



Repositorio Institucional de la Universidad Autónoma de Madrid

<https://repositorio.uam.es>

Esta es la **versión de autor** de la comunicación de congreso publicada en:
This is an **author produced version** of a paper published in:

Photonic Network Communications 15.2 (2008): 159-169

DOI: <http://dx.doi.org/10.1007/s11107-007-0107-9>

Copyright: © Springer-Verlag Berlin Heidelberg 2008

El acceso a la versión del editor puede requerir la suscripción del recurso
Access to the published version may require subscription

On capacity planning for the GMPLS network control plane

Daniel Morató*, Javier Aracil†, Juan P. Fernández-Palacios‡, Oscar González de Dios‡, Jesus F. Lobo Poyo§

* Universidad Pública de Navarra, Campus de Arrosadía s/n, E-31006 Pamplona, Spain e-mail:
daniel.morato@unavarra.es

† Universidad Autónoma de Madrid, Campus de Cantoblanco, E-28049 Madrid, Spain e-mail:
javier.aracil@uam.es

‡ Telefónica I+D, Emilio Vargas 6, 28043 Madrid (Spain) e-mail: {jpfpg, ogondio}@tid.es

§ Telefónica S.A., Gran Vía 28, 28013 Madrid (Spain) e-mail: jesus.loba@telefonica.es

Abstract

This paper presents capacity planning rules for the control plane of all-optical networks featuring GMPLS and RSVP-TE as a connection setup protocol. As per RSVP standard, a refresh message mechanism is incorporated to RSVP such that the state is periodically refreshed on a link per link basis. We provide analytical expressions for the bandwidth and buffer sizes to be provided such that no flows are torn down due to lack of refresh messages. Our findings show that small buffers (several KBytes) suffice to sustain the signaling load for as much as 400 RSVP flows per link, with the simplest RSVP refresh mechanism (neither using link bundling nor acknowledgments). On the other hand, we also find the packet drop probability per link for a given network topology for the case that the flow survival probability is larger than a given threshold. We provide numerical examples based on the COST 239 european network topology and real RSVP traffic traces from early-commercial switching equipment.

1. INTRODUCTION

The advent of dynamical Optical Circuit Switching (OCS) networks calls for a signaling protocol for lightpath establishment and tear down. Within the framework of Generalized Multiprotocol Label Switching (GMPLS [1]), the RSVP-TE protocol [2] and [3] can be used to this end, with its extensions for GMPLS [4] and [5]. With RSVP-TE, lightpaths are established in the same way as LSPs (Label Switched Paths) are established in MPLS networks.

According to the RSVP protocol, the state must be continuously updated in the routers along the path from source to destination, namely RSVP is a soft-state protocol. Once the lightpath has been established, periodic refresh messages are sent on a hop-by-hop basis to signal that the LSP is alive. The lack of those messages is interpreted as a path failure and the lightpath is torn down. The goal of this paper is to analyze the case of *erroneous lightpath disconnections*, which may happen due to consecutive refresh message losses, and not path failure. On the other hand, we provide capacity planning rules for the optical network control plane. First, we calculate the flow duration distribution as a function of the control plane loss probability. Then, we provide expressions that serve to obtain the queue size and bandwidth per link, such that the flow duration is larger than a given value with a certain probability objective (e.g., 99%).

A. State of the art

The RSVP protocol has been extensively analyzed. A markovian model is proposed in [6] that serves to analyze inconsistencies in the RSVP endpoints. The paper compares a number of approaches that depart from the pure soft-state paradigm, considering different packet loss, delay and refresh timer values. A simulation analysis has been performed in [7] in which the effect of different control message delivery mechanisms is evaluated. The authors show that RSVP performs poorly using best-effort delivery for its control messages. Actually, the best results are obtained when a separate bandwidth is reserved for the RSVP messages. Furthermore, the timing parameters of the protocol have also a strong influence on RSVP performance. The optimisation of such timing parameters is subject of analysis in [8], where a multi-objective evolutionary optimisation approach is adopted. The optimized parameters include connection setup time, protocol overhead and end-to-end delay, and losses experienced by the data packets. Additionally, the performance of hardware and software implementations of RSVP are analyzed in [9]–[11], with emphasis on the router capacity (in terms of CPU and memory) that is required to sustain a certain number of RSVP flows.

B. Problem statement

The case of optical networks differs from previous studies on RSVP performance in the following issues. Out-of-band capacity is provided for the RSVP signaling flows, possibly through separate wavelengths. The delay performance is not an issue since dedicated bandwidth is provided and delay is expected to be bounded. Most importantly, the lightpath duration is very large in comparison to the typical duration of other RSVP flows. In fact, a 40 Gbps wavelength may be allocated for several days, whereas this is very unlikely to happen for other Internet RSVP flows.

Precisely, the objective of this paper is to analyze what are the bandwidth and buffer requirements such that a lightpath is not erroneously torn down due to loss of RSVP refresh messages. To the best of our knowledge, previous works on RSVP have never studied the probability of erroneous disconnection of LSPs. Furthermore, we provide expressions for the queue size and bandwidth that ensure that a lightpath will remain switched on for a given period of time with a certain probability. The results are extended for any given network topology and a numerical example is provided with the COST 239 european network topology. Consequently, this paper provides considerable insight concerning the applicability of simple RSVP signaling for lightpath provisioning in optical networks. It also provides capacity planning rules for the control plane.

The paper is structured as follows. In Section 2 we present the modeling assumptions. In Section 3 the analysis for a one-link scenario is presented, together with the results. Section 4 is devoted to the analysis and results for the multiple hop scenario. Finally, Section 5 presents the discussion, followed by conclusions and future work.

2. MODELING ASSUMPTIONS

According to RFC [3] on RSVP-TE and the corresponding extensions for GMPLS [4] and [5], the refresh messages are called *Path* and *Resv* depending on whether they are transmitted in the forward or reverse direction of the flow establishment. They contain several mandatory RSVP objects, but they can also contain optional objects for different purposes (confirmation request, message identification, protection information, etc.). For example, the objects <SESSION>, <RSVP_HOP>, <TIME_VALUES>, <LABEL_REQUEST> and <sender descriptor> are mandatory for RSVP-TE Path messages. On the other hand, the objects <SESSION>, <RSVP_HOP>, <TIME_VALUES>, <STYLE> and <flow descriptor list> are mandatory for Resv messages. Assuming unicast data flows, the total minimum IP packet size is around 120 bytes for Path and Resv messages. Table 1 shows the mandatory RSVP objects in a Path and Resv message for an IPv4 LSP tunnel.

Path message		Resv message	
Object	size (bytes)	Object	size (bytes)
IP header	20	IP header	20
Common header	8	Common header	8
SESSION	16	SESSION	16
RSVP HOP	24	RSVP HOP	24
TIME_VALUES	8	TIME_VALUES	8
LABEL_REQUEST	8	STYLE	8
SENDER_TEMPLATE	12	FLOWSPEC	<varies>
SENDER_TSPEC	<varies>	FILTER_SPEC	12
UPSTREAM_LABEL	8	LABEL	8

TABLE 1

MANDATORY OBJECTS AND TYPICAL SIZE FOR PATH AND RESV MESSAGES FOR GMPLS LSP IPV4 TUNNELS

However, typical configurations also include optional objects in the RSVP-TE messages. In order to validate the packet size we track the RSVP packet flows between a pair of early-commercial core routers¹. We analyze two different traces with 126 and 378 LSPs, respectively. They were captured for a time period of approximately 60 minutes, late in the evening, when no new configuration events took place in the network.

According to our trace, the Path and Resv messages were 332 Bytes long, which shows that the optional objects are intensively used in commercial RSVP implementations. In what follows, we will assume the packet sizes observed in the trace, without loss of generality because the packet size is a parameter in our analytical expressions.

A. Simulation architecture

In order to assess the validity of our analytical results, an event-driven simulator has been implemented that mimics the behavior of RSVP signaling agents. The simulation setup consists of a one-hop bidirectional signaling channel between two RSVP nodes. The channel is modeled as two fixed-rate servers working at the link line rate (one for each direction). A buffer can be specified for each link direction. The buffer size is measured in number of packets, and, since packets are constant-size, it is straightforward to obtain the buffer size in bytes. We run our simulations using the default recommended values for the timer parameters and a uniform distribution for the random refresh period. The number of flows traversing the link is the main parameter in the simulations, as it clearly relates to the control plane load. We configure the simulations for the range of hundreds of LSPs. It is worth noting that the case of multicast LSPs is not considered in this paper and that the LSPs are assumed to be bidirectional.

On the other hand, a bundling mechanism for the refresh messages is defined in [12], with the purpose of reducing signaling load. To do so, a single refresh message serves as a keepalive message for several LSPs. However, it has been shown in [9] that a commercial PC is able to sustain the RSVP signaling load for as much as 50,000 simultaneous flows, with no bundling at all. Thus, we purposely consider the case of no link bundling, as it is simpler to implement. On the other hand, we also consider that the refresh messages are not acknowledged. As a result, we analyze the *simplest* case of RSVP keepalive signaling. Our findings show that a very small buffer and bandwidth per flow becomes necessary using the simplest RSVP refresh mechanism.

¹The router brand and model cannot be disclosed for confidentiality reasons

3. SINGLE-HOP SCENARIO

We consider a single link between two adjacent routers, with RSVP signaling and unicast flows. As mentioned before, once the reservation is in place, the state installed in the routers is a *soft-state*. The soft-state is erased after the expiration of a *cleanup timeout* from the arrival of the last Path or Resv message for that flow. In Fig. 1 the bold arrows represent the direction of the LSP establishment, from “initiator” to “terminator”. Fig. 1 (a) matches the case of an LSP establishment in the direction from node A to node B, while Fig. 1 (b) the opposite one. As shown, the Path messages will travel in the flow establishment direction and the Resv messages in the opposite one. Since flows can be established in both directions, Path and Resv messages will travel in both directions.

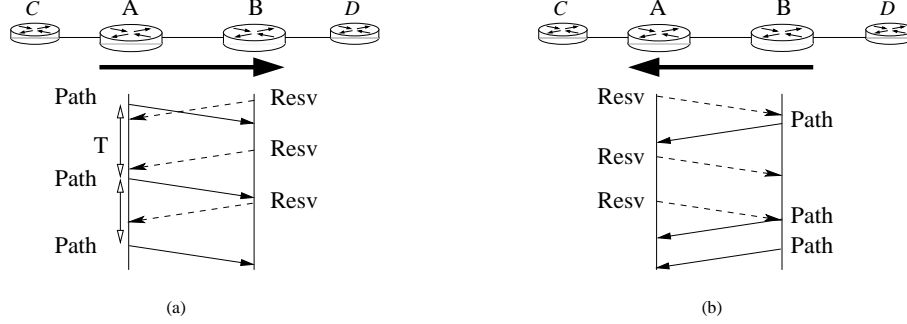


Fig. 1. RSVP message flows

Such RSVP messages are sent periodically by each router to its adjacent nodes in the flow’s path, in order to *refresh* the soft state. Hence, they are called *refresh messages*. Most importantly, the transmission period is smaller than the cleanup timeout. In fact, if R is the refresh period, the timeout L must satisfy $L \geq (K + 0.5) \cdot 1.5R$. The values suggested in [13] are $R = 30$ secs and $K = 3$. The refresh period R takes a random value in the range $[0.5R, 1.5R]$, in order to avoid message synchronization. Each refresh message contains the $\langle \text{TIME_VALUES} \rangle$ object [13] with the R value that was used to generate it. The value of L used for a given flow in a router is based on the value of R stated in such object.

Routers A and B in Fig. 1 will also generate Path and Resv messages to their respective other adjacent routers in the path (routers C and D in the figure). However, neither those messages use the link under analysis nor they trigger the transmission of any other message in such link [13]. Thus, links are completely independent and can be analyzed separately. Concerning the router architecture, it will be assumed that there is a control wavelength for RSVP signaling traffic and a single E/O converter at the router output port for this wavelength. Thus, there is a single queue in which the RSVP signaling traffic is statistically multiplexed for transmission. This is the most cost-effective router architecture, due to the cost of O/E converters.

First, an arrival process for the multiplexed traffic will be proposed, and validated with real traces. Secondly, the output queue size will be obtained, for a given loss probability objective. Based on this loss probability, the distribution of the time elapsed until an erroneous disconnection happens will be derived.

A. Arrival process

Fig. 2 shows the complementary cumulative distribution function (ccdf) or survival probability function of the refresh message inter-arrival times obtained by simulation, compared to a Poisson process with the same rate and number of samples (log-natural axes). The number of active flows is 100. We observe that the multiplexing of 100 flows worth of RSVP refresh

messages produces inter-arrival times that are very close to exponential. This behavior has also been observed in the real traffic trace (Fig. 3).

This finding suggests that a $M/D/1/K$ system may be a good model for the queueing system (packet service time is constant due to the constant packet size assumed). However, we also need to check the independence of the inter-arrival times. For this purpose, Fig. 4 presents the autocorrelation from the simulated packet arrival process (more than 3 million samples) compared to the 95% confidence interval around zero for a Poisson process. There is some substantial correlation in the arrival process for lags lesser than 200. Thus, the hypothesis of independent packet arrivals cannot be accepted.

We find a similar behaviour for a scenario with 200 and 400 flows (Fig. 4). However, we note that the absolute value of the autocorrelation drops as we increase the number of flows. Indeed, the larger the number of independent flows being multiplexed the lower the autocorrelation and the more independent the increments of the resulting traffic process. Therefore, we expect a better fit to the $M/D/1/K$ model as the number of signaling flows increases.

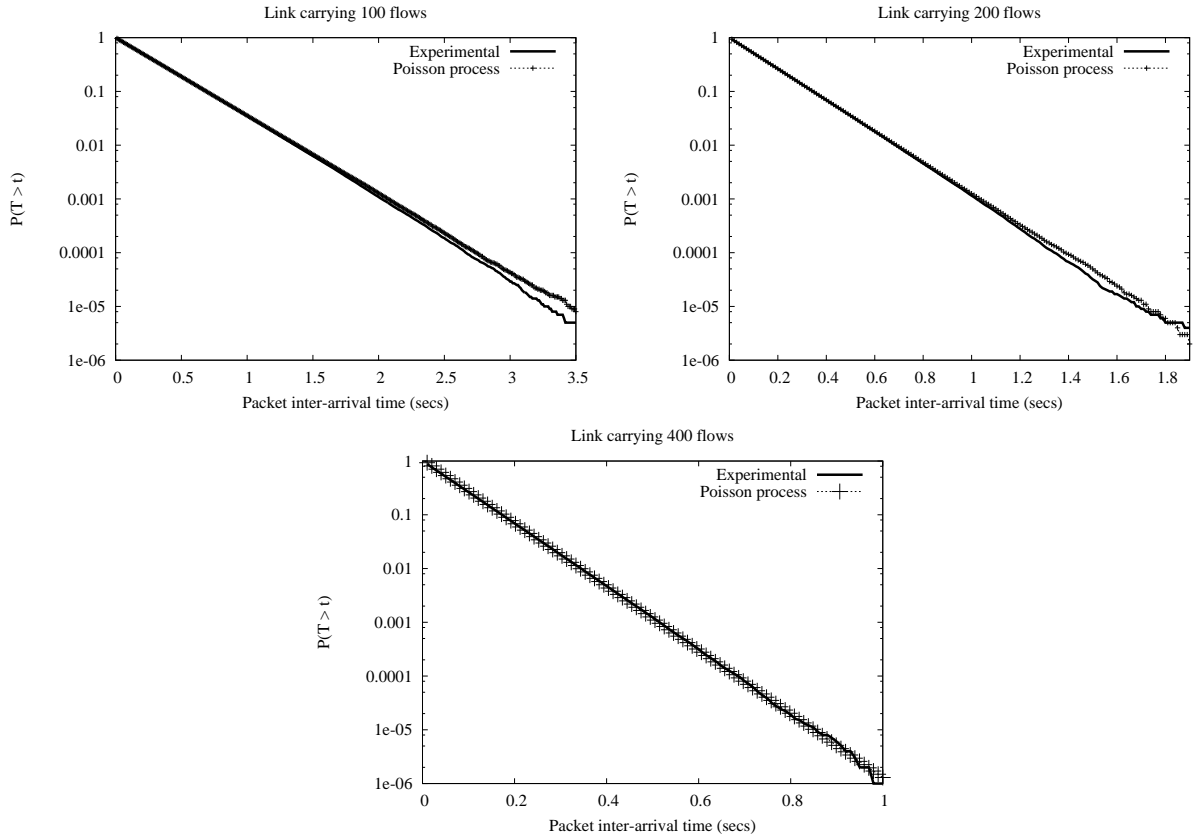


Fig. 2. Survival function for the refresh message interarrival time (100, 200 and 400 flows)

B. Queueing model for the RSVP router

As follows from the previous results, an $M/D/1/K$ model is an approximation to the real system, the better the larger the number of flows in the link. Assuming that the total number of flows is $NFlows$, the link bandwidth is BW , the queue size is Q_{size} packets and $\rho = \lambda/\mu = \frac{NFlows \cdot PacketSize}{R \cdot BW}$ is the utilization factor, the queue length distribution follows from the $M/G/1/K$ system analysis ($K = Q_{size} + 1$) [14]. We first obtain the following recursive formulas for the queue length distribution $\pi_i^{(\infty)}$ of an infinite buffer system

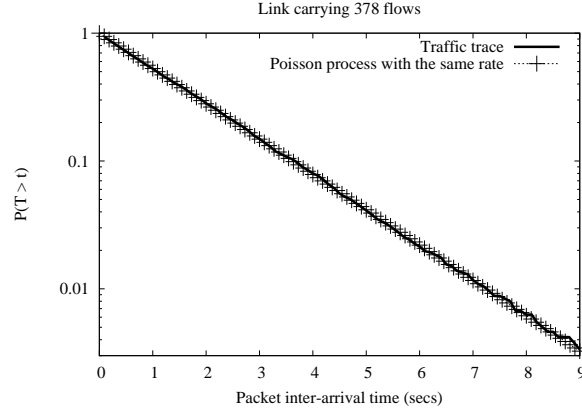


Fig. 3. Survival function of the refresh PATH message interarrival times (traffic trace)

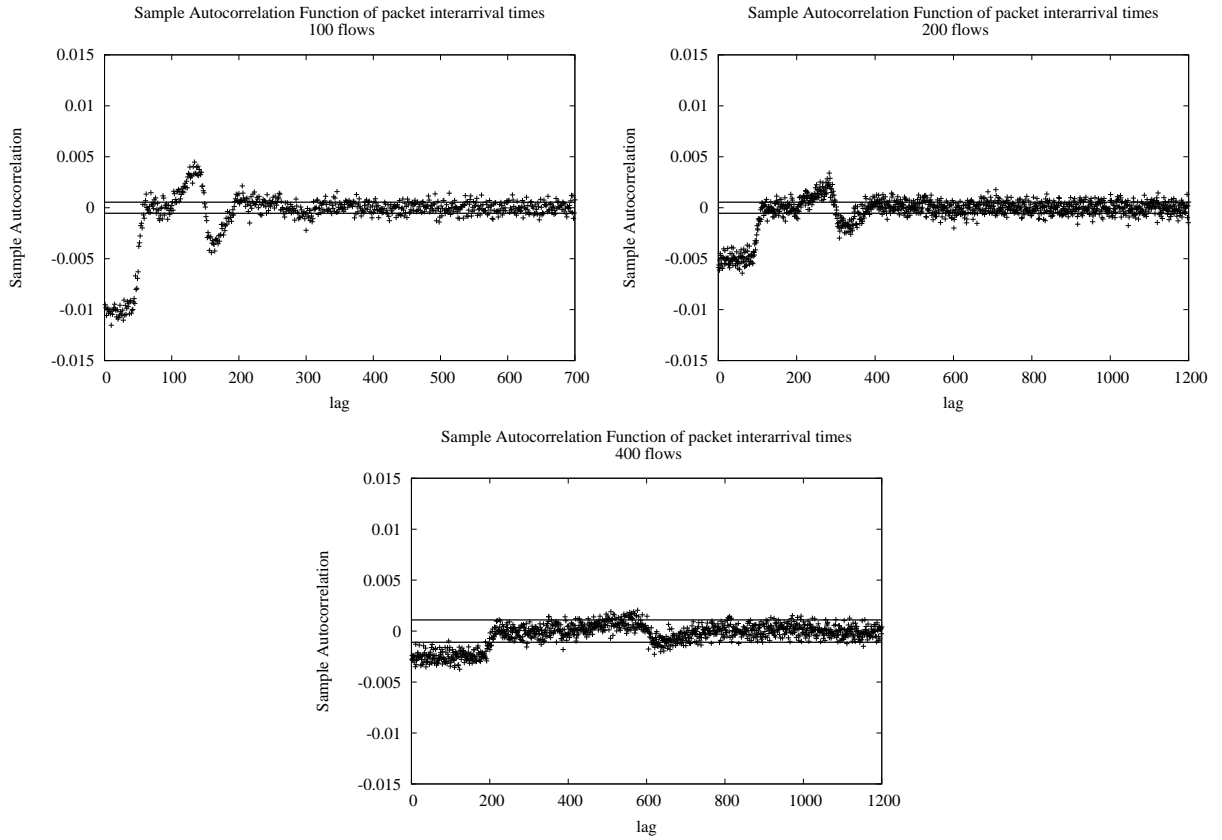


Fig. 4. Autocorrelation of the refresh message interarrival times (100, 200 and 400 flows)

$$\pi_i^{(\infty)} = \begin{cases} 1 - \rho & \text{if } i = 0 \\ \frac{1}{1 - a_0} \left(a_{i-1} \pi_0^{(\infty)} + \sum_{j=1}^{i-1} a_{i-j} \pi_j^{(\infty)} \right) & \text{if } i > 0 \end{cases} \quad (1)$$

where the values a_i are the probability that during a service time there are at least $i + 1$ arrivals, namely

$$a_i = \int_0^\infty G_S(x) \frac{(\lambda x)^i}{i!} \lambda e^{-\lambda x} dx, \quad i \geq 0 \quad (2)$$

where $G_S(x)$ is the survival function of the service time. In our scenario, a constant service time $\bar{s} = \frac{PacketSize}{BW}$ is assumed because Path and Resv packet sizes are constant,

$$G_X(x) = P(S > x) = \begin{cases} 1 & \text{if } 0 \leq x < \bar{s} \\ 0 & \text{if } x \geq \bar{s} \end{cases} \quad (3)$$

Hence, the values of a_i are easily computed recursively as follows

$$a_i = \int_0^{\bar{s}} \frac{(\lambda x)^i}{i!} \lambda e^{-\lambda x} dx = \begin{cases} a_{i-1} - \frac{(\lambda \bar{s})^i e^{-\lambda \bar{s}}}{i!} & \text{if } i > 0 \\ 1 - \frac{(\lambda \bar{s})^i e^{-\lambda \bar{s}}}{i!} & \text{if } i = 0 \end{cases} \quad (4)$$

Once the queue length distribution $\pi_i^{(\infty)}$ for the infinite buffer system has been obtained from equations 1 and 4, we compute the probability

$$q_K = 1 - \sum_{i=0}^{Q_{size}} \pi_i^{(\infty)}. \quad (5)$$

Finally, the loss probability for the $M/D/1/K$ system can be expressed as follows

$$p_{loss} = \frac{(1 - \rho)q_K}{1 - q_K \rho}. \quad (6)$$

Fig. 5 presents the packet loss probability from the above mentioned model compared to the simulation results. The analytical results provide an upper bound for the loss probability. Hence, they can be used in worst-case approximations. The deviation of the analytical results from the simulation results is due to the correlation between refresh messages, that makes the arrival process not being Poisson. As the number of flows multiplexed in the link grows the traffic gets closer to a Poisson process and the difference between analytical and simulation results is reduced. This effect can be observed in Fig. 6. We have presented the error percentage in the packet loss probability estimation for the same utilization factor and queue size as the number of flows increases. We observe that the error decreases significantly with the multiplexing degree.

From the results in Fig. 5 it turns out that very small buffers (in the range of Kbytes) are necessary to achieve packet losses smaller than 10^{-3} . However, note that the number of flows is in the order of hundreds, and this is expected to grow as the optical bandwidth increases.

C. Survival time of a lightpath

In this section, we analyse the probability of erroneous disconnection and the time it takes for a flow to (eventually) be torn down due to loss of refresh messages. In order for a flow to be torn down no refresh message must be received during the *cleanup timeout*.

1) *Derivation of the erroneous disconnection probability:* Let us assume that the packet loss probability with a number of flows $NFlows$, bandwidth BW and queue size $Qsize$ is p_{loss} .

Fig. 7 shows that the number of losses to turn down the flow is not always the same, as explained in the following example. First, recall that the interarrival times of refresh messages are uniformly distributed between $0.5R$ and $1.5R$. Concerning the timeout value L , we note that $L = (K + 0.5) \cdot 1.5R = 5.25R$. Let us consider eleven interarrival times of refresh messages, which are equal to $0.5R$, as seen in Fig. 7a. Then, the RSVP signaling agent does not receive refresh messages for $5.5R$ and the flow is turned down. If, however, the interarrival times happen to be equal to $1.5R$ then only three consecutive losses suffice to turn down the flow (Fig. 7b). Finally, Fig. 7c shows a case in between the previous extreme cases.

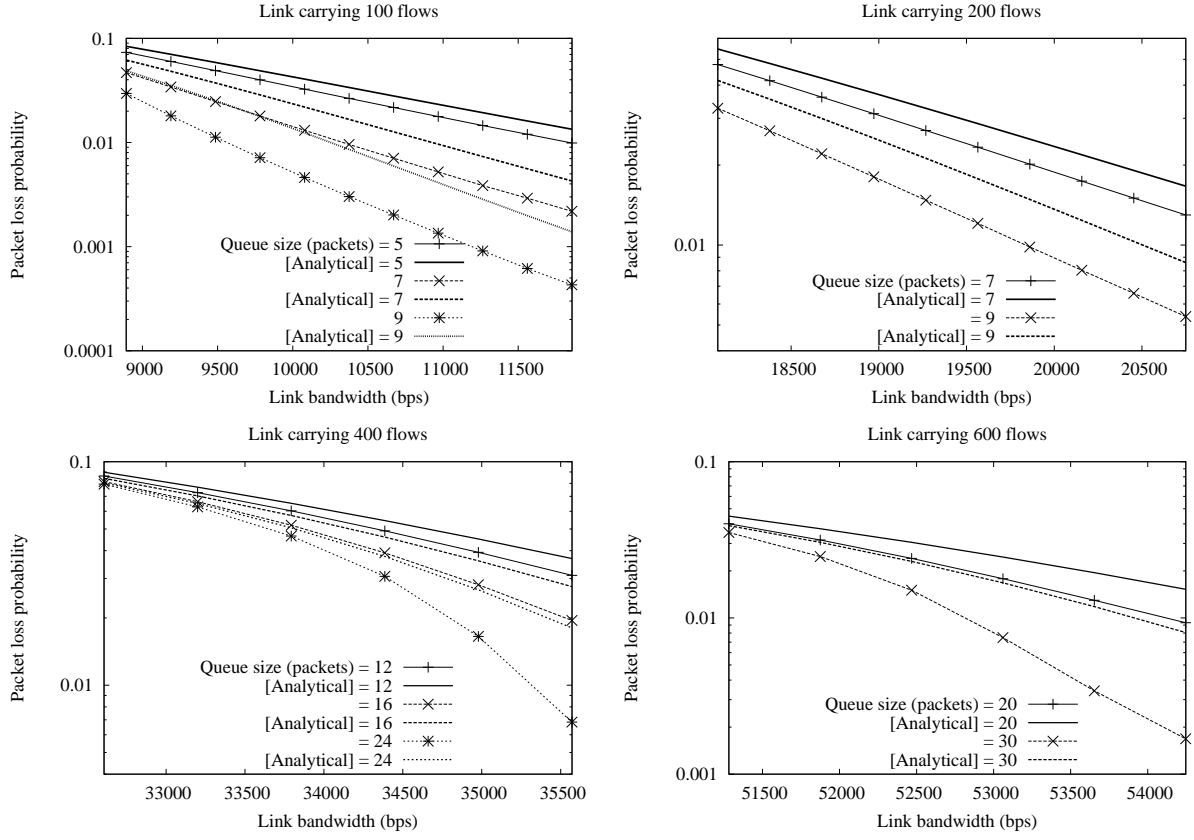


Fig. 5. Packet loss probability from simulation and the analytical model

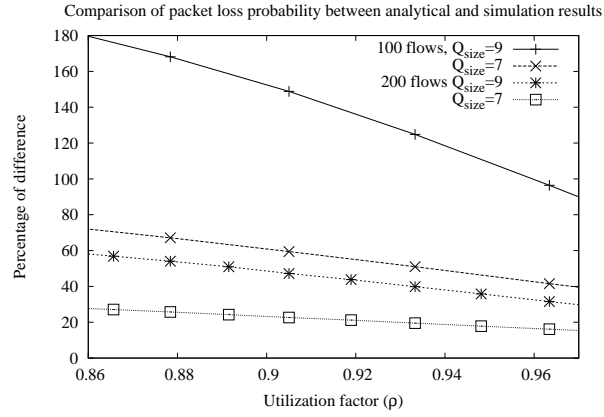


Fig. 6. Model improvement with the multiplexing degree

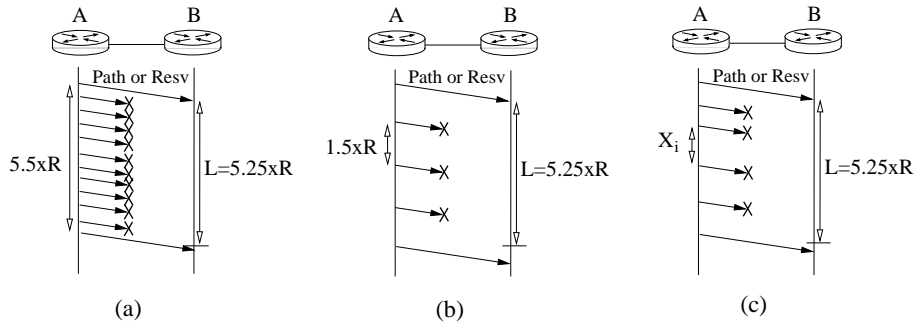


Fig. 7. Disconnections with different number of losses. (a) Minimum inter-arrival time. (b) Maximum inter-arrival time. (c) Random case

Let X_0, \dots, X_n refer to the interarrival times of n consecutive lost refresh messages. X_0 represents the time elapsed between the last successful arrival and the first loss. X_i is the time elapsed between the i -th loss and the $(i+1)$ -th loss and X_n is the time from the n -th loss to the first message successfully transmitted. It is worth remarking, as seen in Fig. 7, that a series of n consecutive losses can either turn down the associated flow or have no effect at all. Let $T^{(n)} = \sum_{i=0}^n X_i$. Then,

$$p_{down} = \sum_{n=1}^{\infty} P(T^{(n)} > L)(p_{loss})^n(1 - p_{loss}) \quad (7)$$

The random variables X_i are independent and identically distributed as uniform random variables in the range $[0.5R, 1.5R]$. The values of $P(T^{(n)} > L)$ can be computed numerically. For example, they are shown in Table 2 for $R = 30$ secs.

Number of losses (n)	$P(T^{(n)} > L)$
2	0
3	0.0131836
4	0.353467
5	0.851581
6	0.990212
7	0.999846
8	~ 1

TABLE 2

VALUES OF $P(T^{(n)} > L)$ FOR $R = 30$ SECS

In order to derive p_{down} we take the approximation

$$p_{down} \approx \max_{n=1 \dots \infty} P(T^{(n)} > L)(p_{loss})^n(1 - p_{loss}) \quad (8)$$

Note that the function to maximize is well defined as a map $\mathcal{N} \rightarrow [0, 1]$ and $(p_{loss})^n(1 - p_{loss})$ is a decreasing function with n . On the other hand,

$$\begin{aligned} P(T^{(n+1)} > L) &= P\left(\sum_{i=1}^n X_i > L\right) = P\left(\sum_{i=1}^{n-1} X_i > L - X_n\right) \\ &> P\left(\sum_{i=1}^{n-1} X_i > L\right) = P(T^{(n)} > L) \end{aligned} \quad (9)$$

And, thus, $P(T^{(n)} > L)$ is an increasing function with n (as seen for example in Table 2). Furthermore, it can be easily shown that $P(T^{(2)} > L) = 0$ but $P(T^{(n)} > L) \neq 0, n \geq 3$ and $P(T^{(10)} > L) = 1$. Therefore, $P(T^{(n)} > L)(p_{loss})^n$ has a zero in $n = 2$ and a zero in $n \rightarrow \infty$, and takes values in the range $(0, 1)$ in between. Thus, a maximum exists in the interval $[3, \infty] \in \mathcal{N}$.

To calculate such maximum we follow an iterative procedure. If the maximum is attained at $n = 3$ losses then

$$P(T^{(3)} > L)(p_{loss})^3 > P(T^{(n)} > L)(p_{loss})^n, \quad \forall n > 3. \quad (10)$$

This means that $p_{loss}^{n-3} < P(T^{(3)} > L)/P(T^{(n)} > L), \forall n > 3$. From the values in Table 2 it is easily checked that $p_{loss} < P(T^{(3)} > L)/P(T^{(4)} > L) \approx 0.0373$. For larger loss probability the case with 4 losses is more likely. Following the induction argument this case will be the most likely as long as $p_{loss} < P(T^{(4)} > L)/P(T^{(5)} > L) \approx 0.415$ and the general rule is that

the maximum is attained at n losses iff $P(T^{(n-1)} > L)/P(T^{(n)} > L) < p_{loss} < P(T^{(n)} > L)/P(T^{(n+1)} > L)$, $n \geq 4$. Table 3 shows the most likely number of losses, for a range of packet loss values and $R = 30$ secs.

Most likely number of losses (n)	Packet loss range
3	$p_{loss} < 0.0373$
4	$0.0373 < p_{loss} < 0.415$
5	$0.415 < p_{loss} < 0.86$
6	$0.86 < p_{loss} < 0.99$

TABLE 3

LOSS PROBABILITY AND MOST PROBABLE DISCONNECTIONS

For a case with $N_{flows} = 100$, $BW = 9500$ bps and $Q_{size} = 5$ the experimental loss probability is $p_{loss} = 0.049$ and it turns out that the maximum is attained at $n = 4$ losses.

2) *Derivation of the flow survival probability for a given time interval:* We may now derive the probability that a flow remains connected after an arbitrary period of time of t seconds. To make the problem analytically tractable, the time axis is slotted in constant length refresh intervals, whose length is approximated by the mean $E[R]$. Then, we calculate the probability that a connection lifetime is greater than a given number of such intervals $n_t = \lceil t/E[R] \rceil$.

Now, assumming packet losses are i.i.d. we are interested in the event that a signaling flow suffers enough consecutive losses. To this end, we use a discrete-time discrete Markov chain (DTDMC) in which the states represent the number of consecutive losses. In our numerical example ($N_{flows} = 100$, $BW = 9500$ bps and $Q_{size} = 5$), the chain states lie in the set $\{0,1,2,3,4\}$. The state four is an absorbent state that represents four consecutive losses and the connection removal due to timeout. Consequently, we wish to derive the distribution of the chain hitting time for state four. Equation (11) shows the one-step transition probability matrix for this chain.

$$\mathbf{Q} = \begin{pmatrix} 1 - p_{loss} & p_{loss} & 0 & 0 & 0 \\ 1 - p_{loss} & 0 & p_{loss} & 0 & 0 \\ 1 - p_{loss} & 0 & 0 & p_{loss} & 0 \\ 1 - p_{loss} & 0 & 0 & 0 & p_{loss} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (11)$$

Now, if D represents the flow survival time then

$$P(D \leq t) \approx \mathbf{Q}^{n_t}(1, 5) \quad (12)$$

where $n_t = \lceil t/E[R] \rceil$. Fig 8 shows the survival flow duration probability $P(D > t)$ for different degrees of multiplex in the link (100, 200 and 400 flows) and different link bandwidths. We focus on the range of very high survival flow probability (above 0.99). The figures compare the analytical result presented above with the values for the flow duration obtained from simulation. Both analytical and simulation values are close and typically the analytical approximation behaves as a lower bound to the flow duration, mainly due to the worst-case loss probability.

Finally, Table 4 summarizes the steps for analytically finding the bandwidth BW and queue size Q_{size} in order for a flow to survive a time larger than t with a probability $p_{objective}$.

Single-hop network scenario: $NFlows$, R and $PacketSize$

Survivability objective: $p_{objective}$ and t such that $P(D > t) > p_{objective}$

Target results: BW and Q_{size} for the two link directions

Procedure:

1. Compute the approximate number of refresh intervals $n_t = \lceil t/E[R] \rceil$
2. Find p_{loss} as a zero of $f(p_{loss}) = Q^{n_t}(1,5) - (1 - p_{objective})$, for example with an iterative algorithm.
3. Find an $M/D/1/(Q_{size} + 1)$ system such that its loss probability is p_{loss} . This can be done with a plot similar to the ones displayed in Fig. 5 with number of flows $NFlows$ and values of $BW > NFlows \cdot PacketSize/R$. The value p_{loss} is a horizontal line that intersects the curves in the figure (each one for a different Q_{size}). There is a solution for each given Q_{size} , the smaller the value of Q_{size} the larger the value of BW .

TABLE 4

LINK DIMENSIONING PROCEDURE

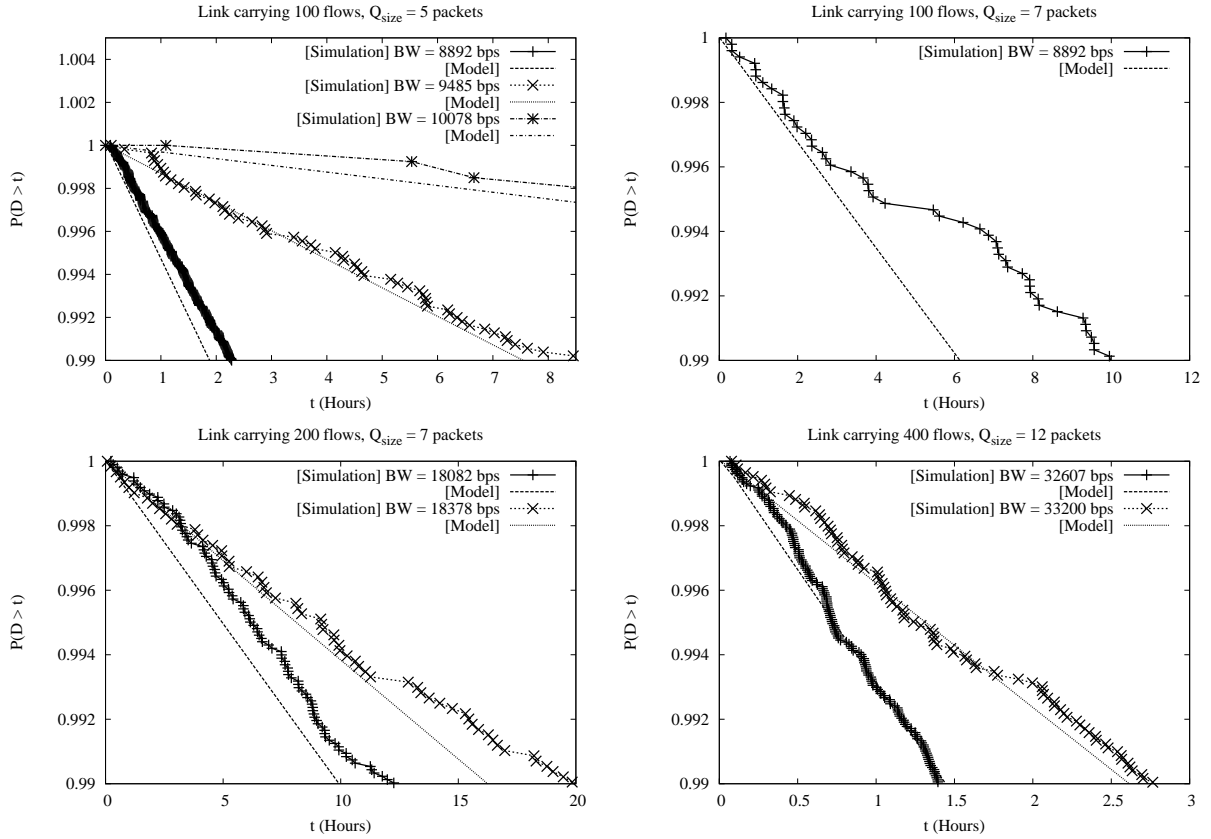


Fig. 8. Simulated and analytical results comparison for the flow duration probability

4. MULTI-HOP SCENARIO

The previous sections provided the capacity planning rules for a single link. Bandwidth and queue size values were found such that, with a given probability, the flows crossing the link will not suffer an erroneous disconnection in less than a certain amount of time. It is worth noting that the links are completely independent, i.e., neither the arrival nor the loss of refresh messages in a link affects any other link. Let us now consider an optical network represented by the graph (N, V) where N is the set nodes and V is the set of vertices. Let us also consider an RSVP flow that is routed across nodes $N_1, \dots, N_m \in N$

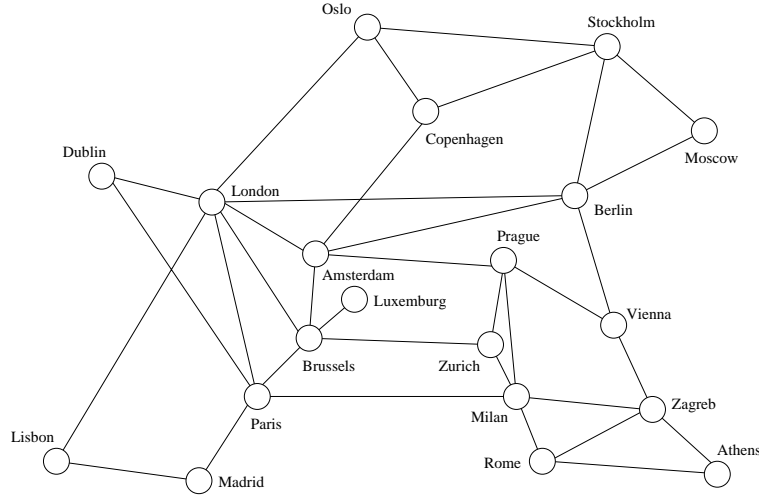


Fig. 9. COST239 network

without loss of generality. Let $P(D_{N_i \rightarrow N_{i+1}} > t)$ refer to the probability that the flow lifetime is longer than t , assuming that the distance between nodes N_i and N_{i+1} is one-hop. Such probability is given by (12). Then, for a multi-hop flow, the flow survival time probability $P(D_{\text{multihop}} > t)$ is given by

$$P(D_{\text{multihop}} > t) = \prod_{i=1}^{m-1} P(D_{N_i \rightarrow N_{i+1}} > t) \quad (13)$$

The above equation provides a means to derive the signaling plane loss probability in terms of a given flow survival probability objective, which will be denoted by $p_{\text{objective}}$. We note that there are multiple solutions to the equation $P(D_{\text{multihop}} > t) > p_{\text{objective}}$, which yield different loss probabilities in each link $N_i \rightarrow N_{i+1}, i = 1, \dots, m-1$. Nevertheless, assuming that all links have the same loss probability then (13) reduces to

$$P(D_{\text{multihop}} > t) = P(D > t)^{m-1} \quad (14)$$

where $P(D > t)$ is obtained from (12). Since $m-1$ is the number of hops for the flow, we consider a worst-case capacity planning by letting m be equal to the longest path in number of hops, whose length will be referred to as m_{max} . If that is the case,

$$P(D > t) > (p_{\text{objective}})^{1/(m_{\text{max}}-1)} \quad (15)$$

and we only have to solve numerically equation (16) for the value of p_{loss} .

$$1 - \mathbf{Q}^{n_t}(1, 5) > (p_{\text{objective}})^{1/(m_{\text{max}}-1)} \quad (16)$$

A. Numerical experiment with the COST239 network

In this section, we provide the maximum allowable loss probability per link in the COST239 network, which is shown in Fig. 9.

In the COST 239 topology and using shortest paths based on the hop count, the maximum path length is five hops. We solve equation (16) and obtain the required loss probability versus $P(D_{\text{multihop}} > t)$ for t equal to one and two weeks, as

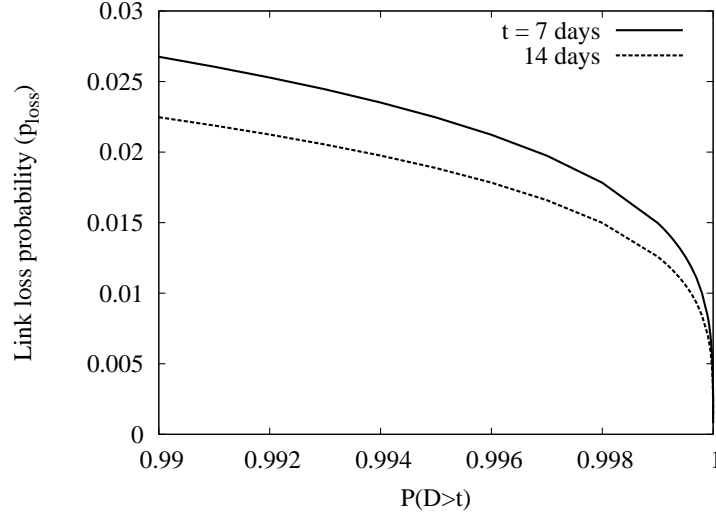


Fig. 10. Loss probability as a function of the quality requirements

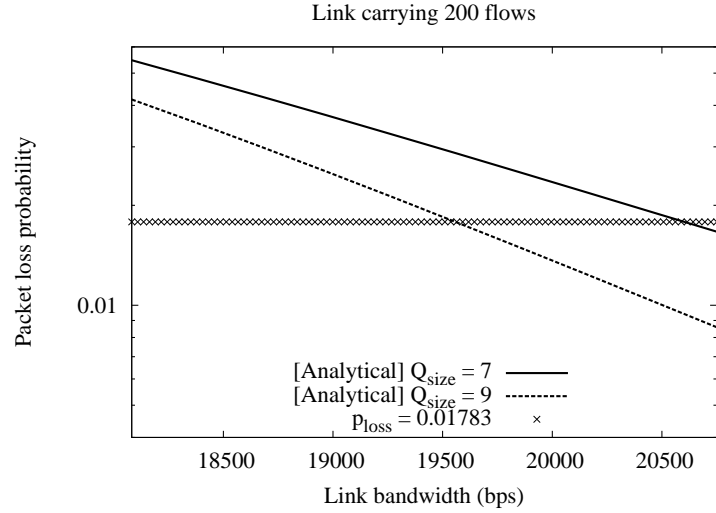


Fig. 11. Example of minimum link resources computation

shown in Fig. 10. Then, assuming for example that our target is $P(D_{multihop} > 7 \text{ days}) = 0.99$ it follows from (14) that $P(D > 7 \text{ days}) \approx 0.998$ and from Fig. 10 the loss probability should be less than 0.01783. Now, let us assume that our data path is carrying 200 flows. The minimum bandwidth required will be $BW > NFlows \cdot PacketSize / R = 200 \cdot 332 \cdot 8 / 30 \text{ bps} = 17,707 \text{ bps}$. From Fig. 11 we identify a couple of possible solutions for the link resource requirements. We may set up a minimum bandwidth of 19,500 bps with a buffer of at least nine packets (less than 3 KBytes). A second possibility would be to set up a smaller buffer size, with only seven packets (less than 2.3 KBytes), but a link bandwidth of at least 20,500 bps.

5. CONCLUSIONS

This paper provides capacity planning rules for the control plane of a GMPLS network. We have presented analytical rules to obtain the maximum allowable loss probability on a link in order to ensure a survival time for a given lightpath. The analysis also provides the link resources needed, in term of reserved bandwidth and buffer space. Our findings show that a very small buffer size and bandwidth can sustain the signaling load for hundreds of flows, even in a scenario with the simplest RSVP refresh mechanism. The analytical procedure has also been extended from single links to complex network topologies. Thus,

the results presented in this paper can be used to accurately dimension the GMPLS network control plane.

REFERENCES

- [1] A. Banerjee, J. Drake, J. Lang, B. Turner, D. Awduche, L. Berger, K. Kompella, and Y. Rekhter, "Generalized Multiprotocol Label Switching: An overview of signaling enhancements and recovery techniques," *IEEE Communications Magazine*, vol. 39, no. 7, pp. 144–151, July 2001.
- [2] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, "RSVP: A new resource reservation protocol," *IEEE Network*, vol. 7, no. 5, pp. 8–18, 1993.
- [3] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP tunnels," RFC 3209, December 2001.
- [4] L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) signaling functional description," RFC 3471, January 2003.
- [5] L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) signaling resource reservation protocol-traffic engineering (RSVP-TE) extensions," RFC 3473, January 2003.
- [6] P. Ji, Z. Ge, J. Kurose, and D. Towsley, "A comparison of hard-state and soft-state signaling protocols," in *Proceedings of the 2003 conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, Karlsruhe, Germany, August 2003, pp. 251–262, ACM Press.
- [7] O. Komolafe and J. Sventek, "An evaluation of RSVP control message delivery mechanisms," in *Proceedings of Workshop on High Performance Switching and Routing (HPSR'04)*, Phoenix, AZ, USA, April 2004, pp. 43–47.
- [8] O. Komolafe and J. Sventek, "RSVP performance evaluation using multi-objective evolutionary optimisation," in *Proceedings of IEEE INFOCOM*, Miami, FL, USA, March 2005, vol. 4, pp. 2447–2457.
- [9] M. Karsten, J. Schmitt, and R. Steinmetz, "Implementation and evaluation of the KOM RSVP engine," in *Proceedings of IEEE INFOCOM*, Anchorage, AK, USA, April 2001, vol. 3, pp. 22–26.
- [10] A. Neogi, T.-C. Chiueh, and P. Stirpe, "Performance analysis of an RSVP-capable router," *IEEE Network*, vol. 13, no. 5, pp. 56–63, September/October 1999.
- [11] H. Wang, R. Karri, M. Veeraraghavan, and T. Li, "A hardware-accelerated implementation of the RSVP-TE signaling protocol," in *Proceedings of IEEE International Conference on Communications*, Paris, France, June 2004, vol. 3, pp. 1609–1614.
- [12] L. Berger, D. Gan, G. Swallow, P. Pan, F. Tommasi, and S. Molendini, "RSVP refresh overhead reduction extensions," RFC 2961, April 2001.
- [13] R. Braden (Ed.), L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource reservation protocol (RSVP) – version 1 functional specification," RFC 2205, September 1997.
- [14] D. Gross and C. M. Harris, *Fundamentals of Queueing Theory*, John Wiley and Sons, 2nd edition, 1985.