Provisioning for Probabilistic Failures in Multi-Domain DWDM Networks

M. Esmaeili[1], **K. Kazi**[2], **S. Khan**[3], **A. Rayes**[4], **N. Ghani**[1]
[1]University of New Mexico, [2]Defense Photonics Group, [3]North Dakota State University, [4]Cisco Systems

Abstract: *This paper addresses multi-domain lightpath provisioning within the context of correlated multi-failure events. The work jointly incorporates both risk minimization and traffic engineering objectives, and develops a novel graph theoretic scheme for distributed operation in realistic optical network settings. Detailed performance evaluation results are then presented to gauge the effectiveness of the proposed multi-domain provisioning solution.*

I. INTRODUCTION

Optical *dense wavelength division multiplexing* (DWDM) networks have matured rapidly in the last decade and now offer unprecedented bandwidth scalabilities. As a result of this growth, network survivability is now a key concern and a full range of solutions have been proposed here, i.e., including protection (pre-provisioning) as well as restoration (post-provisioning) strategies. However, many of these strategies are only geared for single network domains. Indeed, given expanding DWDM deployments, the topic of *multi-domain* resilience is also becoming important [1].

To date, various solutions have been proposed for multi-domain (optical) network survivability. These include extensions of "localized" SONET/SDH domain-to-domain link protection interconnection strategies [2],[3], e.g., multi-trunked/multi-hubbed, as well as broader "end-to-end" path protection schemes. In particular, the latter have leveraged detailed inter-domain topology state (inter-domain routing) to compute diversified working/backup path routes [4],[5], i.e., in *multi-protocol label switching* (MPLS) and *generalized multi-protocol label switching* (GMPLS) networks.

Nevertheless, most optical survivability schemes have only been designed to handle single failures. Now given the massive geographical scale of multi-domain DWDM backbones, these infrastructures are indeed very vulnerable to correlated multi-failure events, e.g., such as those resulting from large power outages, natural disasters, *weapons of mass destruction* (WMD) attacks, etc. Along these lines, only a handful of solutions have looked at such scenarios [6],[7]. Most notably, [7] extends the *shared risk link group* (SRLG) concept to a *probabilistic SRLG* (p-SRLG) one to handle correlated multi-failure events. Although these solutions do yield improved path reliabilities, they are only evaluated in "non-optical" single-domain settings where computational entities have full knowledge of all link resources, i.e., global views. Clearly, this will not be the case in realistic *multi-domain* (DWDM) settings, where only a subset of nodes may have partial/dated "global" state [1]. Moreover, the schemes in [6] and [7] strictly focus on risk minimization objectives, and hence may have adverse effects on *traffic engineering* (TE) efficiencies as well.

In light of the above, this paper proposes a novel "risk-aware" provisioning scheme for multi-domain DWDM *routing and wavelength assignment* (RWA). The goal here is to handle multiple correlated failures and achieve a measure of TE provisioning as well. The solution leverages the GMPLS protocols framework for hierarchical inter-domain routing and path computation/setup signaling [1],[8] and this paper is organized as follows. Section II first presents the psuedocode for the scheme, whereas Section III details the performance results. Conclusions are then presented in Section IV.

II. BACKGROUND

Many solutions have been proposed for multi-domain survivability. For example, [2] looks at different types of DWDM border node interconnections to protect working/ backup paths traversing a common set of "end-to-end" domains. Albeit very fast, these are localized strategies and hence very susceptible to multi-failure events affecting a set of co-located nodes, e.g., such as a WMD attacks. Meanwhile [3] studies "end-to-end" working/protection path pair routing in MPLS networks and proposes several sequential and parallel strategies for improved "domain diversity". In addition, some researchers have even developed advanced topology abstraction schemes to help compute non-overlapped inter-domain protection routes, see [4],[5]. Although the blocking gains here are good, these abstractions impose significant overheads at the inter-domain routing level.

Several multi-failure recovery solutions have also been studied in recent years. For example, some researchers have focused on specific dual near-simultaneous failure scenarios. A good example is the work in [9], which proposes post-fault re-computation of backup routes in optical DWDM networks. Overall findings show notable increases in backup lightpath availabilities. In addition, [10] performs pre-emptive re-provisioning after dual link failures in DWDM networks, but results show some key deficiencies for the case of fiber faults. Meanwhile [11] develops various pre-provisioned shared protection schemes to improve efficiencies under dual-link failures. The work in [12] takes a different angle and evaluates the performance of active post-fault restoration schemes versus pre-provisioned protection strategies for dual-link DWDM failures. Other dual link failures studies are also presented in [13] and [14] for IP-tunneling/rerouting networks. However, all of these schemes only address single domain settings with full topology state.

As mentioned earlier, some new studies have also emerged for more generalized multi-failure scenarios. For example, [6] studies protection for correlated failure events, and introduces the concepts of local and global reliability. Namely, local reliability selects routes that span the lowest number of failure events, whereas global reliability selects routes that fail a minimum number of connections. Meanwhile [7] also studies recovery under correlated probabilistic failures and extends the basic SRLG definition with a probabilistic variant, i.e., *probabilistic SRLG* (p-SRLG). Namely, the authors propose several schemes to maximize the reliability of pre-computed protection routes and assume that all links within a p-SRLG fail independently. Nevertheless, these probabilistic models are only developed and analyzed for smaller single domain network settings.

In light of the above, there is a pressing need to look at multi-failure analysis in real-world distributed multi-domain networks. This requirement is given impetus by the fact that most multi-domain networks are inherently large (in a geographical sense) and hence more vulnerable to natural and/or man-made failure events. Along these lines, this paper now proposes a novel solution to improve the reliability of multi-domain lightpath routes, i.e., working-mode only. Here, it is postulated that the use of *a-priori* link risk state can likely improve reliability performance over existing TE-based provisioning schemes.

III. JOINT TE AND RISK MINIMIZATION

In this paper we propose a novel scheme which improves lightpath reliability in multi-domain DWDM networks with correlated failures. To aim here is to use a-priori probabilistic risk state for susceptible links and improve the reliability of the provisioned lightpaths. The solution assumes “all-optical” domains with full opto-electronic conversion/regeneration at border gateway *optical cross-connect* (OXC) nodes. As per the GMPLS framework, all OXC nodes at the intra-domain level run link-state routing and have full domain visibility, e.g., via the ubiquitous *open-shortest-path first* (OSPF-TE) protocol. Meanwhile, border nodes OXC nodes also run another “hierarchical” link-state routing level to propagate updates for inter-domain links and maintain abstracted multi-domain views, i.e., simple node abstraction [1],[8]. Now at both the intra-/inter-domain routing levels, link state updates are generated using *significance change factor* (SCF) threshold policies. In addition, routing *hold-down* (HT) timers are also used to limit excessive overheads [8]. Associated domain-level *path computation elements* (PCE) [1] then use this condensed information to compute “skeleton” *loose route* (LR) inter-domain paths over the abstract graph. Finally, RSVP-TE signaling is leveraged to expand these paths into full end-to-end *explicit route* (ER) sequences, i.e., intra-domain RWA expansion done using *fixed alternate routing* (FAR) with *most-used* wavelength selection, as in [8]. The solution is now presented.

A multi-domain network is comprised of D domains, with the i-th domain denoted by sub-graph, $\boldsymbol{G}^i(\boldsymbol{V}^i, \boldsymbol{L}^i)$, $1 \leq i \leq D$, where $\boldsymbol{V}^i = \{ v^i_1, v^i_2, \dots \}$ is the set of OXC nodes and $\boldsymbol{L}^i = \{ l^{ij}_{km} \}$ is the set of links, i.e., l^{ii}_{jk} is intra-domain link from v^i_j to v^i_k in domain i, and l^{ij}_{km} is inter-domain link between k-th border node v^i_j in domain i and m-th border node v^j_k in domain j, where $i \neq j$. The total and free/available wavelengths on link l^{ij}_{km} are also given by w^{ij}_{km} and c^{ij}_{km}, respectively. Hence at the global inter-domain routing level, assuming basic simple node topology abstraction [8], the abstract multi-domain network graph is denoted by $\boldsymbol{H}(\boldsymbol{A},\boldsymbol{E})$, where $\boldsymbol{A}$ is set of vertices representing the D domains and $\boldsymbol{E}$ is the set of physical inter-domain links, i.e., $\mathbf{E} = \{ l^{ij}_{km} \} \surd i \neq j$.

Now further consider failure modeling. Here, it is assumed that a static set of risk profiles are being modeled, and these are captured by M events, each of which can randomly affect multiple inter-domain links, i.e., akin to p-SRLG [7] but at inter-domain level. This set is denoted by $\boldsymbol{R} = \{ \boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_M \}$, where the n-th event r_n occurs with probability φ_n and causes link l^{ij}_{km} to fail with probability $p_n(l^{ij}_{km})$. A link *risk vector* is also then defined to capture susceptibility to all failure events as $\boldsymbol{x}(l^{ij}_{km}) = \{ p_1(l^{ij}_{km}), p_2(l^{ij}_{km}), \dots, p_M(l^{ij}_{km}) \}$.

Now with regards to multi-domain RWA computation, most existing solutions use a variety of techniques (mostly graph-theoretic) to achieve some sort of TE objective. For example, this can include minimizing resource usages (hop counts), load-balancing, and other metrics [1],[8]. Along these lines, the work herein extends these approaches by further incorporating a-priori multi-failure probabilistic information to build *joint* TE and risk-based inter-domain solutions. Specifically, the following schemes are considered:

```
1. Compute $K$-shortest path from src-dest domain in H(A,E)
2. Compute cost of each path using Eq. (1), i.e.,
       LB cost $= \sum_{\text{path links}} \alpha^{ij}_{km}$
3. Rank paths by increasing LB cost
4. Select path with minimum cost
```

Figure 1: Load Balancing (LB) LR Algorithm

A. Load-Balancing (LB)

The end-to-end skeleton LR sequence is chosen to even out load distributions across inter-domain links. Namely, each link is assigned a weighted cost that is inversely proportional to its free wavelengths:

$$\alpha^{ij}_{km} = \frac{1}{c^{ij}_{km} + \varepsilon} \qquad \text{Eq. (1),}$$

where ε is a small quantity to avoid floating-point errors. The LB scheme then computes the K shortest paths between the source and destination domains over the abstract multi-domain graph $\boldsymbol{H}(\boldsymbol{A},\boldsymbol{E})$, and then selects the path with the lowest cost (see Figure 1).

B. Risk Minimization (RM)

This scheme tries to minimize the “end-to-end” risk of the inter-domain skeleton path. First, the maximum “risk exposure” of link l^{ij}_{km} to any multi-failure event is pre-computed as:

$$\beta^{ij}_{km} = \max_n \{ \phi_n p_n(l^{ij}_{km}) \} \qquad \text{Eq. (2).}$$

Assuming a single multi-failure attack occurrence, the probability of the above link not being affected is then bounded by (1- β^{ij}_{km}). Using these risk exposure values, the RM scheme computes the K shortest LR paths between the source and destination domains and selects the one with the minimum risk sum, see Figure 2. Note that this is basically a *static* algorithm, i.e., all inter-domain route combinations can be pre-computed based upon a-priori risk information in the link risk vectors $\boldsymbol{x}(l^{ij}_{km})$.

```
1. Compute $K$-sp from source-dest domains in $\boldsymbol{H}(\boldsymbol{A},\boldsymbol{E})$
2. Compute risk product of each path using Eq. (2), i.e.,
       $\text{Risk Sum} = \frac{\sum_{\text{path links}}(\beta^{ij}_{km})}{\text{path links}}$
3. Select path with minimum risk
```

Figure 2: Risk Minimization (RM) LR Algorithm

C. Joint Strategy (JS)

This scheme achieves a tradeoff between the above two strategies by incorporating both risk and TE objectives. Namely, the goal here is to improve lightpath reliability, as pure load-balancing may yield more susceptible routes. Namely, the JS solution first computes the K shortest LR paths between the source and destination domains over $\boldsymbol{H}(\boldsymbol{A},\boldsymbol{E})$. These routes are then sorted by both their LB costs (Figure 1) and risk sums (Figure 2), with the respective ranks being denoted by χ^{i}_{LB} and χ^{i}_{RM}, $1 \leq \chi^{i}_{\text{LB}}, \chi^{i}_{\text{RM}} \leq K$. Using these values, the K shortest path with the minimum total rank is selected, i.e., $\chi^{i}_{\text{LB}} + \chi^{i}_{\text{LB}}$, as shown in Figure 3.

```
1. Compute $K$-sp from source-dest domains in $\boldsymbol{H}(\boldsymbol{A},\boldsymbol{E})$
2. Compute cost of each path using Eq. (1), i.e.,
       $\text{LB cost} = \sum_{\text{path links}} \alpha^{ij}_{km}$
3. Rank paths by increasing LB cost, i.e., $\chi^{i}_{LB}$ ordering
4. Compute risk product of each path using Eq. (2), i.e.,
       $\text{Risk Sum} = \frac{\sum_{\text{path links}}(\beta^{ij}_{km})}{\text{path links}}$
5. Rank paths by increasing product risk, i.e., $\chi^{i}_{RM}$ ordering
6. Select path with min total rank, i.e., $min_i\{\chi^{i}_{LB} + \chi^{i}_{RM}\}$
```

Figure 3: Joint Strategy (JS) LR Algorithm

IV. PERFORMANCE EVALUATION

The LB, RM, and JS inter-domain lightpath provisioning strategies are tested using the *OPNET Modeler*^{TM} simulation tool. In particular, detailed models are coded to implement all routing, path-computation, and signaling functionalities. Here, two different multi-domain topologies are used as well, including a sample 10-domain topology with 25 inter-domain links (Figure 4) and a modified NSFNET topology (with nodes replaced by domains) with 16 domains/25 inter-domain links (Figure 5). Overall, the latter topology has higher inter-domain connectivity levels, as measured by the ratio of inter-domain links/domains. Meanwhile, all links have 32 wavelengths and the average domain size in each network is set to about 10 nodes. Note that multi-homed interconnectivity is also implemented in the 10-domain network.

Furthermore, all lightpath requests are randomly generated between domains/nodes and have exponential holding times with mean 600 seconds (inter-arrival times varied according to desired load). Meanwhile, the SCF thresholds for intra-/inter-domain routing protocols are both set to 10%, and all inter-domain HT values are set to 120 seconds. Furthermore, in terms of risks, 10 pre-defined multi-failure attack profiles are defined, i.e., M=10, and their associated failure probabilities uniformly distributed to ensure $\sum_n \varphi_n = 1$. Furthermore, the conditional failure probabilities, $p_n(l^{ij}_{km})$, of all links within a domain radius of the n-th multi-failure attack profile are uniformly distributed between (0, 0.01). Finally, the number of inter-domain routes used for LR skeleton path computation is also fixed at K=5 and each run is averaged over 500,000 lightpath requests.

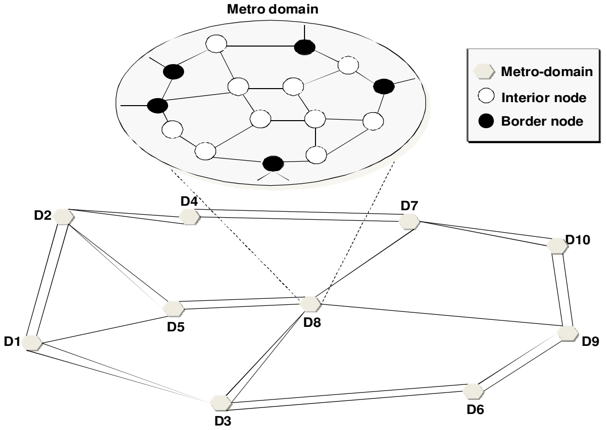

Figure 4: Sample 10-domain topology

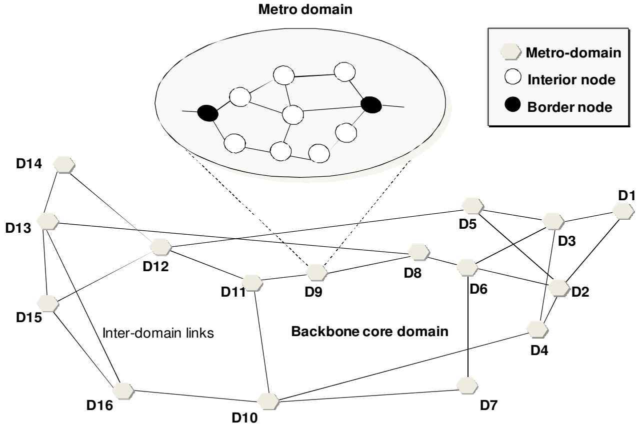

Figure 5: 16-domain modified NSFNET topology

Figures 6 and 7 plot the respective inter-domain lightpath blocking probabilities for the NSFNET and 10-domain topologies. These results indicate that the LB scheme consistently gives the lowest blocking rates, albeit the proposed JS scheme is also quite competitive. For example, considering a sample run of 500,000 requests, the differential between the two strategies is generally limited to under 20 blocked lightpaths. By contrast, the "pure" RM strategy gives the worst blocking of all, as it does not account for any network load information.

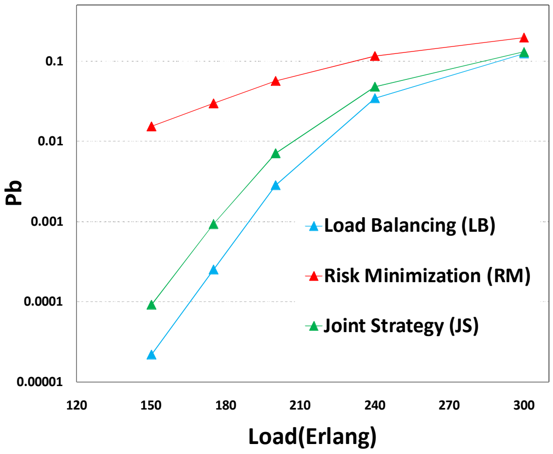

Figure 6: Blocking probability for NSF topology

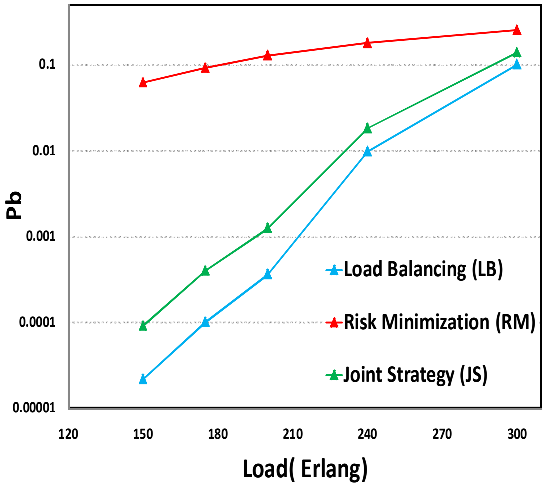

Figure 7: Blocking probability for 10-domain topology

Post-attack failure rates are also measured and plotted in Figures 8 and 9 for the two respective topologies. This metric provides an indication of how robust a particular scheme is to a multi-failure attack. Overall, the findings here show that the proposed JS strategy is very effective indeed, consistently giving under 50% failure rates and even outperforming the pure RM risk-minimization scheme. By contrast, the TE-only LB scheme has very high failure rates, averaging about 100% more than those for the JS scheme. As such, this scheme is clearly not well-suited for multi-failure environments.

Now it is also desirable to measure the resource efficiencies of the various schemes. To do this, the average inter-domain path lengths are plotted in Figures 10 (modified NSFNET) and 11 (10-domain topology). These plots indicate that the JS scheme actually gives the lowest resource usages of all (i.e., highest efficiency), a key finding. In fact, the related path lengths are consistently 5-10% lower than those for the pure TE LB scheme. As expected, the pure RM scheme is the most inefficient here, and all schemes give decreasing path lengths at higher load settings.

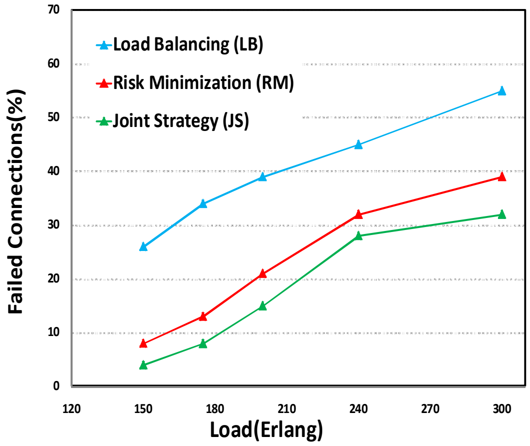

Figure 8: Failed lightpaths for NSF topology

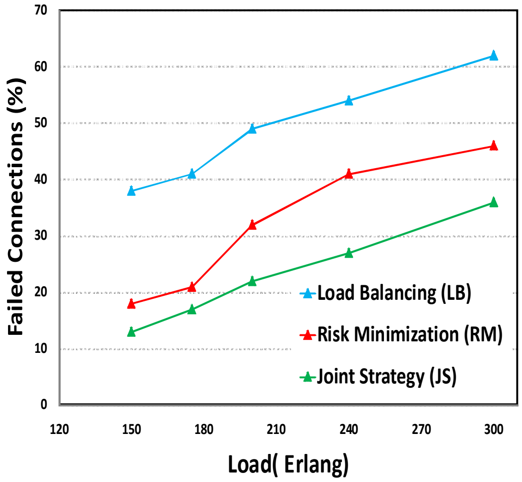

Figure 9: Failed lightpaths for 10-tomain topology

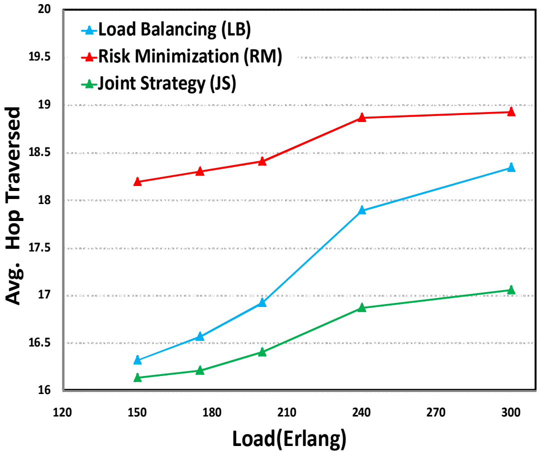

Figure 10: Average path length for NSF topology

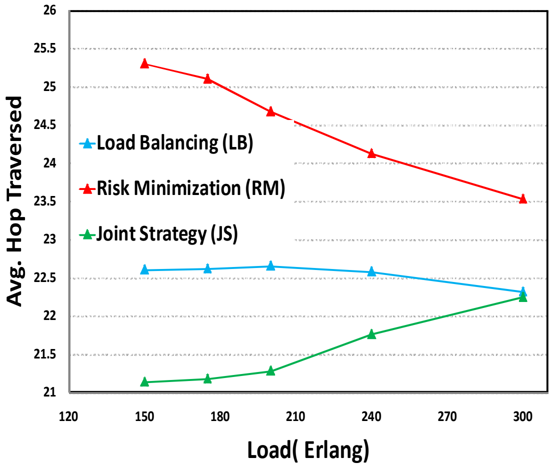

Figure 11: Average path length for 10-domain topology

V. CONCLUSIONS

This paper develops a novel "risk-aware" provisioning solution for lightpath routing in multi-domain optical networks experiencing multiple correlated failures, i.e., as caused by WMD-type attacks. The findings here show that the joint incorporation of TE and risk objectives is very effective in improving overall lightpath reliability. More importantly, the corresponding TE performance is also quite good, yielding minimal increases in blocking and lowest overall resource consumption levels. Future efforts will extend these joint provisioning strategies to multi-domain lightpath protection strategies.

VI. ACKNOWLEDGMENTS

This work has been funded in part by the US Defense Threat Reduction Agency (DTRA, Basic Research Program). The authors are very grateful for this support.

VII. REFERENCES

[1] N. Ghani, *et al*, "Control Plane Design in Multidomain/ Multilayer Optical Networks," *IEEE Communications Magazine,* Vol. 46, No. 6, 2008, pp. 78-87.

[2] D. Staessens, *et al*, "Enabling High Availability Over Multiple Optical Networks," *IEEE Communications Magazine*, Vol. 46, No. 6, June 2008, pp. 120-111.

[3] F. Ricciato, U. Monaco, D. Ali, "Distributed Schemes for Diverse Path Computation in Multidomain MPLS Networks," *IEEE Communications Magazine*, Vol. 43, No. 6, June 2005, pp. 138-146.

[4] A. Sprintson, *et al*, "Reliable Routing with QoS Guarantees for Multi-Domain IP/MPLS Networks," *IEEE INFOCOM 2007*, Anchorage, AL, May 2007.

[5] D. Troung, B. Thiongane, "Dynamic Routing for Shared Path Protection in Multidomain Optical Mesh Networks," *OSA Journal of Optical Networking*, Vol. 5, No. 1, January 2006, pp. 58-74.

[6] S. Stefanakos, "Reliable Routings in Networks with Generalized Link Failure Events," *IEEE/ACM Transactions On Networking*, Vol. 16, No. 6, December 2008, pp. 1331-1339.

[7] H. Lee, E. Modiano, "Diverse Routing in Networks with Probabilistic Failures" *IEEE Globecom 2009*, Honolulu, HI, December 2009.

[8] Q. Liu, *et al*, "Hierarchical Routing in Multi-Domain Optical Networks," *Computer Communications*, Vol. 30, No. 1, 2006, pp. 122-131.

[9] J. Zhang, *et al*, "A Comprehensive Study on Backup Re-Provisioning to Remedy the Effect of Multiple-Link Failures in WDM Mesh Networks*," IEEE ICC 2004,* Paris, France, June 2004.

[10] R. Ramamurthy, *et al*, "Pre-Emptive Reprovisioning in Mesh Optical Networks," *OFC 2003*, Atlanta, GA, March 2003.

[11] H. Choi, S. Subramaniam, "On Double-Link Failure Recovery in WDM Optical Networks," *IEEE INFOCOM 2002*, New York City, NY, June 2002.

[12] D. Schupke, R. Prinz, "Performance of Path Protection and Rerouting for WDM Networks Subject to Dual Failures," *OFC 2003*, Atlanta, GA, March 2003.

[13] S. Kini, *et al*, "Fast Recovery from Dual Link Failures in IP Networks," *IEEE INFOCOM 2009*, Rio De Janeiro, Brazil, April 2009.

[14] K. Xi, J. Chao, "IP Fast Reroute for Double-Link Failure Recovery", *IEEE Globecom 2009*, Honolulu, HI, December 2009.