On the Difficulty of Establishing Interdomain LSPs

Cristel Pelsser, Steve Uhlig, Olivier Bonaventure

CSE Dept., Université Catholique de Louvain, Belgium
Email: cpe,suh,Bonaventure@info.ucl.ac.be

***Abstract*— Nowadays, the success of MPLS is mostly due to the increasing demand for BGP/MPLS VPNs. Even though the need for interdomain LSPs is growing, no ISP today proposes the dynamic establishment of LSPs across AS boundaries. In this paper, we investigate the complexity of establishing end-to-end interdomain LSPs with QoS guarantees, based on the BGP routes locally available at a router.**

We explain the main issues of relying on BGP for the computation of interdomain constrained paths. To illustrate our point, we compare two LSP establishment techniques. Our benchmark technique is centralized and assumes the complete knowledge of the intradomain topologies. The second path computation technique is decentralized and relies on the BGP routes locally available by each router. Our simulations confirm that the difficulty in designing BGP-based interdomain LSP establishment techniques lies within the trade-off between the scalability of the computation technique and the quality of the path found in terms of the considered metrics.

I. Introduction

The initial motivation for introducing Multiprotocol Label Switching (MPLS) in the 1990s was the low performance of IP routers compared to ATM and Frame Relay switches [1]. MPLS allowed IP networks to use higher bandwidth links thanks to the closer integration with ATM or Frame Relay. Today's routers and switches are very different from those available ten years ago. Due to the improvements in packet forwarding capabilities, routers are now able to route normal IP packets at line rates of 10 Gbps and 40 Gbps. Today, the main motivation for running an MPLS network is to provide MPLS-based services such as BGP/MPLS Virtual Private Networks (VPNs) [2] , supporting traffic engineering [3] and to allow the network to recover quickly from failures by using detours or bypass tunnels [4] .

Those services are currently used inside large networks [5]. Inside a single domain, several techniques can be used to establish Label Switched Paths. Some rely on a specific signalling protocol such as LDP [6] or RSVP-TE [3], others piggyback label information inside route advertisements as with BGP [7]. LDP is used to create best-effort LSPs while RSVP-TE allows to specify constraints such as bandwidth or delay for the establishment of traffic engineered LSPs.

Due to the success of MPLS-based services, users of those services are urging network providers to cooperate in order to support BGP/MPLS VPN networks between sites attached to different ASes [8]. This is a common requirement for large multinational companies with sites spread worldwide. In addition to inter-AS VPNs, interdomain LSPs can be used to provide various types of services. For example, a lab inside a university could establish a LSP toward another lab in another country across several transit ASes to transmit large amounts of experimental data, or, interdomain LSPs can be used by ISPs to install remote Point of Presences (POPs) outside their region of operation, by cooperating with other ISPs [9].

This paper addresses the problem of interdomain constrained path computation for the establishment of LSPs. We describe a technique to compute interdomain constrained paths in a distributed manner based on the routing information available with BGP and, inside each AS, on the topology distributed by the IGP. This technique is applicable to the computation of LSPs crossing an arbitrary number of ASes.

In the short term, we expect that the first motivation for using MPLS across interdomain boundaries will be to provide multi-AS VPN services or to interconnect large telephone switches in different domains (VoIP traffic). We expect that those services will initially be deployed between ASes that are directly connected and likely managed by the same company [10].

II. Related Work

The problem of establishing LSPs inside a network has attracted a lot of interest during the last five years [11], [12]. Most of the solutions proposed to solve this problem have assumed that the LSPs were established inside a single domain in which the routers were all in the same IGP area. A consequence of this assumption is that each router knows the entire network topology. By using the traffic engineering extensions to the IGP [13], [14], the routers may also know the amount of reserved bandwidth on each link. In this case, the layout of the LSPs inside the domain becomes an optimization problem that can be solved by considering various objectives such as load-balancing, protection in case of failure and minimizing the end-to-end delay.

Although many large ISPs are organized as a single IGP area, some rely on multiple areas. In that case, the problem becomes more complex than an optimization problem because a router knows the complete topology of its own area but has a limited view of the topology of the other areas. Establishing LSPs in this environment can be centralized by using a Path Computation Element (PCE) [15]. A PCE could collect information distributed by the IGP, in all areas, to compute the path for LSPs upon request from other routers. The main limitation of the PCE solution is its scalability as the path for all inter-area (inter-domain) LSPs must be computed by the PCE. A distributed approach is also possible by allowing each

Area Border Router to select the egress border router of the area. However, in this case the chosen path is not necessarily optimal.

Several researchers have proposed extensions to RSVP-TE to ease the establishment of interdomain LSPs. In [16], we propose extensions to enable the use of RSVP-TE across AS boundaries for primary and end-to-end backup paths with respect to requirements formulated by ISPs [9]. [17] proposes extensions for local link, node and SRLG protection of interdomain LSPs. Protocol extensions have also been proposed to BGP in order to advertise QoS information along with the BGP reachability information [18]. However, no study has been published on the use of these extensions to established QoS constrained LSPs. To our knowledge, the computation of interdomain paths with QoS and disjointness constraints has not been addressed in the literature.

III. Interdomain constrained path computation and BGP

When considering VPN services across domains, the limited topological information available through BGP interdomain routes is crucial. BGP is the current interdomain routing protocol [19]. It is a path vector protocol that allows each domain to define its own routing policies. A router attached to another AS via a peering link establishes an eBGP session over the peering link with the BGP neighboring router. This eBGP session is used to advertise the routes that are reachable by each AS. A BGP router advertises its best route to reach each destination prefix. When a BGP router receives a route over an eBGP session, it determines whether this route is its best route towards the destination. If so, it advertises the route to the other BGP routers of the AS. This is done by means of iBGP sessions.

The routing information distributed by BGP is very different from the topology information distributed by IGPs such as OSPF or IS-IS. BGP is much more scalable than a link-state IGP in that it only distributes reachability information subject to routing policies that limit the routes announced to neighboring ASes. The price for this scalability is the lack of information available on the Internet topology [20]. For each prefix, each peer only advertises its best route over BGP sessions. This route is selected based on criterions that are independent of the quality of the route in terms of end-to-end metrics like delay and reservable bandwidth.

BGP was initially designed assuming a full mesh of iBGP sessions between all the border routers of an AS, to exchange the best eBGP routes in the AS and allow each router to compute its best route towards any reachable destination. Due to this assumption, a BGP router does not advertise, over iBGP sessions, a route received over an iBGP session. If a router selects a route received via iBGP as best route, it will not advertise routes learned on eBGP sessions inside the AS. As a consequence there may be many available interdomain paths that are never learned by the routers and thus never used for packet forwarding.

If there are N border routers in the AS, a full mesh of iBGP sessions corresponds to $\frac{N\times(N-1)}{2}$ iBGP sessions. This is a severe scalability problem in networks containing more than a few tens of border routers. Two solutions have been proposed to solve this problem : confederations [21] and route reflectors (RRs) [22]. We do not consider the confederations in this paper as they are not frequently used.

A route reflector is a router that is allowed to re-advertise, over iBGP sessions, routes that it received over other iBGP sessions. The simplest way of deploying RRs is to replace a full mesh of iBGP sessions with a single RR. When a single RR is connected to all other BGP routers of the domain, each BGP router receives only one route from the RR instead of the $N-1$ routes received in the case of a full mesh of iBGP sessions.

The placement of RRs inside a domain might create problems [23], [24] that can be avoided by following the recommendations of [25].

IV. Path computation techniques

For the purpose of illustrating the issues in interdomain constrained LSPs computation, this section presents two alternative techniques. The first technique relies on the availability of the complete topology at one point in the network. This technique is only applicable when the administrators running the different ASes are willing to share topology information. It is an ideal situation that may not occur in practice except eventually between 2 ASes that belong to the same company. We use this technique as a benchmark. It consists of a centralized approach where the node possessing the intradomain topology of the ASes is responsible for the computation of interdomain paths. The second approach is applicable in a more general framework. It is a decentralized technique where each node on the path of the LSP completes the path computation toward the destination based on local routing information. This technique is applicable for the establishment of LSPs crossing any number of ASes.

The LSPs considered in this paper are subject to end-to-end delay and bandwidth guarantees as well as link and node disjointness constraints.

A. Centralized computation with CSPF

A centralized path computation can only be envisaged for LSPs crossing ASes that belong to the same company as ISP topology information is often considered strategic to the functioning of ISPs and kept secret. In that case, a Path Computation Element (PCE) [15] that centralizes the topology information of both ASes can compute the path of the inter-AS LSPs.

The PCE collects the link state packets advertised by the IGP in both ASes and thus possesses the complete topology of the two ASes with the TE information, if either IS-IS TE or OSPF-TE is used. For the purpose of this paper we assume that both delay and reservable bandwidth are distributed by the IGP. Based on this information, the PCE runs a CSPF algorithm. It prunes the links with insufficient remaining

reservable bandwidth, runs Dijkstra algorithm with costs set to the delay of the links and finally sends the computed path to the source of the LSP, if the path respects the delay constraint. For the disjoint path computation, the PCE first prunes the links and nodes that are on the primary path from the topology. Then, it runs the computation as for the primary path.

B. BGP-based Distributed Path Computation (DPC)

Since it may not be possible or desirable that a single node knows the complete intradomain topologies of several ASes, we now look at a decentralized constrained path computation approach.

Our Distributed Path Computation technique relies on the routing information distributed by BGP. Each router uses a single best BGP route to forward IP packets toward each distant destination prefix. These routes are stored in its Local Routing Information Base (Loc-RIB). However, a router may receive one route toward each prefix from each of its peers. If they pass the import filters, these routes are stored in its Adj-RIB-Ins. We use these routes to compute our constrained paths. As a consequence, the computed paths respect the BGP policies of the ASes that are enforced by the import and export filters inside the BGP routers.

The DPC of a primary LSP is illustrated in figure 1. Inside the source AS $AS1$, the source (PE) router selects, from all the routes toward the destination PE present in its Adj-RIB-Ins, the route with the Next-Hop (NH) that is reachable through a path with enough reservable bandwidth and smallest delay. This consists in performing a CSPF inside the source AS toward all the NHs advertised with the destination prefix, with the delay as metric. Once the NH $R4$ is selected, the LSP is established toward this NH using RSVP-TE with an ERO containing the computed constrained path segment $R2 - R4$. The NH $R4$, i.e. the egress AS Border Router (ASBR)[1], then selects a NH in the neighboring AS from the NHs of the routes to the destination PE, in the local Adj-RIB-Ins. Therefore, $R4$ evaluates the reservable bandwidth and the delay toward each of these NH, $R5$ and $R6$. $R3$ is not evaluated to avoid routing loops. Finally, the ingress ASBR $R6$, inside the downstream AS $AS2$, computes the path toward the PE[2] by running a CSPF on the topology of the destination AS.

We note that if a node needs to complete the path computation but does not have routes in its Adj-RIB-Ins, with NHs that can be joined by a path segment respecting the constraints, cranckback takes place [26]. A RSVP Path Error message is sent back to the source. An upstream node on the path, the previous ASBR in our case, computes an alternative path toward the destination, based on interdomain route advertisement toward the PE destination prefix that have not been tried.

[1]In this paper, we assume the use of Next-Hop self. A BGP router replaces the NH of a route by its own IP address before readvertising the route inside the AS. This option is commonly used because it avoids having to advertise the peering routers of neighboring ASes inside the IGP of the AS. However, the DPC technique is also applicable if NH-self is not used.

[2]If the PE does not belong to this AS, the ingress ASBR selects a NH from the routes in its Adj-RIB-Ins.

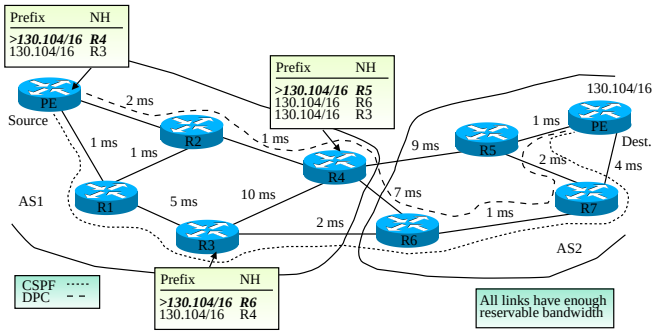

Fig. 1. Distributed Path Computation of a primary LSP

We observe that the path computed with DPC, on figure 1, has a larger delay than the CSPF path. This is due to the limited information available locally for the route selection. The DPC technique makes a local choice that may not lead to the globally optimal path. Another solution would be to evaluate end-to-end paths through all the NHs available for the destination, not only through the locally best NH. However, such exploration grows exponentially with the network size and connectivity [27].

Moreover, in figure 1, once the primary LSP is established, an end-to-end link and node disjoint path cannot be found. In order to establish a disjoint path, the nodes that complete the backup path, i.e. the ASBRs in our case, need to know the links and nodes crossed by the primary path. For this purpose, the nodes along the primary path can be recorded in the Record Route Object [3]. Then the source of the LSP stores these nodes in the eXclude Route Object (XRO) defined in [28]. This object is used by intermediate nodes to compute path segments that avoid the nodes stored in this object. Based on the XRO, the source PE router selects a NH that does not belong to the primary LSP and that is reachable with a path segment respecting the delay, bandwidth and disjointness constraints ($R3$ in figure 1). However, router $R3$ cannot continue the establishment of the disjoint LSP. The two NHs available for prefix $130.104/16$ are already on the path of the primary LSP, hence cranckback takes place. A Path Error message is sent to the PE router. The PE router does not possess any other route with a NH that has not already been explored. Thus, the backup LSP cannot be established.

C. Simulations

Our simulation environment contains two ASes because of the first technique. Each AS contains several interconnected routers. Furthermore, the routers in each AS are grouped in POPs as in most networks. A small POP may contain a single router while a large POP may be composed of a few tens of routers. The ASes are interconnected with one peering link in each city where both ASes have a POP. To establish interdomain LSPs, we consider the case of inter-AS VPNs where each AS may offer VPNs services toward the POPs of the other AS. For this reason, we attach a Provider Edge (PE)

router to each POP containing more than one router. This PE router is connected to two different routers inside the POP for redundancy reasons. We establish a full mesh of traffic engineered LSPs between those PE routers.

The AS topologies, with link delays and routers grouped in POPs, used for this purpose, have been collected by the rocketfuel project [29]. We assigned a bandwidth of 10 Gbps to each link. Moreover, each link connecting a PE router to other routers has a delay set to 1 ms. The same delay of 1 ms is assigned to the inter-AS links that we added to interconnect the ASes two by two. A router in each POP is configured as a route reflector, all the routers inside the POP are fully meshed from an iBGP viewpoint, for optimal intra-POP routing, and the route reflectors themselves are fully-meshed as recommended by [25].

In table I, we find the ASes involved in each topology with the number of nodes as well as the number of intra and inter-domain links. The last column indicates the number of LSPs to be established. We note that the number of inter-domain links varies from 3 to 14 links. The topology, "topo3", with most inter-domain links does not contain the largest number of nodes and links. The biggest topologies in terms of links and nodes are "topo4" and "topo7". Not all the ASes could be interconnected because they did not all have POPs in common locations.

Topology	ASes		Nodes	Links			LSPs
	ASN1	ASN2		intra	inter	total	
topo0	3257	3967	281	557	3	560	828
topo1	1239	3967	443	1217	5	1222	1116
topo2	3967	6461	246	577	5	582	396
topo3	1755	3257	291	575	14	589	920
topo4	1239	3257	530	1408	9	1417	1426
topo5	3257	6461	333	768	4	772	506
topo6	1239	1755	453	1235	6	1241	1240
topo7	1239	6461	495	1428	8	1436	682

TABLE I

PROPERTIES OF THE COMBINED ROCKETFUEL TOPOLOGIES

To illustrate the techniques described in section IV, we compute primary and backup paths with a 100ms delay constraint, with or without 100Mbps bandwidth reservations. That is, for each primary path, we compute an end-to-end link and node disjoint path with the same constraints as for the primary path, for protection purposes. The existence of backup paths is used as an indication of the diversity of the paths available to the centralized and the distributed techniques.

Figure 2 shows the number of LSPs that could not be established for each topology and each path computation technique. For each topology, the total number of LSPs to be established is indicated by a point. The first and third bars show the number of primary and, respectively, backup LSPs that could not be established with the CSPF algorithm. The second and fourth bars represent the same values for the DPC technique.

The top left portion of figure 2 presents the number of LSPs that could not be established for the simulations with a full mesh of iBGP sessions in the ASes and no bandwidth reservations associated to the LSPs. We note that all the primary and backup CSPF LSPs could be established for most of the topologies. Thus, a more elaborate disjoint path computation algorithm than CSPF is not necessary. CSPF is a good approximation of a k-SPF algorithm [30]. However, DPC could not always find a feasible path for the backup LSPs. The results of the same simulations but with RRs, instead of a full mesh of iBGP sessions, are provided in the bottom left portion of figure 2. Here, we observe that paths could not be found for most backup LSPs with the DPC technique. This illustrates the fact that RRs hide part of the BGP routes to their clients.

The right part of figure 2 concerns simulations of the establishment of LSPs with bandwidth reservations and with a full iBGP mesh (top) or with RR (bottom) in the ASes. We note that some LSPs cannot be established with the CSPF algorithm due to the limitation on the link capacities in the topologies and the structure of the rocketfuel topologies themselves. The same observation applies to the LSPs computed with DPC.

These figures confirm that RRs have a large impact on the possibility to find alternative paths. The difference between the number of primary LSPs that could not be established with CSPF and DPC lies in the limited number of routes available with BGP. We performed the same simulations with different orderings of the LSPs and observed the same behavior. Moreover, we did not observe a big difference in the number of established LSPs when removing the full mesh of iBGP sessions inside the POPs when using RRs. The difference mostly lies in the presence of the RRs inside POPs not in the way iBGP sessions are established in the POP.

Figure 3 shows the distribution of the difference in delay between CSPF and DPC LSPs. One curve compares the delay of the primary paths and the other curve compares the delay of the backup paths. Positive values indicate that the CSPF path has a shorter delay than the respective DPC path. Negative values occur when the DPC path has a shorter delay than the CSPF path between the same source and destination. This figure only shows the LSPs for which both the CSPF and the DPC paths could be computed. The results of figure 3 concern the establishment of LSPs without/with reservations, on the left (right, respectively), on topology "topo4" with RRs inside the ASes.

First, we observe that there are a large number of LSPs with the same delay for the primary CSPF and DPC paths. This indicates that most of the paths have the same quality independently from the path computation technique. Most paths computed based on the information available with BGP (DPC technique) have a delay comparable to the paths obtained with CSPF. *Even though the path found by DPC is often of a similar quality than the CSPF path, for large topologies, the former is never found on the first try, i.e. cranckback is used for every path computed by DPC.* On the left part of figure 3, we see that some CSPF backup paths have a higher delay than their respective DPC paths (negative values). This behavior results from the lack of information available on the quality of the BGP routes and the local search of the DPC technique. The

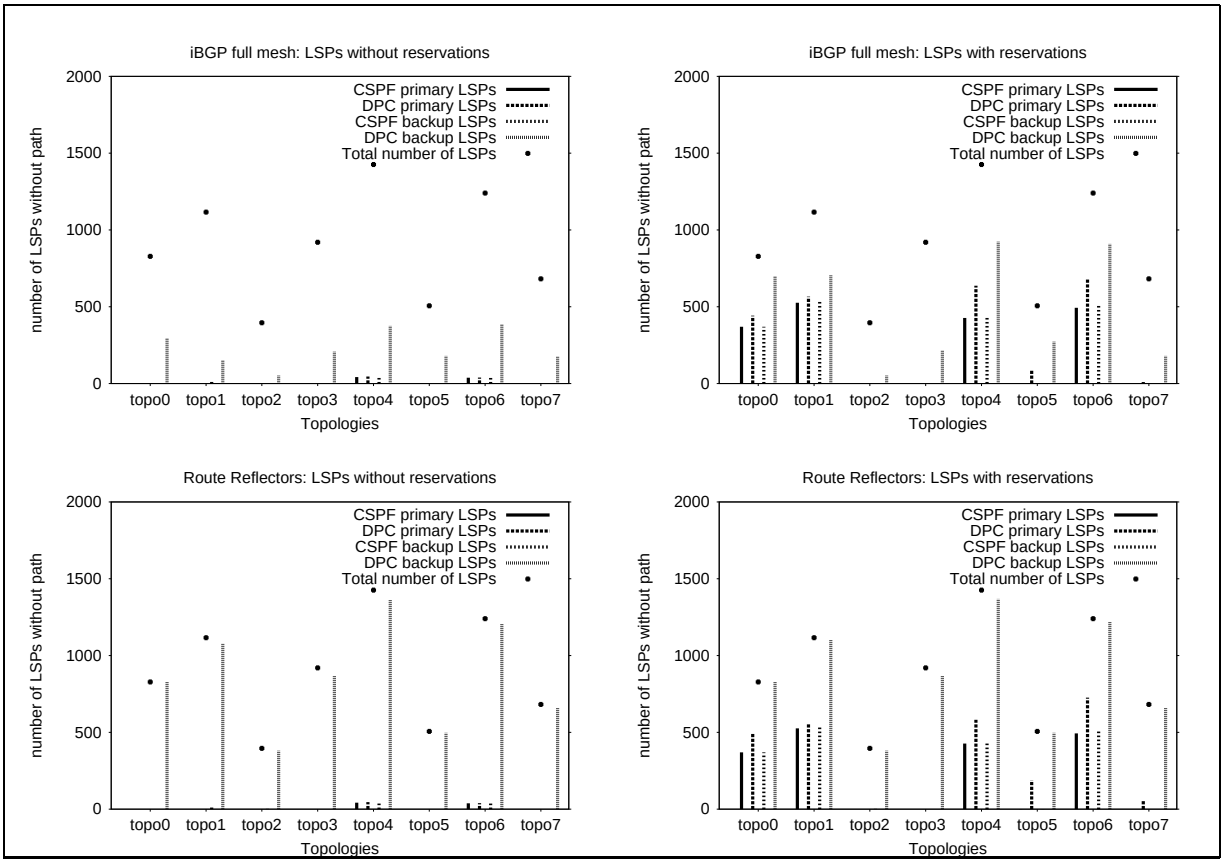

Fig. 2. Number of paths that could not be established

DPC algorithm chooses the NH reachable with the smallest delay. This is a local choice that may not be appropriate to minimize the end-to-end delay. For the backup path, the NHs used on the primary path are pruned from the topology. Bad NH choices, in terms of delay, made for the primary path, leave better alternatives for the backup path. Thus, the backup path may eventually follow the same path as the primary CSPF path.

In the right part of figure 3, we note that some DPC primary paths may have a shorter delay than the respective CSPF primary path when LSPs with bandwidth reservations are established. This results from the different distribution of the paths on the topologies with both computation techniques. With CSPF, the links with low delay will be used first. When there is no bandwidth left on these links, links with higher delay will be used resulting in a degradation of the end-to-end delay of the paths. Since DPC may perform bad choices based on local search, links with low delay may not be used by the first LSPs to be established. This leaves paths with low delays for following LSPs. The problem of balancing end-to-end delay and bandwidth constrained LSPs inside a single domain, with the complete knowledge of the topology, is still unsolved [27]. Thus, finding a solution to the same problem for interdomain LSPs is out of reach today.

V. Conclusion

We evaluated in this paper the difficulty of establishing interdomain LSPs. We showed that BGP-related limitations make the problem of computing constrained end-to-end LSPs difficult, namely the topological information hiding and the unawareness of end-to-end metrics by BGP when choosing its best route.

We illustrated our case by comparing two different LSP computation techniques. The first technique, a centralized one, is based on CSPF and assumes that the intradomain topology of the ASes crossed by the LSP is known. The second technique, fully decentralized, relies on the BGP routes present locally in the routers as well as on the topology of the local domain.

Our simulations show that the decentralized technique is not able to provide end-to-end link and node disjoint paths only based on the BGP routes. Moreover, in large topologies, the establishment of the constrained LSPs with the DPC technique always requires to cranckback. Thus, designing BGP-based interdomain LSPs computation techniques with guarantees will always face the fundamental trade-off between the scalability of the interdomain path computation and the quality of the paths found in terms of the considered metrics.

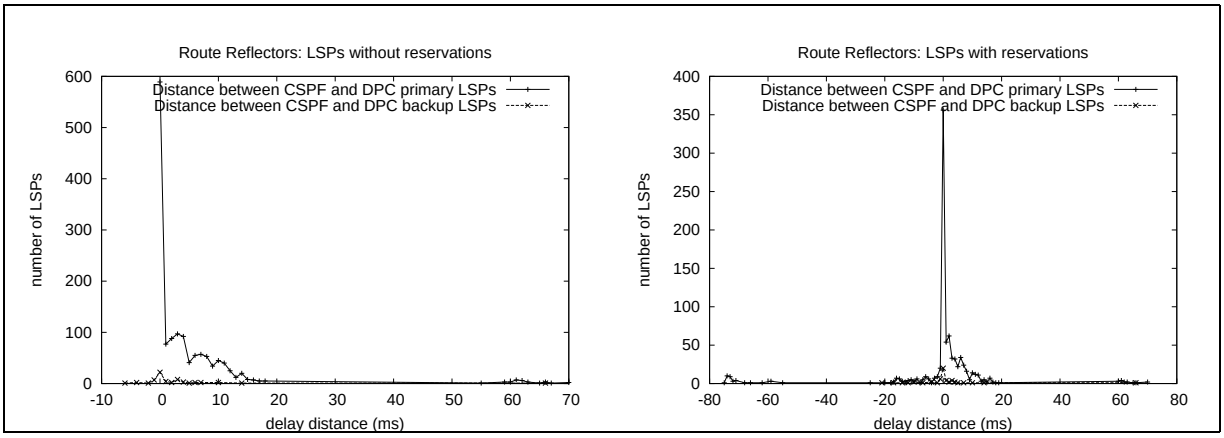

Fig. 3. Delay distance (topo4)

VI. Acknowledgements

This work was supported by the Waloon Government (DGTRE) within the TOTEM project (`http://totem.info.ucl.ac.be`). The authors thank Stefaan De Cnodder and Bruno Quoitin for their comments. The authors are grateful to Bruno Quoitin for the C-BGP tool [31] used to produce the results in this paper.

References

[1] B. Davie and Y. Rekhter, *MPLS Technology and Applications*. Morgan Kaufmann Series in Networking, 2000.

[2] E. Rosen and Y. Rekhter, "BGP/MPLS IP VPNs," September 2003, internet draft, draft-ietf-l3vpn-rfc2547bis-01.txt, work in progress.

[3] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," December 2001, rFC 3209.

[4] P. Pan, G. Swallow, and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels," March 2004, internet draft, draft-ietf-mpls-rsvp-lsp-fastreroute-05.txt, work in progress.

[5] X. Xiao, A. Hannan, B. Bailey, and L. Ni, "Traffic engineering with MPLS in the Internet," *IEEE Network Magazine*, pp. 28–33, March 2000.

[6] L. Andersson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas, "LDP specification," January 2001, internet RFC3036.

[7] Y. Rekhter and E. Rosen, "Carrying label information in BGP-4," May 2001, rFC 3107.

[8] L. Fang, "Meeting VPN customer requirements: Lessons from real world deployments," February 2004, mPLS World Congress 2004.

[9] R. Zhang and J. Vasseur, "MPLS Inter-AS traffic engineering requirements," June 2004, internet draft, draft-ietf-tewg-interas-mpls-te-req-07.txt, work in progress.

[10] M. Carugi and J. D. Clercq, "Virtual Private Network Services: Scenarios, Requirements and Architectural Constructs from a Standardization Perspective," *IEEE Communications Magazine*, vol. 42, no. 6, June 2004.

[11] P. Aukia, M. Kodialam, P. Koppol, T. Lakshman, H. Sarin, and B. Suter, "RATES: A server for MPLS Traffic Engineering," *IEEE Network Magazine*, pp. 34–41, March/April 2000.

[12] F. Blanchy, L. Mélon, and G. Leduc, "An efficient decentralized on-line traffic engineering algorithm for MPLS networks," in *18th International TELETRAFFIC CONGRESS - Providing QoS in Heterogeneous Environments*, vol. 5a, Berlin, Germany, August 31st - September 5th 2003, pp. 451–460.

[13] H. Smit and T. Li, "IS-IS extensions for Traffic Engineering," August 2003, internet draft, draft-ietf-isis-traffic-05.txt, work in progress.

[14] D. Katz, K. Kompella, and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2," September 2003, rFC 3630.

[15] J. Vasseur, C. I. (Editors), R. Zhang, X. Vinet, S. Matsushima, and A. Atlas, "RSVP Path computation request and reply messages," July 2004, internet draft, draft-vasseur-mpls-computation-rsvp-05.txt, work in progress.

[16] C. Pelsser and O. Bonaventure, "Extending RSVP-TE to support inter-AS LSPs," in *2003 Workshop on High Performance Switching and Routing (HPSR 2003)*, Turin, Italy, June 24-27th 2003.

[17] S. D. Cnodder and C. Pelsser, "Protection for inter-AS MPLS tunnels," July 2004, internet draft, draft-decnodder-ccamp-interas-protection-00.txt, work in progess.

[18] G. Cristallo and C. Jacquenet, "Providing quality of service indication by the BGP-4 protocol : the QoS_NLRI attribute," June 2003, internet draft, draft-jacquenet-qos-nlri-04.txt, work in progress.

[19] J. Stewart, *BGP4 : interdomain routing in the Internet*. Addison Wesley, 1999.

[20] H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "Towards Capturing Representative AS-Level Internet Topologies," *Computer Networks Journal, Elsevier*, vol. 44, no. 6, pp. 737–755, April 2004.

[21] P. Traina, "Autonomous system confederations for BGP," June 1996, internet RFC 1965.

[22] T. Bates, R. Chandra, and E. Chen, "BGP route reflection - an alternative to full mesh iBGP," April 2000, internet RFC 2796.

[23] T. Griffin and G. Wilfong, "On the correctnes of iBGP configuration," in *SIGCOMM'02*, Pittsburgh, PA, USA, August 2002, pp. 17–29.

[24] A. Basu, C. L. Ong, A. Rasala, F. B. Shepherd, and G. Wilfong, "Route oscillations in I-BGP with route reflection," in *SIGCOMM'02*, Pittsburgh, PA, USA, August 2002.

[25] T. Bates, R. Chandra, and E. Chen, "BGP route reflection - an alternative to full mesh IBGP," November 2004, internet draft, draft-ietf-idr-rfc2796bis-01.txt, work in progress.

[26] A. Farrel, A. Satyanarayana, A. Iwata, N. Fujita, G. Ash, and S. Marshall, "Crankback Signaling Extensions for MPLS Signaling," July 2004, internet Draft, draft-ietf-ccamp-crankback-02.txt, work in progress.

[27] K. Gopalan, T. Chiueh, and Y. Lin, "Load Balancing Routing with Bandwidth-Delay Guarantees," *IEEE Communications Magazine*, vol. 42, no. 6, June 2004.

[28] A. Farrel and S. D. Cnodder, "Exclude Routes - Extension to RSVP-TE," July 2004, internet draft, draft-ietf-ccamp-rsvp-te-exclude-route-02.txt.

[29] R. Mahajan, N. Spring, D. Wetherall, and T. Anderson, "Inferring link weights using end-to-end measurements," in *2nd Internet Measurement Workshop (IMW2002)*, Marseille, France, November 6-8th 2002. [Online]. Available: http://www.cs.washington.edu/research/networking/rocketfuel/

[30] W. D. Grover, *Mesh-Based Survivable Networks: Options and Strategies for Optical, MPLS, SONET and ATM Networking*. Pearson Education, 2003.

[31] B. Quoitin, "C-BGP, an efficient BGP simulator," http://cbgp.info.ucl.ac.be/, March 2004.