Diverse Routing in Multi-Domain Optical Networks With Correlated and Probabilistic Multi-Failures

F. Xu[1], N. Min-Allah[2], S. Khan[3], N. Ghani[1]
[1]*University of New Mexico,* [2]*COMSATS Institute of Information Technology,* [3]*North Dakota State University*

Abstract: *Multi-failure network survivability is a key concern for operators owing to the recent spate of natural and man-made disasters. Hence this paper presents a new diverse lightpath protection scheme for random correlated failure recovery in multi-domain optical networks. The solution leverages topology abstraction and hierarchical inter-domain routing and assumes a-priori link failure probability state. The path computation process also takes into account both traffic engineering and risk minimization objectives. The proposed scheme is analyzed using network simulation for a variety of multi-failure scenarios.*

Keywords: *Multi-domain survivability, diverse routing, lightpath protection, correlated/probabilistic failures*

I. INTRODUCTION

Network survivability remains one of the key issues facing operators today. Owing to the highly disruptive impact of natural and man-made catastrophes, there is a pressing need to develop effective recovery techniques to secure services under multiple *correlated* failures, i.e., both temporal and spatial. These faults can arise during the occurrences such as floods, earthquakes, tsunamis, and *weapon of mass destruction* (WMD) attacks. Now in general, it is very difficult to provide rapid guaranteed recovery for all multi-failure combinations. This challenge is further compounded in multi-domain networks due to the reduced visibility across global domains. However, it is still highly desirable to achieve some level of increased recovery under such conditions, i.e., as compared to single link failure recovery schemes. This forms the key motivations for this effort.

In recent years, a range of proactive protection schemes have been proposed for IP *multi-protocol label switching* (MPLS) and optical *generalized MPLS* (GMPLS) networks [1],[2]. These strategies use diverse routing to compute primary/backup path pairs and provide backup routes to resume data transmission in case of primary path failures. Nevertheless, most protection schemes only focus on single-domain settings, as they assume complete topology and resource state information. Moreover, most of these algorithms are only designed to handle single link failures, as they focus on achieving link-disjoint primary/backup paths. However, recent interest in disaster recovery has resulted in some new schemes for handling multiple correlated failure events. In particular, the concept of *shared risk link group* (SRLG) [3] has been generalized into the *probabilistic SRLG* (p-SRLG) notion [4] to facilitate the modeling of multiple failures and design of new protection schemes that minimize failure probabilities. However, these efforts tend to focus on single-domain settings and cannot be applied directly to larger and more complex multi-domain infrastructures which, as a matter of fact, are more vulnerable to expansive correlated failures.

As a result, there is still much scope to develop new survivability schemes for multi-domain networks experiencing correlated/probabilistic multi-failures. To address this need, a novel distributed pre-provisioned protection scheme is presented in this paper. This framework leverages the *path computation element* (PCE)-based architecture [5] introduced by the *Internet Engineering Task Force* (IETF) and implements enhanced topology abstraction and hierarchical routing to achieve both risk mitigation and *traffic engineering* (TE) objectives in the path pair computation.

The rest of this paper is organized as follows. Section II presents a survey of related work on multi-domain and correlated/probabilistic protection. Next, Section III introduces the multi-failure model and Section IV details the proposed inter-domain protection solution. Numerical results analysis is then provided in Section V, followed by the conclusions in Section VI.

II. BACKGROUND

Multi-domain protection is gaining increased focus today. In general, related strategies can be classified into two categories according to the level of “domain diversity” provisioned between the primary and backup routes, i.e., *per-domain* and *end-to-end* [2]. Specifically, per-domain schemes select the same sequence of domains for each path pair from the source to the destination [6],[7]. For example, [6] provides a comprehensive look at the per-domain protection and then proposes a new link-disjoint path pair computation solution using sequential signaling along the pre-assigned domain sequence. Nevertheless, despite their simplicity and the ability to prevent intra-domain alarm propagation across domain boundaries, these solutions require increased domain-to-domain connectivity and are more susceptible to multiple failures. By contrast, the end-to-end protection strategies provide better domain diversity between the primary and backup routes and can achieve load balancing as well. However, these solutions require some level of abstract “global” topology and resource information and use a “two-step” distributed path computation approach. Namely, a *loose route* (LR) pair is first computed at the inter-domain level to resolve the end-to-end primary and backup domain sequences. These sequences are then expanded using *explicit route* (ER) signaling via the *resource reservation protocol-TE* (RSVP-TE) [8] and/or *PCE protocol* (PCEP) [9]. For example, [10] has studied inter-domain diverse path pair routing in IP/MPLS networks, using full-mesh topology abstraction to extract domain traversal characteristics and Suurballe's algorithm [11] to compute two link-disjoint paths. This scheme is formally shown to be able to find two diverse paths with minimum total cost, but the abstraction overheads are excessive, i.e., O(N^4) for N border nodes in a domain. Meanwhile, [12] develops a modified scheme that requires less detailed abstract state, i.e., O(N^2) overheads, and ensures the intra-domain route disjointness during ER expansion. However, substantial work is still needed to address multi-failure cases.

As mentioned earlier, new studies on multi-failure recovery have also been done and these efforts follow along two directions, i.e., network design and service reliability. Specifically, the former types analyze the impact of potential geographical/regional failures on the given network topologies. For example, [13]-[15] study network vulnerability under various failure models and identify critical locations yielding maximum capacity or connectivity disruption. These findings can be used to improve the design and maintenance of fault-tolerant network infrastructures. Meanwhile, service reliability schemes focus on improved path protection (in single domains). For example, [16] presents a dual-link failure recovery scheme for IP-tunneling networks, whereas [17] develops heuristic solutions for networks with multiple independent link failures. However, these efforts assume independent link failure models and do not capture any inter-failure correlation effects.

To address these concerns, a new p-SRLG model is proposed in [4] by defining *a-priori* probabilities for specific events and link failures (resulting from these events). An optimal *integer linear programming* (ILP) formulation and heuristic strategies are then developed to minimize the joint failure probabilities of primary/ backup path pairs under these p-SRLG events. However, these schemes only focus on risk minimization, and tend to yield lengthy paths (with likely much lower TE efficiencies). Hence there is a further need to study the multi-failure protection schemes in multi-domain settings. In particular, the p-SRLG concept needs to be defined across domain boundaries. Furthermore, there is also a need to incorporate TE constraints into the (LR pair) computation process in order to improve resource efficiencies. These challenges are now addressed.

III. CORRELATED/PROBABILISTIC FAILURE MODEL

The proposed correlated multi-failure model used in this work is based upon and extension of the single-domain p-SRLG framework in [4]. Consider the requisite notation first. A multi-domain optical network is assumed to comprise D domains with the i-th domain having n^i *optical cross-connect* (OXC) nodes and b^i border OXC nodes, $1 \leq i \leq D$. This network is represented as a set of domain sub-graphs, $\boldsymbol{G}^i(\boldsymbol{V}^i,\boldsymbol{L}^i)$, where $\boldsymbol{V}^i = \{v^i_1, v^i_2, \ldots\}$ is the set of OXC nodes in domain i and $\boldsymbol{L}^i = \{l^{ii}_{jk}, \forall\, v^i_j$ and v^i_k if they are connected$\}$ is the set of intra-domain links ($1 \leq i \leq D$, $1 \leq j,k \leq n^i$) with available and maximum numbers of wavelengths λ^{ij}_{km} and Λ^{ij}_{km}, respectively. The inter-domain link l^{ij}_{km} is defined similarly ($1 \leq i,j \leq D$, $1 \leq k \leq b^i$, $1 \leq m \leq b^j$).

Meanwhile, the predefined set of correlated p-SRLG events is represented by $\boldsymbol{E}$, where each event $e \in \boldsymbol{E}$ has an occurrence probability of $\pi_e \in [0, 1]$. Furthermore, it is assumed that these large-scale events are sufficiently rare so as to be mutually exclusive, i.e., $\sum_{e \in E} \pi_e = 1$. Finally, when a p-SRLG event e occurs, it is assumed to affect a *circular* failure region within which all links fail *independently* with predefined probabilities, i.e., link l fails with probability $p_e(l) \in [0,1]$ (link subscripts/ superscripts removed for simplicity). Since a link can be located in multiple overlapping p-SRLG regions, a related risk vector is also defined, $\bar{\boldsymbol{p}}_l = [p_1, p_2, \ldots, p_{|\boldsymbol{E}|}]$.

IV. MULTI-DOMAIN PROTECTION SCHEMES

Using the above multi-failure model, a novel inter-domain protection scheme is developed to compute link-disjoint lightpath pairs with reduced *joint* failure risk probabilities. Since the computation of the optimal path pair with minimum joint failure probability is an NP-*complete* problem [4], this effort focuses on a heuristic approach to reduce both risk and TE cost. The performance is then compared versus other heuristic strategies that only account for one of these two objectives (Section V). Consider some assumptions.

The proposed inter-domain protection scheme assumes realistic distributed optical GMPLS-based settings in which OXC nodes and domain PCEs have full visibility of intra-domain link state, e.g., via *open shortest path first-TE* (OSPF-TE) routing. Domain PCEs also have partial inter-domain views, as provided by PCEP. The framework also assumes "all-optical" domains with full wavelength conversion at border OXC nodes, i.e., a realistic modeling of operational settings with regeneration and bit-monitoring at boundaries. Finally, all setup signaling is done using the RSVP-TE protocol.

A. *TE-Only Diverse Routing (TE-Only) Algorithm*

A baseline TE-only diverse routing strategy from [12] is first presented, i.e., one that does not consider link failure probabilities. The algorithm uses topology abstraction to generate partial "global" views, thereby enabling domain PCEs to compute end-to-end "skeleton" inter-domain routes. Specifically, two well-studied abstraction schemes are used here:

Simple-Node (SN) Abstraction: Each domain is condensed into an abstract node emanating all physical inter-domain links. This abstraction entails no added inter-domain routing overheads.

Full-Mesh (FM) Abstraction: Each domain is reduced into a mesh of intra-domain *abstract* links connecting border OXC node pairs to provide domain traversal costs. The abstract link capacity is computed as the maximum of the bottleneck capacities of the *k-shortest paths* (*k*-SP) between the corresponding border OXC nodes.

Detailed link-state information on physical inter-domain links and abstract intra-domain links (for full-mesh abstraction only) is exchanged between the domain PCEs to build and maintain an abstract topology. Hence, upon the arrival of a lightpath setup request, the source domain PCE performs diverse LR pair routing over the abstract topology using Suurballe's algorithm [11]. ER expansion then proceeds along the LRs to set up the end-to-end lightpaths. Note that Suurballe's algorithm is re-run at the intra-domain level in the case where both the primary and backup LRs traverse the same domain, i.e., to guarantee intra-domain path diversity; see [12].

Overall, "pure" TE-based path computation uses two common link weighting schemes. Namely, the link "cost", $\omega(l)$, is defined as either:

$$\omega(l) = 1, \tag{1}$$

or

$$\omega(l) = 1 / (u \cdot \lambda(l) / \Lambda(l)), \tag{2}$$

where u is a constant for scale purpose. Essentially, Eq. (1) minimizes the inter-domain *hop count* (HC) of a route to improve resource efficiency, whereas Eq. (2) pursues *load balancing* (LB) over links with fewer reserved wavelengths. The overall TE-only approach is summarized below.

TE-Only Algorithm

Given: Simple-node or full-mesh topology abstraction

1. Set link weights with $\omega(l) = 1$ or $1 / (u \cdot \lambda(l) / \Lambda(l))$
2. Find LR pair with minimum total cost via Suurballe's algorithm
3. Expand LRs via RSVP-TE signaling
4. Apply Suurballe's algorithm in domains traversed by both LRs

B. *Risk Minimization-Only (RM-Only) Algorithm*

This approach extends the heuristic risk-minimization scheme in [4] for multi-domain operation. In particular, LR path pair computation is done using inter-domain link failure probabilities. A new topology abstraction scheme is developed to introduce "p-SRLG-aware" state into the full-mesh topology abstraction. Specifically, the weight of each abstract link is set to the failure probability of the minimum-overall-risk path, L, between the corresponding border OXC node pair. However, computing such a path is known to be a convex maximization (or concave minimization) problem that is generally NP-*hard* [4]. Hence an approximation is proposed here by assuming low failure probabilities, i.e., $p_e(l) \ll 1$. Namely, the link weights are set to:

$$\omega(l) = \sum_{e \in E} (\pi_e \cdot p_e(l)), \tag{3}$$

Using Eq. (3), the path L can be determined by applying Dijkstra's shortest path algorithm. Next, the risk vector of the abstract link can be derived from L as:

$$\overline{\boldsymbol{p}} = [1 - \prod_{l \in L} (1 - p_e(l)),\ \forall e \in \boldsymbol{E}], \tag{4}$$

Now given a lightpath request, the weights of the abstract links and inter-domain links are assigned according to Eq. (3) and the primary LR, L_{pr}, is first computed. Subsequently, all links used by L_{pr} are pruned from the abstract topology for path disjointness, and a backup LR is computed over the remaining network with the link weights re-adjusted as:

$$\omega(l) = \sum_{l_{pr} \in L_{pr}} \sum_{e \in E} (\pi_e \cdot p_e(l) \cdot p_e(l_{pr})), \tag{5}$$

where l_{pr} is one of the links on L_{pr}; see [4]. Here, if the LR pair is successfully found, at the intra-domain level, each ER segment between the given ingress and egress border OXC nodes on the primary lightpath is expanded in exactly the same manner as the abstract link computation. The full primary route and its link risk vectors are then substituted into Eq. (5) for the expansion of end-to-end backup lightpath.

Overall, the RM-only heuristic is a greedy algorithm which first computes a primary lightpath with minimum overall path failure probability and then likewise computes a backup path (with adjusted link weights). The overall RM algorithm is summarized below.

RM-Only Algorithm

Given: Simple-node or full-mesh p-SRLG-aware topology abstraction

1. Set link weights with $\omega(l) = \sum_{e \in E} (\pi_e \cdot p_e(l))$
2. Find the primary LR with minimum overall risk
3. Remove all the links used by the primary LR and set $\omega(l) = \sum_{l_{pr} \in L_{pr}} \sum_{e \in E} (\pi_e \cdot p_e(l) \cdot p_e(l_{pr}))$ for the backup LR
4. With $\omega(l) = \sum_{e \in E} (\pi_e \cdot p_e(l))$, expand intra-domain primary ER segments with minimum overall risk
5. Substitute primary lightpath link risks into $\omega(l) = \sum_{l_{pr} \in L_{pr}} \sum_{e \in E} (\pi_e \cdot p_e(l) \cdot p_e(l_{pr}))$ and expand backup ER

C. *Joint Risk Minimization & TE (RM+TE) Algorithm*

The proposed joint scheme implements path pair computation via two key steps. Namely, 1) k candidate route pairs are first computed using the TE link weights, and 2) the optimal pair is then selected from the candidate pairs in terms of lowest failure risk. Hence this algorithm incorporates both objectives with reduced complexity. Consider the details.

Foremost, the proposed RM+TE scheme uses a modified p-SRLG-aware topology abstraction approach. Specifically, an intra-domain route is chosen for each abstract link by selecting from the k-SP (between the respective border nodes) with minimum overall risk, P_L:

$$P_L = \sum_{e \in E} (\pi_e \cdot (1 - \prod_{l \in L} (1 - p_e(l))) \, , \qquad (6)$$

The risk vector of the selected path, given by Eq. (4), is then assigned to the corresponding abstract link. Using this, inter-domain path pair computation is done by computing k_1 primary LR candidates over the abstract topology with TE-based link weights, Eq. (1) or (2). Next, for each primary LR, up to k_2 backup LRs are computed under the link-disjoint requirement. From these $k_1 \cdot k_2$ candidate combinations, the LR pair with lowest failure probability *dot-product* is then selected. Specifically, this value is defined as follows:

$$\sum_{e \in E} (\pi_e^2 \cdot (1 - \prod_{l \in L_1} (1 - p_e(l))) \cdot (1 - \prod_{l \in L_2} (1 - p_e(l)))) \, , \quad (7)$$

where L_1 and L_2 are the paired primary/backup LRs. Finally, ER expansion is done using sequential signaling. Now if a domain is traversed by a single LR, dot-products are computed between the intra-domain candidate routes and the other LR. Otherwise, if a domain lies on both the primary and backup LRs, the intra-domain path pair is chosen from $k_1 \cdot k_2$ intra-domain candidates by comparing their dot-products.

Overall, Eq. (7) quantifies the level of "p-SRLG diversity" between primary/backup LRs and helps avoid path pairs with high risk correlation. Carefully note that the choice of the k, k_1, and k_2 parameters can also affect the tradeoff between risk minimization or TE-based selection. For example, increased values may improve selection of lower risk paths but may compromise TE efficiency (and vice versa). The RM+TE algorithm is summarized below.

RM+TE Algorithm

```
Given: Modified simple-node or full-mesh p-SRLG-aware abstraction for RM and TE
1. Set link weights with $\omega(l) = 1$ or $1 / ( u \cdot \lambda(l) / \Lambda(l) )$
2. Find $k_1 \cdot k_2$ link-disjoint LR pair candidates
3. Select LR pair with lowest dot-product expressed as:
   $\sum_{e \in E} (\pi_e^2 \cdot (1 - \prod_{l \in L_1} (1 - p_e(l))) \cdot (1 - \prod_{l \in L_2} (1 - p_e(l))))$
4. Expand LRs by selecting intra-domain paths yielding
   the lowest dot-product
```

V. PERFORMANCE EVALUATION

The multi-failure protection schemes are evaluated using *OPNET Modeler*^{TM} with a modified multi-domain version of the NSFNET topology, i.e., nodes replaced by domains, Figure 1. All link capacities are fixed to 16 wavelengths and lightpath requests are randomly generated between domains/intra-domain nodes with exponential holding times (mean 600 sec) and varying exponential inter-arrival times (as per the traffic loads). Meanwhile, to achieve sufficient route variability, a value of k=3 is chosen for the full-mesh abstraction, along with values of k_1=3 and k_2=2 for the RM+TE scheme. Finally, five mutually exclusive p-SRLGs are tested on the network in Figure 1, i.e., $|\boldsymbol{E}|$=5. Here, each simulation is averaged over 1,000 failure occurrences, and link failures within each p-SRLG region are independently generated with random probabilities.

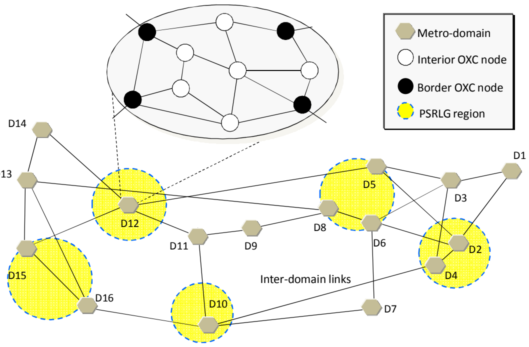

Fig. 1: Modified NSFNET topology with correlated failures

Initially, all five p-SRLG events are assumed to be equiprobable, i.e., π_e=0.2, $\forall \, e \in \boldsymbol{E}$, and tests are done to measure the recovery performance of varying combinations of abstraction and link weighting schemes. In particular, Figure 2 plots the *protection failure rates* of the various schemes, defined as the percentage of connections experiencing both primary *and* backup lightpath failures during a p-SRLG event. These results indicate that the TE-only schemes yield notably higher failure rates, as they do not incorporate link failure probabilities. Meanwhile, the new RM+TE algorithms give the best recovery, even outperforming the RM-only schemes. The reasons are twofold. First, the RM-only scheme uses a greedy approach, which may not result in the lowest path pair failure probability. Second, the RM-only heuristic works under the assumption of low failure probability, and hence its performance may abate in cases where this does not hold very well. Note that the use of p-SRLG-aware full-mesh abstraction also gives better multi-failure recovery versus more basic simple-node aggregation. This is expected as enriched intra-domain risk information allows PCEs to find path pairs with lower failure risks, especially when some p-SRLGs only affect intra-domain regions.

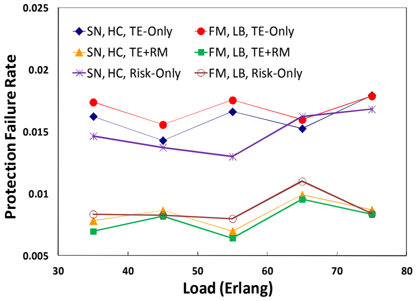

Fig. 2: Protection failure rate

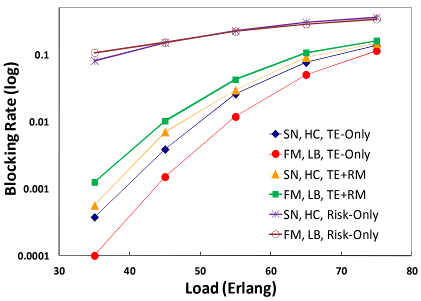

Fig. 3: Lightpath blocking rate

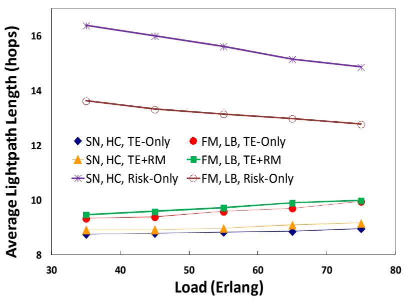

Fig. 4: Average lightpath length

Lightpath blocking rates are also measured to gauge the "load-carrying capacity" of the schemes, Figure 3. As expected, the TE-only algorithms give the lowest blocking, especially with full-mesh abstraction and load-balancing weighting. By contrast the RM-only schemes give unacceptably high blocking rates, barely dropping below 10% at very light loads. Meanwhile, the proposed RM+TE solution achieves a nice tradeoff between these two regimes, following more closely to the TE-only strategies. Finally, resource usages are evaluated by measuring the average end-to-end lightpath lengths, as plotted in Figure 4. Again, the proposed RM+TE algorithm gives very competitive results as compared to the TE-only strategy, i.e., less than 5% increase in path length. On the contrary, the RM-only solutions are much less efficient, averaging anywhere from 30-65% higher hop-count values.

Note that additional tests are also conducted for *uneven* p-SRLG event occurrence probabilities. However, since the results of these cases are very similar to those of even occurrence probabilities, they are omitted here.

VI. CONCLUSIONS

This paper presents a novel multi-domain lightpath protection solution to recover from multiple correlated (probabilistic) link failures. The solution implements diverse intra-domain routing and combines both risk reduction and TE objectives into the path pair selection process. This approach also introduces some novel risk-based full-mesh topology abstraction schemes. Overall simulation results show that the proposed solution yields very good failure recovery along with highly competitive request blocking and resource utilization behaviors, i.e., as compared to the pure TE strategy.

ACKNOWLEDGEMENTS

This work has been supported by the Defense Threat Reduction Agency. The authors are very grateful for this support.

REFERENCES

[1] P. Cholda, *et al*, "A Survey of Resilience Differentiation Frameworks in Communication Networks", *IEEE Communications Surveys & Tutorials*, 4th Quarter, 2007.

[2] N. Ghani, M. Peng, A. Rayes, "Service Provisioning and Survivability in Multi-Domain Optical Networks", *WDM Systems and Networks: Modeling, Simulation, Design and Engineering*, Springer, 2012.

[3] D. Papadimitriou, *et al*, "Inference of Shared Risk Link Groups", *Internet draft*, November 2001.

[4] H. Lee, *et al*, "Diverse Routing in Networks With Probabilistic Failures", *IEEE/ACM Transactions on Networking*, Vol. 18, No. 6, May 2010, pp. 1895-1907.

[5] A. Farrel, *et al*, "A Path Computation Element (PCE)-Based Architecture", *IETF RFC 4655*, August 2006.

[6] T. Takeda, *et al*, "Diverse Path Setup Schemes in Multi-Domain Optical Networks", *IEEE BROADNETS 2008*, London, England, September 2008.

[7] T. Takeda, *et al*, "Analysis of Inter-Domain Label Switched Path (LSP) Recovery", *IETF RFC 5298,* August 2008.

[8] D. Awduche, *et al*, "RSVP-TE: Extensions to RSVP for LSP Tunnels", *IETF RFC 3209,* December 2001.

[9] J. Ash, J. Le Roux, "A Path Computation Element (PCE) Communication Protocol Generic Requirements", *IETF RFC 4657*, September 2006.

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

[11] J. Suurballe, R. Tarjan. "A Quick Method for Finding Shortest Pairs of Disjoint Paths", *Networks*, Vol. 14, pp. 325–336, 1984.

[12] F. Xu, *et al*, "Novel Path Protection Scheme for Multi-Domain Networks", *IEEE GLOBECOM 2011 Joint Workshop on Complex Networks and Pervasive Group Communications*, Houston, TX, December 2011.

[13] A. Sen, *et al*, "Region-Based Connectivity - A New Paradigm for Design of Fault-Tolerant Networks", *IEEE HPSR 2009*, Paris, France, June 2009.

[14] S. Neumayer, *et al*, "Assessing the Vulnerability of the Fiber Infrastructure to Disasters", *IEEE INFOCOM 2009*, Rio de Janeiro, Brazil, April 2009.

[15] P. Agarwal, *et al*, "The Resilience of WDM Networks to Probabilistic Geographical Failures", *IEEE INFOCOM 2011*, Shanghai, China, April 2011.

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

[17] Q. She, X. Huang, J. Jue, "Maximum Survivability under Multiple Failures", *OFC/NFOEC 2006*, Anaheim, CA, March 2006.