A New NS2 Simulation Module for Bandwidth Constraints Models in DS-TE Networks

D. Adami¹, C. Callegari², S. Giordano², and M. Pagano²

¹CNIT Research Unit - Dept. of Information Engineering - University of Pisa (ITALY)
E-mail: davide.adami@cnit.it
²Dept. of Information Engineering - University of Pisa (ITALY)
E-mail: {christian.callegari, s.giordano, m.pagano}@iet.unipi.it

Abstract — In recent years, network operators rolled out new revenues-generating services by using IP/MPLS networks. MPLS traffic engineering (TE) allows the creation of end-to-end paths with guaranteed bandwidth as well as the optimization of transmission resources, but it does not provide end-to-end QoS to specific traffic flows. DiffServ-aware MPLS-TE (DS-TE) improves the MPLS-TE model by allowing bandwidth reservations to be performed on a per-class basis. In particular, the Bandwidth Constraints (BCs) model is a building block of the DS-TE architecture, since it establishes how bandwidth is allocated to different traffic classes. This paper deals with the development of a new software module for the simulation of DS-TE networks and BCs models within the Network Simulator 2 (NS2) environment. The results of some simulations are reported to show the operational working of the developed modules and the benefits coming from the use of the DS-TE architecture.

Index Terms — DS-TE, MAM, NS2

I. INTRODUCTION

In the last years, Internet has become the cheapest and most attractive communication facility to interconnect different sites of public institutions as well as private enterprises. Meanwhile, Internet Service Providers (ISPs) have started offering new Quality of Service (QoS)-sensitive network services, such as Voice-over-IP (VoIP), IPTV, Virtual Private Networks (VPN). To increase their revenues, while at the same time keeping capital and operational expenditures down, ISPs have rolled out IP/MPLS [1] networks. However, even if MPLS plays a key role in enabling QoS, QoS is not a fundamental feature of MPLS. More accurately, MPLS provides a connection-oriented environment that enables Traffic Engineering (TE) [2] of packet networks. Traffic engineered networks can guarantee bandwidth for various flows, but to control latency and jitter of time-sensitive applications, MPLS-TE must be combined with functions and mechanisms that provide traffic flows with their class specific treatment (e.g., Differentiated Services [3]).

DiffServ-aware MPLS TE (DS-TE) [4] makes MPLS-TE aware of Class-of-Service (CoS), thus allowing resource reservation with CoS granularity and providing the fault tolerance properties of MPLS at a per-CoS level. By combining the functionalities of both DiffServ and TE, DS-TE is particular useful to achieve the following objectives:

1. Maintaining relative proportion of traffic.
2. Limiting proportion of classes on a link.
3. Providing guaranteed bandwidth services.

For example, DS-TE can be used to ensure that traffic is routed over the network so that, on every link, there is never more than 40 percent (or any percentage) of the link capacity assigned to guaranteed traffic (e.g., voice), while there can be up to 100 percent of the link capacity assigned to other traffic flows. Assuming QoS mechanisms are also used on every link to serve guaranteed traffic separately from the other traffic, it then becomes possible to enforce different "overbooking" ratios for each traffic class. Therefore, for the guaranteed traffic it becomes possible to enforce no overbooking at all, or even an underbooking, so that strict QoS can be achieved end-to-end for that traffic. Also, through the ability to enforce a maximum percentage of guaranteed traffic on any link, the network administrator can directly control the end-to-end QoS performance parameters without having to rely on over-provisioning or on expected shortest path routing behaviour. This is essential for transport of applications that have strict QoS requirements, especially when over-provisioning cannot be assumed everywhere in the network. One of the main building blocks of the DS-TE architecture is the bandwidth constraints (BC) model, which fixes how bandwidth is allocated to different traffic classes. In particular, [5] and [6] define two alternative BC models, called Maximum Allocation Model (MAM) and Russian Dolls Model (RDM), respectively. The complexity of the DS-TE architecture and the great number of available options (TE, traffic recovery strategies, QoS mechanisms, etc.) make necessary to assist network designers and administrators with powerful and flexible simulation tools. Indeed, simulation tools allow to evaluate the performance of DS-TE networks, before their deployment, in terms both of reliability and performance.

Since, up to now, a full open-source simulation tool has not been realized, our efforts have been focused on the implementation of new DS-TE modules for the Network Simulator (NS2) [7]. In previous works, some of the authors of this paper developed several modules for the simulation of MPLS-TE networks. The research activity, described in this paper, will mainly be focused on the extensions and enhancements introduced in NS2 to support the DS-TE architecture and, more specifically, on bandwidth allocation models. The paper is organized as follows. Section 2 provides an overview of the DS-TE architecture, focusing on MAM and RDM bandwidth constraints models. Next, section 3 describes some features of the new developed software module, whereas
Section 4 shows the results of some simulations carried out to show the application of the bandwidth constraints module in a specific case study. Finally, section 5 contains some concluding remarks.

II. DS-TE ARCHITECTURE

As already stated in the introduction, DS-TE is emerging as a powerful architecture to provide QoS on a per-CoS level. To achieve such differentiated treatment, RFC 3564 [4] introduces the concept of Class-Type (CT), defined as “the set of Traffic Trunk crossing a link, that is governed by a specific set of Bandwidth Constraints. CTs are used for link bandwidth allocation, constraint based routing and admission control. A given traffic trunk belongs to the same CT at all links”.

RFC 3564 requires support of up to 8 CTs, from CT0 to CT7. It also introduces a pre-emption mechanism, based on two distinct attributes, the set-up and holding priorities, which may be assigned a value from 0 to 7. The combination of a CT and a pre-emption priority defines a TE-Class. Since each traffic class may be associated to all the pre-emption priority values, we can have up to 64 distinct TE-classes. This means that the link-state Interior Gateway Protocols (IGPs) must advertise the available bandwidth for each CT at each priority level on every link. However, the IETF decided to limit the number of classes to eight out of the possible 64 values [8]. Thus, it is necessary to select the TE-classes through configuration (note that in the following we will use the terms CT, class and TE-class without distinction, always referring to the TE-class). Once C TE-classes (CTi i=0,…C-1) have been defined and the different traffic trunks have been associated to a given class, it is possible to allocate the LSPs, which are reserved for such classes. The overall bandwidth allocated to the LSPs, belonging to the class i and simultaneously established, is indicated as Ni.

A key point in the LSP allocation is given by the constraints on the bandwidth that can be reserved to each class. The amount of the link bandwidth that a class or a group of classes may reserve is called Bandwidth Constraint (BC), and the relationship between the classes and the BCs is called “Bandwidth Constraint Model”. In the following sub-sections, we describe the two BC models that have been standardized by the IETF.

a) Maximum Allocation Model

MAM is the most intuitive BC model, since it maps one BC to one CT. From a practical point of view, the link bandwidth is simply divided among the different classes. Figure 1 shows how MAM works (for simplicity only three CTs are shown).

![Figure 1 MAM (C=3) Bandwidth allocation](image)

Formally, MAM may be defined as follows:

- for each \( i \in \{0, N - 1\} \)

\[
N_i \leq BC_i \leq M
\]

where M is the maximum reservable bandwidth;

- with the constraint

\[
\sum_{i=0}^{C-1} N_i \leq M
\]

- The sum of all reservable bandwidth can be greater than M:

\[
\sum_{i=0}^{C-1} BC_i \geq M
\]

The benefit of MAM is that it perfectly isolates different traffic classes and guarantees bandwidth to CTs without the need for pre-emption.

The main drawback of this model is that bandwidth may be wasted instead of being used for carrying other CTs, since it is not allowed to share the unused bandwidth.

b) Russian Dolls Model

RDM improves bandwidth efficiency in comparison with MAM by allowing CTs to share bandwidth. In this model, CT7 is the class type with the strictest QoS requirements, whereas CT0 is the class type for best effort traffic.

With reference to figure 2, which shows how RDM works when three CTs are active, the RDM bandwidth constraints may be explained in the following way:

- All LSPs from CT2 use no more than BC2.
- All LSPs from CT2 and CT1 use no more than BC1.
- All LSPs from CT2, CT1, CT0 use no more than BC0.

![Figure 2 RDM (C=3) Bandwidth allocation](image)

RDM is defined as follows:

- for each \( i \in \{0, N - 1\} \)

\[
\sum_{j \in i} N_j \leq BC_i \leq M
\]

where M is the maximum reservable bandwidth;

- with the constraint

\[
\sum_{i=0}^{C-1} N_i \leq M
\]
Since RDM maps one BC to one or more CTs, it is more difficult to manage than MAM. Moreover, RDM can be used in conjunction with pre-emption to achieve isolation among CTs (so that each CT is guaranteed its share of bandwidth no matter the level of contention by other classes), bandwidth efficiency, and protection against QoS degradation of all CTs. The use of RDM is not recommended in networks where pre-emption is precluded.

III. THE SIMULATION MODULE

NS2 is an open-source discrete-event simulator, specifically targeted at networking research, but it does not offer any support for the design and performance analysis of DS-TE networks. To support the DS-TE architecture, it has been necessary to extend RSVP-TE and OSPF-TE protocols, previously developed by the authors of this paper for MPLS-TE networks, so that Label Switched Paths (LSPs) are signalled and allocated according to CTs requirements. As regards OSPF-TE, in addition to the overall available link bandwidth, opaque LSAs also advertise the available bandwidth for each CT. Moreover, the structure of the Traffic Engineering Database (TED) has been modified and constraint-based routing takes this more complex advertised information into account during path computation.

We have also developed a new module which provides the functionalities of MAM, and that can be easily extended, thanks to its modularity and flexibility, so as to support other bandwidth constraints models.

First of all, we have implemented a new link type that can be created by means of the command

```
$ns duplex-dste-link <Node1> <Node2> <Bandwidth> <Latency> <Res-Bw> <RSVP-Bw> <Queue-limit> <Admission control> <Estimator> <BC1> <BC2> <BC3> <BC4> <BC5> <BC6> <BC7>
```

where
- `<Bandwidth>`: total link bandwidth
- `<Latency>`: link propagation delay
- `<Res-Bw>`: reservable bandwidth
- `<RSVP-Bw>`: bandwidth reserved to the RSVP messages
- `<Queue-limit>`: buffer size
- `<Admission-control>`: selected admission control algorithm
- `<Estimator>` estimator of the admission control algorithm
- `<BC1>...<BC7>`: bandwidth constraints corresponding to CT1 ... CT7

Basically, the main functionalities added to the simulator are related to the set-up, release, and rerouting of LSPs, based on the bandwidth constraints imposed by MAM.

The new commands, which allow such events to be scheduled in a simulation, are the following:

- `<Ingress-LSR>` create-crlsp-ospfte-mam
- `<Source>` `<Egress-LSR>` `<SessionID>` `<FlowID>` `<TunnelID>` `<Bandwidth>` `<Buffer>` `<TTL>` `<Class-Type>`
- `<Ingress-LSR>` release-LSP `<SessionID>` `<FlowID>`
- `<Ingress-LSR>` reroute-prealloc-mam `<Source>` `<Egress-LSR>` `<Destination>` `<OldSID>` `<SessionID>` `<FlowID>` `<TunnelID>` `<Rate>` `<Bucket>` `<TTL>` `<Er>` `<Class-Type>`
- `<Ingress-LSR>` reroute-precalc-mam `<Source>` `<Egress-LSR>` `<Destination>` `<OldSID>` `<FlowID>` `<TunnelID>` `<Rate>` `<Bucket>` `<TTL>` `<Er>` `<Class-Type>`

More details about the listed parameters are reported in [9]. It is worth noticing that, as far as LSP rerouting is concerned, both protection switching and on-the-fly fast rerouting have been implemented.

IV. SIMULATION RESULTS

In this section, we show the results of two different sets of simulations carried out to demonstrate:

1. the effectiveness of the new developed module for network design purposes;
2. the main advantages intrinsic to the use of the DS-TE architecture.

![Figure 3 Network Topology (NS2 NAM)](image)

In our simulation scenario, the network topology (figure 3) matches the Géant Network [10], a multi-gigabit pan-European data communications infrastructure, specifically deployed for research and education services.

All the network nodes, except for the traffic sources and destinations, are considered as part of a unique MPLS domain. Table I lists the bandwidth $B$ associated to each link.
For sake of simplicity, we have considered only three bandwidth constraints (BC0, BC1, BC2) and the percentage of resources that each CT can allocate on each link has been set as follows:

- **BC1**: 30%
- **BC2**: 45%
- **BC0**: 100%

The total amount of reservable link bandwidth has been set in the following way:

- 25% for the 10 Gbps links
- 50% for the 5 Gbps links
- 99% for the 2.5 Gbps links
- 99% for the 155 Mbps links

Three CBR (Constant Bit Rate) traffic sources have been directly connected to node 9 (IT), while three traffic destinations have been connected to node 3 (NO). Therefore, three distinct CTs have been taken into account:

- **CT0** corresponding to Best Effort traffic
- **CT1** corresponding to data traffic
- **CT2** corresponding to voice traffic

**a) First Simulation Scenario**

In the first set of simulations, we have considered the following events:

1. **T = 8.0s.** A first LSP, associated to data traffic (CT1), with a reserved bandwidth of 730 Mbps, is set-up between nodes 28 and 29. Since over 10 Gbps links we can allocate to CT1 a maximum of 750 Mbps (10 Gbps * 0.25 * 0.30), there is enough available bandwidth for the first LSP along the path 9_19_2_3, which is also the shortest path.

2. **T = 13.0s.** A second LSP, associated to data traffic (CT1), with a reserved bandwidth of 720 Mbps, is set-up. Since for CT1, now there is not enough bandwidth along the shortest path, the LSP is established along the path 9_8_7_22_2_25_3.

3. **T = 22.0s.** A third LSP is set-up for voice traffic (CT2), with a reserved bandwidth of 1.1 Gbps. Since over 10 Gbps links we can allocate to such CT a maximum of 1.125 Gbps (10 Gbps * 0.25 * 0.45), while over 2.5Gbps links we can allocate a maximum of 1.113 Gbps (2.5 Gbps * 0.99 * 0.45), there is enough available bandwidth on the shortest path 9_19_2_3.

If we consider the same events in a not DS-TE network, the first two LSPs (associated to data traffic) are both allocated on the shortest path 9_19_2_3, while the third LSP, associated to voice traffic, is established along the path 9_8_7_22_2_25_3.

Thus, as shown in figure 4, the end-to-end delay experienced by the voice traffic, strongly differs from the DS-TE case to the not DS-TE one. Indeed, before t = 22s, corresponding to the set-up of the third LSP, all the traffic is forwarded along the shortest path computed by the layer 3 routing protocol, while after the LSP has been set-up, the traffic is bound to the corresponding LSP and thus it experiences a different end-to-end delay in two cases: in the DS-TE, since the LSP is created along the shortest path the delay remains almost constant (31.2 ms), while in the not DS-TE architecture the delays arises to about 125ms.
b) Second Simulation Scenario

In the second simulation set, we have considered the following traffic sources:
- Node 0: CBR traffic source with a bit rate of 2.4 Gbps, towards node 16. The traffic flow is classified as CT_0 (best effort).
- Node 28: CBR traffic source with a bit rate of 730 Mbps, towards node 29. The traffic flow is classified as CT_1.
- Node 26: CBR traffic source with a bit rate of 1.1 Gbps, towards node 6. The traffic flow is classified as CT_2.

Moreover, we have scheduled at t = 16.0 s the set-up of the LSP corresponding to CT_1, and at t = 25.0 s the set-up of the LSP corresponding to CT_2.

As shown in figures 5-6, before t = 16.0 s, corresponding to the set-up of the LSP bound to CT_1, all the three traffic flows experience several packet losses, since the overall traffic rate exceeds the link capacity. Losses for traffic CT_1 and CT_2 stop after the establishment of the corresponding LSPs, while those related to the best effort traffic get worse after t = 25.0 s, since only the residual bandwidth can be used by such traffic.

As far as the throughput is concerned, note that the throughput of traffic CT_1 increases after the first LSP has been created, so as the throughput of traffic CT_2 increases after t = 25.0 s, while the throughput of best effort traffic get worse and worse after t = 16.0 s and t = 25.0 s.

This simulation set shows that, coherently with the theory, MAM allows the best effort traffic to only use the residual link bandwidth, when other CTs reserve bandwidth.

V. CONCLUSION

This paper presents a new software module for the simulation of DS-TE networks in the NS2 simulation environment. Our efforts have focused on the implementation of functions and protocols typical of the DS-TE architecture and, more specifically, of the MAM BC model. Thanks to its modularity and flexibility features, the implemented module can be easily extended so as to support further BC models, such as the RDM. Moreover, we have shown the results of different sets of simulations carried out to simultaneously demonstrate the operational behaviour of the implemented module and its effectiveness for network design and planning.

ACKNOWLEDGMENTS

This paper presents an implementation of the MAM module and its effectiveness for network design and planning. Our efforts have focused on the implementation of functions and protocols typical of the DS-TE architecture and, more specifically, of the MAM BC model. Thanks to its modularity and flexibility features, the implemented module can be easily extended so as to support further BC models, such as the RDM. Moreover, we have shown the results of different sets of simulations carried out to simultaneously demonstrate the operational behaviour of the implemented module and its effectiveness for network design and planning.

REFERENCES