

# Spare Capacity Reprovisioning for Shared Backup Path Protection in Dynamic Generalized Multi-Protocol Label Switched Networks

Pin-Han Ho\*, János Tapolcai<sup>†</sup>, and Anwar Haque\*

\* Dept. of Electrical and Computer Engineering, University of Waterloo, Canada, pinhan@bcr.uwaterloo.ca

<sup>†</sup>Dept. of Telecommunications and Media Informatics, Budapest University of Technology, tapolcai@tmit.bme.hu

**Abstract**—Spare capacity allocation serves as one of the most critical tasks in dynamic Generalized Multi-Protocol Label Switching (GMPLS) networks to meet the stringent network availability constraint stipulated in the Service Level Agreements (SLAs) of each connection. In this paper, an availability-aware spare capacity reconfiguration scheme based on shared backup path protection (SBPP) is proposed, aiming to guarantee the end-to-end (E2E) availability of each Label Switched Path (LSP). We first provide an E2E availability model for SBPP connections in presence of all possible single and dual simultaneous failures. Partial restoration is identified to further improve the capacity efficiency, and achieve finer service differentiation. For this purpose, *restoration attempt* is defined as a parameter for each connection that can be manipulated at the source node when the spare capacity of each link is scheduled. Based on the developed model, a Linear Program (LP) is formulated to perform inter-arrival spare capacity reconfiguration along each pre-determined shared backup LSP to meet the availability constraint of each connection. Simulation is conducted to verify the derived formulation, and to demonstrate the benefits gained in terms of the spare capacity saving ratio, where the conventional SBPP scheme that achieves 100% restorability for any single failure is taken as a benchmark. We will show that the simulation results validate the proposed E2E availability model, where a significant reduction on the required redundancy can be achieved in the effort of meeting a specific availability constraint for each SBPP connection.

## I. INTRODUCTION

As the Internet evolves to a connection-oriented environment that addresses various quality of service (QoS) requirements, the Generalized Multi-Protocol Label Switching (GMPLS) based bandwidth provisioning [1], [2] is envisioned to be the most promising platform that can greatly facilitate traffic engineering, multiple classes of service (CoS), end-to-end (E2E) QoS guarantee, and interoperability of heterogeneous network environments. As it has gradually come to a usual case that an availability requirement is stipulated in the Service Level Agreement (SLA) by the end users, the E2E availability for each Label Switched Path (LSP) supporting a specific type of service (e.g., VoIP, TCP, or real-time multimedia streaming, etc.) is of great interest to the network control and management organizations.

To improve E2E availability of a connection, it has been well proved that allocating redundant network resources for the connection is the best policy in the network layer when the physical availability of each network component is constant. In

the dynamic GMPLS-based bandwidth provisioning scenario, a working LSP could be equipped with one or multiple Shared Risk Group (SRG) -disjoint backup LSPs (or path segments) such that one or multiple simultaneous unexpected failures that affect the connection could be automatically restored. A number of protection schemes have been reported and extensively investigated in the past, including shared backup path protection (SBPP) [3]–[6], 1+1 protection [1], shared segment protection (SSP) [7], [8], and dual-failure protection, etc. All of them have a design goal of reducing or /minimizing the allocated spare capacity subject to different constraints, and failure scenarios, such as a recovery time constraint, availability constraint, SRG-disjointness constraint, or a guarantee in terms of survivability under one or two simultaneous failures, etc.

SBPP with a single backup LSP for a working LSP is a type of widely adopted scheme for achieving dynamic GMPLS-based recovery due to its simplicity, and dynamicity. Compared with 1+1 protection, SBPP has been considered as a more aggressive spare capacity allocation strategy that can be significantly more capacity-efficient by enabling spare resource sharing among different backup LSPs while yielding a similar level of E2E availability. In general, with SBPP, or 1+1 protection, the E2E availability of a connection can be significantly improved by an order of two or more compared to the case with only a single working LSP. The service provisioned by the working LSP and the SRG-disjoint backup LSP can only be impaired when both paths are interrupted simultaneously.

A significant number of previously reported studies on network availability and survivability assume traffic uniformity (i.e., each connection carries the same amount of bandwidth), and connection indivisibility (i.e., the working bandwidth of a connection must be provisioned either all or none), along with 100% restorability to a specific number of simultaneous failures. The availability constraint of each connection, nonetheless, is either considered separately from the cost optimization process [4], [9]–[13], or totally ignored [3], [5]–[8], [14]. In some cases, such as all-optical bandwidth provisioning with lightpaths in Wavelength Division Multiplexing (WDM) networks, keeping the traffic uniformity and connection indivisibility in the restoration process is inevitable. However, because the GMPLS control plane supports different switching granularities and restoration capacities, the

assumptions of traffic uniformity and connection indivisibility may not always be necessary. Also, the effort of achieving 100% restorability for any number of simultaneous failures may deviate from the design premise that the E2E availability of each connection is interesting to the customers, and should be taken as an ultimate design goal.

The representative studies that addressed the efforts of availability modeling/evaluation can be seen in [4], [9]–[17]. The authors in [4] investigated the availability impairment due to shared protection on SBPP connections compared with the case of 1+1 protection. The study in [9] conducted availability analysis on WDM networks using both Integer Linear Programming (ILP), and heuristic approaches, where a connection is protected by either none, or a single dedicated protection path. In [10], case studies were conducted on a number of network topologies and protection scenarios for availability evaluation and modeling were conducted. The study in [11] explored the restorability under the dual-failure scenario in APS rings and mesh networks with span-protection originally designed for achieving 100% restorability in the single failure scenario.

In [12], the availability of a long-haul point-to-point optical transmission system was evaluated, where an availability-aware link-state packet is devised and disseminated to facilitate dynamic routing under an availability constraint. In [13], ILPs were formulated to perform span protection on each connection to fit into a specific design objective under the dual-failure scenario. The paper has tackled the cases where the optimization on the availability of each connection subject to a capacity constraint, and the minimization of total spare capacity subject to an availability constraint. In [15], two traffic grooming algorithms were introduced to guarantee the E2E availability of each connection based on dedicated protection.

The evaluation of E2E availability was discussed in [16] based on a number of rules of thumb, such as the E2E availability under a protection scheme for all single failures, or all single and & dual failure events, etc. In [14], a precise approach was introduced in estimating the unavailability of each failure pattern using Markov chains, where the sequence of failures in each failure pattern is considered and modeled in order to correctly evaluate the E2E availability with shared protection. By the same authors in [14], the evaluation of E2E availability for SBPP connections was conducted in [18] based on the stationary probability of pre-defined failure patterns, where each connection is assumed to have uniform and , indivisible bandwidth.

Note that none of the above mentioned studies discussed the possibility of partial restoration, which has been investigated in [19]–[21]. The study in [19] adopted a partial protection strategy to achieve a deterministic Quality of Protection (QoP) paradigm. The study in [20] conducted extensive simulation, and concluded that the partial restorability could lead to smaller resource consumption than that in the full restorability case. The authors in [21] demonstrated that partial restorability on the video streams in SONET/SDH rings leads to smaller capacity demand. The research concluded that the consumed resources are a linear function of a fraction of restorability.

Nonetheless, the above three papers have never touched the availability evaluation, and have not provided any information on how the source node randomly drops a portion of the working bandwidth in the restoration phase.

Spare capacity reconfiguration is another focus in the paper, and has also been extensively studied in the past [13], [22]–[24]. In [13], spare capacity reconfiguration is performed on the framework of p-cycle with a focus on how to improve the dual-failure restorability, which was originally designed for achieving 100% restorability for any single failure. The study in [22] introduced a decent approach in calculating the minimum spare capacity along each link to achieve 100% restorability for any single failure. The proposed algorithm, Successive Survivable Routing (SSR), can effectively solve the spare capacity reconfiguration by sequentially rerouting the backup path of each connection. However, the E2E availability and the inference by double simultaneous failures have not been considered. In our previous work of [23], inter-arrival spare capacity reconfiguration is performed by investigating into the computation efficiency and grouping policies of network traffic, where each lightpath is prepared with backup path segments for achieving 100% restorability in the single failure scenario. In [24], a new link-state metric in rerouting each backup path through a wavelength channel is proposed. The goal of the backup path rerouting is to evacuate all the backup lightpaths traversing through a specific wavelength link to claim that the wavelength link is free. A complete work on availability-aware spare capacity reconfiguration has never been reported.

It is clear that the network control and management requires an integrated strategy to perform availability-aware spare capacity reconfiguration in a dynamic network environment. Instead of following specific policies such as achieving 100% restorability under a single or double (or even triple) simultaneous failures, it is envisioned that a more general framework is of a high significance to the network design, and resource allocation. Thus, this paper is committed to investigating the availability-aware spare capacity allocation problem, and providing a general model for evaluating the E2E availability for SBPP connections; this is – a simple and , efficient protection scheme that is being widely adopted by the current carrier networks. Distinguished from the previous studies, the paper explores the best design generality by highlighting the concept of partial restoration, where contention among different backup LSPs under a common failure event is considered. To exercise this concept, and formulate the problem, the E2E unavailability of each connection is modeled by enumerating the availability impairments due to all the related single and dual failure events.

To perform a global optimization in spare capacity reconfiguration, a novel iterative linear program (LP) is introduced to guarantee the unavailability of each connection while minimizing the consumed redundancy. This is done by performing inter-arrival reconfiguration on the spare capacity along each link and the *restoration attempt* (i.e., the percentage of a connection’s working bandwidth intended to be restored). We demonstrate that most of the policies taken by the previous studies can be categorized as a special case of the submitted

TABLE I  
NOTATIONS

$N, E$	The node set, and link set of the input topology, respectively
$j$	The index of links taken by backup LSPs and working LSPs
$f_j, v_j, q_j$	The free, spare, and working capacity on link $j$ , respectively
$D, d$	The set of connection in the network, and its notation as index
$b^d, U_{sla}^d$	The bandwidth, and the unavailability constraint of $d$
$W^d, P^d$	The working, and shared backup LSP of connection $d$
$R, r$	The set of failure patterns under consideration, and its index
$m, n$	indices of SRGs
$\{m\}$	The failure pattern of a SRG $m$
$\{m, n\}$	The failure pattern of an SRG duplet, where $m$ is the earlier failed SRG, and $n$ the latter one
$\pi_r$	The stationary probability of failure pattern $r$
$u_r^d$	The availability impairment on connection $d$ due to $r$
$\hat{R}^d$	The subset of $R$ such that $u_r^d > 0$
$\theta_k^d$	The fraction of the total bandwidth switched on the protection route of $k^{th}$ backup segment of connection $d$
$q_r^d$	The fraction of working bandwidth be randomly dropped due to $r$
$R_{non}^d$	The intrinsically non-restorable failure patterns of $d$
$R_{w\bar{p}}^d$	The failure patterns that hit both $W^d$ , and $P^d$ , but $\notin R_{non}^d$
$R_{\bar{w}p}^d$	The failure patterns which interrupt $W^d$ , but not $P^d$
$R_w^d$	The failure patterns not influencing the availability of $d$
$u^d$	The E2E unavailability of $d$
$u_{\bar{w}p}^d, u_{w\bar{p}}^d$	the availability impairment due to $r \in R_{\bar{w}p}^d$ , and $r \in R_{w\bar{p}}^d$
$\hat{u}_{\bar{w}p}^d, \hat{u}_{w\bar{p}}^d$	the value of $u_{\bar{w}p}^d$ for single, and dual failures, respectively
$s_{j,r}$	The required spare capacity along link $j$ to accommodate the restoral bandwidth due to $r$
$s'_{j,r}$	A working variable to make $s_{j,r}$ linear in the ILP formulation
$y_{j,r}$	The non-restorability density on link $j$ due to $r$
$P^d$	The set of backup LSPs that traverse through any link taken
$\equiv \{m\}$	by $P^d$ that will be activated due to the failure on SRG $\{m\}$

model. Simulation is conducted to validate the proposed model, and verify the developed spare capacity reconfiguration scheme by solving the iterative LP.

The rest of the paper is organized as follows. Section II formulates the problem, and defines the system parameters. Section III introduces the concept of partial restoration, and models the resultant availability impairment, as well as the random dropping mechanism caused by contention. Section IV presents the proposed E2E availability model for SBPP connections. Section V introduces our dynamic availability-aware survivable routing architecture, where a Linear Program (LP) for performing spare capacity reconfiguration is formulated based on the derived E2E unavailability model. Section VI presents the simulation results, including the validation of the proposed model, and verification of the efficiency in the spare capacity reconfiguration. Section VII concludes the paper.

## II. PROBLEM FORMULATION

Let the network be represented by  $G(E, N)$ , where  $E$  is the set of links, and  $N$  is the set of nodes, launched with a set of connections denoted as  $D$ . Each connection  $d \in D$  is defined with a source, and a destination node; along with the required bandwidth  $b^d$ , and unavailability constraint  $U_{sla}^d$ . The working LSP of  $d$  is denoted as  $W^d$ , which is protected by a shared backup LSP (denoted as  $P^d$ ) to meet the unavailability constraint. In this study, a Shared Risk Group (SRG) is defined as one or a set of links and nodes in the network topology that could be hit by a single failure event, and a failure event on an SRG is independent of that on the others. Let all the SRGs be

contained in a set denoted as  $SRG$ . A failure pattern is defined as one or a number of SRGs that could be subject to failure simultaneously. It is clear that the number of total possible failure patterns in the networks is exponentially increasing as the number of SRGs increases. Because some failure patterns are subject to a really low probability to occur (such as the ones with a large number of simultaneous failed SRGs), it is feasible to only consider the failure patterns with a large enough probability to occur in the availability evaluation in spite of some imprecision introduced by ignoring the failure patterns with very little likelihood.

Let  $R$  denote the set of failure patterns under consideration, and the number of total failure patterns be denoted as  $|R|$ . Each failure pattern  $r$  has its stationary probability, which is denoted as  $\pi_r$ , and can be derived by solving the Markov chain model introduced in [14]. With all the  $|R|$  failure patterns defined, and their stationary probabilities solved, we can evaluate the E2E unavailability of connection  $d$  by enumerating the failure patterns which impair the E2E availability of  $d$ . Let the availability impairment on connection  $d$  due to failure pattern  $r$  be denoted as  $u_r^d$ . Since each failure pattern stands for a state in the developed Markov chain, the E2E unavailability of the connection can be evaluated by summing up the stationary probability corresponding to each failure pattern weighted by the availability impairment due to the failure pattern:

$$u_R^d = \sum_{\forall r \in R} \pi_r \cdot u_r^d = \sum_{\forall r \in \hat{R}^d} \pi_r \cdot u_r^d \quad (1)$$

where  $\hat{R}^d$  denotes the subset of  $R$  such that  $u_r^d > 0, \forall r \in R$ .

With a given set of failure patterns  $R$  under consideration, we define that a protection scheme for  $d$  can achieve *full restorability* (or  $u_R^d = 0$ ) in case  $\hat{R}^d = \emptyset$ . An example for full restorability is that connection  $d$  with SBPP can restore all possible single failures. In this case,  $R$  only contains all the failure patterns with a single SRG, and  $\hat{R}^d = \emptyset$  such that  $u_R^d = 0$  for any SBPP connection  $d$ .

A protection scheme with *partial restorability* is in contrast to the case of full restorability with  $\hat{R}^d \neq \emptyset$ , and  $u_R^d > 0$ . By taking the same example,  $u_R^d$  would become nonzero if dual simultaneous failures are considered. In this case,  $W^d$  is *partially restorable* if the expected restoral bandwidth in presence of  $\forall r \in R$  is only a fraction of the bandwidth provisioned by  $W^d$ .

Partial restorability on connection  $d$  can be achieved by allocating one or multiple backup path/segments to  $W^d$  such that the expected restorable bandwidth in presence of  $\forall r \in \hat{R}^d$  is a fraction of  $b^d$ . In this case, the switching node of the  $k^{th}$  backup segment switches a fraction  $\theta_k^d$  of the total bandwidth while the rest of the bandwidth is disregarded when the corresponding failure occurs. The pre-allocated spare capacity along the backup segment, thus, could be  $\theta_k^d \cdot b^d$  instead of  $b^d$ . Such a backup segment is termed  $\theta_k^d$ -restorative to  $W^d$ . Here,  $\theta_k^d$  is termed the *restoration attempt* of the  $k^{th}$  backup segment assigned to  $d$ . With the above definition on partial restoration, the conventional SBPP serves as a special case with a single 100% - restorative backup path.

Most of the previously reported approaches on availability evaluation had  $R$  to include all the possible failure patterns. The availability analysis becomes very complicated when resource sharing is enabled, and/or segment protection is adopted. However, some failure patterns have fairly low probabilities to occur, such as the ones with triple or more simultaneous failed SRGs. Thus, this study assumes that the network can only be hit by up to two simultaneous failures since because the stationary probabilities of the failure patterns with more than two simultaneous failed SRGs could be subject to very low probability compared with the stipulated availability requirement.

To perform the availability evaluation in Eq. (1), the failure patterns specific to SBPP connection  $d$  can be categorized into the following four groups: (1) intrinsically non-restorable failure patterns  $R_{non}^d$ , such as SRGs containing the source or destination node; (2) the failure patterns denoted as  $R_{w\bar{p}}^d$  that hit both  $W^d$ , and  $P^d$ , but not in  $R_{non}^d$ ; (3) the failure patterns denoted as  $R_{\bar{w}p}^d$  which interrupt  $W^d$  but not  $P^d$ ; and (4) the failure patterns that do not influence the availability of the connection ( $R_w^d$ ). Obviously, the four sets of failure patterns form a partition of  $R$ .

The E2E unavailability of  $d$  is denoted as  $u^d$ , and can be written as:

$$u^d = \sum_{\forall r \in R} \pi_r \cdot u_r^d = u_{non}^d + \sum_{\forall r \in R_{w\bar{p}}^d} \pi_r \cdot u_r^d + \sum_{\forall r \in R_{\bar{w}p}^d} \pi_r \cdot u_r^d = u_{non}^d + u_{w\bar{p}}^d + u_{\bar{w}p}^d \quad (2)$$

where  $u_{w\bar{p}}^d$ , and  $u_{\bar{w}p}^d$  are the availability impairments due to the failure patterns  $r \in R_{w\bar{p}}^d$ , and  $r \in R_{\bar{w}p}^d$ , respectively. It is clear that  $u_r^d = 100\%$  with  $r \in R_{w\bar{p}}^d$ , while  $0 \leq u_r^d < 1$  if  $r \in R_{\bar{w}p}^d$ . Note that the first group  $R_{non}^d$  is composed of the failure patterns that isolate the source and destination of the connection after its occurrence, and cannot be restored through any networking approach. Therefore, the terms  $u_{w\bar{p}}^d$ , and  $u_{\bar{w}p}^d$  in Eq. (2) determine the availability impairment on  $d$  will be the main focus in this study. Our objective is to perform SBPP for connection  $d$ ,  $\forall d \in D$ , such that  $u^d \leq U_{sla}^d$ , and the minimum amount of network capacity is consumed, where  $U_{sla}^d$  is the unavailability bound on connection  $d$ , and  $R$  contains all the single and dual simultaneous failures.

### III. PROPOSED SCHEME

#### A. Partial Restoration with SBPP

SBPP with a partially restorative backup LSP can be implemented in the IP/MPLS networks, in the sense that a working LSP may be composed of numerous independent nested LSPs such that dropping any/some of them would not affect the others. With a partially restorative backup LSP under SBPP, only a proportion of randomly selected nested LSPs of the working LSP could be restored in response to a failure event. In this case, the required spare capacity along link  $j$  to accommodate the restoral bandwidth due to failure event  $r$  (which hits a group of working LSPs) is denoted as  $s_{j,r}$ . By considering the possible contention between multiple restoral

flows for the working LSPs affected by a single failure event, only a fraction of the total bandwidth of each affected LSP can be restored. We assume that some randomly selected nested LSPs of the backup LSPs taking link  $j$  are simply dropped in proportion to the bandwidth of the contending backup LSPs. Thus, the effective restoral bandwidth of  $d$  could be even less than  $b^d \cdot \theta^d$ . Let the working bandwidth be randomly dropped by a fraction of  $q_r^d$ . The effective restoral bandwidth upon  $r$  is:

$$s_{j,r} = \sum_{\forall d | r \in R_{w\bar{p}}^d, j \in P^d} b^d \cdot \theta^d \cdot (1 - q_r^d) \quad (3)$$

The term  $\theta^d \cdot (1 - q_r^d)$  is the *effective restorability* for  $d$  at the occurrence of  $r$ , where  $\theta^d$  is a parameter implemented at the source node of  $d$ , and  $q_r^d$  is the result of random contention among the restoral flows due to the failure event. Thus, the terms  $s_{j,r}$  depends not only on  $\theta^d$ , but also on the bottleneck of each backup LSP when the failures occur. Both factors are due to the random dropping at the source and the intermediate nodes of the backup LSPs.

Fig. 1 exemplifies the random dropping of restoral flows at intermediate node C.  $K$  working LSPs originally pass through  $s$ - $b$ - $t$ , and the corresponding  $K$  backup LSPs with restoration attempt  $\theta_k$  for  $k = 1, 2, \dots, K$ , respectively, are activated along  $b$ - $c$ - $t$  when a failure on  $s$ - $b$  occurs. Because the spare capacity along  $b$ - $c$  and  $c$ - $t$  may not be sufficient to satisfy all the restoration attempts issued by node  $s$ , the exact restoral flow of each backup LSP is bottlenecked by one of the two links. In case the random dropping mechanism at an intermediate node is such that the bandwidth of each contending LSP is dropped in proportion to the requested bandwidth, the percentage of non-restorable bandwidth (or termed *non-restorability density*) at link  $s$ - $c$  and  $c$ - $t$  due to the failure of link  $s$ - $b$  would be:

$$y_{sc, sb} = \frac{\sum_k b_k \cdot \theta_k - v_{sc}}{\sum_k b_k \cdot \theta_k} \quad \text{and} \quad y_{ct, sb} = \frac{\sum_k b_k \cdot \theta_k - v_{ct}}{\sum_k b_k \cdot \theta_k}$$

respectively.

The bottleneck of the backup LSPs will lie in the link with the largest value of non-restorability density along each link taken by the backup LSPs. Thus, the link with the largest non-restorability density taken by backup LSP  $k$  due to the failure on  $s$ - $b$  can be expressed as  $q_{k, sb} = \max\{0, y_{sc, sb}, y_{ct, sb}\} \quad \forall k \in \{1, 2, \dots, K\}$ , which is also called the *bottleneck non-restoration density* of each of the  $K$  LSPs. Obviously, the term  $b_k \cdot \theta_k \cdot q_{k, sb}$  is the bandwidth of backup LSP  $k$  dropped due to the bottleneck at the occurrence of failure on link  $s$ - $b$ . Thus, the effective restorable bandwidth for backup LSP  $k$  becomes  $b_k \cdot \theta_k \cdot (1 - q_{k, sb})$ . Note that in case  $\sum_{k=1}^K b_k \cdot \theta_k \leq v_{sc}$ , and  $\sum_{k=1}^K b_k \cdot \theta_k \leq v_{ct}$ , all the restoral flows launched by the source node can go through the backup path without any dropping at node C.

The example in Fig. 1 is a special case because all the backup LSPs passing through  $s$ - $c$  also go through  $c$ - $t$  such that all the backup LSPs have the same bottleneck link. The situation could become much more complicated when

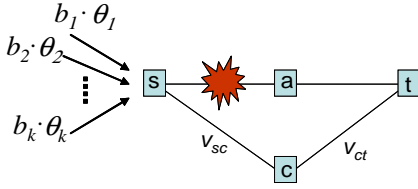


Fig. 1. An example showing the bottleneck, and the density of non-restorable bandwidth for a backup LSP: A special case where all the backup LSPs follow the same route.

we consider the contention among backup LSPs at a link with different source nodes, and physical routes. In this case, each backup LSP may have a different bottleneck, and the a corresponding non-restorability density, which is in turn subject to different effective restorability.

Fig. 2 demonstrates an example with five connections.  $W_1$ ,  $W_2$ , and  $W_3$  go through  $s$ - $b$ - $t$ ; while  $W_4$ , and  $W_5$  go through  $b$ - $t$ .

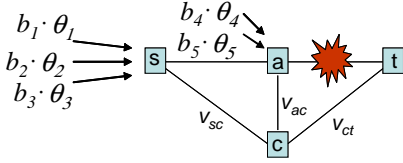


Fig. 2. An example showing the bottleneck, and the density of non-restorable bandwidth for a backup LSP: A general case.

When the five working LSPs are interrupted at link  $b$ - $t$ ,  $P_1$ ,  $P_2$ , and  $P_3$  going through  $s$ - $c$ - $t$ ; and  $P_4$ , and  $P_5$  going through  $b$ - $c$ - $t$  are activated with restoration attempt  $\theta_k$  where  $k = 1, 2, 3, 4$ , and  $5$ , respectively. The bottleneck non-restorability density, and the bottleneck link for  $P_1$ ,  $P_2$ , and  $P_3$  are determined not only by their restoration attempts, and the spare capacity on each link along  $s$ - $c$ - $t$ , but also by the bottlenecks of the other backup LSPs traversing through any link of  $s$ - $c$ - $t$ . Let the bottleneck link for  $P_4$ , and  $P_5$  be on link  $b$ - $c$  with a bottleneck non-restorability density  $Y_{4,bt} = Y_{5,bt}$ ; and the bottleneck link for  $P_1$ ,  $P_2$ , and  $P_3$  be on  $s$ - $c$  with a bottleneck non-restorability density  $q_{1,bt} = q_{2,bt} = q_{3,bt}$ . The non-restorability density on link  $c$ - $t$  thus can be expressed as:

$$y_{ct} = \frac{\sum_{k=1}^3 b_k \theta_k \cdot (1 - q_{1,bt}) + \sum_{k=4}^5 b_k \cdot \theta_k \cdot (1 - q_{4,bt}) - v_{ct}}{\sum_{k=1}^3 b_k \cdot \theta_k \cdot (1 - q_{1,bt}) + \sum_{k=4}^5 b_k \cdot \theta_k \cdot (1 - q_{1,bt})} = \frac{s_{ct,bt} - v_{ct}}{s_{ct,bt}} \quad (4)$$

where  $s_{ct,bt} = \sum_{k=1}^3 b_k \cdot \theta_k \cdot (1 - q_{1,bt}) + \sum_{k=4}^5 b_k \cdot \theta_k \cdot (1 - q_{4,bt})$  is the amount of effective restoration on link  $c$ - $t$ .

Fig. 3 exemplifies the definition of the bottleneck link along  $P$ . Obviously, link 4 is the bottleneck link for  $P$  with the largest value of  $y_{j,r} = (s_{j,r} - v_j)/s_{j,r}$ , where  $v_j$  is the amount of spare capacity  $\forall j \in \{1, 2, 3, 4\}$ . In other words, only a

fraction of the restoration attempt issued by the source node of the connection along  $P$  can go through, and is bottlenecked at link 4.

To generalize, the effective restorability of a working LSP is specific to the corresponding backup LSP  $P$  and the failure event  $r$ , and is the exact amount of restorable bandwidth for the working LSP. However, to determine effective restorability  $P$  on failure event  $r$ , we need to derive the effective restorable bandwidth of each link along  $P$ , which is in turn determined by the effective restorability of all the other backup LSPs interrupted by  $r$ , which traverse through any one of the links of  $P$ . In the next section, we will develop an approach to deriving the effective restoration attempt for each backup LSP when a failure event occurs.

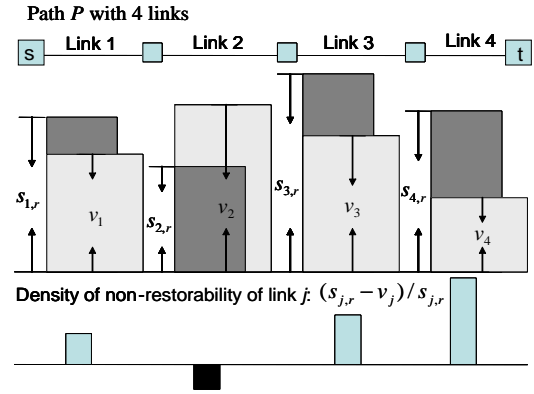


Fig. 3. An illustration of the bottleneck link along  $P$ .

### B. Random Dropping Mechanism in the MPLS Layer

The random dropping mechanism serves as a key in enabling the contention-based partial restoration in the MPLS core. The implementation of the scheme at each node taken by a backup LSP must be fast enough in order not to increase the restoration time. Meanwhile, the scheme should be operated in the MPLS control plane, and should not go through an even higher layer.

One of the most distinguished characteristics in the MPLS control plane is on the hierarchy of LSPs, where multiple nested labels corresponding to a group of LSPs bundled together can be popped, or pushed as the traffic is disassembled, or groomed, respectively. To perform random dropping, the node must be able to recognize the bandwidth of each individual flow nested in the backup LSPs traversing through the node. This can be done by maintaining a *dropping table* at each node, which bears the information of minimum required bandwidth corresponding to each nested label of the backup LSPs. This table is updated whenever a backup LSP is established through the node. Thus, when a dropping ratio on a backup LSP at a node is given, the node can randomly drop some of the packets of the LSP by identifying the nested labels defined in the dropping table such that the remaining nested LSPs can go through the node with their minimum required bandwidth satisfied.

Some overheads will certainly be introduced by employing the abovementioned random dropping mechanism in both the

protection, and restoration phases. In the protection phase, the setup of the backup LSPs involves in the effort of maintaining the dropping table in the source, and each intermediate node. In the restoration phase, the overhead is due to the extra processing time on each packet for the label check. To reduce the processing time, high-efficiency packet classifiers could be equipped.

#### IV. AVAILABILITY MODEL FOR SBPP CONNECTIONS

Based on the study in [18], the stationary probability of each failure pattern can be derived. Our goal is to evaluate the E2E unavailability of connection  $d$  by enumerating all failure patterns that affect the connection.

By Eq. (2), the E2E unavailability of  $d$  can be expressed as  $u^d = u_{non}^d + u_{\bar{w}p}^d + u_{\bar{w}p}^d$ . The term  $u_{\bar{w}p}^d$  is contributed by all the dual-failure events defined in  $R_{\bar{w}p}^d$  each interrupting both  $W_d$ , and  $P^d$ . Because the failure patterns totally block the connection, we can evaluate  $u_{\bar{w}p}^d$  by simply summing up the stationary probabilities of all the states  $r \in R_{\bar{w}p}^d$ :

$$u_{\bar{w}p}^d = \sum_{r \in R_{\bar{w}p}^d} \pi_r \quad (5)$$

The term  $u_{\bar{w}p}^d$ , on the other hand, is contributed by all failure events belonging to  $R_{\bar{w}p}^d$  that cause interruption on  $W^d$ , while  $P^d$  is not affected. With such failure patterns, contention among the restoration attempts of  $P^d$ , and the restoral flows of the other backup LSPs may occur at any link along  $P^d$ .

Let the failure pattern  $r$  be represented by a SRG duplet  $\{m, n\} \in R_{\bar{w}p}^d$ , where  $m$  is the earlier failed SRG, and  $n$  the latter one. The evaluation of the term  $u_{\bar{w}p}^d$  is divided into the following two parts. The first is the case where  $W^d$  traverses through SRG  $m$  which fails earlier; while the second is the case where  $W^d$  traverses through SRG  $n$  which fails latter. The former is equivalent to the case where a single failure on  $m$  occurs because the restoral flows due to the second failure can only compete for the residual of the spare capacity of each link. In the latter case, we need to consider the residual capacity along each link after the restoration of the first failure.

**Case 1:**  $\{m, n\} \in R_{\bar{w}p}^d$ , and  $W^d$  traverses through SRG  $m$  which failed earlier

In this case, because the restoral flows due to the latter, occurring failure will not contend with the restoral flows due to the earlier failure, the analysis only needs to consider the first failed SRG  $m$ . Contention among restoral flows may happen, which causes random dropping at link  $j$  if the pre-allocated spare capacity  $v_j$  is not sufficient to support the intended restoral flows.

Let  $d$  denote a tagged connection in the availability evaluation, and the set of backup LSPs that traverse through any link taken by  $P^d$  that will be activated due to the failure on  $\{m\}$  be denoted as  $\underline{P}_{\{m\}}^d$ . In other words,  $P_k \in \underline{P}_{\{m\}}^d$  if and only if the backup LSP of connection  $k$  activated after failure on  $\{m\}$  traverses through any common link with  $P^d$ . The effective restorability of  $d$  is determined by the bottleneck link<sup>1</sup> capacity along  $P^d$ , which, in turn, is determined by

<sup>1</sup>The bottleneck link of connection  $k$  is the link with the bottleneck (i.e. maximum of) non-restorability density (i.e.  $y_{j,\{m\}}$ ) along  $P^k$ .

the effective restorability of each backup LSP  $P_k \in \underline{P}_{\{m\}}^d$ . Therefore, in order to solve the effective restorability of  $P^d$ , the effective restorability of  $P_k \in \underline{P}_{\{m\}}^d$  have to be jointly solved because they are correlated with each other.

Formally, the bottleneck non-restorability density of  $P^k$ , where  $P^k \in \underline{P}_{\{m\}}^d$ , can be expressed as:

$$q_{\{m\}}^k = \max_{j \in P_k} \{0, y_{j,\{m\}}\}, \forall m \in W^d, \forall k | P_k \in \underline{P}_{\{m\}}^d \quad (6)$$

where  $y_{j,\{m\}} = (s_{j,\{m\}} - v_j) / s_{j,\{m\}}$ , and the term  $m \in W^d$  represents that SRG  $m$  is involved in  $W^d$  (i.e.  $m \cap W^d \neq \emptyset$ ).

Because  $s_{j,\{m\}}$  is the summation of the restoral flows along link  $j$  due to the failure of SRG  $m$ , we have:

$$s_{j,\{m\}} = \sum_{\forall k | m \in W^d} b^k \cdot \theta^k \cdot (1 - q_{\{m\}}^k) \quad \forall j \in P_k \quad (7)$$

where the term  $b^k \cdot \theta^k \cdot (1 - q_{\{m\}}^k)$  is the effective restorability of backup LSP  $k$  on the single failure of  $m$ .

We can derive  $s_{j,\{m\}}$ , and  $q_{\{m\}}^k$  for  $\forall j \in P_k, \forall P_k \in \underline{P}_{\{m\}}^d$  by jointly solving Eq. (6) and (7) if  $v_j$  and  $\theta_k$  are constant. With  $s_{j,\{m\}}$ , and  $q_{\{m\}}^k$ , the effective non-restorability of the tagged connection  $d$  (denoted as  $q_{\{m\}}^d$ ) can be expressed by:

$$q_{\{m\}}^d = \max_{j \in P^d} \{0, y_{j,\{m\}}\} \quad (8)$$

Based on Eq. (8), the non-restorable bandwidth of connection  $d$  due to failure on  $m$  can be expressed as:

$$\dot{b}_{d,\{m\}} = (1 - \theta^d) \cdot b^d + q_{\{m\}}^d \cdot \theta^d \cdot b^d \quad (9)$$

In the RHS of Eq. (9), the term  $(1 - \theta^d) \cdot b^d$  is the non-restorable bandwidth of  $d$  due to the partial restoration attempt  $\theta^d$  launched by the source node of  $d$ . The second term  $q_{\{m\}}^d \cdot \theta^d \cdot b^d$  is the non-restorable bandwidth due to the contention of the restoration attempts issued by all the interrupted connections, which is determined by the bottleneck link along  $P^d$ . Based on Eq. (9), the unavailability of the spare capacity along  $P^d$  seen by connection  $d$  when  $W_d$  traverses through  $m$ , and failure pattern  $\{m, n\}$  occurs, is:

$$u_{\{m\}}^d = \dot{b}_{d,\{m\}} / b^d = (1 - \theta^d) + q_{\{m\}}^d \cdot \theta^d \quad (10)$$

The overall unavailability can be derived by averaging  $u_{d,m}$ :

$$\begin{aligned} \dot{u}_{\bar{w}p}^d &= \sum_{\forall \{m,n\} \in \dot{R}_{\bar{w}p}^d} u_{\{m,n\}}^d \cdot \pi_{\{m,n\}} = \\ &= \sum_{\forall m \in W^d} u_{\{m\}}^d \cdot \sum_{\forall n | \{m,n\} \in \dot{R}_{\bar{w}p}^d} \pi_{\{m,n\}} \quad (11) \end{aligned}$$

Note that after solving the joint Eqs. (6) and (7) jointly, we can use Eqs. (8) – (10) to derive the E2E unavailability for each connection with a backup LSP belonging to  $\underline{P}_{\{m\}}^d$  such that Eq. (11) can be calculated. Also note that in Eq. (11),  $n$  could be null to represent the cases where a single failure on  $m$  occurs. In addition, the failure pattern  $\{m, n\}$  where  $W^d$  traverses through both  $m$  and  $n$  are also included in this case.

**Case 2:**  $\{n, m\} \in R_{\bar{w}p}$ , and  $W^d$  traverses through SRG  $m$ , which is the latter failed one

In this case, an approach similar to that in Case 1 is developed, with the only difference in that the restoration is performed on the residual capacity of each link after the first failure. With the first failure on SRG  $n$ , the term  $s_{j,\{n\}}$ , and the bottleneck link of  $P_k$  (i.e.,  $q_{\{n\}}^k$ ) can be solved using the method introduced in Case 1. At this moment, the amount of spare capacity along each link could become less to accommodate the restoral flows due to the first failure. Let the updated spare capacity along each link be denoted as  $v'_j = v_j - s_{j,\{n\}}$ . When the second failure on  $m$  occurs before the repair of  $n$ , we are interested in deriving  $s_{j,\{n,m\}}$ , and the bottleneck non-restorability density of each activated backup LSP (i.e.,  $q_{\{n,m\}}^d$ ).

Let the set of failure patterns of  $R_{\bar{w}p}^d$  where the second failure hits  $W^d$  be denoted by  $\ddot{R}_{\bar{w}p}^d$ . Let  $\underline{P}_{\{n,m\}}^d$  denote the set of backup LSPs that traverse through any link taken by  $P^d$ , and will be activated due to the dual-failure event  $\{n,m\}$ . Thus, the set of backup LSPs that will be activated in the dual-failure event  $\{n,m\}$ , and traverse through any link taken by  $P^d$  with their working LSPs traversing through  $m$ . is denoted as  $\ddot{P}_{\{n,m\}}^d = \underline{P}_{\{n,m\}}^d - \underline{P}_{\{n\}}^d$ . To derive  $s_{j,\{n,m\}}$ , a similar approach to that in Case 1 can be developed by solving the following two equations:

$$q_{\{n,m\}}^k = \max_{\forall j \in P_k} \{0, y_{j,\{n,m\}}\} \quad \forall \{n,m\} \in \ddot{R}_{\bar{w}p}^d, \forall k | P_k \in \ddot{P}_{\{n,m\}}^d \quad (12)$$

where  $y_{j,\{n,m\}} = (s_{j,\{n,m\}} - v'_j) / s_{j,\{n,m\}}$

$$s_{j,\{n,m\}} = \sum_{\forall k | \{n,m\} \in R_{\bar{w}p}^k} b^k \cdot \theta^k \cdot (1 - q_{\{n,m\}}^k) \quad \forall j \in P_k \quad (13)$$

Note that the term  $b^k \cdot \theta^k \cdot (1 - q_{\{n,m\}}^k)$  is the effective restorability of backup LSP  $k$  on the failure event of  $\{n,m\}$ . Similarly, we can derive  $s_{j,\{n,m\}}$ , and  $q_{\{n,m\}}^k$  by jointly solving Eqs. (12), and (13) if  $v'_j$ , and  $\theta^k$  are known. Therefore,  $q_{\{n,m\}}^d$  can be derived by:

$$q_{\{n,m\}}^d = \max_{j \in P^d} \{0, y_{j,\{n,m\}}\} \quad \forall \{n,m\} \in \ddot{R}_{\bar{w}p}^d \quad (14)$$

Thus, the non-restorable bandwidth of connection  $d$  can be expressed as:

$$\ddot{b}_{\{n,m\}}^d = (1 - \theta^d) \cdot b^d + q_{\{n,m\}}^d \cdot \theta^d \cdot b^d \quad (15)$$

The availability impairment on  $P^d$  seen by connection  $d$  when  $W^d$  traverses  $m$ , and failure pattern  $\{n,m\}$  occurs, is:

$$u_{\{n,m\}}^d = \ddot{b}_{\{n,m\}}^d / b^d = (1 - \theta^d) + q_{\{n,m\}}^d \cdot \theta^d \quad (16)$$

The overall unavailability of connection  $d$  can be derived by averaging  $u_{\{n,m\}}^d$  for all  $\{n,m\} \in R_{\bar{w}p}^d$ :

$$\ddot{u}_{\bar{w}p}^d = \sum_{\forall \{n,m\} \in \ddot{R}_{\bar{w}p}^d} u_{\{n,m\}}^d \cdot \pi_{\{n,m\}} \quad (17)$$

By considering Eq. (11), and Eq. (17), the availability impairment due to failure patterns  $\{n,m\} \in R_{\bar{w}p}^d$  on  $P^d$  can be simply expressed as:

$$u_{\bar{w}p}^d = \dot{u}_{\bar{w}p}^d + \ddot{u}_{\bar{w}p}^d \quad (18)$$

Therefore, Eq. (2) can be evaluated by combining Eq. (11) and Eq. (18).

## V. AVAILABILITY-AWARE SPARE CAPACITY RECONFIGURATION ARCHITECTURE

Based on the developed E2E unavailability model for a SBPP connection, a LP is formulated for performing spare capacity reconfiguration with availability guarantee for each connection. The following two sections describe our dynamic bandwidth provisioning architecture, and the proposed LP for spare capacity reconfiguration.

### A. Dynamic Bandwidth Provisioning Architecture

At the arrival of a connection request, it is firstly allocated by using an arbitrary routing approach. In this study, two previously reported dynamic survivable routing algorithms are considered, including Successive Survivable Routing (SSR) [?], and the asymmetrically weighted diverse routing algorithm [25]. The newly arrived connection is temporarily protected with SSR for near 100% restorability under any single failure before the spare capacity reconfiguration is done on the backup routes. The LP is solved once per a number for each connection arrival & departure events to determine the required restoration attempt of each connection (i.e., the parameter of  $\theta^d$ ,  $\forall d \in D$ ), and reconfigure the spare capacity along each link (i.e.,  $v_j$ ,  $\forall j \in E$ ) in order to meet the availability constraint for each connection. If a reconfiguration process cannot be completed before the arrival of the next network event, the reconfiguration process will be dropped, while another reconfiguration process on the new network state will be initiated.

To enable the network-wide reconfiguration on spare capacity, a centralized computation process is adopted, where the following information/link-states are required from each connection  $d$ :

- The unavailability constraint for connection  $\forall d \in D$  is denoted as  $U_{sla}^d$ .
- The working LSP of connection  $\forall d \in D$  (denoted as  $W^d$ ).
- The backup LSP of connection  $\forall d \in D$  (denoted as  $P^d$ , which is SRG-disjoint with  $W^d$ ).

### B. Linear Program (LP) for Spare Capacity Reconfiguration

To increase the E2E availability requirement for each connection, we can either increase the spare capacity along some links, or increase the restoration attempt of some connections, or do both of the above. In the proposed availability model, and LP formulation (which is introduced in the following section), both of the approaches in improving the E2E availability will be exercised in the proposed method.

A LP is formulated to derive  $\theta^d$ , and  $v_j$  for  $d \in D$ , and  $j \in E$  such that the availability requirement of each connection is met, and the minimal amount of spare capacity is allocated along each link. Because the original formulation is not linear, an iterative approach is devised, in which the LP is solved iteratively by assigning  $\theta^d = 1$  in the first iteration. The target function of the LP is:

$$\text{Minimize : } \sum_{\forall j \in E} v_j$$

A set of working variables  $y_{j,r}$  is defined for all link  $j$ , and failure pattern  $r \in R_{\bar{w}p}^d$  with the following constraints:

$$1 \geq y_{j,r} \geq 0 \quad \forall j \in E, \forall r \in R_{\bar{w}p}^d \quad (19)$$

$$y_{j,r} \geq 1 - v_j / s_{j,r} \quad \forall j \in E, \forall r \in R_{\bar{w}p}^d \quad (20)$$

where

$$s_{j,r} = \sum_{\forall k|r \in R_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k \cdot (1 - q_r^k) \quad (21)$$

$$1 \geq q_r^d \geq y_{j,r} \quad \forall j \in P^d, \forall d \in D, \forall r \in R_{\bar{w}p}^d \quad (22)$$

To make Eqs. (20), (21) linear, we set:

$$s'_{j,\{m\}} = \sum_{\forall k|\{m\} \in \dot{R}_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k.$$

It is clear that the following two equations hold:

$$v_j / s_{j,\{m\}} \leq 1 \quad (23)$$

$$s'_{j,\{m\}} - s_{j,\{m\}} = \sum_{\forall k|\{m\} \in \dot{R}_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k \cdot q_{\{m\}}^k \geq 0 \quad (24)$$

Thus, instead of Eq. (20) for any single failure event, we have:

$$\begin{aligned} y_{j,\{m\}} &\geq 1 - \frac{v_j}{s_{j,\{m\}}} \geq 1 - \frac{v_j + (s'_{j,\{m\}} - s_{j,\{m\}})}{s_{j,\{m\}} + (s'_{j,\{m\}} - s_{j,\{m\}})} \\ &= 1 - \frac{v_j + \sum_{\forall k|\{m\} \in \dot{R}_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k \cdot q_{\{m\}}^k}{\sum_{\forall k|\{m\} \in \dot{R}_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k} \end{aligned} \quad (25)$$

With a similar approach, we set

$$s'_{j,\{n,m\}} = \sum_{\forall k|k \in \dot{R}_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k, \quad (26)$$

and

$$v'_j = v_j - s_{j,\{n\}}. \quad (27)$$

Thus, for any dual failure event, we have:

$$\begin{aligned} y_{j,\{n,m\}} &\geq 1 - \frac{v'_j + \sum_{\forall k|k \in \dot{R}_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k \cdot q_{\{n,m\}}^k}{\sum_{\forall k|k \in \dot{R}_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k} \\ &\quad \forall j \in E, \forall \{n,m\} \in \dot{R}_{\bar{w}p}^k \end{aligned} \quad (28)$$

where

$$s_{j,\{n,m\}} = \sum_{\forall k|k \in \dot{R}_{\bar{w}p}^k, j \in P_k} b^k \cdot \theta^k \cdot (1 - q_{\{n,m\}}^k) \quad (29)$$

Based on Eqs. (2), (5) (11), and (18), the E2E availability of each connection must meet the given availability requirement:

$$\begin{aligned} u^d &= u_{non}^d + u_{\bar{w}p}^d + \dot{u}_{\bar{w}p}^d + \ddot{u}_{\bar{w}p}^d = \sum_{\forall r \in R_{non}^d} \pi_r + \sum_{\forall r \in R_{\bar{w}p}^d} \pi_r \\ &+ \sum_{\forall m \in W^d} \left( (1 - \theta^d) + q_{\{m\}}^d \right) \cdot \sum_{\forall n|\{m,n\} \in \dot{R}_{\bar{w}p}^d} \pi_{\{m,n\}} + \\ &\sum_{\forall \{n,m\} \in \dot{R}_{\bar{w}p}^d} \left( (1 - \theta^d) + q_{\{n,m\}}^d \cdot \theta^d \right) \cdot \pi_{\{n,m\}} \quad \forall d \in D \end{aligned} \quad (30)$$

$$u^d \leq U_{sla}^d \quad \forall d \in D \quad (31)$$

Note that the approximation made in Eq. (25) and & (28) is to remove any variable in the denominator such that the iteration can be solved linearly. After solving an iteration of the LP, a new suite of  $\theta^d$  for all  $d \in D$  in the next iteration is determined by setting:

$$\theta^d \leftarrow \text{avg}_{\forall \{m\} \in R_{\bar{w}p}} q_{\{m\}}^d \cdot \theta^d \quad \forall d \in D \quad (32)$$

In Eq. (32),  $\{m\}$  stands for the case where only SRG  $m$  is in the failure state.

### C. A General Framework of Survivable Routing with Partial Restoration

We claim that the proposed model serves as a general case for a number of previously reported designs. In our model, the effective restorability of connection  $d$  is determined by Eq. (21), and Eq. (28) when the failure event is on a single SRG, and dual SRGs, respectively, in which  $q_{\{n,m\}}^d$ , and  $\theta^d$  are two variable parameters ranged in  $[0,1]$  that can be manipulated in the LP formulation.

A special case is seen when  $\theta^d = 1$ , and  $q_{\{n,m\}}^d$  is binary (i.e., either 1 or 0) for a connection. In this case, the connection is treated as indivisible with 100% restoration attempt (e.g., a lightpath in WDM networks). In this case, the formulation turns out to become an Integer Linear Program (ILP) that will yield a solution for each connection either 100% restorable or non-restorable in the presence of each failure pattern. In other words, for the connection, the occurrence of some failure patterns is restorable, and the others are not for the connection, where the E2E unavailability can be gained by summing up all the stationary probabilities of those non-restorable failure patterns.

In the case where  $q_{\{n,m\}}^d = 0$ , and  $\theta^d = 1$  for all possible failure patterns  $\{n,m\}$ , it simply degrades to the case where every connection achieves 100% restorability for any single failure. In such a circumstance, no random dropping of any nested LSP caused by contention could happen.

Another special case is when  $q_{\{n,m\}}^d = 0$ , and  $0 \leq \theta^d \leq 1$ , where no contention could happen, and each restoration flow is throttled only by the source node based on the restoration attempt (or the parameter  $\theta^d$ ). In the case, the restoration of each connection becomes failure dependent, and the solution of the LP will be the same as that without the constraint of  $q_{\{n,m\}}^d = 0$  by simply assigning  $\theta_{\{n,m\}}^d = \theta^d \cdot (1 - q_{\{n,m\}}^d)$ . It



is failure dependent restoration because  $\theta_{\{n,m\}}^d$  is specific to not only the working and backup LSPs of each connection, but also which failure pattern occurs. Because the source node needs to localize the failures before it knows how much restoration flow it will launch, the implementation could be subject to more complexity, and is unrealistic.

With the proposed approach, the E2E availability of connection  $d$  can be guaranteed by inputting  $U_{sla}^d$ , and the stationary probability of each failure pattern. The result by solving the iterative LP includes  $\theta^d$ , which is the restoration attempt of connection  $d$ , and  $v_j$ , the spare capacity allocated along link  $l$ .

## VI. SIMULATION

Simulation is conducted to (1) validate the proposed availability Model, and (2) verify the effectiveness of the spare capacity reconfiguration strategy in terms of the spare capacity saving ratio.

Two network topologies are adopted in the simulation: a 16-node Pan-European network, and a 17-node German reference network, as shown in Fig. 4.

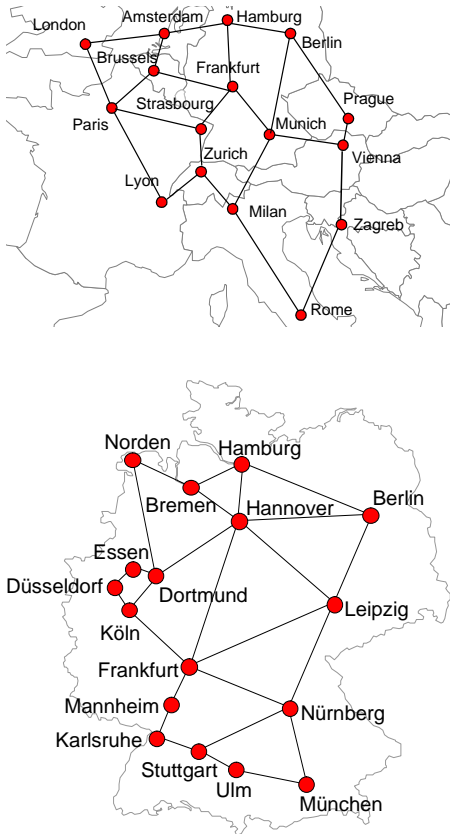


Fig. 4. (a) the Pan-European network. (b) The German network.

The Mean Time To Repair (MTTR), and Mean Time To Failure (MTTF) values are assigned to each link and node of the networks [17]. The SRG of the networks are the links, and nodes. Each network link is described by two values: the ‘link interface’, which does not depend on the length of the fiber, and the ‘link/km’, where the MTTF is in proportion to

the length of the fiber (in km). The MTTR is defined by two values: MTTR of the fiber, and MTTR of the link interface, which do not depend on the length of the fiber in any case. The following table Table II shows the values used in the simulation.:

TABLE II  
THE MTTF, AND MTTR VALUES IN THE SIMULATION.

MTTF			MTTR		
node	link/km	interface	node	fiber	interface
$MTTF_{node}$	$MTTF_{link/km}$	$MTTF_{intfa}$	$MTTR_{node}$	$MTTR_{fiber}$	$MTTR_{intfa}$
20000	2380000	57000	1.4	11.4	6.0

Finally, a residual MTTF (and the MTTR) values for a link can be modeled as a serial system [17].

In the simulation, the convergence criterion is defined as that  $\theta^d$  for all  $d \in D$  was changed by less than 5% in the previous iteration, or the number of iteration is larger than 5. We made it completed 5 iterations in order to speed up the solution. With such a policy, we have seen in the simulation (which will be described in next section) that most of the connections can reach a very good solution with just a few iterations, especially when the availability requirement for each connection is high, where  $\theta^d$  for each connection is close to 1.

### A. Availability Model Validation

A discrete-event continuous time simulation was conducted on the Pan-European, and German networks. To simulate a more realistic situation, each link and a or node fails following a Poisson process with an arrival rate  $1/MTTF_{link}$ , and  $1/MTTF_{node}$ , respectively, where no restriction on the number of simultaneous failures has been addressed. Failure holding time for each link, and node follows a negative exponential distribution with the rate of  $1/MTTR_{link}$ , and  $1/MTTR_{node}$  respectively.

To validate the proposed availability models, our approach is to evaluate the E2E availability of a group of connections in the network topology in the presence of the arrival and departure of random failure events. For this purpose, firstly a set of 135 SBPP connections were randomly allocated among all the node pairs using the Successive Survivable Routing (SSR) algorithm [22]. After allocating the connections, the spare capacity is reconfigured by solving the proposed LP formulation, by which the theoretical E2E availability for each connection using the developed model can thus be derived. On the other hand, the simulated value of the E2E availability of connection  $d$  is derived simply using the ration:  $a_{simulated}^d = 1 - \frac{\sum_{r \in R} t_r^d}{t_{total}}$ , where  $t_{total}$  is the total simulated time,  $R$  is the set of failure events simulated in the experiment, and  $t_r^d$  is the down time of connection  $d$  due to failure event  $r$ .

Fig. 5 shows the simulated, and theoretical E2E availabilities for each connection when the restoration attempt of all the connections is virtually 100%. It is observed that the theoretically derived E2E availability is no less than the corresponding simulated value for most of the connections due to the fact that no more than two links could fail simultaneously in the theoretical availability model, whereas the simulation

model allows any number of simultaneous failures to occur. Thus, a slightly higher E2E availability could be derived in the theoretical model than that of the realistic one.

We also investigated the impacts by having different values of MTTF for each link and node on the resultant E2E availability of each connection. Fig. 6 shows the difference between the simulated, and theoretical cases by scaling the MTTF from 0.4 to 2.8 with 100% restoration attempt. Both the theoretical, and the simulated values are averaged over all the connections. We can see that the difference is small when MTTF is low, and increases when MTTF is increased. The smallest difference is 0.001597% when the scaling factor is 0.4. When the theoretical connection availability drops to 99.6362% with the scaling factor of 2.8, the difference is only 0.016514%, which indicates a high accuracy in our E2E availability model. Two different SBPP routing methods based

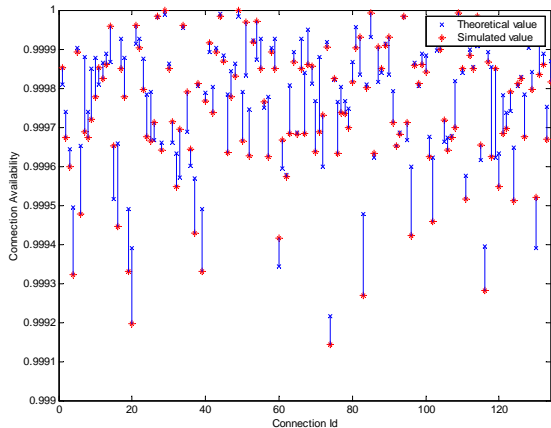


Fig. 5. The theoretical vs. simulated E2E availabilities with 100% restoration attempt for each connection.

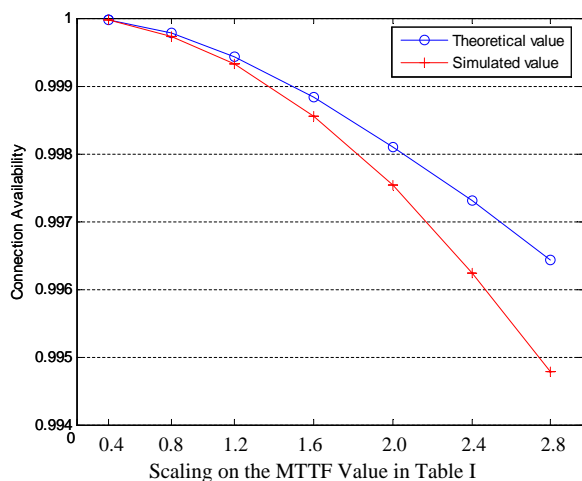


Fig. 6. Difference between the theoretical, and simulated connection availabilities for different scaling on MTTF values with 100% restoration attempt of each connection.

on the Active Path First approach (APF) are considered. The first method is denoted as “2D”, where the working LSP is simply the shortest path, while the backup path is derived by

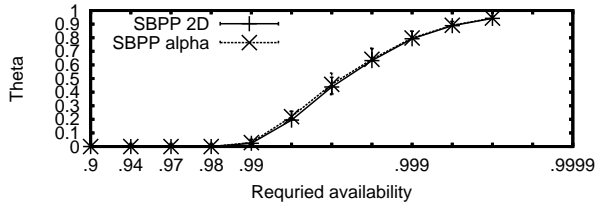
using the shortest path first algorithm again on the residual graph with the working path excluded [25]. The cost function in the survivable routing process is simply the number of hops. The second method is denoted as “alpha”, which is a little bit different from the first method in terms of the way in deriving the working LSP. With “alpha”, a disjoint path-pair is calculated in the same way as that in “2D” with the cost function as  $2 \cdot c_w + c_p$ , where  $c_w$ , and  $c_p$  are the cost of the working, and backup paths, respectively. The method is also referred to as asymmetrically weighted diverse routing with  $\alpha = 2$  [16], and has been proved to be very efficient to avoid the so called “trap topology”. In most of the cases in the simulation, the two SBPP methods frequently yield very different results in the presence of the same link state.

The minimum amount of spare capacity by the conventional SBPP for achieving 100% restorability under any single failure is taken as a benchmark because the study is rather interested in the improvements that can be made by the proposed reconfiguration mechanism. Also, we are interested in observing the performance impacts due to the following two important design factors: the diverse routing algorithm, and the spare capacity sharing policy. The former one concerns the route selection for both working, and backup paths, while the latter determines the way of choosing  $\theta$ , and the amount of spare capacity for each connection (e.g.,  $v_j$ ). Each connection is equipped with a specific E2E availability requirement in the range of [0.9, 0.9999].

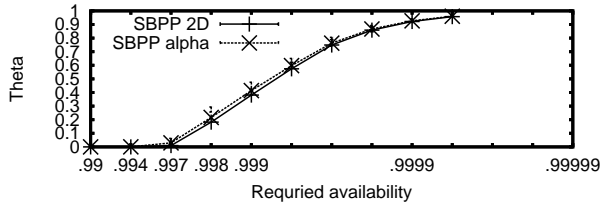
Dynamically arriving connection requests are generated with Poisson arrival, and departure following an exponential distribution. Because the network capacity saving by using the proposed spare capacity reconfiguration is of interest, we adopted the infinite link capacity assumption, where no blocking of any connection request could occur. The proposed LP formulation is solved on the current link state for every 10 arrival and departure events to evaluate the possible savings on the total amount of consumed spare capacity. In the experiment, the network state is not updated according to the LP solution, while the capacity savings of each spare capacity reconfiguration is kept.

Fig. 7 shows the average  $\theta$  derived in each reconfiguration process with a specific availability constraint for each connection. We can easily observe that the restoration attempt of each connection (i.e.  $\theta_d$ ) can be manipulated in the proposed spare capacity reconfiguration process to meet the E2E availability constraint of each connection. Note that the conventional SBPP scheme always has  $\theta_d = 1$ , and 100% restorability for any single failure. An error bar corresponding to each data in the experiments shows the minimum and maximum values of  $\theta$ , which is very small and hardly recognized in the graphs. Fig. 7 also shows that no significant difference is made by using the two survivable routing algorithms even if “2D” yields much higher blocking probability due to the topological traps. Note that in the German reference network, “four-nines” availability can hardly be achieved with a single backup LSP according to the given MTTR, and MTTF data. That means that we may need to add one more backup LSP since even 1+1 protection may not be sufficient to achieve four-nine availability.

It is also interesting to observe in Fig. 7 that with the



(a) Results of the Pan-European network.



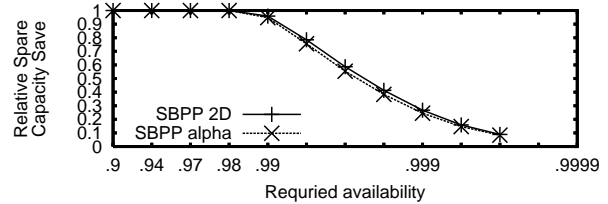
(b) Results of the German network.

Fig. 7. Restoration attempt (Theta) vs. different availability constraints defined for each connection.

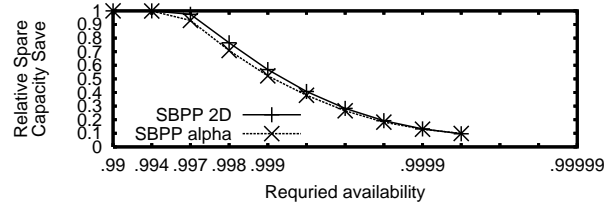
restoration attempt of each connection always to be 1 may not yield better overall availability than that by our scheme with the restoration attempt less than 1. In the Pan-European network, the conventional SBPP scheme yields an average availability of 0.9994. After the LP optimization with the required availability set as 0.99968, we get a feasible solution with the average restoration attempt value less than 1. This is due to the fact that any connection that has a larger-than required restoration attempt would impair the availability of the other connections sharing the common spare capacity. This observation also demonstrates the importance of further exploring the design dimension of restoration attempts in developing spare capacity reconfiguration algorithms.

Fig. 8 shows the average saving ratio on spare capacity, which is defined as the average amount of spare capacity saved from that by the conventional SBPP in a single reconfiguration process normalized by the result of the conventional SBPP. It is observed that, when the availability requirement is loose (i.e.,  $a_{sla}^d = 1 - U_{sla}^d = 0.9$  to  $0.99, \forall d \in D$ ), the average spare capacity saving ratio is close to 100% since because a backup LSP with a small amount of spare capacity could be enough to meet the availability constraint for some connections. On the other hand, the average saving ratio is approaching to 0 when the availability constraint is going higher, where the amount of spare capacity taken by each connection becomes similar to the case of the conventional SBPP.

Fig. 9 illustrates the total consumed spare capacity along each link versus the average value of the restoration attempt of each connection. It is observed that the consumed spare capacity along each link is getting increased when the restoration attempt is increased, which meets our intuition. It is notable that the “2D” design consumed significantly lower total spare capacity in the German network because much

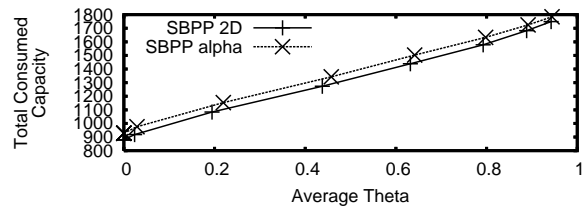


(a) Results of the Pan-European network.

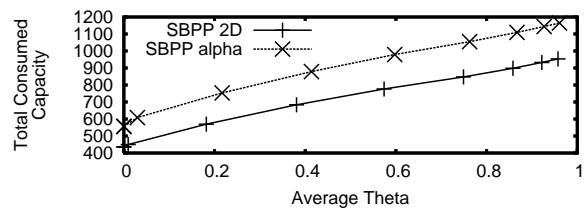


(b) Results of the German network.

Fig. 8. The spare capacity saving ratio vs. different availability constraints.



(a) Results of the Pan-European network.



(b) Results of the German network.

Fig. 9. The total consumed spare capacity versus the average value of theta with respect to different availability constraints for all the connections.

more connection requests were blocked due to the smaller average nodal degree, which causes topological traps.

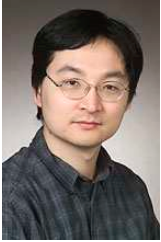
In the simulation, the LP was solved using by LP Solver, where a high-end Dell workstation with dual Xeon 2.8GHz processors, and 1GB memory was adopted. In the experiments on both of the network topologies, the computation time for performing a single spare capacity reconfiguration process is generally a few minutes. Because solving the LP takes polynomial computation time in each iteration, the proposed scheme is considered to be computationally efficient.

## VII. CONCLUSIONS

This paper has introduced a novel survivable routing architecture with availability constraints, and is expected to create a new design paradigm in availability-aware survivable routing problems. An E2E availability model on SBPP connections considering up to two simultaneous failures was proposed, which is characterized by a novel contention-based partial restoration framework with a random dropping mechanism in the MPLS layer. Based on the developed model, a Linear Program was formulated to perform spare capacity allocation in order to determine the amount of redundancy along each link, and the restoration attempt of each connection under the given E2E availability requirement for each SBPP connection. Simulation was conducted to validate the availability model, and verify the proposed spare capacity allocation scheme with two dynamic survivable routing algorithms. From the simulation results, we have seen merits in the proposed scheme that can significantly improve the conventional SBPP designed for 100% restorability under any single failure in terms of capacity efficiency. We found that the restoration attempt  $\theta$  for each connection serves as a key parameter that can determine the E2E availability of each connection, and the resultant spare capacity saving. Also, due to the pre-calculated stationary probability for each failure pattern, the computation could be efficient and scalable.

## REFERENCES

- [1] D. Papadimitriou and E. Mannie, "Analysis of generalized mpls-based recovery mechanisms," Internet Draft, Sep. 2003.
- [2] I. Hussain, "Understanding High Availability of IP and MPLS Networks," white paper, *Ciscopress*, Jan. 2005.
- [3] P.-H. Ho, J. Tapolcai, and H. T. Mouftah, "On optimal diverse routing for shared protection in mesh WDM networks," *IEEE Transactions on Reliability*, vol. 53, no. 6, pp. 2216–2225, Jun. 2004.
- [4] J. Doucette, M. Clouqueur, and W. D. Grover, "On the availability and capacity requirements of shared backup path-protected mesh networks," *Optical Networks Magazine, Special Issue on Engineering the Next Generation Optical Internet*, vol. 4, no. 6, pp. 29–44, Nov./Dec. 2003.
- [5] M. T. Frederick, P. Datta, and A. K. Somani, "Evaluating dual-failure restorability in mesh-restorable wdm optical networks," in *Proc. IEEE ICCCN*, 2004, pp. 309–314.
- [6] A. Kodian and W. Grover, "Multiple-quality of protection classes including dual-failure survivable services in p-cycle networks," in *Broadband Optical Networking Symposium (BroadNets)*, vol. 1, Boston, MA, Oct. 2005, pp. 231–240.
- [7] P.-H. Ho, J. Tapolcai, and T. Cinkler, "Segment shared protection in mesh communication networks with bandwidth guaranteed tunnels," *IEEE/ACM Transactions on Networking*, vol. 12, no. 6, pp. 1105–1118, December 2004.
- [8] G. Shen and W. D. Grover, "Extending the p-cycle concept to path-segment protection for span and node failure recovery," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 21, no. 8, pp. 1306–1319, Oct. 2003.
- [9] J. Zhang, H. Zang, and B. Mukherjee, "A new provisioning framework to provide availability-guaranteed service in wdm mesh networks," in *Proc. IEEE ICC*, Anchorage, Alaska, May 2003, pp. 1484–1488.
- [10] D. Arci, D. Petecchi, G. Maier, A. Pattavina, and M. Tornatore, "Availability models for protection techniques in wdm networks," in *Proc. DRCN*, Banff, CA, Oct. 2003, pp. 158–166.
- [11] M. Clouqueur and W. D. Grover, "Availability analysis of span-restorable mesh networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 20, no. 4, pp. 810–821, May 2002.
- [12] Y. Huang, W. Wen, J. Heritage, and B. Mukherjee, "A generalized protection framework using a new link state availability model for reliable optical networks," *IEEE/OSA J. Lightwave Technol.*, vol. 22, no. 11, pp. 2536–2547, Nov. 2004.
- [13] D. A. Schupke, W. D. Grover, and M. Clouqueur, "Strategies for enhanced dual failure restorability with static or reconfigurable p-cycle networks," in *Proc. IEEE ICC*, Paris, France, Jun. 2004, pp. 1628–1633.
- [14] D. A. A. Mello, J. U. Pelegrini, R. P. Ribeiro, D. A. Schupke, and H. Waldman, "Dynamic provisioning of shared-backup path protected connections with guaranteed availability requirements," in *Proc. GOSP co-located with BroadNets*, Boston, MA, Oct. 2005, pp. 397–404.
- [15] W. Yao and B. Ramamurthy, "Survivable traffic grooming with differentiated end-to-end availability guarantees in wdm mesh networks," in *Proc. IEEE LANMAN*, San Francisco, California, 2004 2004, pp. 87–90.
- [16] W. D. Grover, *Mesh-based Survivable Networks: Options and Strategies for Optical, MPLS, SONET and ATM Networking*. Upper Saddle River, New Jersey: Prentice Hall PTR, 2003.
- [17] S. Verbrugge, D. Colle, P. Demeester, R. Huelsermann, and M. Jaeger, "General availability model for multilayer transport networks," in *Proc. DRCN 2005*, Lacco Ameno, Italy, Oct. 16-19, 2005.
- [18] D. A. A. Mello, D. A. Schupke, and H. Waldman, "A matrix-based analytical approach to connection unavailability estimation in shared backup protection," *IEEE Communications Letters*, vol. 9, no. 9, pp. 844–846, Sep. 2005.
- [19] O. Gerstel and G. Sasaki, "Quality of Protection (QoP): A Quantitative Unifying Paradigm to Protection Service Grades," *Opt. Netw. Mag.*, vol. 3, no. 3, pp. 40–49, May/June. 2002.
- [20] J. Fang, M. Sivakumar, A. Somani, and K. Sivalingam, "On partial protection on groomed optical wdm networks," in *Conference on Dependable Systems and Networks (DSN)*, Yokohama, Japan, 2005, pp. 228–237.
- [21] O. Gerstel and G. Sasaki, "A new protection paradigm for digital video distribution networks," in *Proc. IEEE ICC*, Istanbul, Turkey, Jun. 2006.
- [22] Y. Liu, D. Tipper, and P. Siripongwutikorn, "Approximating optimal spare capacity allocation by successive survivable routing," *IEEE/ACM Trans. Networking*, vol. 13, no. 1, pp. 198–211, Feb. 2005.
- [23] P. H. Ho and H. T. Mouftah, "Reconfiguration of spare capacity for mpls-based recovery for the internet backbone networks," *IEEE/ACM Transactions on Networking*, vol. 12, no. 1, pp. 73–85, Feb. 2004.
- [24] L. Shen, X. Yang, and B. Ramamurthy, "A load-balancing spare capacity reallocation approach in service-rich sonet metro mesh networks," in *Broadband Optical Networking Symposium (BroadNets)*, 2004, pp. 269–278.
- [25] J. Tapolcai and T. Cinkler, "On-line routing algorithms with shared protection in WDM networks," in *Proc. ONDM*, 2003, pp. 351–364.



**Professor Pin-Han Ho** received his Ph.D. from ECE Department at Queen's University in 2002. He joined the Electrical and Computer Engineering department at the U of Waterloo, Canada, as an assistant professor in the same year. His current research interests cover a wide range of topics in wired and wireless communication networks. He is the recipient of the Early Researcher Award in 2005; and the Best Paper Awards in SPECTS'02, ICC'05 Optical Networking Symposium, and ICC'07 Security and Wireless Communications symposium.



**János Tapolcai** received his M.Sc. ('00 in Technical Informatics), and Ph.D. ('05 in Computer Science) degrees in Technical Informatics from Budapest University of Technology and Economics (BME), Budapest, Hungary. Currently he is an associate professor at the High-Speed Networks Laboratory at the Department of Telecommunications and Media Informatics at BME. His research interests include applied mathematics, combinatorial optimization, linear programming, linear algebra, routing in circuit switched survivable networks, availability analysis,

grid networks, and distributed computing. He has been involved in a few related European and Canadian projects (IP NOBEL; NoE e-Photon/ONe; BUL). He is an author of over 30 scientific publications, and is the recipient of the Best Paper Award in ICC'06.



**Anwar Haque** received the B.S., M.Sc., and M.Math degrees in Computer Science from the North South University, Bangladesh, the University of Windsor, Canada, and the University of Waterloo, Canada in 1997, 2001, and 2007 respectively. He has published more than 25 referred technical papers. He has served as a reviewer, and TPC member for many international journals and conferences. His research interests include IP network QoS (Quality of Service), network design and planning with a focus on survivability, and network security. He is

currently employed at Bell Canada as Manager - Network Planning.