Cooperation of Control and Management Plane for Provisioning in MPLS networks

E. Grampín
Universidad de la República,
Julio Herrera y Reissig 565,
11300 Montevideo,
Uruguay
grampin@fing.edu.uy

J. Serrat
Universitat Politècnica de Catalunya,
Jordi Girona 1- 3, Campus Nord,
08034 Barcelona,
Spain
serrat@tsc.upc.es

Abstract

This paper presents an alternative method to setup Label Switched Paths (LSPs) with Quality of Service (QoS) constraints in Multiprotocol Label Switching (MPLS) enabled networks[*]. The motivation behind this work is to gather the advantages of the Control Plane based approach and the Management Plane based approach to establish LSPs. In other words we aim for a mechanism with setup times as shorter as those realizable by Control Plane signalling and with Traffic Engineering capabilities as powerful as in management systems. The concept has been prototyped and initial tests have been performed in a simulated network environment. The outcome of these tests show that our system can be nearly as fast as conventional Control Plane setup mechanisms and by construction it can take advantage of all the network information got by the Management Plane for Constrained-Based Routing (CBR) purposes. Therefore, a trade-off between speed and resource optimisation is feasible. The implementation of the concept in real networks requires platforms with standard control and management interfaces, that have the role of peer network nodes and management agents at a time.

Keywords

MPLS, Constraint-Based Routing, Cooperation of Control and Management Plane.

1. Introduction

Carriers and Service Providers (SPs) are adopting the MPLS architecture augmented with the Differentiated Services model as a base for the convergence of multiple services and supporting networks, allowing for automatic or

---

[*] This work is part of a project undertaken by Universidad de la República (UdelaR) and ANTEL, the public Telco of Uruguay, for development of a Metropolitan Multiservice Network with MPLS over optical transport.

semiautomatic connectivity setup. Standards like ASON/ASTN [1][2] and Generalized MPLS (GMPLS) [3] have been proposed to extend the network Control Plane capabilities for this purpose, while other initiatives have also explored the feasibility of using the Management Plane [4]. Anyhow, SPs willing to deploy multiservice networks need an operational framework to support Service Level Agreement (SLA) guarantees, QoS assurance and overall interworking service management, supported by Traffic Engineering (TE) tools.

TE involves evaluation and optimization of network operation. Among the objectives of TE is worthy to mention the control and optimization of the routing function, which should satisfy the QoS requirement for every admitted connection, achieving global efficiency in sharing resources. Constraint-Based Routing (CBR) is the TE mechanism for computing a feasible network path based on a traffic description and a set of constraints. CBR, as a generalization of QoS routing, evolves from current topology driven hop-by-hop Internet Interior Gateway Protocols (IGPs).

CBR is usually qualified as offline and online. Offline CBR performs path computation outside network elements, in a Path Computation Server (PCS). It takes as input a known static traffic matrix and, based on a detailed and accurate topology map (built with information gathered from the network), it computes the optimal network paths for that given traffic matrix. The drawback of such approach is that a detailed traffic matrix has to be known in advance. The solution is valid for the given static input, but it cannot satisfy new traffic demands.

Online CBR is a routing mechanism embedded on network elements intelligence. Such a routing process receives, as input, dynamic traffic requests and has no knowledge of future requirements. Given this traffic demand and based on a dynamic (and possibly incomplete) network state it computes feasible paths for that demand. The drawback of such approach is that it has to be performed under strict operational requirements (e.g., computational complexity, algorithm convergence time) and has to be resilient to transient network conditions.

Solutions for connectivity setup by the Management Plane usually perform off-line route computation on a PCS. This solution, while logically correct and robust for static networks, cannot meet the timing requirements for dynamic provisioning. The update of network information often requires several interactions between the management application and network devices, and inaccuracies are very likely to happen. Delay is also relevant in provisioning time, when the management application need to configure every Label Switched Router (LSR) along the current LSP path. Another problem is caused by equipment vendors, which usually provide non-standard management applications, as a market differentiator from competitors. Is very unlikely to manage certain equipment with other vendors' management software.

The traditional Control Plane-based setup of LSPs based on IGP next-hop routing cannot ensure the fulfilment of QoS and policy/administrative constraints. This behaviour is slightly improved by some commercial routers which provide some basic online CBR, usually regarded as Constraint Shortest Path First (CSPF). Since CBR with more than one restriction is a well-known NP-complete problem, computation power is unbounded. This prevents the introduction of full CBR capabilities into network devices, since they have scarce computational resources mainly devoted to packet forwarding. Moreover, different implementations of CBR algorithms lead to the impossibility of fulfilling network-wide TE objectives. The result is that only local optimization can be performed, but the network, considered as a whole, waste resources and consequently may reject connectivity requests, leading SPs to revenue losing.

This paper presents a combined, synergetic approach, between a purely Control Plane and a purely Management Plane provisioning approach. An entity called Routing and Management Agent (RMA) performs online CBR and acts as a peer network node for signalling purposes but excluding packet forwarding. Being a logically centralised platform, the computational resources of the RMA are practically unbounded and it can also take advantage of its centralised position to cooperate with the Management Plane to carry out global optimization tasks or to apply routing policies. Different RMAs can cooperate together at the Management Plane for long term optimization or inter-provider connectivity setup.

The paper is structured as follows. In Section 2 we present the proposed solution, highlighting the process to setup an LSP and summarising its main characteristics. This Section continues describing how the system makes use of standardised signalling protocols and ends with a description of the RMA system architecture. In Section 3 we present the most relevant related work while the evaluation results obtained in a simulated environment are presented in Section 4. The last section make some concluding remarks and point out ongoing work.

2. LSP setup by means of a Routing and Management Agent

The proposal stands for decoupling packet forwarding and path establishment from path computation, based on the functionality of an entity called Routing and Management Agent (RMA), which interacts with both the Control Plane and the Management Plane.

Network nodes (LSRs) main function is packet forwarding, and they can also timely perform LSP setup using the standard signalling protocol RSVP-TE[5]. We propose to migrate the CBR computation from network devices to a specific-purpose server, the RMA. This agent can perform online/offline CBR with arbitrary constraints using arbitrary large computation power. To achieve this goal, the RMA can make use of existing algorithmic and computing techniques; for example, classic High Performance Computing (HPC) strategies (i.e. problem parallelization) can be applied to solve the problem.

As shown in Figure 1, the RMA is a peer node at the routing and signalling plane trough the appropriate interfaces, but avoiding traffic forwarding. Also, we assume that in order to fulfil its main objective, the RMA has enough computation power to solve the CBR problem in near real time. This entity, while enabling a Control Plane-based provisioning, can be used as a complementary Traffic Engineering tool by management applications, using its interface towards the Management Plane.

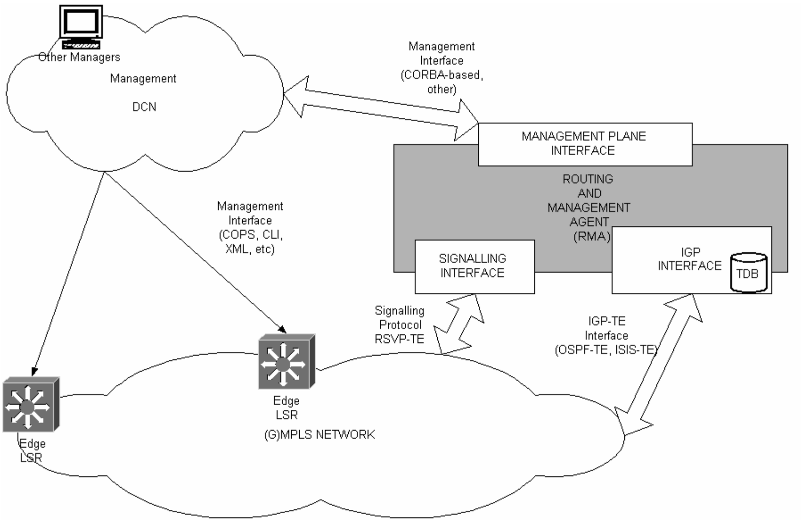

Figure 1 - A Routing and Management Agent in the MPLS Network

This approach requires an innovative usage of the signalling protocols in order to enable Ingress LSRs which need to satisfy a connectivity request, to use the RMA as a computation tool. Current status of RSVP-TE signalling protocol enables this functionality without any further extensions as shown in section 2.3.

2.1. LSP setup steps

Let's consider an LSP setup upon reception of a request at an Ingress LSR. A basic pre-requisite is that all nodes on the network, including the RMA, run the RSVP-TE protocol. Figure 2 shows the numbered sequence of events as described hereafter:

(1) The management application (1), or the client (1'), configures the ingress LSR by means of a suitable protocol like COPS-PR [6]. This implies the specification of the QoS constraints and a "dummy" Explicit Route (ER) towards destination. The first hop of this dummy ER must be the RMA.
(2) The ingress LSR initiates a LSP setup issuing a RSVP-TE Path Message towards the RMA, the first loose hop of the Explicit Route Object (ERO), as configured in the previous step.
(3) The RMA receives the Path Message, and computes an Explicit Route based on the QoS descriptors carried in.
(4) Once computed, the RMA replaces the "dummy" ERO object by the calculated ERO and sends the modified Path Message downstream to the Egress LSR.
(5) The Egress LSP issues a Resv Message upstream to the Ingress LSP. This message, while passing through the network, signal the reservation of the resources needed by the Traffic Engineered LSP.

(6) Once the Resv Message reaches the Ingress LSR, the LSP is established and traffic can be assigned to the appropriate Forwarding Equivalent Class (FEC).

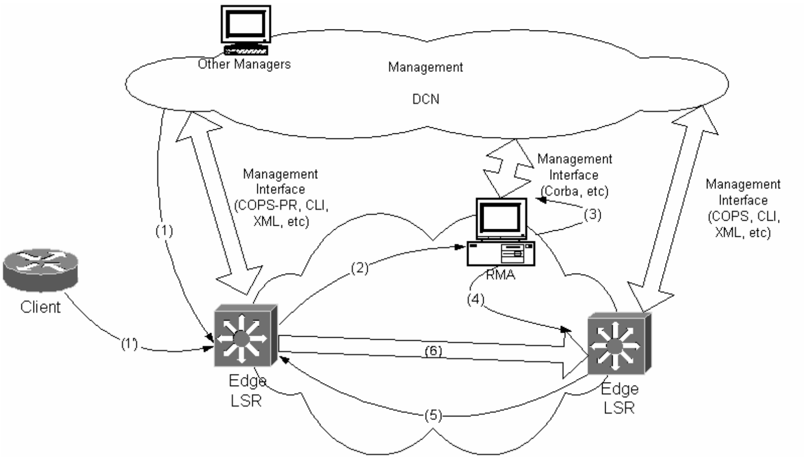

Figure 2 - LSP establishment using the RMA

The main advantage of this proposal is a timely LSP setup provided by the Control Plane, and the CBR accurate computation provided by the Management Plane, since the RMA works like a standard PCS but with dynamic knowledge of the network status and provisioning demands, due to the peering at the signalling and routing protocol level. Multi-vendor router interoperability is also guaranteed because a standard signalling protocol is used. The RMA can be thought as a logical centralized entity. Nevertheless, it is possible to have a number of RMAs in a network domain for load sharing, cooperating using their management interfaces.

2.2. Additional process advantages

- Global optimization of network resources

When the CBR process is performed online, distributed in the CPUs of network nodes, is not possible to obtain a global optimal solution. Nevertheless, this goal can be achieved using a PCS with complete knowledge of the network status and administrative/policy constraints (the RMA CBR engine is a PCS).

Online optimization of network resource usage should be performed carefully, following a "make-before-break" approach not to cause service disruption. In addition, the proposed RMA can be used as a general Traffic Engineering (TE) tool. For example, classic off-line CBR computing can be performed once in a while (i.e. in a weekly basis) to re-optimize the usage of network resources.

Classical off-line route computing experiences synchronization problems between the Topological DataBase (TDB) and the actual network status. The RMA is a peer at the routing level, avoiding inaccuracies caused by delays of the management communication protocols. Nevertheless, inaccuracies inherent to

IGPs still have to be considered and techniques like the discussed in [7] and [8] can be used.

- Offload router processing

The proposed solution boosts the router main function: packet forwarding, offloading the CBR processing from routers' CPU. This enables the usage of arbitrary complex algorithms for CBR computing without effect on network nodes performance. Besides this, freeing the routers' CPU and memory will improve overall network performance, decreasing packet latency and congestion.

- Operational advantages for Configuration Management

In a network with hundreds of routers, software updates needed when a new feature is added to the routers' operating system is a tedious an error prone task. In the approach described, any update on routing strategies and algorithms is done only in the RMA(s), reducing the aforementioned risk.

- Multi-vendor provisioning support based on standards

The RMA operates as a signalling peer in the MPLS network, intercepting RSVP-TE messages for explicit route specification. Furthermore, network status and topology are gathered as a routing peer. Since both the signalling and routing protocols are industry standards, no proprietary extensions are needed to use it as a CBR computation engine in multi-vendor environments.

- Support for inter-area LSP setup

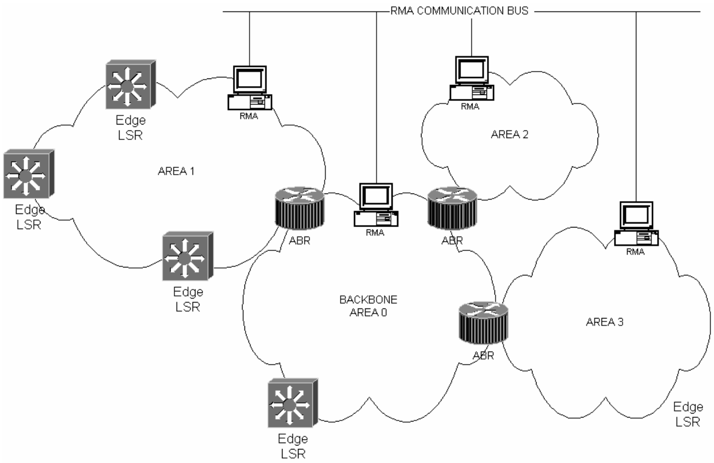

Figure 3 - RMA support for Inter-area LSP Provisioning

As stated in [9], the current set of MPLS Traffic Engineering mechanisms have been to date limited to use within a single IGP area. It would be useful to extend MPLS TE capabilities across IGP areas to support inter-area resource optimization. The RMA architecture can be augmented to provide a solution to inter-area TE-optimized LSPs.

In the example network with OSPF areas shown in Figure 3, an LSP traversing areas 1 and 2 will use an Explicit Route with strict hops within AREA 1, a loose hop between ABRs in the backbone area and another loose hop to destination. Each area has to compute and transform these loose hops into strict hops for LSP setup. If each area solves the LSP establishment using a RMA, is feasible to coordinate RMAs to have a complete (i.e., explicit) knowledge of the whole network, allowing the specification of inter-area end-to-end path computation, based in the collaboration between RMAs.

Note that not only local per-area optimization is needed, but a global optimization based on network-wide traffic engineering objectives. This imply that the LSP setup process should be more complex than simple concatenation of per-area computed LSPs.

2.3. RMA usage of MPLS signalling protocols

Following IETF consensus to focus its efforts on RSVP-TE as the MPLS signalling protocol for Traffic Engineering applications, we consider only RSVP-TE as the signalling protocol for TE-LSP setup in the context of this work.

A detailed analysis of the relevant documents (Section 2.2 of [5]), leads to the conclusion that the claimed RMA behaviour can be implemented within the standard. In fact, when the EXPLICIT_ROUTE Object (ERO) is present in a Path Message, it is forwarded towards its destination following this ERO, and nodes may also modify the ERO before forwarding the Path Message. This enables the RMA to replace the ERO upon reception of a Path Message from an Ingress LSR before it is forwarded to the next hop. In practice this means that an Ingress LSR that need to setup an LSP can send a Path Message to the RMA specifying the QoS parameters and the egress point with a "dummy" ERO, and let the RMA to compute a feasible Explicit Route using the proper CBR algorithms. The RMA, in turn, will replace the "dummy" ERO by this Explicit Route, and the rest of the reservation will follow the standard procedures specified in [10] and [5].

The Resv Message (reverse path) will flow from the Egress LSR towards the Ingress following the newly created ERO, as specified in the standard.

2.4. RMA functional components

The RMA is built using a component-based framework, with basic scheduling and other supporting components needed to build the described functionality. The interfaces and "core" components shown in Figure 4 are described below:

Signalling Interface

This component implements the RSVP-TE signalling. Its basic function is to intercept the Path Message from Ingress LSRs and replace the "dummy" ERO by an ERO computed by the appropriate RMA component before forwarding the message downstream towards the Egress LSR. Error control shall be implemented to prevent LSP setup throughout the RMA.

IGP Interface

This component implements the communication at the routing protocol level, performing network topology and state gathering. The RMA is like any node in the routing graph but, as said before, actions should be taken to prevent the forwarding of valid data traffic through it. The goal of this component interface is to maintain its Topology DataBase (TDB), which in turn is used to feed the Traffic Engineering DataBase (TE-DB), the basic information source for CBR computation. As mentioned above, to prevent update information inaccuracies in the TDB, some of the referenced strategies can be applied. Any of the TE-enabled IGPs can be used: OSPF-TE [11] and ISIS-TE [12].

Management Plane Interface

This component implements the interaction with management applications, which enables the RMA to be used as a Traffic Engineering component for high-level applications. Other possibilities of this interface include the usage of the RMA as a monitoring agent. This interface shall be based in standards, enbling a distributed RMA architecture to be deployed.

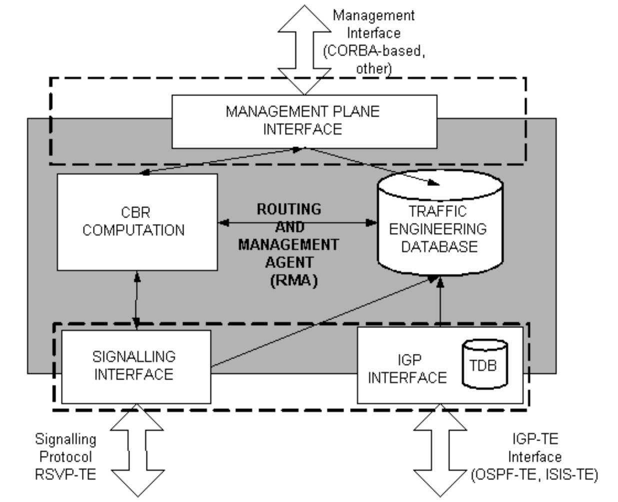

Figure 4 - RMA Internal Architecture[1]

CBR Computation Component

This is the core of the RMA, which provides the intended functionality: a computation engine for Constraint-Based Routing. The component implements the needed algorithms to solve the Path Computation problem with multiple restrictions. Well-known algorithms and heuristics can be used to accomplish the

---

[1] The dotted rectangles represent physical interfaces towards the Management DCN and the network nodes (LSRs).

intended goal [13], making sure that route computation time is limited (i.e. by the usage of polynomial-time CBR algorithms); also, the applicability of HPC concepts to this problem will be explored.

Traffic Engineering DataBase Component
The TE-DB contains the up-to-date information regarding link states in the network, gathered by the IGP Interface. Additional information, like constraints and administrative policies can also be made persistent in the TE-DB. This information, which defines the TE objectives of the network, will typically come from Policy-Based Management applications. Besides this, the RMA could also implement a monitoring interface (i.e. using SNMP) to gather information not provided by the IGP, as mentioned in [14]. The design of the TE-DB is vital in order to speedup CBR computation with minimal inaccuracies.

3. Related Work

Many proposals have been formulated regarding TE path provisioning:

RATES [15] is a centralized solution for Traffic Engineering in MPLS networks presenting a component-based, expandable architecture. Also it is worth mentioning an interesting innovation proposed by RATES: the usage of a Network Topology and State Discovery component to gather network information using the IGP. This idea is used in the RMA design; nevertheless, RATES provisioning is based in a standard management component. No implementation and/or simulation results are presented.

Wise<TE> is a proposal of a Traffic Engineering server for MPLS networks [16] with a similar approach to RATES. The basic difference is that Wise<TE> provisioning is done vendor-specific for Cisco and Juniper routers (special purpose mediators are provided for these vendors). No quantitative evaluation is presented, although underway implementation is shown.

Both RATES and Wise<TE> suffer from the drawbacks already stated for Management Plane-based solutions. Nevertheless, the RMA implementation can benefit from their distributed architecture.

4. Proof of concept

The objective of this proof of concept is to have a first hint on the performance of the proposal in a controlled environment. A basic assumption is that the route computation time is bounded. For each considered network topology, explicit routes will be pre-computed and made available to the RMA component, so in this stage of testing only the signalling interface has been implemented.[2]

A comparative performance evaluation has been undertaken over representative Internet topologies, using the Waxman random graph model [17].

In order to compare LSP setup time by the Control Plane, the Management

---

[2] A complete implementation is foreseen, as mentioned in Section 5.

Plane and the hybrid approach used by the RMA the following setup is used:

- The Control Plane approach signal the LSP using RSVP-TE, following a network path computed by the IGP.
- The Management approach assigns the manager role to the head-end node, which communicates with each node along a pre-computed path.
- The RMA approach defines a node with the RMA role; in this case the LSP is signalled as described in Section 2.1. The simulation avoid the problem of bounding routing computation time and accuracy providing pre-computed paths to the RMA element.

The three approaches measures the whole setup time, including communication time between the considered entities.

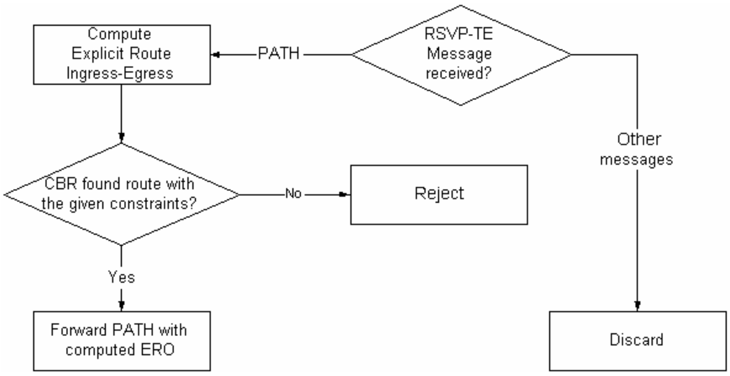

Figure 5 - RMA algorithm

Figure 5 depicts the basic algorithm carried out by the node with the role of RMA. Basic control is performed to prevent LSP setup throughout the RMA (i.e. discard Resv Messages). This and other necessary network functionality has been built over the NS simulator [18] using MPLS and RSVP-TE extensions [19].

The results displayed in Figure 6 and Figure 7 were obtained averaging a hundred realisations of 100 and 1000 node topologies respectively.

The parameters considered in these simulations are the location of the RMA in the network graph and the number of hops contained in the LSP. The tests are performed under controlled network traffic without congestion. The selection of the node with the RMA role is based on node connectivity, i.e. the total number of links of the considered node towards its neighbours in the network graph (node degree). Note that the global network routing table shall be considered for such selection.

Dotted lines in Figure 6 and Figure 7 represent average times for LSPs setup by the Control Plane and continuous lines represent average time for LSPs setup by the RMA. Note that for a given RMA node degree and a given number of hops per LSP, the setup time for "pure" Control Plane is lower than our RMA.

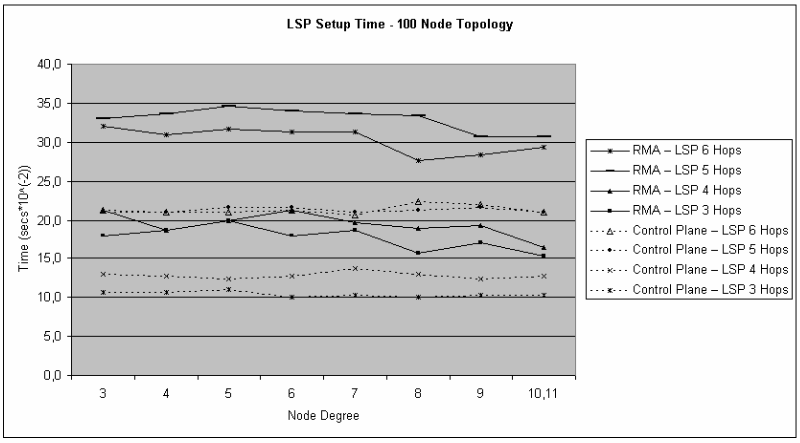

Figure 6 - LSP Setup Time - 100 Node Topology

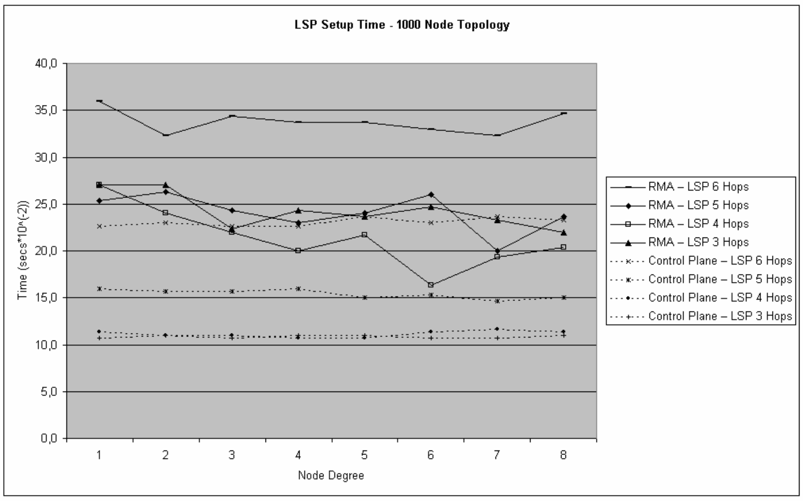

Figure 7 - LSP Setup Time - 1000 Node Topology

Nevertheless, the penalty paid to run the RMA in the worst case is by 150 ms, that is more than acceptable in a real time setup process. In fact, note that the minimum setup time is already 100 ms. Also, the time required to setup a LSP increases with the number of hops. These are consistent results for both network sizes. Moreover, the performance of the RMA approach is practically independent of the location of the RMA in the network.

Figure 8 and Figure 9 show the average setup time for LSPs by the Management Plane, the Control Plane, and the RMA, as a function of the number of hops in the LSP. The averages for these results were obtained with a number of a hundred realisations of 100 and 1000 nodes.

Note that the management solution increases linearly with the number of hops of the configured LSP, since the management application must communicate and configure each node along a given LSP. On the other hand, the RMA has the same moderate increase trend as the pure Control Plane approach and should converge to the same value as the number of hops increases.

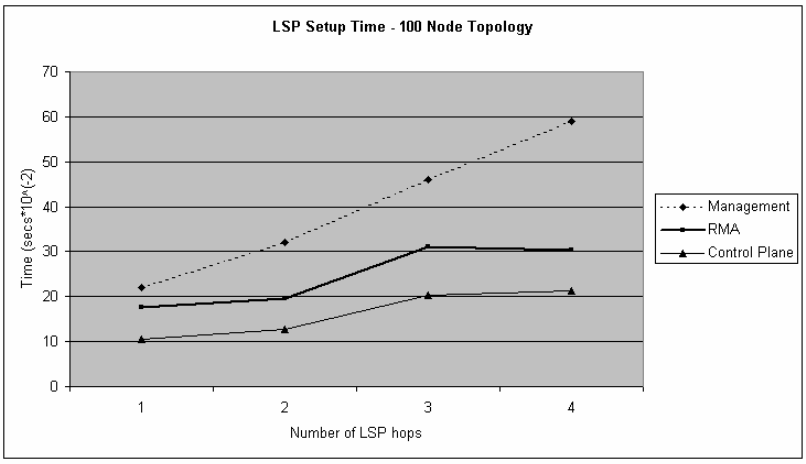

Figure 8 - LSP Setup Time as a function of the number of hops in the LSP

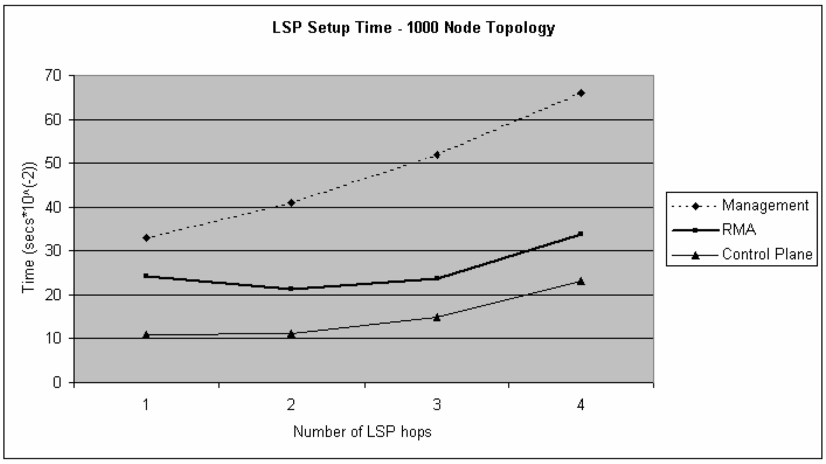

Figure 9 - LSP Setup Time as a function of the number of hops in the LSP

These results show the "halfway" nature of the RMA solution: it clearly performs better than the Management Plane, and a bit worse than the Control Plane. Moreover, note that the RMA solution performance tends to be independent from the RMA node degree (that is, the location of the agent in the network). Also note that the size of the network affects RMA performance by a very small factor. The timing difference between the RMA and the Control Plane solutions, which remains almost constant under any condition, is due to the processing time at the RMA. A real implementation shall take this fact into account for optimization.

The quantitative evaluation reveals the trade-off between a slightly slower response of the RMA solution vs. the aforementioned improvement in the routing function. These preliminary results validate the promising perspectives of the RMA approach and recommend additional research as described hereafter.

5. Conclusion and future work

This paper presents a proposal that combine the strategies of the Control and Management Plane for LSP provisioning into a hybrid process integrated into the RMA, an entity that peers with both network devices and management systems.

The performance of the proposed solution has been obtained by simulation and compared with the pure Management and Control Plane solutions in the same environment. Results show a very reasonable degradation in comparison with Control Plane LSP provisioning, because our solution enables a global optimization of network resources by means of an improved routing function outside LSRs, which also offload router processing, boosting the packet forwarding functionality. The RMA solution is entirely based on a standards protocols, and enables network administrators to control the routing strategies (i.e. network-wide TE objectives). Moreover, support for inter-area LSP setup is seen feasible.

The simulation results are promising but care must be taken to get the RMA solution closer to pure-Control Plane performance.

Further validation is being conducted, using alternative network topology models. Regarding computation accuracy with time constraints, different exact algorithms and heuristics in a parallel computational environment are being tested in the context of the aforementioned project for development of a Metropolitan Multiservice Network with MPLS over optical transport in Uruguay. A real testbed composed of Linux-based and commercial routers is being installed for field evaluation of the RMA architecture. Current work involves adaptation of the RSVP-TE engine used in the simulations to a Linux-based MPLS implementation [20] and installation of the routers and optical infrastructure.

Acknowledgements

The authors would like to express their gratitude to Josep Solé, Salvatore Spadaro and Marcelo Yannuzzi from the Department of Computer Architecture (DAC) of UPC for their valuable comments and to Javier Baliosian for his support.

References

[1] ITU-T Recommendation G.8080/Y.1304, “Architecture for the Automatically Switched Optical Network (ASON)”. July 2001.

[2] ITU-T Recommendation G.807/Y.1302, “Requirements for Automatic Switched Transport Networks (ASTN)”. July 2001.

[3] Eric Mannie (Editor), “Generalized Multi-Protocol Label Switching Architecture”. RFC 3945, October 2004.

[4] L. Raptis, G. Hatzilias, F. Karayannis, K. Vaxevanakis, E. Grampin, “An integrated network management approach for managing hybrid IP and WDM networks”, IEEE Network, Vol.17, Issue 3, May-June 2003.

[5] D. Awduche et al, “RSVP-TE: Extensions to RSVP for LSP Tunnels”. RFC 3209, December 2001.

[6] K. Chan et al, "COPS Usage for Policy Provisioning (COPS-PR)", RFC 3084, March 2001.

[7] George Apostolopoulos, Roch Guerin, Sanjay Kamat, Satish K. Tripathi, “Improving QoS Routing Performance Under Inaccurate Link State Information”, In Proceeding of 16th International Teletraffic Congress, June 1999.

[8] X. Masip-Bruin, “Mechanisms to Reduce the Routing Information Inaccuracy Effects: Application to MPLS and WDM Networks”, PhD Thesis, Departament d’Arquitectura de Computadors, Universitat Politècnica de Catalunya (UPC), 2003.

[9] JL Le Roux, JP Vasseur, Jim Boyle, (Editors), “Requirements for Inter-area MPLS Traffic Engineering”, Internet draft <draft-ietf-tewg-interarea-mpls-te-req-03.txt>. Work in Progress. Expiration Date: April 2005.

[10] R. Braden, (Editor), L. Zhang, S. Berson, S. Herzog, S. Jamin, “Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification”. RFC 2205, September 1997.

[11] D. Katz, K. Kompella, D. Yeung, “Traffic Engineering Extensions to OSPF Version 2”, RFC 3630, September 2003.

[12] H. Smit, T. Li, “IS-IS Extensions for Traffic Engineering”. RFC 3784, June 2004.

[13] F.A. Kuipers, T. Korkmaz, M. Krunz and P. Van Mieghem, “A Review of Constraint-Based Routing Algorithms”, Technical Report, June 2002.

[14] C. Casseti, G. Mardente, M. Mellia, M.Munafò, R. Lo Cigno, “On-line routing optimization for MPLS-based IP networks”. Workshop on High Performance Switching and Routing, 24-27 June 2003.

[15] P. Aukia, M. Kodialam, P.V. Koppol, T.V. Lakshman, H. Sarin, B. Suter, “RATES: a server for MPLS traffic engineering”. IEEE Network, Volume: 14 , Issue: 2 , March-April 2000 Pages:34 – 41

[16] T.S. Choi, S.H. Yoon, H.S. Chung, C.H. Kim, J.S. Park, B.J. Lee, T.S. Jeong, “Wise<TE>: traffic engineering server for a large-scale MPLS-based IP network”. Network Operations and Management Symposium, IEEE/IFIP NOMS 2002, 15-19 April 2002.

[17] Waxman, B.M. “Routing of multipoint connections”. Selected Areas in Communications, IEEE Journal on, Vol.6, Iss.9, Dec 1988. Pages:1617-1622.

[18] The Network Simulator - ns-2: http://www.isi.edu/nsnam/ns/

[19] KOM RSVP Engine - KOM Multimedia Communications Lab at Darmstadt University of Technology: http://www.kom.tu-darmstadt.de/rsvp/

[20] MPLS for Linux project: http://sourceforge.net/projects/mpls-linux/