. (summa cum laude), M.Sc., and D.Sc. degrees in electrical engineering from the Technion — Israel Institute of Technology, Haifa, in 1983, 1985, and 1991, respectively. Since 1994 he has been with the Department of Electrical Engineering at the Technion, where he currently is a full professor. His research interests include network routing, survivability, QoS provisioning, wireless networks, the application of game theory to computer networking, and network pricing. He received the Award of the Chief Scientist in the Ministry of Communication in Israel, a Gutwirth Award for Outstanding Distinction, the Research Award of the Association of Computer and Electronic Industries in Israel, the Jacknow Award for Excellence in Teaching, an ICNP 2004 Best Paper Award, the *Journal of Computer Networks* 2004 Best Editor Award and the 2005 Klein Award for Excellence in Research. He served as Program Co-Chair of IEEE INFOCOM 2002 and is an Editor of *IEEE/ACM Transactions on Networking* and the *Journal of Computer Networks*.

ALEX SPRINTSON (spalex@ece.tamu.edu) is an assistant professor of electrical and computer engineering at Texas A&M University. He received his Ph.D. degree in electrical engineering from the Technion — Israel Institute of Technology in 2003. His research interests lie in the general area of communication networks with a focus on network survivability, QoS routing, network coding, and security of wireless communication networks. He served as a member of the Technical Program Committee for IEEE INFOCOM 2006 and 2007.