

# PERFORMANCE ANALYSIS OF THE CONTROL AND FORWARDING PLANE IN AN MPLS ROUTER

D.Adami<sup>1</sup>, N.Carlotti<sup>2</sup>, S.Giordano<sup>2</sup>, M.Pagano<sup>2</sup>, M.Repeti<sup>2</sup>

<sup>1</sup>*CNIT Research Unit - Dept. of Information Engineering - University of Pisa (ITALY)*  
*E-mail: davide.adami@cnit.it*

<sup>2</sup>*Dept. of Information Engineering - University of Pisa (ITALY)*  
*E-mail: {s.giordano, m.pagano}@iet.unipi.it, {m.repeti, n.carlotti}@netserv.iet.unipi.it}*

**Abstract:** Multiprotocol Label Switching (MPLS) was originally designed to provide higher packet forwarding rates in network equipment. Nevertheless, it was soon realized that it could also provide other advanced features, such as Traffic Engineering and Virtual Private Networks capabilities. The key feature of MPLS is a strict separation between control and forwarding operations, which reflects on the software and hardware architecture of the routers. The paper presents the results of an experimental study aimed at evaluating the performance of a Label Switching Router (LSR). In particular, the behaviour of the LSR control and forwarding planes has been analyzed in different working conditions as it concerns both the processing and computational effort due to the control plane and the data traffic to be forwarded.

## I. INTRODUCTION

Network layer routing can be partitioned in two basic components: control and forwarding. The former is responsible for the construction and maintenance of the forwarding table, the latter is concerned with the forwarding of packets from input to output interfaces, on the basis of the forwarding table maintained by the router and the information carried in the packet itself.

At present, network routers implement the control and forwarding planes in a distributed way. In particular, the control component consists of one or more routing protocols for the exchange of routing information among the routers as well as the algorithms used to convert the collected information into a forwarding table. The forwarding component consists of a set of algorithms used to take a forwarding decision on a packet. Traditional IP routers typically use destination-based forwarding to determine the next hop of a packet. The longest prefix-match, required in IP address look-up to perform destination-based forwarding, was originally implemented in software, but it is too expensive for core routers, because it requires a high computational effort. To improve the performance of a traditional IP router, a trade-off between scalability and flexibility is necessary: on the one hand, coarse forwarding granularity assures system scalability, on the other hand a system supporting only coarse forwarding granularity may be fairly inflexible, making available a number of differentiated treatments which could be insufficient.

In this context, Multiprotocol Label Switching (MPLS) was originally introduced as a fast switching technique. The key architectural principle of MPLS is a clean functional separation of network layer routing into control and forwarding components. The label switching forwarding component makes a forwarding decision on a packet using a label carried in the packet and a forwarding table maintained by the Label Switching Router (LSR). Just like a control component of any routing system, the label switching control component must provide for consistent distribution of routing information among LSRs as well as consistent procedures for constructing forwarding tables from this information. To support label switching, the control component has also other functionalities, such as creating binding between labels and Forwarding Equivalence Classes (FECs), informing other LSRs of the binding created, taking into account of bindings to construct and maintain the forwarding table. Under ideal conditions, a LSR should be able to forward data at whatever speeds the label switching forwarding component runs, regardless of the computational and processing effort required to the control component by routing and signaling protocols. Therefore, the hardware architecture of a LSR should be designed and implemented in order to satisfy also the new functional requirements of the control and forwarding planes. Concerning this issue, the system architecture of Juniper routers has been split in two portions: the Routing Engine and the Packet Forwarding Engine, which have been respectively designed to handle the general routing operations and the forwarding of packets. The experimental analysis, described in this paper, aims at verifying in practice the ideal separation, both in terms of functionalities and performance, between the control and forwarding plane of an MPLS commercial router. To this purpose, an M10 Juniper router has been configured as a LSR and tests have been performed to investigate how the processing and computational load, due to a control plane based on the RSVP-TE signaling protocol, affects the forwarding performance of the router. The paper is organized as follows: section II and III highlight the main features of MPLS and RSVP-TE. Afterwards, section IV describes the RSVP-TE performance tests. Finally, section V contains the experimental results and section VI sums up the work.

## II. MULTIPROTOCOL LABEL SWITCHING

An MPLS domain is a contiguous set of nodes, called Label Switching Routers (LSRs), which are capable of switching and routing packets on the basis of a label appended to each of them. An LSR that interfaces to a traditional router is called an Edge LSR (E-LSR). With respect to the direction of the traffic flow, it is possible to distinguish between Ingress LSR and Egress LSR. An Ingress LSR receives traffic from a non-MPLS router, while an Egress LSR sends traffic to a non-MPLS router. It is relevant to highlight that each traffic flow has its own Ingress and Egress LSR. In a network supporting MPLS, a label-switched path (LSP) is a unidirectional connection through multiple LSR. The Forwarding Equivalence Class (FEC) is defined in [1] as “a group of IP packets which are forwarded in the same manner (e.g., over the same path, with the same forwarding treatment)”. In

general, the granularity of FECs within a router can vary from very coarse to extremely fine, according to the level of information used in assigning an IP packet to a FEC: at one hand, an FEC can be associated with the flow generated by a particular application for a particular source and destination host pair, at the other hand, a FEC can be associated with all the flows destined to an Egress LSR. A packet is assigned to a FEC only at an Ingress LSR, where the label for the packet is also created. Although MPLS was originally designed to speed up the IP packet-forwarding process and retain the flexibility of an IP-based networking approach, the deployment of MPLS enables *Traffic engineering, Resilience and Virtual Private Networks (VPNs) support* capabilities. MPLS is frequently mentioned among major Quality of Service (QoS) technologies for packet networks. It is worthwhile pointing out that MPLS plays a key role in enabling QoS, but QoS is supported only if MPLS is combined with other technologies such as RSVP-TE and Differentiated Services.

### III. THE RSVP-TE PROTOCOL

In an MPLS network, an LSP must be set up and labels assigned at each hop before traffic can be forwarded. Two different classes of LSPs may be established:

- *Control-driven LSPs*: each LSR determines the next interface for the LSP according to its forwarding table and requests the label to the next-hop router;
- *Explicitly-routed LSPs* (or Constraint-based Routed LSPs, CR-LSPs): the set up message specifies the route for the LSP.

To establish an LSP, a signaling protocol for coordinating the label distribution, explicitly routing the LSPs, reserving bandwidth and preventing loops is necessary. Since the MPLS architecture does not suppose the mandatory use of a single signaling protocol, the following protocols have been proposed for label distribution: Label Distribution Protocol (LDP), Resource Reservation Protocol extensions for MPLS (RSVP-TE), Constrained-based Routed LDP (CR-LDP).

In more details, LDP [2] is suitable for establishing a control-driven LSP, but traffic-engineering capabilities are not supported.

RSVP-TE is an extension of RSVP [3] to establish traffic-engineered LSPs: it satisfies the requirements for traffic engineering, but, due to its soft-state nature, it suffers from inherent scalability problems when the number of LSPs per LSR increases. Moreover, the use of an unreliable transport protocol doesn't guarantee fast failure notification to the endpoints affected by the failure, even though an explicit teardown message is sent.

CR-LDP [4] [5] [6] is a hard-state protocol that extends LDP to carry the explicit route information, the traffic parameters for resource reservation and the options for CR-LSP resilience. It is relevant to highlight that CR-LDP has been designed to address and solve the main drawbacks of RSVP-TE, ensuring reliable transport of signaling messages and providing better scaling properties. Nevertheless, some networking devices manufactures (e.g. Juniper Networks), have not yet implemented CR-LDP in their operating systems, so that RSVP-TE is the only

available signaling protocol to be used in MPLS networks. This is the reason why, in our work, we focused on RSVP-TE only.

#### IV. RSVP-TE PERFORMANCE TESTS

The RSVP-TE performance tests have been carried out by using the Adtech AX/4000 from Spirent Communications. The AX/4000 Broadband Test System [8] has been designed for testing the performance and evaluating the QoS level provided by broadband networks. This device is a modular multi-port system capable of testing simultaneously multiple network technologies such as ATM, IP, Frame Relay and Ethernet at speed up to 10 Gbps. Moreover, the system is capable of generating data traffic with different profiles, performing full-rate analysis, stressing the network nodes in critical working conditions, emulating routing and signaling protocol. Our experimental analysis has been performed by means of the RSVP-TE Forwarding Performance During Tunnel Establishment Test [9], a performance test available within the Spirent Connect software package. This test allows measuring the capability of a Device Under Test (DUT) to correctly forward MPLS labeled packets to stable RSVP tunnels, while other tunnels are continuously set up and torn down (flapped). Figure 1 shows the scenario where the performance tests have been carried out. In particular, two Gigabit Ethernet interfaces of the test device, emulating an Ingress and an Egress LSR, have been connected to an M10 Juniper router, which acts as an LSR. The performance test has been organized as it follows:

1. Two different groups of LSPs are established: the first one consists of the LSPs which, after being created, remain stable for all the duration of the test. The second one is the set of LSPs that are flapped.
2. The Ingress LER generates MPLS labeled packets and sends them to the Egress LER using the LSPs belonging to the first stable group.
3. The Egress LER analyses each packet received.

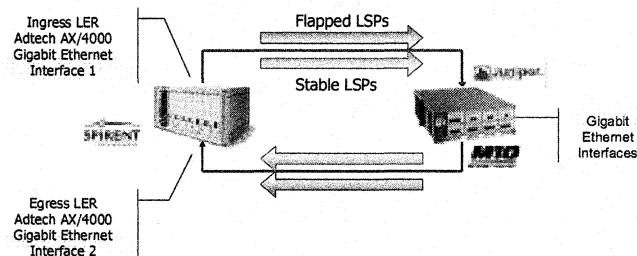


Figure 1: The testbed

In particular, the test consists of two subsequent phases (Fig. 2):

1. **Prior Flapping Phase** (duration time: 15 seconds): a number equal to  $M$  of stable LSPs are established between the Ingress LER and the Egress LER and Constant Bit Rate (CBR) data traffic is sent through the  $M$  LSPs. At the same time,  $N$  LSPs are established and prepared for flapping even if, in this test phase, remain stable.
2. **During Flapping Phase** (duration time: 12 seconds):  $M$  LSPs are stable and data traffic is sent over them, whereas  $N$  LSPs are flapped twice, with the flapping interval set at 5 seconds. During the first flapping event (duration: 5 seconds), the LSPs established for flapping are torn down and, subsequently,  $K$  ( $K \leq N$ ) LSPs are re-established. During the second flapping event (duration: 2 seconds), the group of  $K$  LSPs is torn down and set up again, while, simultaneously, the  $N-K$  LSPs, which have been torn down in the first flapping event, are re-established. At the end of the test, if all the  $N$  LSPs are active, there has not been any failure in the flapping test.

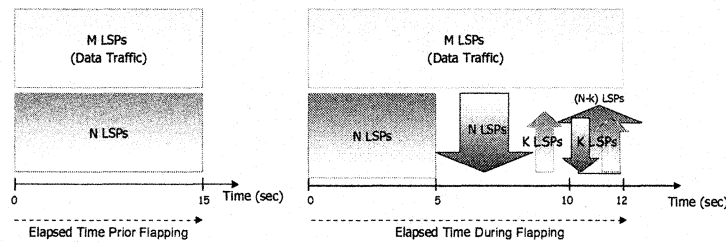


Figure 2: Prior and During Flapping run time

## V. EXPERIMENTAL RESULTS

Two different categories of performance tests have been carried out: the first one (Test 1.x.x and 3.x.x) aims at evaluating the behaviour of the router forwarding plane when the processing and computational effort required to the control plane progressively increases; the second one (Test 2.x.x and 4.x.x) focuses on analysing the router behaviour when the complexity of the operations involving both the control and forwarding plane increases. In more detail, the performed tests are the following:

### 1. Tests 1.x.x and 3.x.x

100 data traffic LSPs ( $M$ ) are set up between the Ingress and the Egress LER. Moreover, the Ingress LER generates and forward labeled CBR traffic into these LSPs. At the same time,  $N$  LSPs are flapped.

The tests have been repeated when  $N$  varies from 100 to 900, with an increment of 100 LSPs from a test to the next and in the following conditions as it concerns the characteristics of the data traffic to be forwarded:

- Total Bit rate: 500 Mbps ( $\rho=0.5$ ) and 900 Mbps ( $\rho=0.9$ ).
- Packet size: 64, 512 and 1500 bytes.

It is relevant to outline that:

1. the overall data traffic is fairly subdivided among  $M$  CBR data sub-stream, so each of the  $M$  LSP carries the same share of traffic to be forwarded;
2. the packet size includes the Gigabit Ethernet header, payload and trailer;
3. for IP over Gigabit Ethernet, when the packet size of the data stream is equal to 64 bytes, the maximum bit rate which can be really reached is limited to 761.9 Mbps, due to the minimum inter-packet gap (96 bits/time) and the preamble (8 bytes) to add to each frame.

## 2. Test 2.x.x and 4.x.x

$M$  data traffic LSPs are set up between the Ingress and the Egress LER. Moreover, the Ingress LER generates and sends labeled CBR traffic into these LSPs. At the same time,  $N$  LSPs are flapped. The tests have been repeated when  $M$  and  $N$  vary from 100 to 500, with an increment of 50 LSPs from a test to the next, in the following operating conditions as it concerns the characteristics of the data traffic to be forwarded:

- Bit rate: 500 Mbps and 900 Mbps.
- Packet size: 64, 512 and 1500 bytes.

### *Analysis of the control plane behaviour*

To characterize the behaviour of the control plane, we have taken into consideration only the first flapping event (5 second).

Figure 3 and 4 respectively show, when the data rate is equal to 500 Mbps and 900 Mbps, the number ( $K$ ) of LSPs that the router is able to re-establish, after tearing down  $N$  LSPs and before the second flapping event.

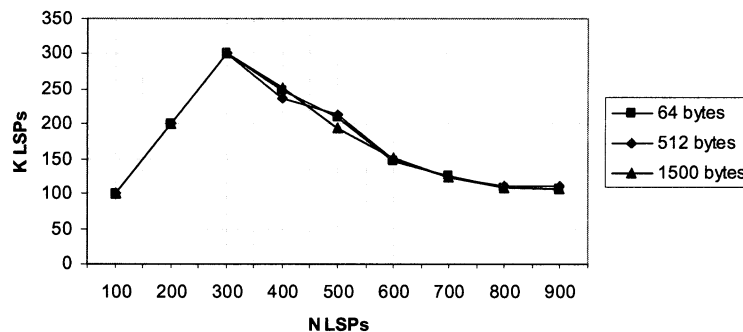


Figure 3: Test 1.x.x - Control Plane behaviour with CBR traffic at 500 Mbps

If  $N \leq 300$ , all the LSPs are torn down and re-established before the second flapping event, regardless of the data streams bit rate and packet size. On the contrary, if  $N > 300$ , the number of LSPs re-established after the first flapping event is always less than the number of LSPs torn down, because the time interval between the two flapping event is not long enough to tear down and re-establish all the LSPs. Moreover, the number of LSPs re-established after the first flapping event slightly changes both with the packet size and data stream bit rate.

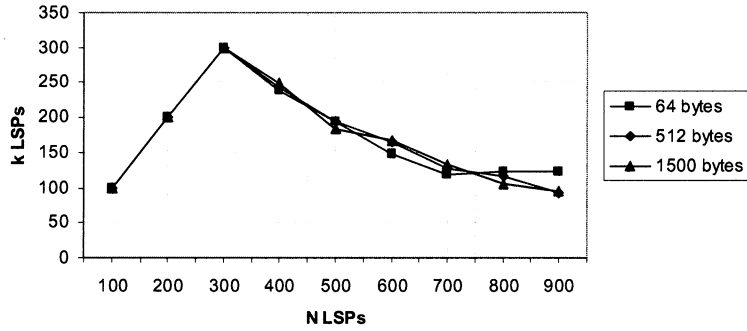


Figure 4: Test 3.x.x - Control Plane behaviour with CBR traffic at 900 Mbps

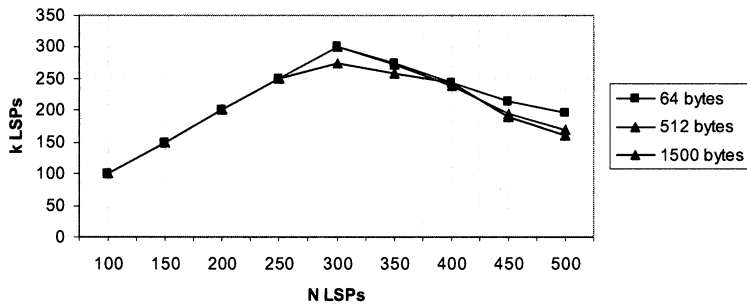


Figure 5: Test 2.x.x - Control Plane behaviour with CBR traffic at 500 Mbps

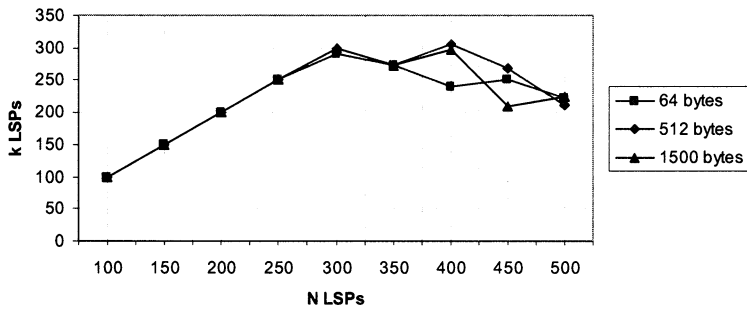


Figure 6: Test 4.x.x - Control Plane behaviour with CBR traffic at 900 Mbps

Figure 5 (data rate equal to 500 Mbps) and 6 (data rate equal to 900 Mbps) respectively show the number ( $K$ ) of LSPs re-established before the beginning of the second flapping event, when the flapped and data traffic LSPs number is

proportionally increased. We can observe that, if  $N \leq 250$ , all the LSPs are torn down and re-established, independently of the bit rate and packet size of the data stream. On the contrary, if  $N$  (and  $M$ )  $> 250$ ,  $K$  is always less than  $N$  and the number of LSPs re-established after the first flapping event changes both with the packet size and bit rate of the data stream.

*Analysis of the forwarding plane behaviour*

To evaluate the performance of the forwarding plane, two metrics have been chosen:

- Packet loss prior and during flapping.
- One-way delay prior and during flapping, computed as the average of the one-way delay measured for the  $M$  data sub-streams.

In all the tests performed, no packet loss was ever experienced.

Table I and II show the results obtained for the one-way delay:

- When the bit rate and the packet size remain equal, the number of flapped LSPs ( $N$ ) doesn't affect the performance of the forwarding plane. For this reason, each table row reports the value of the one-way delay measured for any value of  $N$ .
- In all the tests performed, there is no degradation in the data plane performance prior and during flapping.
- The one-way delay is independent of the test category as well as the bit rate of the data stream, when the packet size is equal, except for 64 bytes packet size.
- When the packet size is equal to 64 bytes and the bit rate is 761.9 Mbps, the one-way delay prior and during flapping dramatically increases. This behaviour may be explained considering that, as previously said, 761.9 Mbps is the maximum theoretical throughput for IP over Gigabit Ethernet in these operating conditions and the packet rate assumes the maximum value (1488085 pkt/s).

First Type Tests		
One-Way Delay		
Test	Prior	During
1.1.x	0.007 ms	0.007 ms
1.2.x	0.016 ms	0.016 ms
1.3.x	0.03 ms	0.03 ms
3.1.x	0.156 ms	0.156 ms
3.2.x	0.016 ms	0.016 ms
3.3.x	0.031 ms	0.031 ms

Table I:  
Average One-Way Delay of the  $M$  data flows

Second Type Tests		
One-Way Delay		
Test	Prior	During
2.1.x	0.007 ms	0.007 ms
2.2.x	0.016 ms	0.016 ms
2.3.x	0.03 ms	0.03 ms
4.1.x	0.157 ms	0.156 ms
4.2.x	0.016 ms	0.016 ms
4.3.x	0.031 ms	0.031 ms

Table II:  
Average One-Way Delay of the  $M$  data flows

The analysis of the control and forwarding plane behaviour clearly shows that the introduction of the MPLS architecture doesn't weigh on the router performance. In particular, the functional separation between the data and the control plane, typical of the MPLS architecture, allows solving the performance issues of traditional IP routers, which are related to the management of signaling protocols and fast forwarding rates.



The hardware and software architecture of next generation routers allows simultaneously handling a high data traffic and a heavy control traffic, due to signaling, exchange of routing information and soft-state refresh messages generated by RSVP-TE.

Finally, we can also state that the introduction of the DiffServ architecture functional components, such as classifier, meter, marker, policer and scheduler shouldn't significantly affect the behaviour of the control and data planes.

## VI. CONCLUSIONS

The paper presents the results of an experimental study aimed at evaluating the performance of the control and forwarding components of an M10 Juniper router that supports Multiprotocol Label Switching. To characterize the behaviour of the control plane, it has been taken into consideration the number of LSPs that the router tears down and re-establishes during the time interval between two subsequent flapping event: the results of the performance tests show that the behaviour of the control plane is not linear and doesn't depend on the bit rate and packet size of the data traffic to forward. As far as the forwarding plane is concerned, using the packet loss and the one-way delay as performance metrics, we found that the forwarding behaviour of the router is independent of the control plane. In fact, no packet loss was ever experienced and the one-way delay of the data streams is independent of the LSP number. Moreover, the RSVP-TE performance tests highlight that there is no degradation in the data plane performance from the prior flapping phase to the during flapping phase. The one-way delay increases with the packet size and is independent of the tests category as well as the bit rate of the data streams for each packet size taken into consideration, except for 64 bytes packets when the bit rate is equal to 761.9 Mbps. In this case, the one-way delay prior and during flapping significantly increases in comparison to the values obtained when the bit rate is equal to 500 Mbps.

## REFERENCES

- [1] E. Rosen, A. Viswanathan, R. Callon "Multiprotocol Label Switching Architecture", RFC 3031, January 2001
- [2] L. Andersson, P. Doolan, N. Feldman "LDP Specifications", RFC 3036, January 2001
- [3] R. Braden, L. Zhang, S. Berson "Resource ReserVation Protocol (RSVP)", RFC 2205, Sept.1997
- [4] B. Jamoussi, L.Andersson, R. Callon "Constraint-Based LSP Setup using LDP", RFC 3212, January 2002
- [5] J. Ash, M. Girish, E. Gray "Applicability Statement for CR-LDP", RFC 3213, January 2002
- [6] J. Ash, Y. Lee, P. Ashwood-Smith "LSP Modification Using CR-LDP", RFC 3214, January 2002
- [7] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001
- [8] <http://www.spirent.com> about "AX/4000 Broadband Test System"
- [9] Spirent Communications Test Methodologies, "RSPV-TE Forwarding Performance During Tunnel Establishment Test"