Optical Switching and Networking 10 (2013) 274–289

Contents lists available at SciVerse ScienceDirect

Optical Switching and Networking

journal homepage: www.elsevier.com/locate/osn

Availability of survivable Valiant load balancing (VLB) networks over optical networks

Wenda Ni [a,*], Changcheng Huang [a], Jing Wu [b], Michel Savoie [b]

[a] Department of Systems and Computer Engineering, Carleton University, Ottawa, Ontario K1S 5B6, Canada
[b] Communications Research Centre Canada, Ottawa, Ontario K2H 8S2, Canada

ARTICLE INFO

Available online 11 February 2013

Keywords:
IP over optical networks
Valiant load balancing (VLB) networks
Logical network embedding
Availability analysis
Correlated failures

ABSTRACT

Valiant load balancing (VLB) network has been proposed as a capacity-efficient solution to handle highly dynamic traffic in future backbone networks. In this paper, we study the availability of VLB networks that are overlaid over an optical infrastructure. The main challenges in such a context arise from the unique routing and protection scheme that goes beyond the definition of conventional connection-level service availability as well as the logical link failure correlation that prohibits the use of traditional analytical methods. We propose a network-level availability model to compute the probability that a VLB network is congestion-free under all traffic patterns. Numerical results show that with a proper truncation level, our calculation on availability can be accelerated significantly by generating tight lower and upper bounds. Our main finding is that physical link sharing in a two-layer setting degrades the network availability drastically by several orders of magnitude due to the full mesh requirement for VLB networks, and may remove the capacity efficiency advantage of VLB networks.

1. Introduction

Backbone networks are designed using a layered approach with packet-switched networks overlaid on top of a common optical network infrastructure. Such a layering method allows more efficient utilization of the optical network infrastructure, and enables flexible and customized internet service at the same time [1]. Internet services are typically carried over packet-switched networks, which provide better bandwidth manageability using IP protocols. IP networks, in turn, are constructed over optical networks by physically establishing an end-to-end lightpath between two routers of each IP link. The setup and teardown of a lightpath correspond to the creation and removal of an IP link, respectively. In this sense, IP links are referred to as logical links and IP network topologies as logical topologies.

Optical networks, on the other hand, are advancing to be highly dynamic and highly resilient [2]. This architecture migration is driven in part by the need to accommodate emerging high-bandwidth applications characterized by highly variable traffic as well as the need to handle frequent failures impacting carrier networks. The ever-increasing flexibility and reconfigurability built into optical networks offer a promising way of fulfilling the twofold need. Ideally, lightpaths would be set up and tore down on demand in instant response to traffic variations and failure events. However, dynamic bandwidth provisioning at a very small timescale is technically prohibitive. Rapid power changes on a fiber link due to the add/drop of wavelength channels result in widespread optical transient effects, which further limit the response time to bandwidth requests [3]. Moreover, finding an on-line provisioning solution subject to various resource constraints and QoS requirements is computationally complex, particularly in a timely and

* Corresponding author. Tel.: +1 613 790 6788.
E-mail addresses: wendani@sce.carleton.ca, wonda.ni@gmail.com (W. Ni), huang@sce.carleton.ca (C. Huang), jingwu@ieee.org (J. Wu), michel.savoie@crc.gc.ca (M. Savoie).

1573-4277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.osn.2013.02.002

cost-effective manner. Additionally, frequent lightpath reconfiguration incurs network control and management overhead, which compromises the benefits of dynamic optical networking.

To bypass the difficulties, an innovative yet simple way of handling highly variable traffic was proposed in [4,5] by introducing the notion of Valiant load balancing (VLB) to backbone network design. Specifically, a two-hop traffic routing scheme is proposed over a full mesh logical topology. In the first hop, traffic load originating from a logical node is evenly split and sent to all the logical nodes, regardless of the packet destinations. In the second hop, a logical node forwards packets received from all the logical nodes to their designated destinations. We refer to logical networks that employ the two-hop routing scheme as VLB networks. VLB networks have the advantage of supporting all possible traffic matrices with no need of dynamic reconfigurations as long as the traffic matrices are valid in the sense that they are consistent with the ingress/egress capacity limits on logical nodes. More importantly, the authors in [4] show that the link capacity required to serve any valid traffic matrix is significantly lower in comparison to traditional static networks with direct source–destination single path routing. Even when survivability feature is incorporated, the spare capacity to be allocated is small relative to the working capacity, but only at the regime where the number of failed links to tolerate is small. From a standalone single-layer viewpoint, which was taken by the above study, the regime for efficient spare capacity allocation can be well appreciated as tolerance up to a small number of failed links can be sufficient for service availability guarantee. However, when the VLB network is laid out on an optical network, such a tolerance level can be far from satisfaction. The seemingly independent logical links can be intimately related by sharing the same fiber link. Consequently, even a single physical link failure can lead to multiple correlated logical link failures. It is observed in real-life carrier networks that at least two logical links fail in an optical failure event, and the number of concurrently failed links can go up to ten [6]. In the case of embedding VLB networks onto optical networks, sharing potential of a physical link can increase substantially among multiple logical links due to the full mesh connectivity required at the IP layer and the sparse connectivity deployed at the optical layer. Therefore, a more realistic study in the context of two-layer architecture is required to evaluate the survivability performance of VLB networks.

In parallel to the above study, the work in [5] (with a full version in [7]) made a step forward by considering how each logical link is routed in the form of a lightpath over an optical network. Once the lightpath routing is decided, a full mesh VLB network is constructed logically at the IP layer. However, no survivability feature is discussed in this original work. Follow-on works in [8–10] filled the gap by protecting VLB networks at the optical layer. Specifically, working lightpaths (i.e., working logical links) are protected against single physical link failures by traditional link or path protection scheme. Due to the large number of working lightpaths to maintain a full mesh logical topology, spare capacity sharing is employed to be capacity-efficient. Note that in [8–10], protection is deployed at the optical layer, while the work in [4] takes a completely different approach by protecting a VLB network at the IP layer with spare capacity allocated on each logical link. Thus, we call the latter approach protection at the IP layer. However, no matter which approach is taken, only a predefined limited set of failure scenarios is tolerated in both cases. Indeed, a more complete and better understanding of these approaches should include the notions of service level agreement and, in particular, network availability [11], which, however, is currently absent to the best of our knowledge.

This motivates us to analyze the availability of survivable VLB networks over optical networks. We consider a highly changing traffic environment. In this context, availability is defined to be the probability that a VLB network is congestion-free to accommodate all valid traffic matrices. By "valid", we mean that a traffic matrix is compliant with the ingress/egress capacity limits on logical nodes. By "congestion-free", we mean that under any possible valid traffic matrix, no link is overloaded, and the network achieves 100% throughput. We focus on protection approach at the IP layer proposed in [4] due to the following reasons: (1) today's carrier networks rely on the IP layer to provide survivable services, and there is generally no lower-layer protection beneath the IP layer [6,11–13]. Protection at the IP layer does not require reconfigurability at the optical layer, and thus can be easily supported in today's carrier networks; and (2) spare capacity is allocated more efficiently when protection is deployed at the IP layer. Specifically, extra capacity for failure tolerance is very small relative to the working capacity by taking advantage of the load balancing property. On the other hand, if protection can be performed at the optical layer, logical links (i.e., lightpaths) are protected by conventional link or path protection. This typically results in the total extra capacity to be over 60% of the total working capacity even if spare capacity sharing is enabled.

Assessing the availability of an overlay logical network is rather challenging due to the correlation of logical link failures arising from physical link sharing. A large body of works [14–19] have dealt with the problem of service availability estimation in a single-layer setting. Analytical models were proposed to compute or bound the steady-state probability that a connection is in an operating state. However, all these models are on a per-connection basis, and assume independent link failures. The work in [20] first considered service availability in a logical network setting. Correlation among logical links is modeled by joint failure probability of two logical links. Based on this model, failure probability (i.e., unavailability) of a connection with two logically link–disjoint paths is approximated. Joint failure probability of two logical links gives a first-order approximation to connection failure probability with an upper bound value. A more accurate model requires introducing joint failure probability of multiple (i.e., four, six, etc.) logical links. Note that the focus of the work is still on availability estimation at a connection level. The network-level availability was recently studied in [21] in a two-layer setting. Given a fixed lightpath

routing of logical topology, availability is defined to be the probability that the logical topology remains connected. Polynomial expression for network availability was established as a function of physical link failure probability, which is required to be homogenous with a global value p. This, however, limits the application scope of the approximation method. Also, the definition of network availability cannot be applied to VLB networks that dictate a unique definition for network availability.

In a VLB network, traffic between any source–destination pair is distributed through multiple paths naturally. There is no need to establish protection paths for higher availability as legacy protection schemes. When a link fails, traffic traversing the link will be distributed to other paths that go to the same destinations. To make sure that all other paths have enough capacity to carry the extra load, links on all other paths need to be engineered with enough capacity. In fact, the link capacity of a VLB network has to be provisioned in such a way that they can accommodate all traffic patterns. In other words, if any single link becomes overloaded under any single traffic pattern, the network is considered to be down. This kind of availability definition is unique in the sense that (1) all paths are mutually protected in a global manner, and (2) it is agnostic of traffic patterns. It was shown in [4] that the extra capacity required to achieve this kind of availability is moderate when the number of concurrent failures is relatively small. Existing studies on network availability cannot be applied to this kind of unique protection scheme as discussed above. In this paper, we develop a new approach for modeling and calculating the availability of VLB networks. To make it easier, we assume that the VLB network is configured to tolerate single logical link failures using the approach developed in [4]. We show that with any multiple (>1) logical link failures, VLB network is congested under certain valid traffic matrices due to link overload, and thus becomes unavailable. This property enables a simple demarcation of network operating states and down states. With a clear classification of network states, we propose an availability model for the VLB network to be overlaid over optical networks. Our model is developed based on the inclusion–exclusion formula, where correlation among logical link failures is modeled as joint failure probability. To address the unique availability definition for VLB networks, we develop a new calculation method through some sophisticated analytical techniques. With a proper truncation level, tight upper and lower bounds on network availability can be calculated much faster than the exact value. Another important aspect of our model is that it does not require the failure probability of physical links to be homogeneous, and is generally applicable to inhomogeneous cases. Our main finding is that once a VLB network is mapped onto an optical network, strong correlation among logical link failures degrades the network availability by several orders of magnitude to further take away the capacity efficiency advantage of VLB networks.

The remainder of the paper is organized as follows. In Section 2, we present the network model, and identify the failure states where a VLB network is unavailable. This provides the basis to derive an availability model. In Section 3, we consider the case of independent link failures to serve as a baseline. In Section 4, availability model under correlated link failures is proposed for VLB networks mapped over optical networks. Numerical results for the correlated failure case are presented in Section 5. We conclude the paper in Section 6.

2. Network model

We model the logical VLB network as an undirected graph $G=(\mathcal{N},\mathcal{L})$, where $\mathcal{N}$ is the node set, and $\mathcal{L}$ is the link set. The nodes and the links are numbered from 1 to $|\mathcal{N}|$ and from 1 to $|\mathcal{L}|$, respectively. Due to the full mesh connectivity, $|\mathcal{L}|=|\mathcal{N}|(|\mathcal{N}|-1)/2$. For simplicity, we also refer to a link between node i and node j by (i,j).

The physical optical network is represented as $G_O=(\mathcal{N}_O,\mathcal{L}_O)$ in the similar fashion. Each physical link fails independently. We assume that each logical node is assigned to one and only one physical node, and logical nodes from the same VLB network are assigned to different physical nodes. Thus, the number of logical nodes in one VLB network should be no greater than that of the physical optical nodes, i.e., $|\mathcal{N}|\le|\mathcal{N}_O|$. The assumption is typical for virtual network embedding [1]. Note that logical nodes from different VLB networks can be mapped to the same physical node.

In the following, we identify the failure scenarios where a VLB network cannot achieve 100% throughput under certain valid traffic matrices due to link congestion. Such a network state is referred to as a down state. To gain sufficient insights, we will first review the capacity allocation results given in the previous literature, starting with the no protection case.

2.1. Capacity without protection

Let λ_{ij} denote the traffic demand from node i to node j. Assume that each logical node has ingress/egress capacity r. A valid traffic matrix is one that follows the access limit, i.e.,

$$\sum_{j\in\mathcal{N},j\neq i}\lambda_{ij}\le r,\quad i\in\mathcal{N}, \tag{1}$$

$$\sum_{i\in\mathcal{N},i\neq j}\lambda_{ij}\le r,\quad j\in\mathcal{N}. \tag{2}$$

If protection is not considered, the ingress/egress capacity decides the capacity on each logical link. Specifically, the capacity on a link is the sum of the maximum traffic load in the first hop and the second hop. In the first hop, ingress traffic from node i to node j', i.e., $\lambda_{ij'}$, is equally split, and then sent to all $|\mathcal{N}|$ nodes (including node i with loopback). Thus, the maximum load on link (i,j) is $\max\{\sum_{j'\in\mathcal{N},j'\neq i}\lambda_{ij'}/|\mathcal{N}|\}=r/|\mathcal{N}|$. As a result of the first hop forwarding, node j receives traffic from any node i' to any node j' ($\neq i'$) with rate $\lambda_{i'j'}/|\mathcal{N}|$. Among them, traffic that terminates at node j is $\sum_{i'\in\mathcal{N},i'\neq j}\lambda_{i'j}/|\mathcal{N}|\le r/|\mathcal{N}|$, which is obviously within the egress limit of node j. More importantly, the total traffic destined for node j' is $\sum_{i'\in\mathcal{N},i'\neq j'}\lambda_{i'j'}/|\mathcal{N}|$. This amount of traffic is forwarded from node j to the destination

node j' in the second hop. Due to the routing symmetry, we know immediately that the load on link (i,j) for the second hop forwarding is $\sum_{i'\in\mathcal{N},i'\neq j}\lambda_{i'j}/|\mathcal{N}|$, with a maximum value of $r/|\mathcal{N}|$. Therefore, the minimum capacity required on any link (i,j) for 100% throughput is $2r/|\mathcal{N}|$. For more discussions on this conclusion, see [4].

2.2. Capacity with single logical link failure protection

The key idea of the above two-hop routing is that every traffic demand λ_{st} is balanced over $|\mathcal{N}|$ paths between s and t, with $|\mathcal{N}|-2$ two-hop paths and 2 one-hop paths (i.e., the direct link (s,t) weighted as two paths). When there are link failures, the number of paths between s and t can be reduced. The original demand share on the disrupted paths will be evenly assigned to the remaining paths. This causes load increase on the surviving logical links. Specifically, the work in [4] showed that given a failure scenario f, the maximum load on a surviving link (i,j) is given by

$$c_{ij}(f)=\max\left\{\frac{2r}{P_{ij}(f)},r\cdot\left(\max_{j'\in\mathcal{N}j'\neq i,j}\frac{A_{jj'}(f)}{P_{ij'}(f)}+\max_{i'\in\mathcal{N},i'\neq i,j}\frac{A_{i'i}(f)}{P_{i'j}(f)}\right)\right\},\quad i,j\in\mathcal{N},\ i\neq j, \tag{3}$$

where $P_{i'j'}(f)$ denotes the number of remaining paths between i' and j', and $A_{i'j'}(f)$ denotes the direct connectivity between i' and j'. Indicator $A_{i'j'}(f)$ equals one if link (i',j') is up; equals zero otherwise. Once a failure scenario is given, both $P_{i'j'}(f)$ and $A_{i'j'}(f)$ are deterministic and known. Clearly, the load on link (i,j) comes from three part as shown in Fig. 1: (1) traffic demand from node i to node j delivered through the one-hop path; (2) traffic demand from node i to node j' $(\neq i,j)$ delivered over link (i,j) in the first hop through a two-hop path; and (3) traffic demand from node i' $(\neq i,j)$ to node j delivered over link (i,j) in the second hop through a two-hop path. Among these source–destination pairs, Eq. (3) indicates that the maximum load on link (i,j) is reached when the maximum demands (i.e., r) are generated for the pair that has the minimum number of remaining paths, provided that link (i,j) is on one of these remaining paths. Note that Eq. (3) is backward-compatible with the no protection case.

By maximizing load $c_{ij}(f)$ over single logical link failures yields the minimum capacity required on link (i,j), which is calculated in [4] as

$$c_{ij}=\frac{r}{|\mathcal{N}|-2}+\frac{r}{|\mathcal{N}|},\quad i,j\in\mathcal{N},\ i\neq j. \tag{4}$$

Due to the symmetry in failure scenarios, the capacity allocated on logical links is uniform. The capacity on link (i,j) is fully utilized under the network state illustrated in Fig. 2, where link (i,j') $(j'\neq j)$ fails, and traffic demands from i to j' and from i' $(i'\neq i)$ to j are r.

2.3. Identifying failure scenarios with network congestion

The calculation of our availability model requires a clear identification of network up states and down states. For a VLB network that tolerates a limited set of failure scenarios, network can be in down states when failures occur outside of the predefined set. However, as capacity allocation on logical links is for the worst case matching the highest load watermark among the predefined failure set, even when failures occur outside of the set, there can be possibility that the allocated capacity still guarantees 100% network throughput.

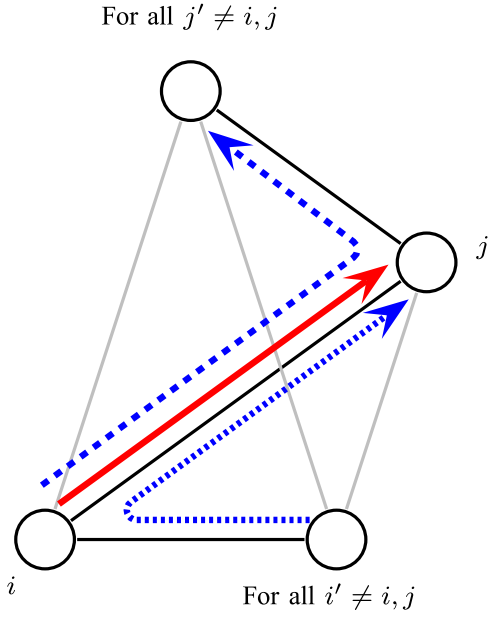

Fig. 1. Load on logical link (i,j).

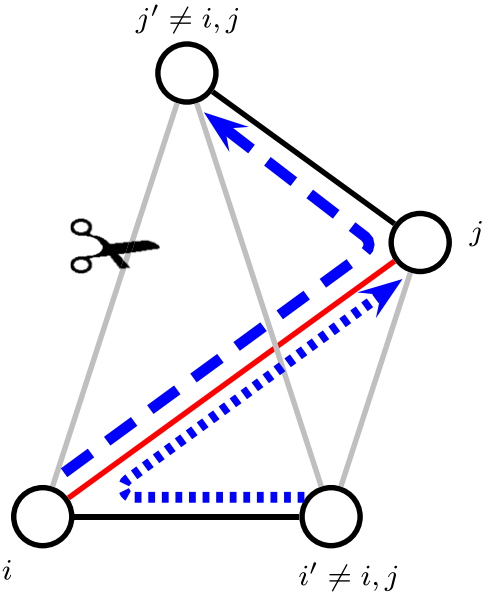

Fig. 2. Network state where capacity on logical link (i,j) can be fully utilized under single logical link failures.

Based on the background discussions in the previous two subsections, we establish the following theorem to identify network down states.

Theorem 1. *For a VLB network that tolerates single link failures with the capacity allocation in (4), any additional link failures will cause network congestion under certain valid traffic matrices.*

Proof. We start with the case of dual link failures. A second failure can happen on two link locations with respect to the first one: (1) links that have a common node with the first failed link; and (2) links that do not. For both cases, we give an example of link capacity violation.

In the case that the two failed links have a node in common, consider the network state in Fig. 3(a) for illustration. Let (i,j) and (i',j) be the two failed links with common node j. Let k be any node other than i, i', or j. We will show that the load on link (i,k) exceeds its capacity.

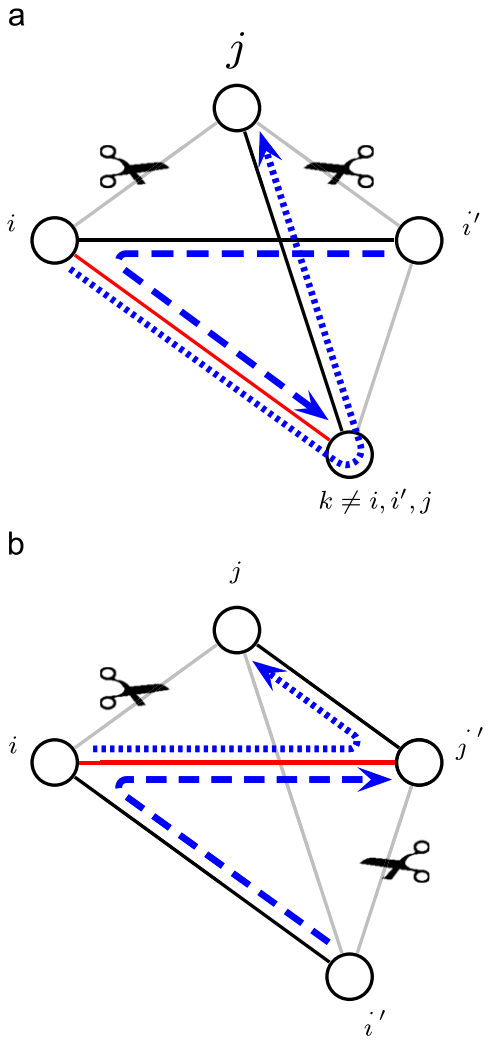

Fig. 3. Network states under dual logical link failures. (a) Two failed links have a common node. (b) Two failed links do not have a common node.

As discussed before, traffic demands that may introduce load to link (i,k) must either originate from node i or terminate at node k. In other words, link (i,k) must be either on the first hop or the second hop of one of the paths that carry these demands. Consider the traffic demands from i to j which originates from node i, and from i' to k which terminates at node k. As link (i,j) is on the one-hop path i–j, and link (i',j) is on the two-hop path i–i'–j, the number of operating paths between i and j is reduced to $|\mathcal{N}|-3$, given the two failed links. (Note that the one-hop path is counted as two paths because it is the degenerated path of both path i–i–j and path i–j–j.) Similarly, due to the failure of two-hop path i'–j–k, the number of operating paths between i' and k is reduced to $|\mathcal{N}|-1$. Among these remaining paths, link (i,k) is on the first hop of path i–k–j between i and j, and is on the second hop of path i'–i–k between i' and k. Thus, when traffic demands from i to j and from i' to k are both r, which forms a valid traffic matrix, load on link (i,k) is $r/(|\mathcal{N}|-3)+r/(|\mathcal{N}|-1)$. The link load exceeds the allocated capacity given in (4) since

$$\left(\frac{r}{|\mathcal{N}|-3}+\frac{r}{|\mathcal{N}|-1}\right)-\left(\frac{r}{|\mathcal{N}|-2}+\frac{r}{|\mathcal{N}|}\right)$$

$$=\frac{r}{(|\mathcal{N}|-3)(|\mathcal{N}|-2)}+\frac{r}{(|\mathcal{N}|-1)|\mathcal{N}|}>0,\quad |\mathcal{N}|\geq 4.$$

In the second case that the two failed links do not have a common node, we illustrate the network state in Fig. 3(b). Let (i,j) and (i',j') be the two failed links. We will show that the load on link (i,j') exceeds its capacity. Again the demands that have impact on link (i,j') either originate from node i or terminate at node j'. Consider the traffic demands from i to j and from i' to j'. Due to the failure of one-hop paths i–j and i'–j', the number of surviving paths are $|\mathcal{N}|-2$ for both traffic demands. Consider link (i,j'), which is on the first hop of path i–j'–j and the second hop of path i'–i–j'. When traffic demands from i to j and from i' to j' are both r, load on link (i,j') is $2r/(|\mathcal{N}|-2)$. The link load violates the capacity limit given in (4) since

$$\frac{2r}{|\mathcal{N}|-2}-\left(\frac{r}{|\mathcal{N}|-2}+\frac{r}{|\mathcal{N}|}\right)=\frac{2r}{(|\mathcal{N}|-2)|\mathcal{N}|}>0,\quad |\mathcal{N}|\geq 4.$$

Clearly, for any more than two logical link failures, the number of paths between source–destination pairs is further reduced to amplify congestion on the surviving links. Note that there can be multiple failure scenarios that disconnect the VLB network, these scenarios are considered as extreme cases of network congestion that throughput among the disconnected node islands is 0%. □

Theorem 1 suggests that for a VLB network that tolerates single link failures, network is in up states when and only when there are no more than one logical link failures. This constitutes our foundation to compute network availability next.

3. Network availability under independent logical link failures

In this section, we consider the special case that each logical link in a VLB network fail independently with probability p. In a two-layer context, this is the case where a VLB network is embedded onto a set of optical nodes that have full mesh connectivity among them, and each logical link is mapped to the direct link (i.e., one-hop path) that connects the corresponding host nodes. We develop the following accurate availability model.

According to Theorem 1, network up states are

(1) All the logical links are up;
(2) Only one logical link fails.

Thus, network availability can be computed as the sum of these state probabilities which are mutually exclusive.

Specifically, let X_0 denote the event that all the logical links are up. Let X_1 denote the event that 1 logical link fails, and the rest of the logical links are up. Probabilities of events X_0 and X_1 are given, respectively, by

$$P(X_0) = (1-p)^{|\mathcal{L}|}, \tag{5}$$

and

$$P(X_1) = |\mathcal{L}|p(1-p)^{|\mathcal{L}|-1}. \tag{6}$$

Thus, network unavailability (complementary to network availability) is computed as

$$\begin{aligned} P(S) &= 1-P(X_0)-P(X_1) \\ &= 1-\sum_{i=0}^{1}\binom{|\mathcal{L}|}{i}p^i(1-p)^{|\mathcal{L}|-i}, \end{aligned} \tag{7}$$

where S denotes the event that network is down.

Fig. 4(a) shows the network unavailability plotted on logarithmic axes. We see that the curves are almost parallel straight lines. This indicates that network unavailability grows in proportion to physical link unavailability. Interestingly, when the physical link unavailability is below 10^{-5}, network unavailability is even lower than the unavailability of a single physical link. This is because when logical link failures are independent, a VLB network is down when at least two physical links fail, which are high-order failure events at the optical layer. This greatly reduces the unavailability of a VLB network.

Also, for a given physical link unavailability, we see that network unavailability increases with the increase in the number of logical nodes, but at a slower pace as the number of logical nodes gets larger. The latter trend can be clearly observed in Fig. 4(b), which plots the network unavailability against the number of logical nodes.

4. Network availability under correlated logical link failures

In this section, we deal with the case of correlated logical link failures. We develop a model that can provide upper and lower bounds on network availability up to arbitrary tightness required. Similar to the independent case in the previous section, we first compute, or more precisely, bound the probabilities of network up states, and then bound the network availability through the sum of them. Unlike the independent case, the model proposed here is general, and it assumes that physical link failure probabilities are known but not necessarily equal. Also, it is immediately applicable to the independent case.

4.1. Bounds on $P(X_0)$

Let E_l denote the event that link l fails. Let $\overline{E}_l$ be the complementary of E_l, denoting the event that link l is up. Recall that X_0 denotes the event that all the links are up. Thus, we have

$$X_0 = \bigcap_{l\in\mathcal{L}}\overline{E}_l. \tag{8}$$

Consequently, probability of event X_0 can be expressed as

$$P(X_0) = 1-P(\overline{X}_0) = 1-P\left(\bigcup_{l\in\mathcal{L}}E_l\right), \tag{9}$$

where $\overline{X}_0$ is the complementary event of X_0. In (9), probability of the union of events E_1 to E_l can be bounded by using the inclusion/exclusion formula [22]. Thus, probability of event X_0 can be lower and upper bounded, respectively, by

$$P(X_0) \geq 1-Z_1(\mathcal{L})+Z_2(\mathcal{L})-Z_3(\mathcal{L})+\ldots-Z_{m'}(\mathcal{L}), \tag{10}$$

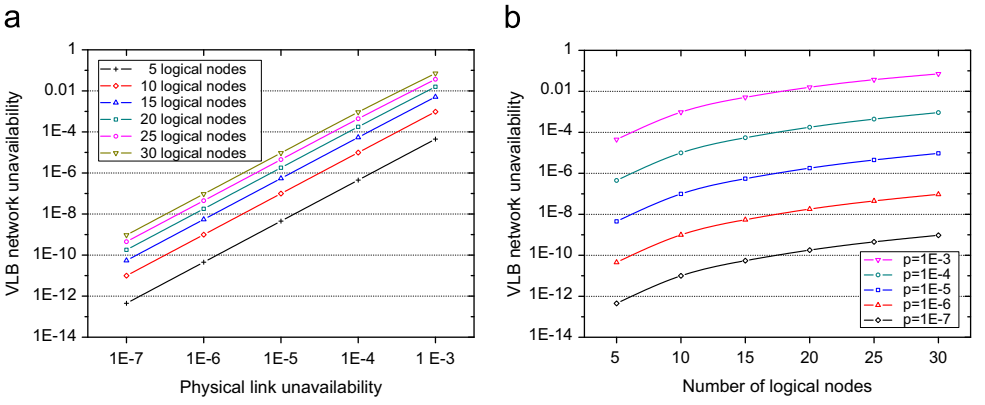

Fig. 4. Network unavailability under independent logical link failures. (a) Versus physical link unavailability. (b) Versus the number of logical nodes.

and

$$P(X_0) \le 1 - Z_1(\mathcal{L}) + Z_2(\mathcal{L}) - \ldots + Z_{m''}(\mathcal{L}), \tag{11}$$

where $Z_m(\mathcal{L})$ $(1 \le m \le |\mathcal{L}|)$ denotes the sum of joint failure probabilities of m logical links, and m' and m'' are odd and even positive integers, respectively. Clearly, larger m' and m'' yield a smaller gap between lower and upper bounds. However, the number of terms in $Z_m(\mathcal{L})$ (i.e., $\binom{|\mathcal{L}|}{m}$) also becomes excessive, leading to significant extra computational complexity. Therefore, values of m' and m'' should be chosen wisely. As two logical link failures are the first-order failures that cause a network down state, joint failure probabilities of two logical links should be included in both (10) and (11) to yield acceptable bounds. In other words, the minimum values for m' and m'' are 3 and 2, respectively. More generally, if in the worst case $\hat{m}$ logical link failures constitute any of the first-order upper-layer failures, values of m' and m'' in (10) and (11) should be no smaller than $\hat{m}$. This is the typical case in a two-layer setting, where multiple logical links can fail in one single physical link failure.

For simplicity, we start with the case of $m' = 3$ and $m'' = 2$ to derive an availability model, and extend the model to consider larger values later in this section. With $m' = 3$ and $m'' = 2$, the derived model assumes that at most two logical links fail in any single physical link failure, i.e., $\hat{m} = 2$. Specifically, we first focus on

$$P(X_0) \ge 1 - Z_1(\mathcal{L}) + Z_2(\mathcal{L}) - Z_3(\mathcal{L}), \tag{12}$$

and

$$P(X_0) \le 1 - Z_1(\mathcal{L}) + Z_2(\mathcal{L}), \tag{13}$$

where $Z_1(\mathcal{L})$, $Z_2(\mathcal{L})$, and $Z_3(\mathcal{L})$ are the sum of joint failure probabilities of one, two, and three logical links, respectively. They are expressed as follows:

$$Z_1(\mathcal{L}) = \sum_{l \in \mathcal{L}} P(E_l) = \sum_{l=1}^{|\mathcal{L}|} P(E_l), \tag{14}$$

$$Z_2(\mathcal{L}) = \sum_{l \in \mathcal{L}} \sum_{l' \in \mathcal{L}, l' > l} P(E_l E_{l'}) = \sum_{l=1}^{|\mathcal{L}|-1} \sum_{l'=l+1}^{|\mathcal{L}|} P(E_l E_{l'}), \tag{15}$$

$$\begin{aligned} Z_3(\mathcal{L}) &= \sum_{l \in \mathcal{L}} \sum_{l' \in \mathcal{L}, l' > l} \sum_{l'' \in \mathcal{L}, l'' > l'} P(E_l E_{l'} E_{l''}) \\ &= \sum_{l=1}^{|\mathcal{L}|-2} \sum_{l'=l+1}^{|\mathcal{L}|-1} \sum_{l''=l'+1}^{|\mathcal{L}|} P(E_l E_{l'} E_{l''}). \end{aligned} \tag{16}$$

Once the node and link mapping is given, $P(E_l)$, $P(E_l E_{l'})$, and $P(E_l E_{l'} E_{l''})$ can be computed based on the link-to-lightpath mapping information and the failure probabilities of physical links. In Appendix A, we propose a general calculation method.

4.2. Bounds on $P(X_1)$

Recall that X_1 denotes the event that 1 logical link fails, and the rest of the logical links are up. Unlike the case in Section 3, here we assume that failure probabilities of logical links can be different and correlated. Therefore, we have to address each failure scenario separately. Let X_1^k be the more specific event that logical link k fails, and the rest of the links in $\mathcal{L}_k = \mathcal{L} \backslash \{k\}$ are up. As events $X_1^1, X_1^2, \ldots, X_1^{|\mathcal{L}|}$ are mutually exclusive, probability of event X_1 can be found through

$$P(X_1) = \sum_{k=1}^{|\mathcal{L}|} P(X_1^k). \tag{17}$$

To calculate $P(X_1^k)$, we introduce Y_1^k to denote the event that links in $\mathcal{L}_k$ are up. The difference between events Y_1^k and X_1^k is that event Y_1^k is regardless of the state of link k while link k fails in event X_1^k, i.e., $X_1^k = Y_1^k E_k$. On the other hand, if link k is up in the event of Y_1^k, then all the links are up, yielding the occurrence of event X_0. In other words, $Y_1^k \overline{E}_k = X_0$. Thus, probability of event Y_1^k can be written as

$$\begin{aligned} P(Y_1^k) &= P(Y_1^k E_k) + P(Y_1^k \overline{E}_k) \\ &= P(X_0) + P(X_1^k), \quad k \in \mathcal{L}. \end{aligned} \tag{18}$$

Rearranging (18) yields

$$P(X_1^k) = P(Y_1^k) - P(X_0), \quad k \in \mathcal{L}. \tag{19}$$

Eq. (19) indicates that probability of event X_1^k can be calculated once the probability of event Y_1^k is established. Therefore, we show next how the value of $P(Y_1^k)$ can be computed.

4.2.1. Bounding $P(Y_1^k)$

Bound calculations of $P(Y_1^k)$ are similar to those of $P(X_0)$. Using the inclusion/exclusion formula, we obtain the lower bound on $P(Y_1^k)$ as

$$P(Y_1^k) \ge 1 - Z_1(\mathcal{L}_k) + Z_2(\mathcal{L}_k) - Z_3(\mathcal{L}_k), \quad k \in \mathcal{L}. \tag{20}$$

Introducing (14), (15), and (16) into (20) yields

$$\begin{aligned} P(Y_1^k) \ge\ & 1 - Z_1(\mathcal{L}) + P(E_k) \\ & + Z_2(\mathcal{L}) - \sum_{l'=k+1}^{|\mathcal{L}|} P(E_k E_{l'}) - \sum_{l=1}^{k-1} P(E_l E_k) \\ & - Z_3(\mathcal{L}) + \sum_{l'=k+1}^{|\mathcal{L}|-1} \sum_{l''=l'+1}^{|\mathcal{L}|} P(E_k E_{l'} E_{l''}) \\ & + \sum_{l=1}^{k-1} \sum_{l''=k+1}^{|\mathcal{L}|} P(E_l E_k E_{l''}) + \sum_{l=1}^{k-2} \sum_{l'=l+1}^{k-1} P(E_l E_{l'} E_k), \quad k \in \mathcal{L}. \end{aligned} \tag{21}$$

Similarly, upper bound on $P(Y_1^k)$ is expressed as

$$\begin{aligned} P(Y_1^k) \le\ & 1 - Z_1(\mathcal{L}) + P(E_k) \\ & + Z_2(\mathcal{L}) - \sum_{l'=k+1}^{|\mathcal{L}|} P(E_k E_{l'}) - \sum_{l=1}^{k-1} P(E_l E_k), \quad k \in \mathcal{L}. \end{aligned} \tag{22}$$

4.2.2. Bounding $P(X_1^k)$

Introducing bounds on $P(X_0)$ and $P(Y_1^k)$ into (19), we establish the lower and upper bounds for $P(X_1^k)$, respectively, as

$$\begin{aligned} P(X_1^k) &= P(Y_1^k) - P(X_0) \\ &\ge P(E_k) - \sum_{l'=k+1}^{|\mathcal{L}|} P(E_k E_{l'}) - \sum_{l=1}^{k-1} P(E_l E_k) \end{aligned}$$

$$-Z_3(\mathcal{L})+\sum_{l'=k+1}^{|\mathcal{L}|-1}\sum_{l''=l'+1}^{|\mathcal{L}|}P(E_kE_{l'}E_{l''})+\sum_{l=1}^{k-1}\sum_{l''=k+1}^{|\mathcal{L}|}P(E_lE_kE_{l''})+\sum_{l=1}^{k-2}\sum_{l'=l+1}^{k-1}P(E_lE_{l'}E_k),\quad k\in\mathcal{L},\tag{23}$$

and

$$P(X_1^k)=P(Y_1^k)-P(X_0)\le P(E_k)-\sum_{l'=k+1}^{|\mathcal{L}|}P(E_kE_{l'})-\sum_{l=1}^{k-1}P(E_lE_k)+Z_3(\mathcal{L}),\quad k\in\mathcal{L}.\tag{24}$$

4.2.3. Bounding $P(X_1)$

Applying (23) and (24) to (17) yields

$$\begin{aligned}P(X_1)&=\sum_{k=1}^{|\mathcal{L}|}P(X_1^k)\\&\ge\sum_{k=1}^{|\mathcal{L}|}P(E_k)-\sum_{k=1}^{|\mathcal{L}|-1}\sum_{l'=k+1}^{|\mathcal{L}|}P(E_kE_{l'})-\sum_{k=2}^{|\mathcal{L}|}\sum_{l=1}^{k-1}P(E_lE_k)-\sum_{k=1}^{|\mathcal{L}|}Z_3(\mathcal{L})\\&\quad+\sum_{k=1}^{|\mathcal{L}|-2}\sum_{l'=k+1}^{|\mathcal{L}|-1}\sum_{l''=l'+1}^{|\mathcal{L}|}P(E_kE_{l'}E_{l''})+\sum_{k=2}^{|\mathcal{L}|-1}\sum_{l=1}^{k-1}\sum_{l'=k+1}^{|\mathcal{L}|}P(E_lE_kE_{l'})\\&\quad+\sum_{k=3}^{|\mathcal{L}|}\sum_{l=1}^{k-2}\sum_{l'=l+1}^{k-1}P(E_lE_{l'}E_k)\\&=Z_1(\mathcal{L})-2Z_2(\mathcal{L})-(|\mathcal{L}|-3)Z_3(\mathcal{L}),\end{aligned}\tag{25}$$

and

$$\begin{aligned}P(X_1)&=\sum_{k=1}^{|\mathcal{L}|}P(X_1^k)\\&\le\sum_{k=1}^{|\mathcal{L}|}P(E_k)-\sum_{k=1}^{|\mathcal{L}|-1}\sum_{l'=k+1}^{|\mathcal{L}|}P(E_kE_{l'})-\sum_{k=2}^{|\mathcal{L}|}\sum_{l=1}^{k-1}P(E_lE_k)+\sum_{k=1}^{|\mathcal{L}|}Z_3(\mathcal{L})\\&=Z_1(\mathcal{L})-2Z_2(\mathcal{L})+|\mathcal{L}|Z_3(\mathcal{L}).\end{aligned}\tag{26}$$

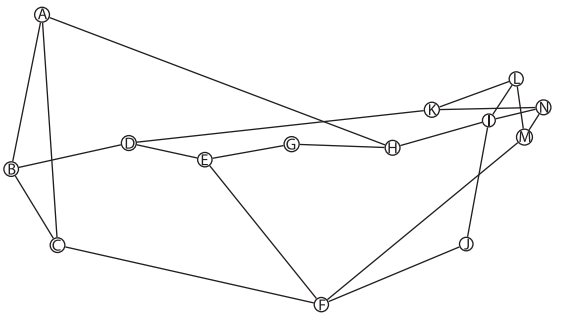

Fig. 5. NSFNET with 14 nodes and 21 links. The average node degree is 3.

Inequalities (25) and (26) give lower and upper bounds on the probability of event X_1, respectively.

4.3. Availability model

With the derived bounds on $P(X_0)$ and $P(X_1)$, network unavailability is lower and upper bounded, respectively, by

$$P(S)=1-P(X_0)-P(X_1)\ge Z_2(\mathcal{L})-|\mathcal{L}|Z_3(\mathcal{L}),\tag{27}$$

and

$$P(S)=1-P(X_0)-P(X_1)\le Z_2(\mathcal{L})+(|\mathcal{L}|-2)Z_3(\mathcal{L}).\tag{28}$$

Note that the availability model is for the case where one single physical link failure can bring down at most two logical links.

4.3.1. Extensions to arbitrary $Z_m(\mathcal{L})$

Once a VLB network is overlaid onto an optical network, physical link sharing can cause multiple, say in the worst

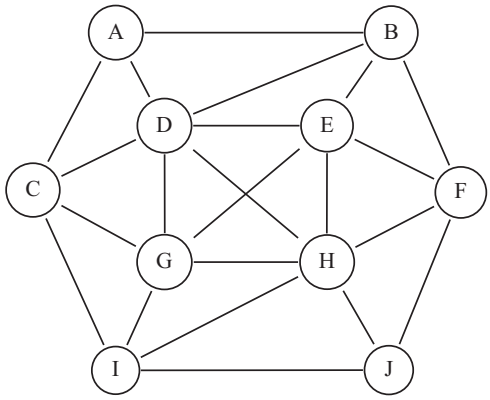

Fig. 6. SMALLNET with 10 nodes and 22 links. The average node degree is 4.4.

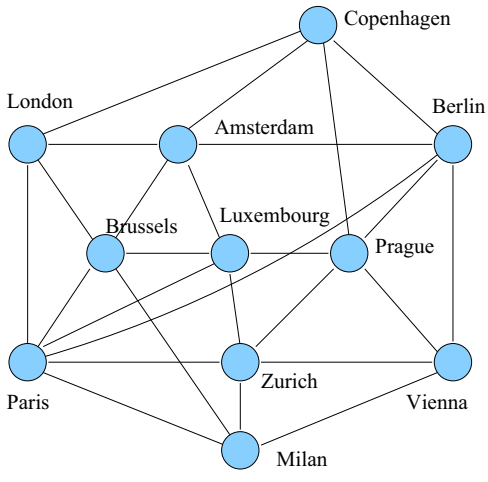

Fig. 7. COST239 with 11 nodes and 26 links. The average node degree is 4.73.

case $\hat{m}$, logical links to fail simultaneously in one single physical link failure. In this case, $\hat{m}$ logical link failures become part of the first-order upper-layer failure events. Accordingly, we will need to include joint failure probabilities of $\hat{m}$ logical links in our availability model to deliver nontrivial bounds.

So far we have considered the case of $\hat{m}=2$ with $m'=3$ and $m''=2$. Taking one step forward, if $\hat{m}=3$, the minimum values of m' and m'' become 3 and 4, respectively. This tightens the upper bound on $P(X_0)$ in (13). Following the same derivation process as $\hat{m}=2$, lower and upper bounds on network availability are given, respectively, by

$$P(S)\geq Z_2(\mathcal{L})-2Z_3(\mathcal{L})-(|\mathcal{L}|-3)Z_4(\mathcal{L}), \tag{29}$$

and

$$P(S)\leq Z_2(\mathcal{L})-2Z_3(\mathcal{L})+|\mathcal{L}|Z_4(\mathcal{L}). \tag{30}$$

To be general, if $\hat{m}$ is an even positive integer with $2\leq\hat{m}\leq|\mathcal{L}|$, m' and m'' take the values of $\hat{m}+1$ and $\hat{m}$, respectively. Network unavailability is lower and upper bounded, respectively, by

$$P(S)\geq\sum_{k=2}^{\hat{m}}(-1)^k(k-1)Z_k(\mathcal{L})-|\mathcal{L}|Z_{\hat{m}+1}(\mathcal{L}), \tag{31}$$

and

$$P(S)\leq\sum_{k=2}^{\hat{m}}(-1)^k(k-1)Z_k(\mathcal{L})+(|\mathcal{L}|-\hat{m})Z_{\hat{m}+1}(\mathcal{L}). \tag{32}$$

On the other hand, if $\hat{m}$ is an odd positive integer with $3\leq\hat{m}\leq|\mathcal{L}|-1$, m' and m'' take the values of $\hat{m}$ and $\hat{m}+1$, respectively. Network unavailability is lower and upper bounded, respectively, by

$$P(S)\geq\sum_{k=2}^{\hat{m}}(-1)^k(k-1)Z_k(\mathcal{L})-(|\mathcal{L}|-\hat{m})Z_{\hat{m}+1}(\mathcal{L}), \tag{33}$$

and

$$P(S)\leq\sum_{k=2}^{\hat{m}}(-1)^k(k-1)Z_k(\mathcal{L})+|\mathcal{L}|Z_{\hat{m}+1}(\mathcal{L}). \tag{34}$$

Once the node and link mapping is given, value of $\hat{m}$ is determined, and the minimum values of m' and m'' are determined accordingly. This is basically how m' and m'' are chosen in our numerical study.

5. Numerical results

We evaluate the proposed availability model over three representative optical topologies shown in Figs. 5–7. Given a VLB network, logical nodes are randomly assigned to optical nodes. Logical links are mapped to the shortest lightpath between the associated optical nodes. If there are multiple shortest lightpaths with equal cost, ties are broken randomly.

Once a VLB network is embedded, we count for a physical link the number of lightpaths it carries, which corresponds to the number of simultaneously failed logical links in a single physical link failure. The maximum number

Table 1
Run time (in s) of our model for physical link unavailability at 10^{-3}.

# of logical nodes		4	5	6	7	8	9
NSFNET	$m'=3, m''=4$	0.2	1.0	5.8	24.1	–	–
	$m'=5, m''=4$	–	2.2[a]	19.8[a]	116.3	–	–
SMALLNET	$m'=3, m''=4$	–	1.3	6.9	28.4	93.3	–
	$m'=5, m''=4$	–	2.5[a]	22.7[a]	134.4[a]	604.7	–
COST239	$m'=3, m''=4$	–	–	13.8	56.2	183.2	512.6
	$m'=5, m''=4$	–	–	59.4[a]	352.0[a]	1593.9	5840.2

[a] Unavailability results are not given in the paper.

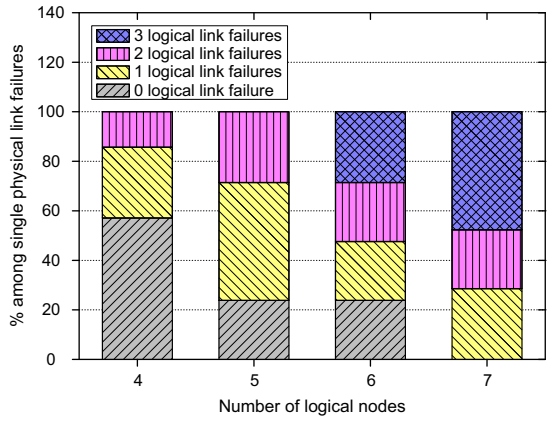

Fig. 8. Failure correlation over NSFNET in terms of the number of failed logical links in single physical link failures.

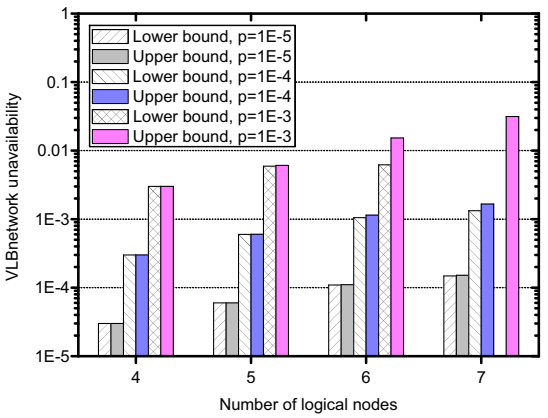

Fig. 9. Network unavailability over NSFNET with $m'=3$ and $m''=4$.

observed is assigned to $\hat{m}$, which further decides the minimum values of parameters m' and m''. For simplicity, we assume homogeneous physical link unavailability to compute the availability model. Our availability model is implemented in MATLAB R2011a, which runs on a Dell Inspiron 620 desktop with Intel Core i5-2310 processor (2.90 GHz) and 8 GB memory.

Figs. 8, 11, and 14 show the number of simultaneously failed logical links in single physical link failures. We see that physical link sharing leads to correlated logical link failures even for a VLB network with a small number of logical nodes. This indicates that once a VLB network is embedded, it is only partially protected against first-order (i.e., single) physical link failures. This immediately raises the network unavailability as shown later. Moreover, we observe that multiple logical link failures become more predominant with the increase in the number of logical nodes. This is because the number of logical links increases at a more rapid pace, which considerably increases the physical link sharing opportunities.

Figs. 10, 13, and 16 show the bounds on network unavailability. We see that with the minimum values of m' and m'', network unavailability is well bounded in most cases. The lower and upper bounds are closely

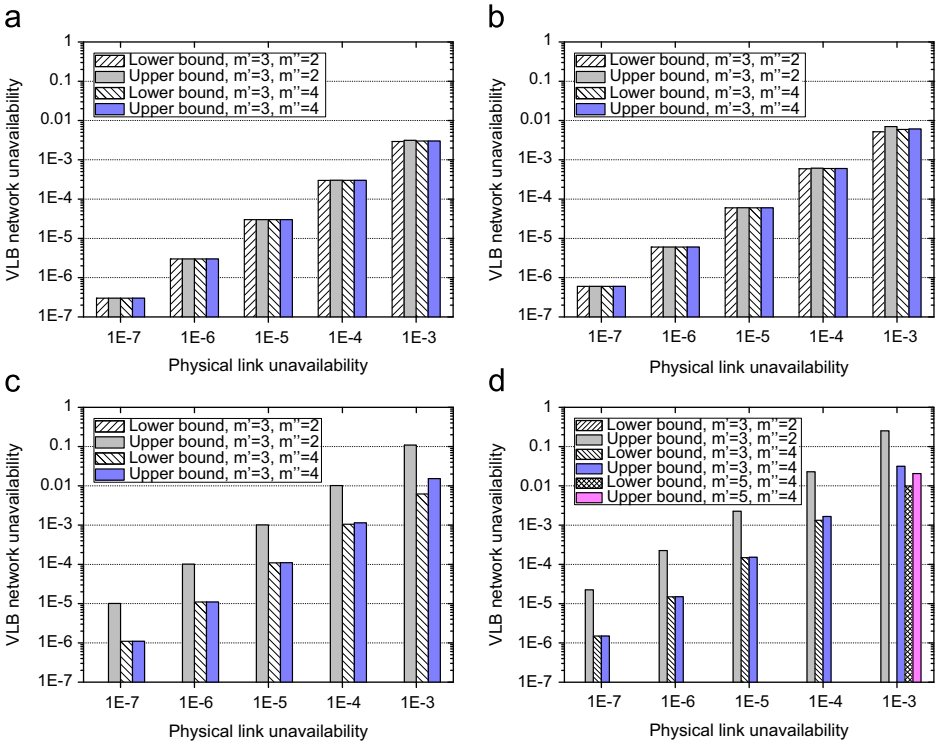

Fig. 10. Network unavailability over NSFNET. (a) 4 logical nodes, $\hat{m} = 2$. (b) 5 logical nodes, $\hat{m} = 2$. (c) 6 logical nodes, $\hat{m} = 3$. (d) 7 logical nodes, $\hat{m} = 3$.

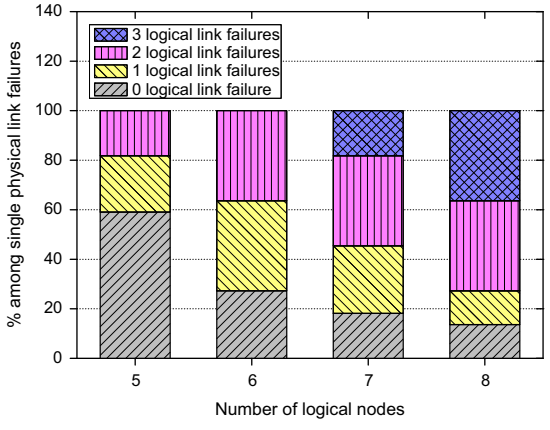

Fig. 11. Failure correlation over SMALLNET in terms of the number of failed logical links in single physical link failures.

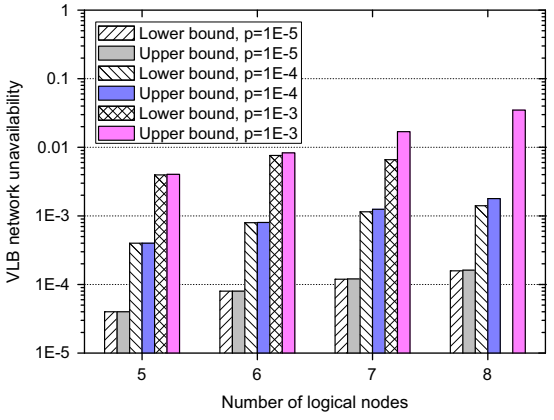

Fig. 12. Network unavailability over SMALLNET with $m' = 3$ and $m'' = 4$.

approximate to each other when the number of logical nodes is small and physical link unavailability is low. Gaps between the two bounds become larger as the physical link unavailability gets higher. In most cases, the two bounds are still close to each other, maintaining small gaps. However, as the number of logical nodes gets larger at the same time, the combined effects can cause the two bounds to be distant to each other. In particular, the lower bound can become negative and trivial. This is observed in Figs. 10(d), 13(d), 16(c), and 16(d), with physical link unavailability at 10^{-3}. In these cases, we increment the value of m' or m'' by two to tighten the bounds with joint failure probabilities of more logical links. In the cases of Figs. 10(d), 13(d), and 16(c), this immediately restores the tightness of the two bounds with small gaps between them. It is also interesting to note that although the lower bounds before increment are trivial, the corresponding upper bounds are only slightly higher than those after increment. In the rest case of Fig. 16(d), the two bounds after increment are still distant to each other by an order of magnitude, which indicates the need to further increase m' or m''.

On the other hand, if m' and/or m'' is smaller than the minimum value required, both bounds become loose as shown in subfigures (c) and (d) of Figs. 10, 13, and 16.

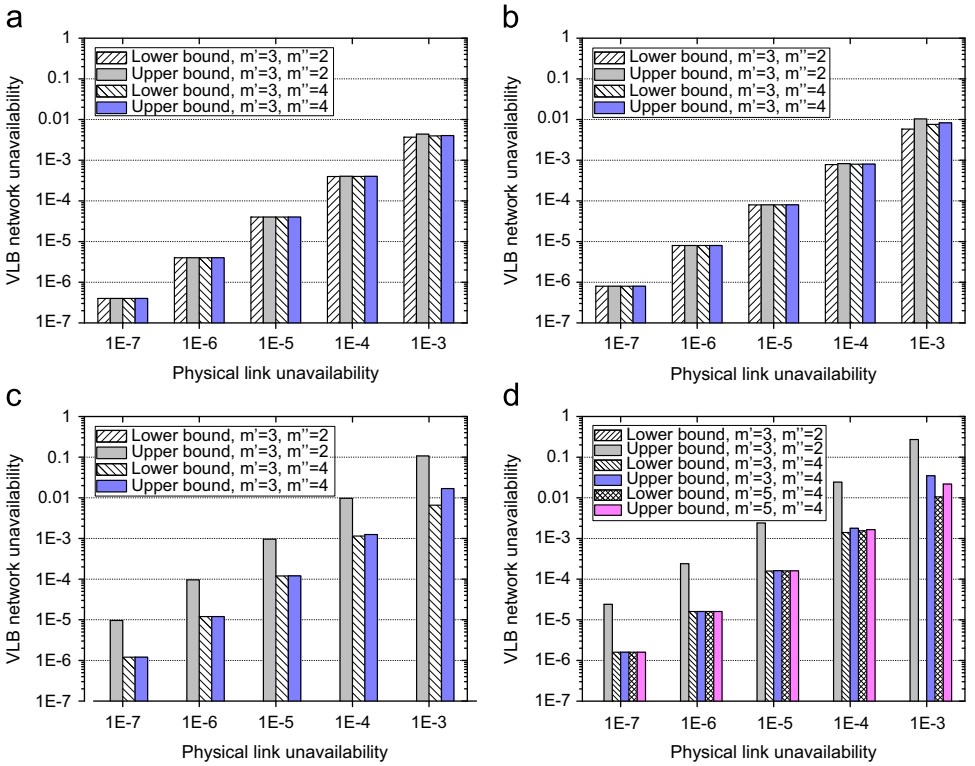

Fig. 13. Network unavailability over SMALLNET. (a) 5 logical nodes, $\hat{m}=2$. (b) 6 logical nodes, $\hat{m}=2$. (c) 7 logical nodes, $\hat{m}=3$. (d) 8 logical nodes, $\hat{m}=3$.

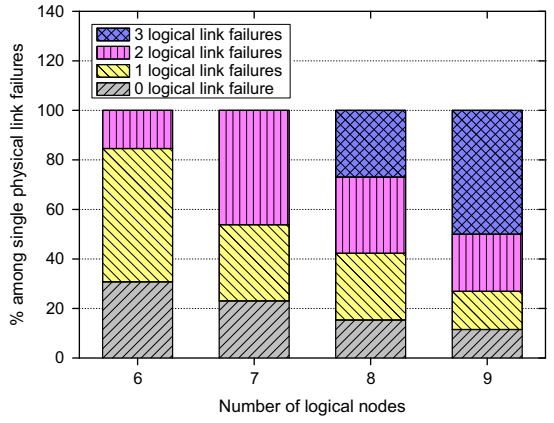

Fig. 14. Failure correlation over COST239 in terms of the number of failed logical links in single physical link failures.

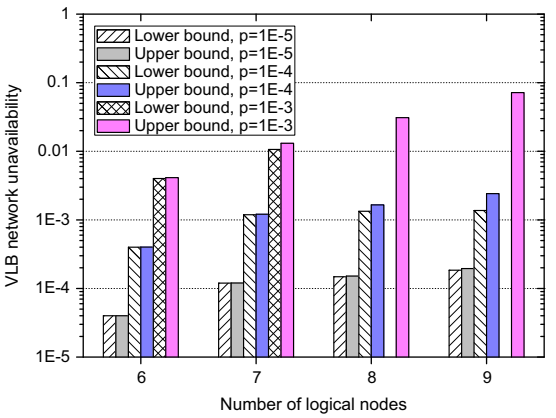

Fig. 15. Network unavailability over COST239 with $m'=3$ and $m''=4$.

The lower bounds are negative and trivial. The upper bounds are an order of magnitude higher than the exact value, which is no greater than the tight upper bounds given with the minimum required m' and m''.

Moreover, for a VLB network with 5 logical nodes, we compare the network unavailability under independent logical link failures in Fig. 4(a) and under correlated logical link failures in Figs. 10(b) and 13(a). We observe that failure correlation deteriorates the network unavailability by several orders of magnitude. For low physical link unavailability at 10^{-7}, network unavailability increases by five orders of magnitude from below 10^{-12} to above 10^{-7}. For high physical link unavailability at 10^{-3}, network unavailability also increases from below 10^{-4} to over 10^{-3}. This indicates that failure correlation has very strong negative effects on network unavailability. Physical link sharing turns multiple logical link failures which are originally high-order failure events to first-order failure events. If these multiple logical link failures are not tolerated, a VLB network is down in first-order failures. This significantly degrades the network unavailability. If, on the other hand, a VLB network is designed to tolerate all first-order failures, the number of concurrent logical link failures to consider in the worst case, i.e., $\hat{m}$, is typically very large due to the sparse physical topology at the optical layer and the full mesh connectivity at the IP layer. Tolerating $\hat{m}$ arbitrary logical link failures leads to significant capacity over-provisioning on links as the work in [4] shows that the capacity efficiency holds only when $\hat{m}$ is small, and when $\hat{m}$ is large, the required capacity goes up exponentially. In this case, the works in [23,24] suggest that link capacity is designed for a set of logical link failure scenarios rather than $\hat{m}$ arbitrary logical link failures to provide better tradeoffs between capacity efficiency and network availability. Specifically, each single physical link failure corresponds to one logical link failure scenario, and all single physical link failures are taken into account to form the set of logical link failure scenarios to protect against. How this new design affects network availability requires future research.

In Figs. 9, 12, and 15, we plot the network unavailability against the number of logical nodes. We observe that network unavailability increases with the number of logical nodes. This trend is clearly delineated when the lower and upper bounds are close to each other at low physical link unavailability.

Given the same VLB networks, Fig. 17 compares the network unavailability over different physical topologies. We see that network unavailability decreases as the average node degree of physical topologies increases, particularly for the case of 6 logical nodes. This is because a higher average node degree indicates a higher network connectivity, which can potentially reduce the physical link sharing opportunities, and thus the failure correlation among logical links. This trend is observed in Fig. 18, which shows that among all single physical link failures, the proportion of failures that cause multiple simultaneous logical link failures becomes less predominant as the average node degree goes higher. Accordingly, the probability that more than one logical link fails becomes lower, leading to reduced network unavailability. However, we also note that the decrease in network unavailability is rather limited. This can be

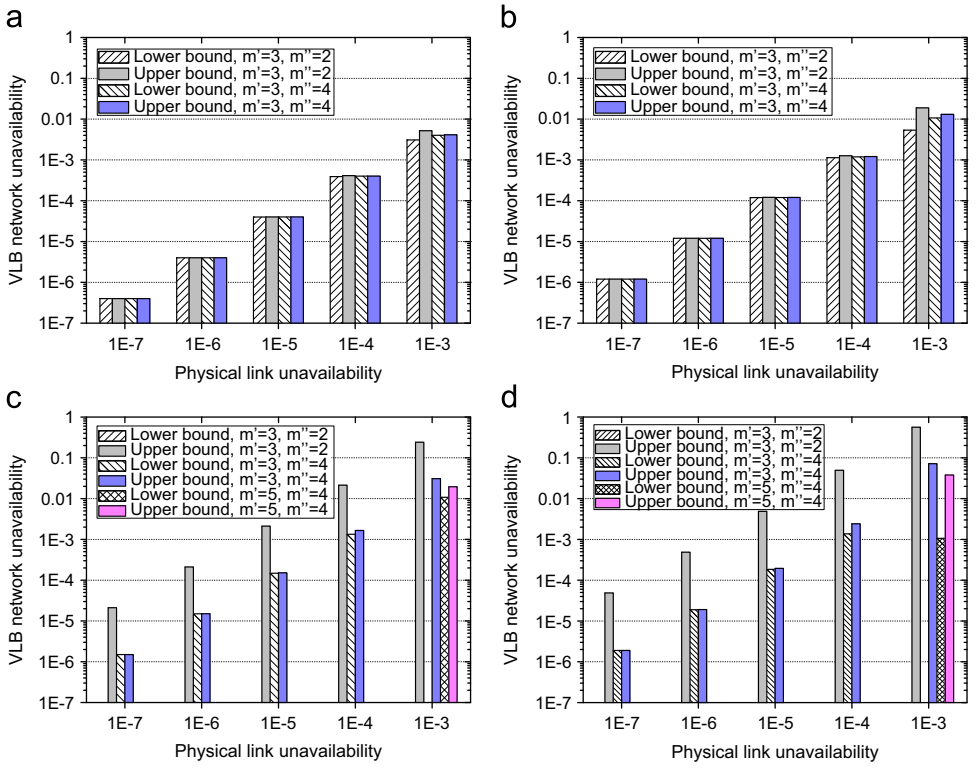

Fig. 16. Network unavailability over COST239. (a) 6 logical nodes, $\hat{m}=2$. (b) 7 logical nodes, $\hat{m}=2$. (c) 8 logical nodes, $\hat{m}=3$. (d) 9 logical nodes, $\hat{m}=3$.

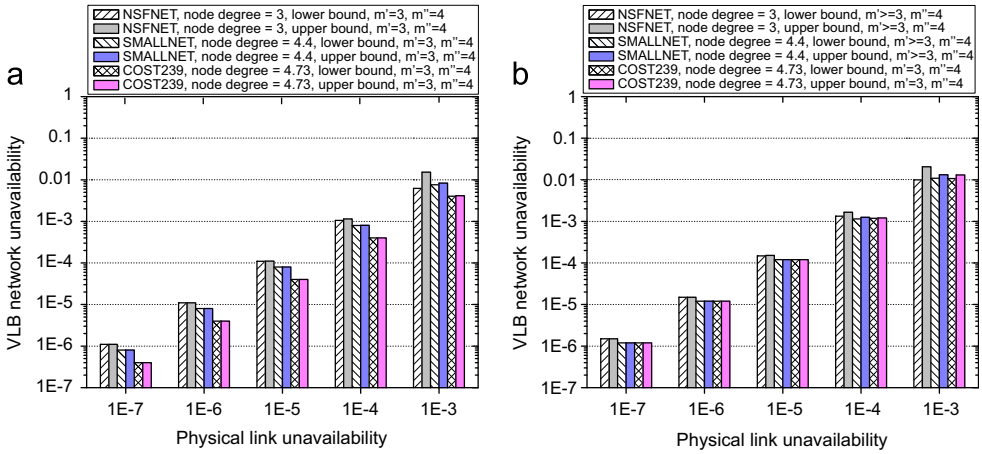

Fig. 17. Network unavailability over different physical topologies. (a) 6 logical nodes. (b) 7 logical nodes.

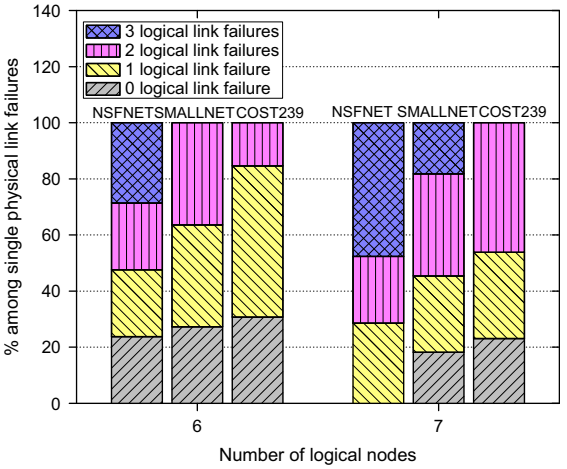

Fig. 18. Failure correlation over different topologies in terms of the number of failed logical links in single physical link failures.

explained by the fact that although physical network connectivity increases as a result of increase in the average node degree of physical topologies, it is still very sparse compared with the full mesh connectivity at the IP layer. Consequently, failure correlations among logical links (and thus the probability that more than one logical link fails) are mainly determined by the dense connectivity of VLB networks (i.e., the number of logical nodes). A slight increase in optical network connectivity does not help much in reducing the physical link sharing and thus the network unavailability, particularly when the number of logical nodes gets larger. In Fig. 17, we observe that when the number of logical nodes increases from 6 to 7, the decrease in network availability becomes even smaller over different physical topologies. In particular, the network unavailability is almost the same on SMALLNET and COST239.

Table 1 gives the run time of our model when the physical link unavailability is 10^{-3}. The run time corresponds to the unavailability results shown in Figs. 8–18. Note that as the value of physical link unavailability does not affect the run time, run time for other values of physical link unavailability is about the same. To give a more complete picture, we present more run time results by increasing the value of m'. These results are denoted as footnote (a) in the table. Note that the run time includes that for computing joint failure probabilities of logical links given in Appendix A. Our observations are as follows: (1) for the same physical topology, m', and m'', the run time increases with the number of logical nodes. This is because the terms of joint failure probabilities of logical links to be computed increase as the number of logical nodes grows; (2) for the same physical and logical topologies, the run time increases considerably for higher value of m'. This is because when the value of m' increases, our model needs to compute the complete set of joint failure probabilities of m' logical links to obtain the corresponding sum; and (3) for the same logical topology, m', and m'', the run time increases with the number of fiber links in the physical topology. This is because the calculation of joint failure probabilities of logical links involves computing more terms with the increase in the number of physical links.

6. Conclusions and future work

We studied VLB networks over optical networks from the availability viewpoint. Protection in a VLB network is fundamentally different from the traditional protection schemes due to its unique routing mechanism to support highly variable IP traffic. Moreover, physical link sharing in a two-layer context leads to multiple correlated logical link failures. These factors prohibit the use of conventional connection-level availability definition and the related analytical methods. In this paper, we proposed an availability model to compute the probability that a VLB network is congestion-free to accommodate all valid traffic matrices. As a first step, we focus on VLB networks that tolerate single logical link failures. We proved that failure of any two or more logical links causes the network not to be supportive to certain valid traffic matrices, resulting in a network down state. This property enables a simple yet accurate identification of network up states and down states. Based on this, we developed a network availability model, where joint failure probability

of logical links is used to represent failure correlation resulting from physical link sharing.

Numerical results show that with a proper truncation level, our model provides upper and lower bounds on network availability up to arbitrary tightness as required. More importantly, we find that once a VLB network is embedded, correlation among logical link failures is so strong that even a single physical link failure can break a large number of logical links. This effect is significantly exaggerated with the increasing number of logical nodes (or equivalently, the increasing number of logical links) and on sparser optical network connectivity. Consequently, network unavailability degrades drastically by several orders of magnitude in comparison to the independent logical failure counterpart, which assumes a standalone single-layer setting. On the other hand, strong correlation indicates that in a two-layer setting tolerating single physical link failures is equivalent to tolerating logical link failures up to a great number, which, however, goes far beyond the regime where spare capacity allocation is efficient. This removes one great advantage of VLB networks.

Our work suggests from the availability viewpoint that for backbone network applications, the number of logical nodes in a VLB network should be kept small to avoid strong failure correlation to further maintain the capacity efficiency at the IP layer. Future work includes: (1) availability analysis for VLB networks that tolerate $k \geq 2$ arbitrary logical link failures, a given set of logical link failure scenarios, or a combination of logical node and link failures; (2) finding the optimal VLB network mapping that maximizes the network availability; and (3) availability analysis for VLB networks with physical links that are subject to shared risk link group failures.

Acknowledgments

The authors would like to thank anonymous reviewers for their constructive comments to improve the paper. Dr. Jing Wu acknowledges the research support from the State Key Laboratory of Advanced Optical Communication Systems and Networks, Shanghai Jiao Tong University, China.

Appendix A. Computing joint failure probabilities of logical links

There are basically two methods for connection-level availability calculation: series–parallel reliability block diagram method [14–16,25] and Markov state–space method [17–19,25].

Series–parallel method is suited when a system can be characterized by a combination of mutually-independent serially and/or parallelly structured subsystems, with each subsystem further consisting of independent components. A serially-structured subsystem fails if any one of its components fails. In contrast, a parallelly-structured subsystem fails only if all of its components fail. Series–parallel method enables simple, fast, and accurate availability calculations. In the following subsections, we will employ this method to compute the failure probability of one logical link and the joint failure probability of two logical, both of which can be modeled by series and/or parallel block diagrams. However, there are many complex systems that cannot be represented in such a form, which is the case with joint failure probability of three or more logical links. In this instance, we resort to the more general state–space method.

State–space method partitions network states according to failure scenarios. Each network state corresponds to one failure scenario, where certain physical links are down. The partitioned failure scenarios are mutually exclusive to ensure that network states do not contain the same failure event. Hence, the unavailability of a system, such as a lightpath, is computed as the sum of the probabilities of network states in which the system is down. In Appendix A.3, we will use this method to compute the joint failure probability of three logical links as an example. Joint failure probability of more logical links can be developed in the same fashion. Also, failure probability of one logical link and joint failure probability of two logical links can be derived using the state–space method, which, however, requires more computational cost than its series-parallel counterpart, and thus is not recommended for use.

A.1. Failure probability of one logical link

In an IP-over-optical architecture, each logical link l is laid out as a lightpath, which can be represented by binary row vector $\boldsymbol{\theta}_l = \{\theta_{li}\}_{1\times|\mathcal{L}_O|}$. Binary element θ_{li} equals one if logical link l uses physical link i; equals zero otherwise. Lightpaths are typical series systems formed by a sequence of physical links, which fail independently in our assumption. Thus, failure probability of logical link l is computed immediately as

$$P(E_l) = 1 - \prod_{i\in\mathcal{L}_O}(1-p_i)^{\theta_{li}}, \quad l \in \mathcal{L}, \tag{A.1}$$

where p_i denotes the failure probability of physical link i.

A.2. Joint failure probability of two logical links

The layout of two logical links, i.e., two lightpaths, can be characterized by shared links and the link–disjoint part as shown in Fig. A1. Two logical links fail concurrently if any one of the shared links fails, or if in the link–disjoint part both lightpaths fail. Consequently, the system can be modeled at

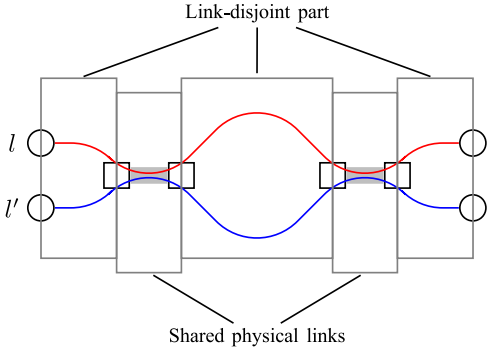

Fig. A1. Physical layout (i.e., lightpath routing) of two logical links l and l'.

the highest level as a series structure consisting of the shared link part and the link–disjoint part. The shared part is further represented as a series subsystem consisting of common physical links used by both logical links. The link–disjoint part is modeled as a parallel subsystem with two arms associated with the two logical links. Each arm consists of physical links that come from the associated logical link, and are not used by the other logical link. Thus, failure probability of two logical links can be expressed as

$$P(E_l E_{l'}) = 1 - A^S_{ll'}[1-(1-A^{D,l'}_{l})(1-A^{D,l}_{l'})], \quad l,l' \in \mathcal{L},\ l<l', \tag{A.2}$$

where $A^S_{ll'}$ denotes the availability of the shared link part, and $A^{D,l'}_{l}$ and $A^{D,l}_{l'}$ denote the availabilities of the arms formed by logical links l and l', respectively, in the link–disjoint part. Availabilities $A^S_{ll'}$, $A^{D,l'}_{l}$, and $A^{D,l}_{l'}$ are given, respectively, by

$$A^S_{ll'} = \prod_{i\in\mathcal{L}_O}(1-p_i)^{\theta_{li}\theta_{l'i}}, \quad l,l' \in \mathcal{L},\ l<l', \tag{A.3}$$

$$A^{D,l'}_{l} = \prod_{i\in\mathcal{L}_O}(1-p_i)^{\theta_{li}(1-\theta_{l'i})}, \quad l,l' \in \mathcal{L},\ l<l', \tag{A.4}$$

and

$$A^{D,l}_{l'} = \prod_{i\in\mathcal{L}_O}(1-p_i)^{(1-\theta_{li})\theta_{l'i}}, \quad l,l' \in \mathcal{L},\ l<l', \tag{A.5}$$

where $\theta_{li}\theta_{l'i}$ equals one if physical link i is used by both logical links; equals zero otherwise, and $\theta_{li}(1-\theta_{l'i})$ and $(1-\theta_{li})\theta_{l'i}$ equal one if physical link i is used by one logical link but not by the other; equal zero otherwise.

A.3. Joint failure probability of three logical links

When the number of logical links goes beyond two, the corresponding physical layout cannot be modeled simply by series–parallel structures. One of such examples can be found in Fig. A2. In this case, the state–space method is employed instead, where network states are enumerated according to physical link failure scenarios. However, a complete enumeration of network states is computationally intractable as the number of all possible failure scenarios, although finite, is on the order of $2^{|\mathcal{L}_O|}$. Thus, we consider a limited but predominant set of network states with no more than three physical link failures. Our numerical studies show that probabilities of higher-order failures (i.e., failures of more than three physical links) are extremely low for nation-wide networks, and are negligible to the computational accuracy.

Specifically, let F_i denote the network state that physical link i fails, and the other physical links are up. Let $\hat{p}_i$ denote the corresponding probability, which can be computed as

$$\hat{p}_i = p_i \prod_{i'\in\mathcal{L}_O, i'\neq i}(1-p_{i'}), \quad i \in \mathcal{L}_O. \tag{A.6}$$

Similarly, we have

$$\hat{p}_{ii'} = p_i p_{i'} \prod_{i''\in\mathcal{L}_O, i''\neq i,i'}(1-p_{i''}), \quad i,i' \in \mathcal{L}_O,\ i<i', \tag{A.7}$$

and

$$\hat{p}_{ii'i''} = p_i p_{i'} p_{i''} \prod_{i'''\in\mathcal{L}_O, i'''\neq i,i',i''}(1-p_{i'''}), \quad i,i',i'' \in \mathcal{L}_O,\ i<i'<i'', \tag{A.8}$$

where $\hat{p}_{ii'}$ denotes the probability of network state $F_{ii'}$ that physical links i and i' fail, and the other physical links are up, and $\hat{p}_{ii'i''}$ denotes the probability of network state $F_{ii'i''}$ that physical links i, i', and i'' fail, and the other physical links are up.

Next, we identify among F_i, $F_{ii'}$, and $F_{ii'i''}$ the states where three logical links fail concurrently. For network state F_i, three logical links fail if physical link i carries all of them. Let $\phi^{ll'l''}_i$ be a binary indicator, which equals one if logical links l, l', and l'' fail in state F_i; equals zero otherwise. Thus, $\phi^{ll'l''}_i$ can be computed as

$$\phi^{ll'l''}_i = \theta_{li}\cdot\theta_{l'i}\cdot\theta_{l''i}, \quad l,l',l'' \in \mathcal{L},\ l<l'<l'',\ i \in \mathcal{L}_O. \tag{A.9}$$

Similarly, we define $\phi^{ll'l''}_{ii'}$ to be a binary indicator, which equals one if logical links l, l', and l'' fail in state $F_{ii'}$; equals zero otherwise. Indicator $\phi^{ll'l''}_{ii'}$ can be found through

$$\phi^{ll'l''}_{ii'} = u(\theta_{li}+\theta_{li'})\cdot u(\theta_{l'i}+\theta_{l'i'})\cdot u(\theta_{l''i}+\theta_{l''i'}),$$
$$l,l',l'' \in \mathcal{L},\ l<l'<l'',\ i,i' \in \mathcal{L}_O,\ i<i', \tag{A.10}$$

where all three logical links fail if each logical link traverses at least one of the two physical links. Function $u(x)$ is a step function defined as

$$u(x) = \begin{cases} 1 & \text{if } x \geq 1, \\ 0 & \text{if } x < 1, \end{cases} \tag{A.11}$$

which equals one when argument x is no less than one; equals zero otherwise.

For network state $F_{ii'i''}$, we introduce binary indicator $\phi^{ll'l''}_{ii'i''}$, which equals one if logical links l, l', and l'' fail in $F_{ii'i''}$; equals zero otherwise. As all three logical links fail in state $F_{ii'i''}$ if each of them traverses at least one of the three physical links, we have

$$\phi^{ll'l''}_{ii'i''} = u(\theta_{li}+\theta_{li'}+\theta_{li''})\cdot u(\theta_{l'i}+\theta_{l'i'}+\theta_{l'i''})\cdot u(\theta_{l''i}+\theta_{l''i'}+\theta_{l''i''}),$$
$$l,l',l'' \in \mathcal{L},\ l<l'<l'',\ i,i',i'' \in \mathcal{L}_O,\ i<i'<i''. \tag{A.12}$$

Then, joint failure probability of logical links l, l', and l'' is simply the sum of the probabilities of network states where all three logical links are down, i.e.,

$$P(E_l E_{l'} E_{l''}) = \sum_{i\in\mathcal{L}_O}\hat{p}_i\phi^{ll'l''}_i + \sum_{i\in\mathcal{L}_O}\sum_{i'\in\mathcal{L}_O, i'>i}\hat{p}_{ii'}\phi^{ll'l''}_{ii'}$$
$$+\sum_{i\in\mathcal{L}_O}\sum_{i'\in\mathcal{L}_O, i'>ii''\in\mathcal{L}_O, i''>i'}\hat{p}_{ii'i''}\phi^{ll'l''}_{ii'i''},$$
$$l,l',l'' \in \mathcal{L},\ l<l'<l''. \tag{A.13}$$

For joint failure probability of four or more logical links, the same calculation process can be followed with

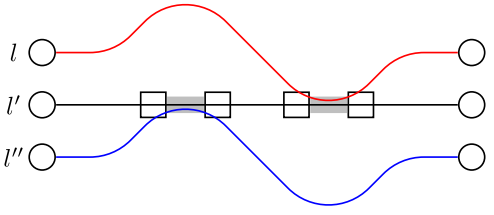

Fig. A2. Physical layout (i.e., lightpath routing) of three logical links l, l', and l''.

more logical link terms included to compute the binary indicators in network states F_i, $F_{ii'}$, and $F_{ii'i''}$.

References

[1] N.M.M.K. Chowdhury, R. Boutaba, A survey of network virtualization, Elsevier Computer Networks 54 (5) (2010) 862–876.

[2] A.L. Chiu, et al., Architectures and protocols for capacity efficient, highly dynamic and highly resilient core networks, IEEE/OSA Journal of Optical Communications and Networks 4 (1) (2012) 1–14.

[3] G. Li, A. Chiu, R. Doverspike, M. Birk, D. Husa, N. Zanki, On transient-constrained wavelength assignment, in: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM), April 2008, pp. 1975–1983.

[4] R. Zhang-Shen, N. McKeown, Designing a predictable internet backbone network, in: Proceedings of the Third Workshop on Hot Topics in Networks (HotNets-III), November 2004.

[5] M. Kodialam, T.V. Lakshman, S. Sengupta, Efficient and robust routing of highly variable traffic, in: Proceedings of the Third Workshop on Hot Topics in Networks (HotNets-III), November 2004.

[6] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C.-N. Chuah, Y. Ganjali, C. Diot, Characterization of failures in an operational IP backbone network, IEEE/ACM Transactions on Networking 16 (4) (2008) 749–762.

[7] M. Kodialam, T.V. Lakshman, J.B. Orlin, S. Sengupta, Oblivious routing of highly variable traffic in service overlays and IP backbones, IEEE/ACM Transactions on Networking 17 (2) (2009) 459–472.

[8] M. Kodialam, T.V. Lakshman, S. Sengupta, Throughput guaranteed restorable routing without traffic prediction, in: Proceedings of the IEEE International Conference on Network Protocols (ICNP), November 2006, pp. 137–146.

[9] M. Kodialam, T.V. Lakshman, J.B. Orlin, S. Sengupta, Resilient routing of variable traffic with performance guarantees, in: Proceedings of the IEEE International Conference on Network Protocols (ICNP), October 2009, pp. 213–222.

[10] M. Kodialam, T.V. Lakshman, S. Sengupta, Locally restorable routing of highly variable traffic, IEEE/ACM Transactions on Networking 17 (3) (2009) 752–763.

[11] R. Doverspike, B. Cortez, Restoration in carrier networks, in: Proceedings of the IEEE Workshop on the Design of Reliable Communication Networks (DRCN), October 2009, pp. 45–54.

[12] G. Iannaccone, C.-N. Chuah, S. Bhattacharyya, C. Diot, Feasibility of IP restoration in a tier 1 backbone, IEEE Network 18 (2) (2004) 13–19.

[13] A.L. Chiu, et al., Network design and architectures for highly dynamic next-generation IP-over-optical long distance networks, IEEE/OSA Journal of Lightwave Technology 27 (12) (2009) 1878–1890.

[14] M. Jaeger, R. Huelsermann, Service availability of shared path protection in optical mesh networks, in: Proceedings of the European Conference on Optical Communication (ECOC), September 2004, We4.P.154.

[15] J. Zhang, K. Zhu, H. Zang, N.S. Matloff, B. Mukherjee, Availability-aware provisioning strategies for differentiated protection services in wavelength-convertible WDM mesh networks, IEEE/ACM Transactions on Networking 15 (5) (2007) 1177–1190.

[16] L. Song, J. Zhang, B. Mukherjee, Dynamic provisioning with availability guarantee for differentiated services in survivable WDM mesh networks, IEEE Journal on Selected Areas in Communications—Supplement on Optical Communications and Networking 25 (4) (2007) 35–43.

[17] D.A.A. Mello, D.A. Schupke, H. Waldman, A matrix-based analytical approach to connection unavailability estimation in shared backup path protection, IEEE Communications Letters 9 (9) (2005) 844–846.

[18] L. Zhou, M. Held, U. Sennhauser, Connection availability analysis of shared backup path-protected mesh networks, IEEE/OSA Journal of Lightwave Technology 25 (5) (2007) 1111–1119.

[19] D.A.A. Mello, H. Waldman, Analytical bounds on the unavailability of protected connections in WDM optical networks, IEEE Communications Letters 11 (11) (2007) 901–903.

[20] W. Cui, I. Stoica, R.H. Katz, Backup path allocation based on a correlated link failure probability model in overlay networks, in: Proceedings of the IEEE International Conference on Network Protocols (ICNP), November 2002, pp. 236–245.

[21] K. Lee, H.-W. Lee, E. Modiano, Reliability in layered networks with random link failures, IEEE/ACM Transactions on Networking 19 (6) (2011) 1835–1848.

[22] S.M. Ross, Introduction to Probability Models, Elsevier Academic Press, 2003.

[23] R. Zhang-Shen, N. McKeown, Designing a fault-tolerant network using Valiant load-balancing, in: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM), April 2008, pp. 301–305.

[24] R. Zhang-Shen, N. McKeown, Designing a fault-tolerant network using Valiant load-balancing, Technical Report TR07-HPNG-070207, Department of Electrical Engineering, Stanford University, July 2007.

[25] W. Ni, J. Wu, C. Huang, M. Savoie, Analytical models of flow availability in two-layer networks with dedicated path protection, Optical Switching and Networking 10 (1) (2013) 62–76. Special Issue on Advances in Optical Networks Control and Management.