Performance Analysis of MPLS TE Queues for QoS Routing

Yihan Li, Shivendra Panwar
Electrical and Computer Engineering Department, Polytechnic University, Brooklyn, NY

C.J. (Charlie) Liu
AT&T Laboratories, Middletown, NJ

***Abstract*— In order to maximize network resources, Multi-Protocol Label Switching (MPLS) Traffic Engineering (TE) is used in IP networks so that traffic can be routed on a path which may not be chosen by a standard routing method. In this paper, the performance of a proposal creating TE queues for configured MPLS TE tunnels in every router the tunnel traverses is presented. The TE queue creation is a new concept to effectively couple the control plane and the data plane for MPLS TE tunnels. The idea takes advantage of an intelligent Constrained Shortest Path First (CSPF) routing mechanism for MPLS TE to enable QoS routing. It will help service providers minimize the task and cost of implementing a complex bandwidth broker Operation Support System to associate allowed tunnel bandwidth and available queues in the network at the time of provisioning. The mechanism will also ensure that an IP network delivers the stringent QoS required to carry real time traffic such as VoIP. The performance is simulated and analyzed with a Generalized Processor Sharing (GPS) system. According to our results, using TE queues leads to lower overflow probabilities for TE tunnel traffic, and better QoS for real time traffic such as VoIP.**

***Index Terms*— Multi-Protocol Label Switching, traffic engineering, Generalized Processor Sharing.**

I. INTRODUCTION

Traffic engineering (TE) refers to techniques and processes to route traffic through a network on a path other than that would have been chosen if standard routing methods had been used. The goal of traffic engineering to a service provider is to maximize the utilization of network resources, and/or enhance the QoS a service provider can offer. To justify the increase in network operational complexity associated with traffic engineering, TE must enable new service offerings, reduce the overall cost of operations, maximize potential revenues and increase customer satisfaction. In a large network, it is possible that available network bandwidth is not efficiently utilized because the intra-domain routing protocol, such as OSPF, finds path based on a single "least-cost" scalar metric for each destination. This least cost route may not have enough resources to carry all the traffic, or satisfy all the SLA (service level agreement) requirements of carried traffic. Congestion, at certain hot spots, can result in sub-optimal use of network resource.

In a large network, it is possible that available network bandwidth is not efficiently utilized because the intra-domain routing protocol, such as OSPF, finds path based on a single "least-cost" scalar metric for each destination. This least cost route may not have enough resources to carry all the traffic, or satisfy all the SLA (service level agreement) requirements of carried traffic. Congestion, at certain hot spots, can result in sub-optimal use of network resources.

Multi-Protocol Label Switching (MPLS) TE provides a more elegant and efficient technique than IP source routing. It allows traffic travel down a path different from conventional Interior Gateway Protocol (IGP) destination based hop-by-hop routing. The path is pre-determined at tunnel setup time. Routers along the path do not have to examine the IP header of every passing packet. The basic idea of MPLS involves assigning short fixed length labels to packets inside an MPLS cloud. Throughout the MPLS domains, the labels attached to packets are used to make forwarding decisions. It allows decoupling of the information used for forwarding (a label) from the information carried in the IP header. MPLS TE, using the RSVP signaling mechanism [1], injects the notion of a connection to connectionless IP through nailed-up label switched paths (LSP). MPLS TE provides capabilities to specify an explicit path for the LSP before it is established. We will refer the nailed-up LSP as an MPLS TE tunnel, or simply tunnel, in this paper.

The tunnel explicit routing capability allows routing flexibility. It allows paths, with unequal OSPF cost, to share traffic load [2]. In addition, the Fast Reroute feature [3] in MPLS TE allows path restoration within 100 ms in case of link or node failure. In this paper we propose a MPLS TE tunnel mechanism for packet forwarding, which can guarantee the service of real time applications such as VoIP and video conferencing.

II. USING MPLS TE TUNNEL IN ROUTERS WHEN FORWARDING PACKETS

MPLS TE tunnel is a connection-oriented entity on top of the conventional connectionless IP network. It has been promoted by router vendors and IETF activists as a valuable tool to maximize utilization of network resources as described in the previous section. However, the MPLS TE admission control mechanism is applied only at the tunnel setup time, not at the packet forwarding time. Bandwidth reservation is policed only at tunnel setup time to limit the number of tunnels traversing a given link. Traffic inside a tunnel has to compete for bandwidth with traffic in other tunnels and regular IP traffic which is not carried by any TE tunnel.

Even though there are benefits of deploying TE tunnels in IP networks, there are concerns about its scalability and extra complexity in network operation. For a facility based ISP which owns the physical links and infrastructure of its IP network, the capacity constraint is a relatively minor issue compared to other ISPs which have to purchase or lease capacity from other providers. It is hard to justify sending all IP traffic into fully meshed TE tunnels ubiquitously deployed for a facility based ISP. Instead, only special traffic, such as VoIP or videoconference traffic transported in an IP network, are candidates for MPLS TE tunnels. This is different from Diffserv-TE [4], [5], [6], where all IP traffic is subjected to the same admission control mechanisms and have to be carried in TE tunnels to ensure the traffic guarantee.

In this paper we investigate the performance enhancement of TE queues creation for configured MPLS TE tunnels in every router the tunnel traverses [7]. Traffic in TE tunnels are preferentially treated by a router's queuing and congestion avoidance mechanism. A TE queue can be used by a single tunnel or shared by multiple TE tunnels. The TE queue is to be created at tunnel set up time based on the MPLS label and bandwidth requests associated with the tunnel. The bandwidth reserved for each queue is to be set according to the bandwidth of configured tunnels sharing the same queue. The TE admission control mechanism ensures that the sum of the TE queue bandwidth will not exceed the configured RSVP bandwidth of the physical link. The reserved bandwidth can only be used by the traffic carried by the tunnels. At packet forwarding time, the top label in the label stack of each packet carried inside the tunnel will be used as the key for the packet to be sent into the TE queue associated with the tunnel.

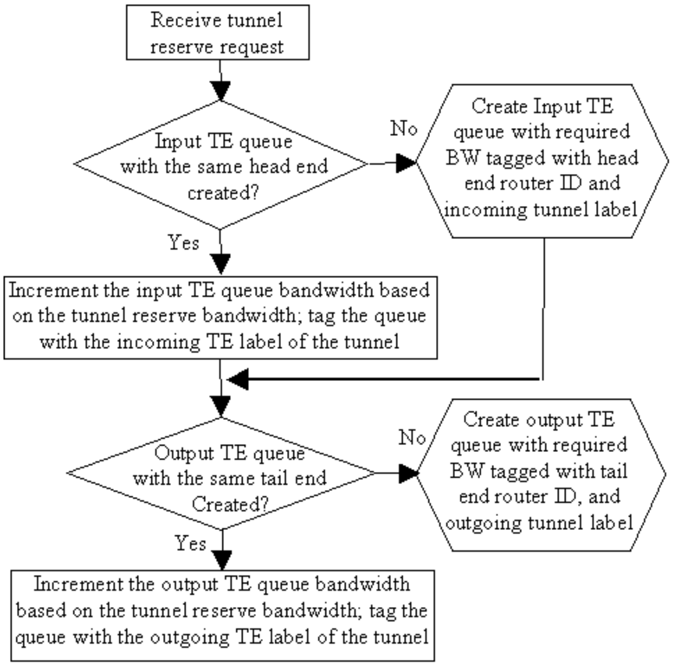

Fig. 1. Create Queues for MPLS TE Tunnels.

A. *TE Queue Creation Process for MPLS TE Tunnels*

The proposed MPLS TE Queue creation mechanism at a router is illustrated in the flow chart as shown in Figure 1. We assume both input queues and output queues are implemented in the router. The process below also assumes tunnels with the same head end will share the same input queue, while tunnels with the same tail end will share the same output queue.

During the TE tunnel setup period, the router will query its database to determine if a TE input queue with the same head end had been created. If no such queue had been assigned, the router will create a TE queue, tagged with the head end router ID and the assigned tunnel label, and with the requested bandwidth. If a TE queue with the same head end router ID has already been created, the bandwidth of the TE queue will be adjusted based on the new tunnel request. The bandwidth adjustment does not have to be the exact increment of the requested bandwidth of the new tunnel. A statistical multiplexing model of tunnel traffic can be incorporated here. The TE queue will also be tagged with one more label, which is the incoming label of the TE tunnel. The output TE queue will be created and set up in a similar fashion.

B. *Switching Process for Packets in MPLS TE Tunnels*

When a packet is received, a router will determine whether the packet is label switched and whether the label is assigned to a TE tunnel. The packet forwarded via a tunnel will be sent into the appropriate input TE queue based on the incoming label. The router will consult its label-forwarding database to determine its outgoing label. The packet will then be label switched to the appropriate output interface based on the router's scheduling mechanism for input TE queues, and be put into an appropriate output TE queue based on its outgoing label. Then the packet will be forwarded to the next hop based on the router's scheduling mechanism for output TE queues. Because tunnels are envisioned for high priority and demanding traffic only, it is recommended TE queues assume scheduling priority over all other non-TE queues.

III. ANALYSIS AND SIMULATED PERFORMANCE

A. The system model

The delay a packet suffered from the time it enters the input interface to the time it is transferred to the destined output interface is determined by the scheduling policy of the switch fabric with input TE queues. Here we only consider the process from the time packets enter the output TE queue to the time they are forwarded to the next hop. We assume that each traffic source can be modeled as a continuous-time Markov process and analyze and simulate the system as a Generalized Processor Sharing (GPS) system [8].

Assume that each output maintains N (output) TE queues and K non-TE queues, as shown in Figure 2(a). All TE queues have the same priority, which is higher than the priorities of non-TE queues. c_n, $1 \le n \le N$, is the guaranteed service rate for TE queue n, and

$$c = \sum_{n=1,N} c_n. \tag{1}$$

When all TE queues are empty, the residual service is distributed to non-TE queues. Each queue n, with instant rate $r_n(t)$, is modeled as a Markov Modulated Fluid Process (MMFP) with state space S_n, rate matrix Λ_n, and infinitesimal generator M_n. The buffer is infinite, and $X_n(t)$ is the occupancy of queue n. For each TE queue, we need to find out the overflow probability with threshold B.

The exact analysis of the system in Figure 2(a) is difficult. To simplify, queue n can be analyzed by a model shown in Figure 2(b) [8].

$$r_n^*(t) = r_n(t), \tag{2}$$

$$r_j^* = \sigma_j^{'} r_j^{'}(t), \tag{3}$$

and

$$c_n^* = c_n + \sum_{j \neq n} \sigma_j c_j, \tag{4}$$

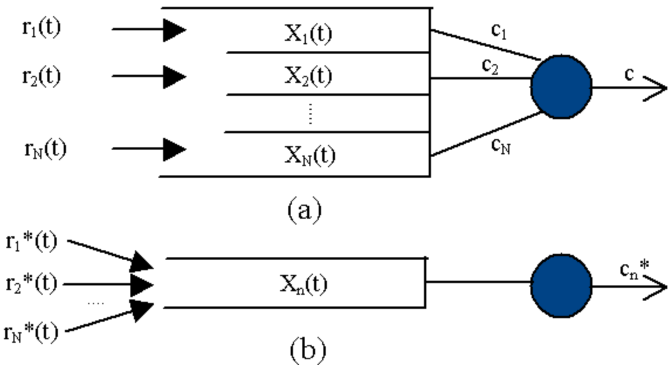

Fig. 2. The analysis model

where for all $j \neq n$,

$$\sigma_j = c_n / \sum_{k \neq j} c_k, \tag{5}$$

and $r_j^{'}(t)$ is the departure process of each queue j obtained while assuming the service rate is c_j. This GPS problem can be resolved using a fluid-flow model as follows.

Consider a system as shown in Figure 2(b) but with a service rate c and N independent general MMFP sources. Each source is characterized by $(M^{(i)}, \Lambda^{(i)})$. Then M and Λ of the aggregate source are

$$M = M^{(1)} \oplus M^{(2)} \oplus \cdots \oplus M^{(L)}, \tag{6}$$

and

$$\Lambda = \Lambda^{(1)} \oplus \Lambda^{(2)} \oplus \cdots \oplus \Lambda^{(L)}. \tag{7}$$

The Kronecker sum ($\oplus$) of matrices $A = [a_{ij}]_{mn}$ and B is defined as follows.

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \tag{8}$$

$$A \oplus B = A \otimes I_m + I_n \otimes B. \tag{9}$$

where I_m and I_n are the identity matrices of order m and n, respectively. Define state probabilities

$$f_s(x,t) = Prob(S(t) = s, X(t) \le x, t). \tag{10}$$

Then

$$(\Lambda - cI)\frac{df(x)}{dx} = Mf(x). \tag{11}$$

TABLE I
SOURCE PARAMETERS

	α	β	p
Source 1	0.4	1.0	1.2
Source 2	0.4	1.0	1.0
Source 3	1.0	1.0	1.2

where I is an identity matrix. $(\Lambda - cI) = D$ is called a drift matrix. If the number of aggregate source states is L, then the result is of the following form

$$f(x) = \sum_{i=1,L} a_i \Phi_i e^{z_i x} \tag{12}$$

where z_i and ϕ_i are eigenvalue-eigenvector pairs for the matrix $D^{-1}M$, and a_i are coefficients. If an eigenvalue is positive, the corresponding a_i is zero. The number of overload states is the same as the number of negative eigenvalues. $f(x)$ is obtained from the following boundary condition. $f_j(0) = 0$ when $j \in S_0$, where S_0 is the set of overload states.

Therefore, the overflow probability, the probability that the buffer occupancy exceeds the threshold B, is as follows.

$$P_{of} = \sum_{j \in S} Prob(S = j, X > B). \tag{13}$$

This analytical method, with its low complexity, can be used for provisioning in real time.

B. Analysis and simulation results

Here we assume that there are three on-off sources, two for TE traffic and one for non-TE traffic. Three cases are considered:

(1) all traffic classes share one queue,

(2) all TE traffic shares one TE queue and non-TE traffic goes to the non-TE queue, and

(3) each TE source traffic goes to its own TE queue and non-TE traffic goes to the non-TE queue.

The source parameters are given in Table I, where α and β are the transition rates from off to on, and on to off, respectively; p is the input rate when the source is on. The guaranteed service rate for source 1 and source 2 are 0.7 and 0.4, respectively. The analysis and simulation results of the overflow probability of TE traffic are shown in Figure 3 and 4. As we can see, with the selected system parameters, using TE queue leads to lower overflow probabilities for TE tunnel traffic, and using multiple TE queues can further differentiate the service of TE tunnel traffic.

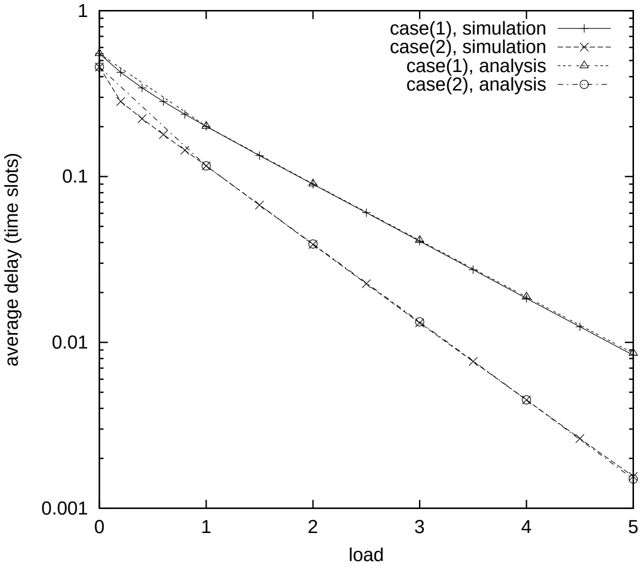

Fig. 3. Tail distributions for case 1 and 2.

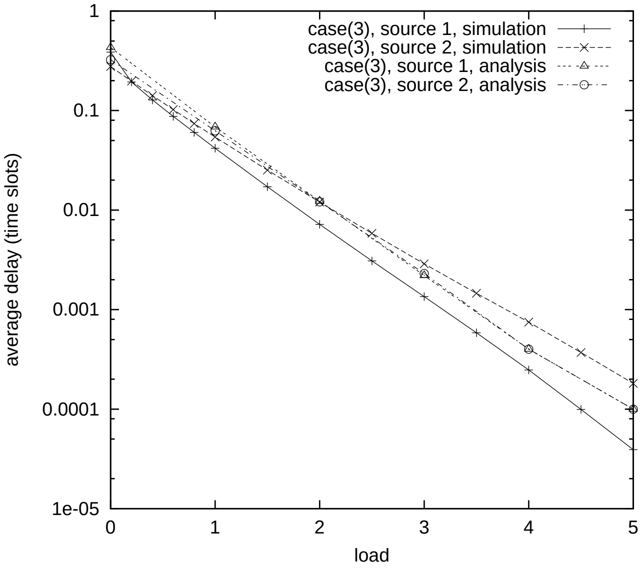

Fig. 4. Tail distributions for case 3.

IV. CONCLUSION

The TE queue creation as described in this paper is a new concept to effectively couple the control plane and the data plane for MPLS TE tunnels. The idea takes advantage of the intelligent CSPF routing mechanism [9] for MPLS TE to enable QoS routing. It will save service providers the task and cost of implementing a complex bandwidth broker Operation Support System to associate allowed tunnel bandwidth and available queues in the network at the time of provisioning. The mechanism will also ensure an IP network to deliver the stringent QoS required to carry real time traffic such as VoIP

[10]. The performance is simulated and analyzed with a Generalized Processor Sharing (GPS) system. According to our results, using TE queues leads to lower overflow probabilities for TE tunnel traffic, and better QoS for real time traffic such as VoIP. In our future work, more complicated system models and traffic models will be considered by analysis and simulation.

REFERENCES

[1] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, “RFC3209, RSVP-TE: Extensions to RSVP for LSP Tunnels.”

[2] D. Awduche, J. Malcolm, J. Agogbua, M. O’Dell, and J. McManus, “RFC 2702, Requirements for Traffic Engineering Over MPLS.”

[3] Ping Pan, Der-Hwa Gan, George Swallow, Jean Phillippe Vasseur, Dave Cooper, Alia Atlas, Markus Jork, “Fast Reroute Extensions to RSVP-TE for LSP Tunnels,” draft-ietf-mpls-rsvp-lsp-fastreroute-00.txt.

[4] F. Le Faucheur (Editor), L. Wu, B. Davie, S. Davari, P. Vaananen, R. Krishnan, P. Cheval, and J. Heinanen, “RFC 3270, Multi-Protocol Label Switching (MPLS) Support of Differentiated Services.”

[5] Francois Le Faucheur and Waisum Lai, “Requirements for support of Diff-Serv-aware MPLS Traffic Engineering,” draft-ietf-tewg-diff-te-reqts-06.txt.

[6] Francois Le Faucheur, Thomas D. Nadeau, Angela Chiu, William Townsend, Darek Skalecki, “Extensions to RSVP-TE and CR-LDP for support of DiffServ-aware MPLS Traffic Engineering,” draft-ietf-mpls-diff-te-ext-01.txt.

[7] C. J. Liu, “MPLS TE queue creation as a mechanism for QoS routing,” *Proceddings of Applied Telecommunication Symposium/ASTC*, pp. 23-27, Orlando, Florida, March 2003.

[8] S. Mao and S. Panwar, “The effictive bandwidth of Markov modulated fluid process sources with a Generalized Processor Sharing server,” *IEEE Globecom 2001*, vol. 4, pp. 2341-2346, November 2001.

[9] David Katz, Kireeti Kompella, and Derek Yeung, “RFC 3630, Traffic Engineering Extensions to OSPF Version 2.”

[10] G. Ash, “Traffic Engineering and QoS Methods for IP-, ATM-, and TDM-Based Multiservice Networks,” draft-ietf-tewg-qos-routing-04.txt.