Network Survivability for Multiple Probabilistic Failures

Oscar Diaz, Feng Xu, Nasro Min-Allah, Mahmoud Khodeir, Min Peng, Samee Khan, and Nasir Ghani

***Abstract*—Network service recovery from multiple correlated failures is a major concern given the increased level of infrastructure vulnerability to natural disasters, massive power failures, and malicious attacks. To properly address this problem, a novel path protection solution is proposed to jointly incorporate traffic engineering and risk minimization objectives. The framework assumes probabilistic link failures and is evaluated against some existing multi-failure recovery schemes using network simulation.**

***Index Terms*—Multi-failure recovery, network survivability.**

I. Introduction

MANY different types of pre-provisioned recovery schemes have been developed for IP *multi-protocol label switching* (MPLS) and optical *dense wavelength division multiplexing* (DWDM) networks [1],[2]. These solutions include dedicated and shared protection strategies [1], and are mostly designed for single node or link failures. However, network deployments across expanded geographic domains are mandating services recovery under highly-challenging multi-failure conditions. In particular, there is much interest in handling multiple *correlated* failures arising from large-scale "stressors" such as natural disasters (earthquakes, hurricanes, floods, snow storms, etc), massive power outages, and malicious *weapons of mass destruction* (WMD) attacks.

Overall, multi-failure recovery is a relatively new research focus area. Although some earlier efforts have studied dual link failures in IP and optical networks [3],[4],[5], these solutions are limited in scope and generally assume independent failures. By contrast, multi-failure scenarios are much more challenging, making it is very difficult (if not impossible) to guarantee pre-provisioned recovery for all possible fault combinations. Hence the design of *probabilistic* recovery schemes is very germane here, and some new solutions have been proposed to minimize path failure probabilities under multiple failures [6],[7]. Nevertheless, these strategies only focus on risk minimization and do not address broader *traffic engineering* (TE) objectives. Therefore it is likely that they will yield very high resource consumption (inefficiency) and lead to increased blocking and reduced revenues.

In light of the above, there is a pressing need to develop improved multi-failure recovery solutions which also incorporate TE constraints. This paper addresses this critical area and is organized as follows. First, Section II presents an overview of some existing work in dual and multi-failure recovery. Next, a novel joint TE and risk mitigation algorithm is proposed in Section III, along with the requisite probabilistic multi-failure model. Section IV then presents detailed simulation analysis results and compares the performance of the proposed scheme against various existing solutions. Finally, conclusions and directions for future work are presented in Section V.

Manuscript received February 19, 2012. The associate editor coordinating the review of this letter and approving it for publication was G. Lazarou.

O. Diaz, F. Xu, and N. Ghani are with the ECE Department, University of New Mexico, USA (e-mail: nghani@ece.unm.edu).

N. Min-Allah is with COMSATS, Islamabad, Pakistan.

M. Khodeir is with JUST, Jordan.

M. Peng is with Wuhan University, People's Republic of China.

S. Khan is with North Dakota State University, USA.

Digital Object Identifier 10.1109/LCOMM.2012.060112.120353

II. Background Survey

As noted above, researchers have looked at dual failure recovery. For example, [3] outlines three shared protection strategies for dual-link failures, two of which assume *a-priori* failure knowledge and one which assumes specialized 3-connected topologies. These schemes are analyzed for three different network topologies and results show that it is possible to achieve full recovery with a modest increase in backup capacity. Meanwhile [4] combines pre-fault protection with post-fault restoration to address dual link failures in DWDM networks. Namely, a protection heuristic scheme (using Bhandari's algorithm) is used to recover the majority of demands via pre-computed backup light-paths. A restoration component is also invoked to dynamically search for backup paths for any unrecovered demands after the first failure. Results for several topologies indicate over 99% recovery with pre-planned protection. Finally, [5] proposes another dual-failure scheme for DWDM networks that tries to achieve complete protection for single-link failures, while also trying to minimize vulnerability to further secondary failures. This method uses sub-graph routing techniques, and overall results show good recovery rates for dual failures, i.e., 70-80%.

However, it is generally difficult to extend dual-failure schemes to handle multiple random failures, i.e., after large stressor or disaster events. Instead, it is much more feasible to use *probabilistic* routing strategies here to try to minimize service disruptions. Along these lines, [6] studies probabilistic routing for multiple *independent* link failures, where each link has a fixed a-priori failure probability. The authors present a detailed optimization model to compute maximally-survivable working/protection path pairs and show it to be NP-complete. A heuristic is then developed using Suurballe's algorithm to compute path pairs with minimum failure probabilities, and results presented for two networks. Nevertheless, the independent failure model in [6] does not accurately reflect realistic stressor/disaster settings, where failures will exhibit very high levels of spatial and temporal correlation, i.e., localized, near-simultaneous [8]. To address this concern, [7] proposes a more realistic multi-failure model that captures regional fault correlation, i.e., *probabilistic shared risk link group* (p-SRLG) concept. Using this model, an optimization

formulation is proposed to compute maximally-reliable primary/backup path pairs. Two greedy heuristic schemes are also developed using approximation techniques, and the results show close agreement between the optimization and heuristic strategies for smaller (tractable) networks.

Although the schemes in [7] provide improved reliability, they only focus on risk minimization. Hence these solutions will likely yield longer paths around higher-risk regions, leading to increased resource usages for network carriers. In light of this, there is a pressing need to develop improved strategies that incorporate both risk and TE resource objectives in the recovery process. Recently the authors in [9] have proposed a graph-theoretic algorithm to take into account both of these concerns. However, this solution only considers single working paths and does not provision added protection paths. As a result, this paper proposes a more comprehensive framework solution for working/protection path pair computation and compares it to existing probabilistic strategies, particularly [7].

Note that some efforts have also studied network topology design under multi-failure settings [10],[11], [12]. Specifically, the goal here is to analyze the impact of large-scale regional stressors on underlying topologies themselves and identify critically vulnerable regions, i.e., those yielding maximum capacity or connectivity disruption. These findings can then be used during network planning phases to help improve the overall fault-tolerance of deployed infrastructures.

III. Joint Risk and Traffic Engineering Approach

A novel multi-failure recovery scheme is proposed to address the above concerns. The solution uses graph-theoretic heuristics and assumes the same probabilistic/stochastic failure model as in [7] (detailed shortly). The framework also assumes bandwidth-provisioning nodes, e.g., IP/MPLS routers, but can be extended to support DWDM networks. Consider the requisite notation first. The network is modeled as a graph, **G(V,L)**, where $\mathbf{V} = \{v_1, v_2, \cdots\}$ is the set of switching nodes and $\mathbf{L} = \{l_{ij} | \forall v_i$ and v_j if they are connected, $i \neq j\}$ is the set of links, i.e., l_{ij} represents a bi-directional link between v_i and v_j with maximum capacity c_{ij} and available/free bandwidth b_{ij}. Meanwhile, all user requests arrive in a dynamic "online" manner and comprise of three parameters, i.e., source node, v_a, destination node, v_b, and desired capacity, r. Hence a request can be summarized by the 3-tuple $\{v_a, v_b, r\}$.

Now the threat environment is modeled by *N* possible multi-failure events given by the set $\mathbf{F} = \{f_1, f_2, \cdots, f_N\}$. Here each event $f_i \in \mathbf{F}$ represents a p-SRLG centered about a given region, with an occurrence probability denoted by ϕ_i, and has an associated set of vulnerable links, $\mathbf{X}_i$, $1 \leq i \leq N$. Overall, this model is sufficiently generic and can incorporate vulnerable nodes by including all their emanating links in $\mathbf{X}_i$. Given the scale of multi-failure events, it is also very likely that only a single stressor will occur at a given time, i.e., mutually-exclusive events, as in [7]. This assumption precludes the need for more complicated inter-stressor dependency modeling and allows occurrence probabilities such that $\sum_i \phi_i = 1$. Like the failure model in [7], it is also assumed that all links within a p-SRLG region fail in an independent manner, i.e., a *non-zero* conditional failure probability, p^i_{jk}, is defined for each link with respect to stressor f_i if $l_{jk} \in \mathbf{X}_i$ ($p^i_{jk=0}$ otherwise). Carefully note that this model allows path failure probabilities to be computed as a series product (detailed later). The path pair computation scheme is now detailed.

1. Given an incoming request 3-tuple $\{v_a, v_b, r\}$.
2. Prune all non-feasible links in $\mathbf{G}(\mathbf{V}, \mathbf{L})$, i.e., $b_{ij} < r$, and generate a modified graph $\mathbf{G}'(\mathbf{V}', \mathbf{L}')$.
3. Assign link weights ω_{ij} via TE load-balancing, Eq.(1)
4. Compute *K*-shortest paths from node v_a to node v_b over $\mathbf{G}'(\mathbf{V}', \mathbf{L}')$, i.e., working paths $\{\mathbf{wp}_i\}$.
5. For each $\mathbf{wp}_i$ prune links used by $\mathbf{wp}_i$ from $\mathbf{G}'(\mathbf{V}', \mathbf{L}')$ and compute its link-disjoint protection path, $\mathbf{bp}_i$.
6. Compute conditional failure prob. for each working or protection path, P_{ij} and P'_{ij}, Eq. (2), $\forall f_i \in \mathbf{F}$.
7. Select path pair w.min. joint failure probability, i.e., $\min_{1 \leq i \leq k} \{\sum_{1 \leq j \leq N} P^{ij} \bar{P}^{ij}\}$.

Fig. 1. Joint path-pair load balancing (JPP-LB) algorithm.

To date researchers have proposed many different path protection heuristics, see [1],[2],[13]. For the most part, these algorithms compute link-disjoint paths to overcome single link failures and also try to achieve some sort of TE resource efficiency in the process. However these schemes yield lower recovery rates under correlated failure scenarios as they do not incorporate a-priori link risk information. Conversely, the newer probabilistic multi-failure schemes in [6] and [7] generate longer detour paths by bypassing higher-risk regions and yield increased resource inefficiencies [9]. As a result, the proposed solution here tries to achieve a balance between these strategies and is termed as the *joint path pair load balancing* (JPP-LB) scheme. Details are now presented.

The overall pseudocode for the JPP-LB scheme is presented in Figure 1. Given an incoming 3-tuple request, the algorithm first prunes all non-feasible links in order to improve blocking performance, i.e., removing all links with insufficient capacity, i.e., $b_{ij} < r$. Next, up to *K-shortest path* (*K*-SP) working routes, $\{\mathbf{wp}_i\}$, are computed between the source and destination nodes v_a and v_b. In particular, a load balancing approach is applied here by defining dynamic link weights as follows:

$$\omega_{ij} = \frac{c_{ij}}{b_{ij} + \xi} \tag{1}$$

i.e., more congested links have higher costs and $\xi \approx 0$ to avoid division-by-zero error. After the working paths are computed, each is processed in isolation to compute a link-disjoint protection path, $\mathbf{bp}_i$, by running Dijkstra's shortest path algorithm (step 5, Figure 1). This step yields *K* working/protection *path pairs*, i.e., $\{\mathbf{wp}_i, \mathbf{bp}_i\}$, from which the final pair is selected based on the lowest risk. Namely, this is done by computing a path failure probability for each working path:

$$P_{ij} = \text{Prob}(\mathbf{wp}_i \text{ fails } | f_i) = 1 - \prod_{l_{mn} \in wp_i} (1 - p^j_{mn}) \tag{2}$$

and similarly, a path failure probability for each protection

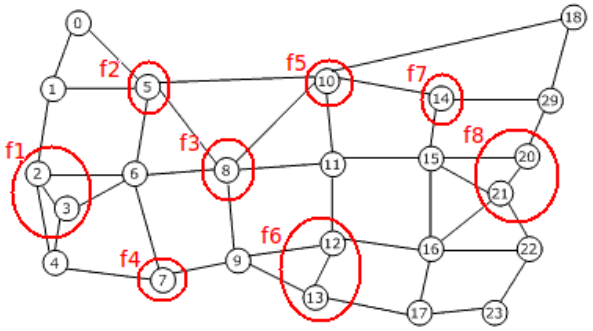

Fig. 2. US IP backbone with N=8 potential stressor regions.

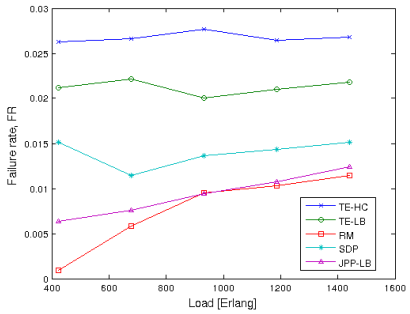

Fig. 3. Failure rate (FR) - working and protection.

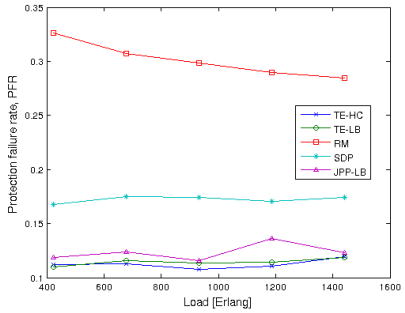

Fig. 4. Protection failure rate (PFR) - protection only.

path, P'_{ij}. Using the above, the path pair with the minimum joint path failure probability is chosen, i.e., computed as a sum of working and protection path failure probabilities (products) across all failure events $f_i \in \mathbf{F}$ (step 7, Figure 1).

Finally, some baseline pure TE and risk minimization algorithms are also included for comparison purposes. Namely, two TE-based protection algorithms are used, *hop count minimization* (TE-HC) and *load-balancing* (TE-LB). These schemes follow the same overall flow as the JPP-LB solution (Figure 1) but omit all probabilistic computations. Namely, K-SP working paths are first computed using either the hop count or load balancing (Eq. 1) link weights. Protection paths are then computed for each individual K-SP path in isolation, and the working/protection path pair with the minimum aggregate (hop count or load-balancing) cost is chosen. Overall, selecting the best *pair* is more effective than using a greedy approach, i.e., sequentially computing the "best" working path followed by the "best" link-disjoint protection path, i.e., $K = 1$. Note that [14] also uses a similar *K*-SP approach for lightpath protection. Moreover, risk minimization comparisons are done using the two greedy heuristic schemes proposed in [7]. In particular, the *shortest disjoint path* (SDP) scheme sets link weights in $\mathbf{G}(\mathbf{V}, \mathbf{L})$ to their average link failure probabilities (across all stressors) and then uses Dijkstras shortest path algorithm to compute the lowest risk working path. This path is then pruned and the process repeated to compute the protection path, i.e., greedy approach. Meanwhile, the second improved heuristic, termed here as *risk minimization* (RM), also uses this greedy Dijkstra-based approach but performs link weight re-adjustments after working path computation, i.e., based upon an approximation to the optimization, see [7] for details.

IV. Performance Analysis

The proposed joint scheme is analyzed using custom-developed *OPNET Modeler™* models. To ensure realistic findings, the US IP network from [7] is used, comprising of 24 nodes and 43 links, see Figure 2. All link rates are set to 10 Gbps and $N = 8$ multi-failure events are defined. These events encompass different numbers of links/nodes and are reflective of varying stressor severities. User demands are uniformly-distributed between 200-1,000 Mbps in increments of 200 Mbps, i.e., to model fractional Ethernet services. Meanwhile, each simulation generates 2,500,000 random connections with exponentially-distributed holding times of mean 600 sec. The actual input loads here are varied by changing the mean request *inter arrival times* (IAT), also exponential. All path pair computation for the JPP, TE-HC, and TE-LB schemes is done using $K = 5$ and results are averaged over 500 randomly-selected stressors. Namely, one of the 8 potential multi-failure events is chosen at a random interval in the simulation using a uniform distribution, i.e., $\phi_i = 1/8$, and failure statistics counted and all failed links restored before the next stressor.

Initial tests are done to gauge recovery effectiveness for the various schemes. Namely, connection *failure rates* (FR) are plotted in Figure 3 by measuring the fraction of connections experiencing *both* working and protection path failures (for varying input loads). Here it is clear that the proposed JPP-LB and optimized risk-only RM schemes give the best reliability and consistently outperform the more basic SDP risk reduction scheme, i.e., slightly over 1% connection failures at very high loads. Note that RM slightly outperforms the JPP-LB solution at very low loads. By contrast, the TE-based strategies yield much lower recovery rates, with at least twice the number of failures versus the JPP-LB and RM schemes. Moreover, the load balancing TE-LB strategy does slightly better than the resource minimization TE-HC scheme.

Next, it is also important to gauge the impact of multiple failures on the reliability of protection paths for *non-failed* working connections. Hence Figure 4 plots the *protection failure rate* (PFR) for different input loads and these results reveal some very interesting findings. Foremost, the RM risk-based scheme gives the highest fraction of failures, averaging almost three times higher than the TE-based strategies. By contrast the JPP-LB solution is much more effective here and closely tracks the improved performance of the TE-based

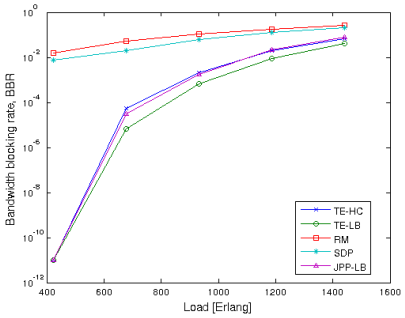

Fig. 5. Bandwidth blocking rate (BBR).

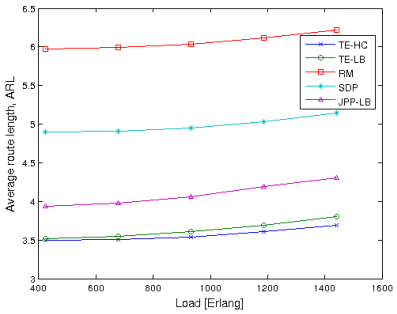

Fig. 6. Average route length (ARL) in hops.

schemes. Meanwhile the SDP solution does better than the RM scheme but still yields significantly higher protection failure rates than the JPP-LB scheme. These findings indicate that the JPP-LB approach can give much better *post-fault* reliability for non-failed connections. In turn, this will mitigate the operational concerns for operators by reducing the need to re-compute/re-provision failed protection paths.

Note that the above tests strictly focus on connection reliability. Hence additional measurements are also done to gauge TE performance. Namely, the average *bandwidth blocking rates* (BBR) are first plotted in Figure 5 using a log-scale to cover a wide range of operational regimes. These results clearly show major shortcomings with the pure risk-based strategies in [7], with neither the SDP nor RM scheme giving under 1% blocking for almost all input loads. By contrast, the pure TE and JPP-LB algorithms are much more effective, yielding orders magnitude lower blocking at low-to-medium loads. In particular, the JPP-LB scheme consistently outperforms the TE-HC heuristic and closely tracks the TE-LB scheme (which gives the lowest blocking). In light of the above, the SDP and RM schemes may be problematic in real-world settings as their increased blocking rates will yield much lower carried loads, i.e., revenues, for carriers.

Finally, the resource usages of working/protection path pairs are also measured by plotting *average route lengths* (ARL) in Figure 6. As expected, the TE-HC scheme gives the lowest resource utilization and is closely followed by the TE-LB scheme. By contrast, the risk-based strategies are much more resource-intensive as they generate longer path pairs to detour around higher-risk regions. For example, the RM scheme gives almost two times longer paths than the TE-LB solution. Meanwhile the JPP-LB solution achieves a very good balance here, with average route lengths closer to the TE-based strategies, i.e., about 12-15% higher.

V. Conclusion

A novel path pair protection solution is proposed to jointly incorporate traffic engineering and risk minimization concerns for multi-failure network scenarios. Simulation results show that the scheme gives very good reliability, surpassing or closely matching existing risk minimization algorithms. At the same time, this solution also yields very competitive blocking rates, almost on par with load-balancing strategies. As a result, this design will have strong appeal in practical environments, as it will allow operators to achieve improved service recovery without compromising resource efficiency (revenues). Future efforts will look at developing more detailed optimization models to derive lower bounds on achievable performance.

Acknowledgment

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

References

[1] P. Cholda *et al.*, "A survey of resilience differentiation frameworks in communication networks," *IEEE Commun. Surveys and Tutorials*, vol. 9, no. 4, pp. 32–55, 2007.

[2] Y. Xin and G. N. Rouskas, "A study of path protection in large-scale optical networks," *Photonic Network Commun.*, vol. 7, pp. 267–278, 2004.

[3] H. Choi, S. Subramaniam, and H.-A. Choi, "On double-link failure recovery in WDM optical networks," in *Proc. 2002 IEEE INFOCOM*, vol. 2, pp. 808–816.

[4] L. Ruan and T. Feng, "A hybrid protection/restoration scheme for two-link failure in WDM mesh networks," in *Proc. 2010 IEEE GLOBECOM*, pp. 1–5.

[5] M. Frederick, P. Datta, and A. Somani, "Evaluating dual-failure restorability in mesh-restorable WDM optical networks," in *Proc. 2004 IEEE ICCCN*, pp. 309–314.

[6] Q. She, X. Huang, and J. Jue, "Maximum survivability under multiple failures," in *2006 IEEE/OSA OFC*.

[7] H.-W. Lee, E. Modiano, and K. Lee, "Diverse routing in networks with probabilistic failures," *IEEE/ACM Trans. Networking*, vol. 18, no. 6, pp. 1895–1907, Dec. 2010.

[8] M. Rahnamay-Naeini *et al.*, "Modeling stochastic correlated failures and their effects on network reliability," in *Proc. 2011 IEEE ICCCN*, pp. 1–6.

[9] M. Esmaeili *et al.*, "Multi-domain DWDM network provisioning for correlated failures," in *Proc. 2011 IEEE/OSA OFC*, pp. 1–3.

[10] A. Sen, S. Murthy, and S. Banerjee, "Region-based connectivity—a new paradigm for design of fault-tolerant networks," in *Proc. 2009 IEEE HPSR*, pp. 1–7.

[11] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the vulnerability of the fiber infrastructure to disasters," in *Proc. 2009 IEEE INFOCOM*, pp. 1566–1574.

[12] P. Agarwal *et al.*, "The resilience of WDM networks to probabilistic geographical failures," in *Proc. 2011 IEEE INFOCOM*, pp. 1521–1529.

[13] N. Ghani and S. Park, "Multi-tiered services in next-generation SONET/SDH networks," in *Photonic Network Commun.*, vol. 13, no. 1, pp. 79–92, Jan. 2007.

[14] C. Xin, Y. Ye, S. Dixit, and C. Qiao, "A joint lightpath routing approach in survivable optical networks," *Optical Networks Mag.*, vol. 3, no. 3, pp. 13–20, 2002.