Multi-Domain DWDM Network Provisioning For Correlated Failures

M. Esmaeili[1], M. Peng[2], S. Khan[3], J. Finochietto[4], Y. Jin[5], N. Ghani[1]
[1]*University of New Mexico*, [2]*Wuhan University*, [3]*North Dakota State University*, [4]*Universidad Nacional de Córdoba*
[5]*Shanghai Jiao Tong University*

Abstract: A novel multi-domain path routing solution is proposed for improving connection reliability under multiple correlated failure events. The formulation assumes probabilistic failures and jointly incorporates traffic engineering objectives.

1. Introduction

Network survivability is a key area and many schemes have been studied for optical *dense wavelength division multiplexing* (DWDM) networks. These include pre-provisioned protection strategies for working/ backup path computation as well as post-fault restoration. With expanding deployments, multi-domain protection schemes have also been evolved for larger backbone networks [1]. For example, SONET/SDH interconnection has been extended for "localized" per-domain protection in [2], whereas others have also tabled more extensive "end-to-end" link/domain disjoint strategies [1],[3],[4].

However, most optical survivability schemes are designed for single node/link failures. Given the proliferation of DWDM backbones, there is growing interest in developing schemes to handle multiple near-simultaneous, i.e., *correlated*, failures such as those resulting from large power outages, natural disasters, *weapons of mass destruction* (WMD) attacks, etc. Now some studies have emerged in this area. For example, [5] details reliable path computation under generalized failures and introduces two notions of reliability, i.e., local and global. The former chooses routes spanning the lowest number of failure events, while the latter minimizes distortion after a failure event. Meanwhile, [6] maximizes the reliability of pre-computed routes against correlated probabilistic failures and extends the static *shared risk link group* (SRLG) concept to a *probabilistic SRLG* (p-SRLG) definition. Although these solutions improve resilience, they are designed for "non-optical" single-domains where computational entities have full knowledge of global resources. Clearly these assumptions are not valid in larger, distributed *multi-domain* (DWDM) networks where only select border nodes have partial/dated "global" state [1]. Moreover, since [5] and [6] focus on risk minimization, these schemes may yield lower *traffic engineering* (TE) efficiencies as well.

Along these lines, this paper proposes a novel multi-domain DWDM *routing and wavelength assignment* (RWA) solution to improve lightpath resiliency (reliability) against multiple correlated (i.e., near simultaneous) failure events and also handle TE concerns. The solution assumes realistic distributed settings using hierarchical inter-domain routing based upon the *generalized multi-protocol label switching* (GMPLS) suite [1]. The paper is organized as follows. Section II presents the notation and algorithmic psuedocode. Detailed simulation results are then presented in Section III along with conclusions.

2. Lightpath Provisioning for Joint TE and Risk Minimization

A novel scheme is proposed to improve lightpath reliability in multi-domain DWDM networks with correlated failures. The goal is to use a-priori probabilistic risk state on susceptible links to jointly model risk and TE objectives. The setup assumes "all-optical" domains with full opto-electronic conversion at border gateway nodes. This is a very valid assumption in larger (inter-carrier) settings, as it mitigates physical layer impairments and also enables bit-level service monitoring at domain boundaries. As per the GMPLS framework, all *optical cross-connects* (OXC) at the intra-domain level run link-state routing and have full domain visibility, e.g., *open-shortest-path first* (OSPF-TE). Meanwhile, at the inter-domain level, select border OXC nodes run "hierarchical" link-state routing to propagate updates for inter-domain links and build an abstract multi-domain graph, i.e., second level of OSPF-TE with simple node abstraction [7]. Here inter-domain link updates are generated using *significance change factor* (SCF) threshold policies as well as *hold-down* (HT) timers to limit excessive overheads, see [7]. Using this inter-domain routing state, end-to-end multi-domain path computation is done using a *path computation element* (PCE) framework [1]. Specifically, source domain PCE entities first compute "skeleton" *loose route* (LR) paths to the destination domains over the abstract graph. RSVP-TE signaling is then used to expand these paths into full end-to-end *explicit route* (ER) sequences, i.e., using *fixed alternate routing* (FAR) with *most-used* wavelength selection in the domains, as in [7]. Herein, the focus is on LR computation for working-only routes.

Now consider the requisite notation. A multi-domain DWDM network is comprised of D domains, with the i-th domain denoted by a sub-graph, $\boldsymbol{G}^i(\boldsymbol{V}^i, \boldsymbol{L}^i)$, $1 \le i \le D$, where $\boldsymbol{V}^i = \{ v^i_1, v^i_2, \dots \}$ is the set of domain OXC nodes and $\boldsymbol{L}^i = \{l^{ij}_{km}\}$ is the set of domain 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 $i \neq j$. In addition, the total and available wavelengths on link l^{ij}_{km} are given by w^{ij}_{km} and c^{ij}_{km}, respectively. Hence at the inter-domain routing level, assuming basic simple node topology abstraction [7], 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., $\mathbf{E}=\{ l^{ij}_{km} \} \surd\ i \neq j$. A pre-determined set of M failure events is also defined, each of which can simultaneously impact a random number of inter-domain links within a given proximity, i.e., akin to p-SRLG [6]. This set is denoted by $\boldsymbol{R}=\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_M\}$, and event r_n occurs with probability φ_n and causes link l^{ij}_{km} to fail with probability $p_n(l^{ij}_{km})$. Hence for each link a *risk vector* for the failure events is defined as $\boldsymbol{x}(l^{ij}_{km})=\{p_1(l^{ij}_{km}), p_2(l^{ij}_{km}), \ldots, p_M(l^{ij}_{km})\}$. In general, these link-specific probabilities can be derived a-priori using off-line analyses (e.g., as per geographical location, weather patterns, geopolitical constraints, etc) and then disseminated using inter-domain link-state routing.

Now consider the design of LR skeleton path computation schemes for inter-domain RWA. To date, most earlier work on multi-domain DWDM network provisioning has focused on TE objectives only, e.g., such as hop-count minimization, load-balancing, other joint cost metrics [1],[7]. In general, related findings show that load-balancing can be very effective at distributing traffic away from critical links and thereby lowering lower blocking (versus pure hop count minimization). Building upon this, the work here proposes to combine the above TE-based strategies with a-priori risk information to build a *joint* LR computation scheme. Specifically, the following graph-theoretic heuristic schemes are considered:

Load-Balancing (LB): Here the PCE computes the end-to-end LR sequence on the global abstract graph with minimum total cost. Namely, each inter-domain link is assigned a cost which is inversely proportional to its set of available (free) wavelengths as:

$$\alpha^{ij}_{km} = \frac{w^{ij}_{km} - c^{ij}_{km}}{w^{ij}_{km}} \qquad \text{Eq. (1).}$$

The LB scheme then computes the K *shortest paths* (K-sp) between the source and destination domains over the abstract multi-domain graph $\boldsymbol{H}(\boldsymbol{A},\boldsymbol{E})$ using a simple hop count metric (excluding links without any wavelengths). This identifies the LR border node sequence for subsequent ER expansion via RSVP-TE signaling. The total cost of each path is then derived as sum of all constituent link costs, i.e., Eq. (1), and the lowest cost path selected (see Fig. 1).

```
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.,
       LB cost $= \sum_{\text{path links}} \alpha^{ij}_{km}$
3. Select path with minimum cost
```

Figure 1: Load Balancing (LB) LR Algorithm

Risk Minimization (RM): Here the PCE computes the end-to-end LR sequence with the minimum failure probability. To determine this path, the maximum risk exposure of each link l^{ij}_{km} to any multi-failure event is first pre-computed as:

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

Hence the probability of l^{ij}_{km} not being affected by a multi-failure attack is upper-bounded by the term $(1-\beta^{ij}_{km})$. Using the this, the RM scheme again computes the K shortest LR paths between the source and destination domains and selects the one with the minimum risk product (maximum reliability), see Fig. 2. Overall, this algorithm implements *static* LR computation at the inter-domain level, i.e., as it only maximizes path reliability according to pre-specified probabilistic risks (with the exception of dynamically removing fully-blocked links in $\boldsymbol{H}(\boldsymbol{A},\boldsymbol{E})$ before the K-sp computation).

```
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.,
       Risk Product $= 1 - \prod_{\text{path links}} (1 - \beta^{ij}_{km})$
3. Select path with minimum risk
```

Figure 2: Risk Minimization (RM) LR Algorithm

Joint Strategy (JS): Generally, the LB scheme may yield routes with higher failure probabilities whereas the RM scheme may yield longer and less resource-efficient routes. Hence to achieve a better balance between the two, the JS scheme is proposed to incorporate both strategies via a ranking method. Namely, the K shortest paths between the source and destination domains are first computed on $\boldsymbol{H}(\boldsymbol{A},\boldsymbol{E})$. These routes are then sorted according to their LB costs (Fig. 1) and also risk products (Fig. 2), with the integral ranks denoted as χ^{i}_{LB} and χ^{i}_{RM}, respectively, where $1 \leq \chi^{i}_{LB}, \chi^{i}_{RM} \leq K$. The LR path with the minimum total rank is then selected, i.e., $\chi^{i}_{LB}+\chi^{i}_{LB}$ (see Fig. 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.,
       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.,
       Risk Product $= 1 - \prod_{\text{path links}} (1 - \beta^{ij}_{km})$
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

3. Performance Evaluation

The various inter-domain RWA strategies are tested using detailed *OPNET Modeler*^{TM} models. Tests are done using a modified NSFNET topology with 16 domains/25 inter-domain links, i.e., average of 1.56 links/domain. Namely, each node in NSFNET is now replaced by a domain with about 15 nodes each. All links (intra/inter-domain) have 32 wavelengths and lightpath requests are randomly generated between domains/nodes, with each run comprising of 500,000 requests with exponential holding times (mean 600 sec). Meanwhile, the inter-domain routing update SCF thresholds are set to 10%, and two inter-domain HT values are tested, 120 sec and 300 sec. In addition, a total of M=10 localized failure scenarios (p-SRLG's) are defined over the global topology and 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 p-SRLG scenario are uniformly distributed between (0,0.01). Finally K=5 inter-domain shortest paths are computed and searched when generating the LR sequences for each scheme (Figs. 1-3).

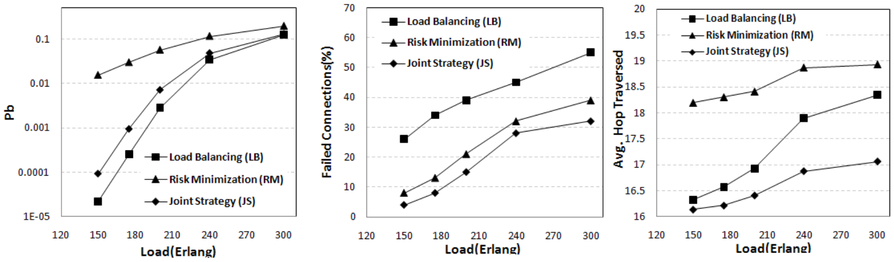

Figure 4: Performance Results w. 120 HT updates: a) blocking, b) failed connections, c)avg. length

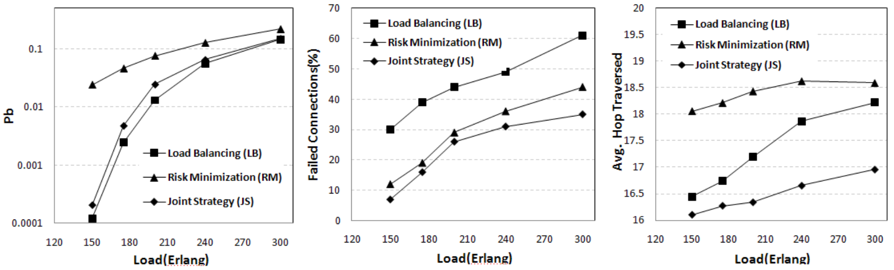

Figure 5: Performance Results w. 300 HT updates: a) blocking, b) failed connections, c)avg. length

Simulation results are presented for the two inter-domain routing timer values, 120 sec (Fig. 4) and 300 sec (Fig. 5). First, the inter-domain blocking results are shown in Fig. 4a/5a and indicate the lowest blocking with the LB scheme, with the JS scheme coming in a close second. Namely, for low blocking regimes in the 10^{-4}-10^{-5} range, the difference between the LB and JS schemes is under 15 failed connections out of 500,000. Meanwhile, the RM scheme gives unacceptably high blocking, over 10^{-2} in all cases. Next, the resiliency of the LR strategies against multi-failure attacks is tested by measuring post-attack failure rates of working lightpaths in Fig. 4b/5b. These results show that the JS scheme actually outperforms the pure risk RM scheme, i.e., under 35% failures even under heavy loads. Moreover, the pure TE-based LB scheme yields the highest failures despite giving lowest blocking, i.e., about 100% higher than the JS scheme. Finally, the resource utilization of these inter-domain RWA strategies is also gauged by plotting the average hop counts in Fig. 4c/5c. Here the JS scheme gives the lowest consumption, consistently below than for pure TE or risk strategies (by 5-10%).

Overall, these results indicate that the joint incorporation of TE and risk minimization objectives in inter-domain RWA is very effective for lowering lightpath susceptibility to large correlated failure events. This approach is shown to give the highest post-fault reliability along with reduced resource consumption, with a very small increase in blocking performance. Moreover, these relative gains tend to hold across a range of inter-domain routing update timer values as well. Future efforts will look at incorporating this strategy for working/protection path pair routing in multi-domain DWDM networks.

4. References

[1] N. Ghani, *et al*, "Control Plane Design in Multidomain /Multilayer Optical Networks," *IEEE Comm. Mag,* Vol. 46, No. 6, 2008, pp. 78-87.
[2] F. Ricciato, U. Monaco, D. Ali, "Distributed Schemes for Diverse Path Computation in Multidomain MPLS Networks," *IEEE Communications Magazine*, June 2005, pp. 138-146.
[3] A. Sprintson, *et al*, "Reliable Routing with QoS Guarantees for Multi-Domain IP/MPLS Networks," *IEEE INFOCOM 2007*, May 2007.
[4] D. Truong, B. Thiongane, "Dynamic Routing for Shared Path Protection in Multi-Domain Optical Mesh Networks," *OSA Journal of Optical Networks*, Vol. 5, Jan. 2006, pp. 58-74.
[5] S. Stefanakos, "Reliable Routings in Networks with Generalized Link Failure Events," *IEEE/ACM Transactions On Networking,* Vol. 16, No. 6, Dec. 2008, pp. 1331-1339.
[6] H. Lee, E. Modiano, "Diverse Routing in Networks with Probabilistic Failures" *IEEE Globecom 2009*, Honolulu, HI, December 2009.
[7] Q. Liu, *et al*, "Hierarchical Routing in Multi-Domain Optical Networks," *Computer Communications*, Vol. 30, No. 1, 2006, pp. 122-131.