Post-Fault Restoration in Multi-Domain Networks with Multiple Failures

F. Xu[1], M. Peng[2], M. Esmaeili[1], M. Rahnamay-Naeini[1], S. Khan[3], N. Ghani[1], M. Hayat[1]
[1]*University of New Mexico,* [2]*Wuhan University,* [3]*North Dakota State University*

Abstract: *Multi-domain networking is an important focus area and crankback signaling strategies offer much potential for post-fault recovery. However, most studies in this space have not addressed multi-failure scenarios, e.g., correlated failures such as those resulting from weapons of mass destruction attacks. Hence this paper proposes a novel solution for intra/inter-domain post-fault crankback restoration in IP/MPLS networks (also extendible to optical DWDM networks). Specifically, inter-domain path/distance-vector routing state information is integrated with dynamic node/link failure and intra-domain link-state information to improve the recovery process. The performance of the proposed solution is then analyzed for a range of multi-failure scenarios.*

Keywords: *Network survivability, multi-domain survivability, multi-failure recovery, crankback signaling*

I. INTRODUCTION

Network survivability is a very well-studied area and a host of solutions have been developed for IP-based *multi-protocol label switching* (MPLS) and optical *generalized MPLS* (GMPLS) networks. Most notably, these include pre-provisioned protection strategies for working/backup path computation as well as post-fault restoration strategies; see detailed survey in [1]. Furthermore, with expanding deployments, a variety of *multi-domain* protection schemes have been developed for larger national backbone networks [2]-[6]. For example, SONET/SDH interconnection strategies have been extended for "localized" dual/multi-homing link protection between border nodes in optical *dense wavelength division multiplexing* (DWDM) networks [3]. Alternatively, "globalized" MPLS/GMPLS algorithms have also been studied for diversified working/backup path routing using sequential/parallel computation [5],[6].

In general, many of the above survivability schemes have been designed to handle single failure events—mostly link failures. However, it is necessary to develop more robust *multi-failure* recovery schemes as well [7]-[12]. In particular, this topic has gained much impetus owing to the increased susceptibility of modern networking infrastructures to large *weapons of mass destruction* (WMD) type stressors/attacks, which can result in multiple *correlated*, i.e., simultaneous, failures at both link and node levels [14]. Along these lines, some recent efforts have proposed various path protection schemes to minimize the impact of such stressors [13],[14]. Nevertheless, even though these solutions yield improved resilience to multi-failure stressors, they are largely designed for single-domain settings in which computational entities have complete knowledge of all network links. Clearly, such assumptions will not hold in larger *multi-domain* backbones where only a limited number of nodes will maintain partial/dated views of the "global" topology, e.g., via distance/ path-vector *exterior gateway protocol* (EGP) state [2]. Moreover, owing to the randomized nature of WMD-induced failures (i.e., location, scope), pre-provisioning strategies may still not be able to guarantee recovery in all cases.

In light of the above, there is a pressing need to develop improved multi-domain recovery solution to handle correlated multi-failure events. This is a very challenging problem given the sheer scale of larger multi-domain networks and generally random nature of (WMD-induced) failure footprints. It is here that dynamic *post-fault* restoration schemes have much potential. Although these strategies exhibit longer recovery times (versus rapid pre-provisioned protection switchovers), they offer a very viable "last-gap" alternative in case of failure of pre-provisioned paths, i.e., as part of a tiered survivability framework. However, existing studies on multi-domain crankback have only addressed regular "working-mode" provisioning, see [16]-[18]. As a result, there is significant scope in developing novel crankback schemes for handling correlated WMD-type attacks. Along these lines, this paper proposes a joint intra/inter-domain crankback solution which supports both intermediate and end-to-end recovery. This solution extends upon a recently-proposed crankback scheme for working-mode provisioning presented in [17]. The proposed framework follows a standardized approach, leveraging the latest MPLS/GMPLS *resource reservation* (RSVP-TE) signaling extensions for crankback support [15] as well as the distributed *path computation element* (PCE) framework [19]. To the best of our knowledge, this is the first study on multi-domain/multi-failure recovery.

This paper is organized as follows. Section II surveys the latest work in multi-failure recovery an also multi-domain survivability. Next, Section III details the proposed crankback restoration scheme. Detailed performance analysis is then conducted in Section IV for a range of multi-failure scenarios to simulate WMD attacks. Finally, conclusions and future directions are presented in Section V.

II. BACKGROUND

A range of studies have been done in the areas of multi-domain survivability and also multi-failure survivability. Consider the former first. Here, [3] proposes several DWDM border node interconnection strategies to protect working/ backup paths traversing the same domain sequence. These solutions are robust to localized link failures but susceptible to concentrated WDM attacks spanning multiple nodes/links. Meanwhile [4] develops sequential and parallel strategies for working/protection MPLS route computation for improved "domain diversity". However, these schemes are susceptible to failures at domain boundaries and/or inter-domain "trap" topologies. Besides, more complex inter-domain path pair schemes have also been proposed to compute non-overlapped inter-domain routes in [5],[6], relying upon advanced topology abstraction and hierarchical routing algorithms. Although the blocking gains provided by these solutions are good, the abstractions impose significant routing overheads.

978-1-4244-8179-8/10/$26.00 ©2010 IEEE 1016

Meanwhile, a range of multi-failure recovery schemes have also been investigated. In particular, earlier studies have looked at specialized *dual near-simultaneous* failures. For example, [7] studies *backup re-provisioning* (BR) in DWDM networks, which notably improves backup availability but entails high computational overheads, whereas [8] details pre-emptive re-provisioning for dual link failures, but findings show key deficiencies for fiber faults. Next, [9] presents three shared protection schemes to handle dual-link failures, and [10] compares post-fault restoration versus pre-provisioned protection for dual-link DWDM. Also, [11] studies dual link failure recovery in IP-tunneling networks, where fast recovery from the first failure is handled via a protection domain around the failed link. Finally, [12] presents a scheme using fast re-routing in the IP layer for double-link failure recovery.

More recently, researchers have started to investigate survivability under probabilistic multi-failure conditions, e.g., as induced by WMD stressors. For example, [13] proposes a scheme for reliable path computation under generalized failure events and introduces two notions of reliability, i.e., *local* and *global*. The former one chooses routes that span the lowest number of failure events. Conversely, the latter selects routes such that a failure affects a minimum number of connections. Meanwhile, [14] studies network recovery under correlated probabilistic failures and tries to maximize the reliability of pre-computed protection routes. In order to model failure susceptibilities, the authors generalize the traditional static *shared risk link group* (SRLG) definition into an expanded *probabilistic SRLG* (p-SRLG) concept. Using this model, all links within a p-SRLG are assumed to fail independently, thereby facilitating the analysis of correlated failures.

Nevertheless, although the above contributions provide many improvements for network survivability, none of them offers a comprehensive multi-domain solution for dynamic WMD threat environments. For example, most multi-domain survivability schemes only handle single failures. On the contrary, most multi-failure studies only address single domain settings. For this reason, significant changes will be required to adapt these algorithms for multi-domain settings. Moreover, nearly all of these respective strategies rely upon *pre-provisioning* of protection resources. Given the highly-random nature of WMD-induced attacks, it is very difficult for them to handle all possible node/link failure combinations.

In light of the above, there is a growing need of dynamic restoration-based solutions to multi-domain/multi-failure recovery, e.g., *signaling crankback* [15]. These strategies perform "active" post-fault path re-computation and hence minimize dependence upon pre-provisioned routes and global state information—ideal for WMD recovery. Now even though some multi-domain crankback schemes have been developed, they tend to focus on basic *traffic engineering* (TE) objectives. For example, [16] defines a basic "per-domain" (PD) scheme which probes egress domain nodes for working routes. Meanwhile [17] builds on this setup by developing improved next-hop domain selection strategies whereas [18] details another *compute while switching* (CWS) scheme which exhaustively searches for shorter length routes. As these schemes do not address survivability, there remains significant scope for developing new and improved crankback algorithms for multi-failure operation. These concerns are now addressed.

III. ENHANCED CRANKBACK RESTORATION

A novel multi-domain crankback restoration solution is now presented, extending upon the "working-mode" crankback solution of [17]. The approach assumes realistic settings with full intra-domain link-state routing and more scalable path/distance-vector routing at the inter-domain level. Each domain is assumed to have a PCE entity [19] with full access to interior and exterior routing databases. This entity plays a key role in crankback recovery as it resolves next-hop domains. Meanwhile, all setup signaling is done using new crankback extensions for RSVP-TE [15].

Overall, three key innovations are introduced in this work to enhance multi-domain restoration, including 1) intelligent per-domain selection, 2) dual intra/inter-domain crankback counters to limit signaling complexity, and 3) intermediate and end-to-end crankback. Details are now presented.

A. Next-Hop Domain Computation

Before presenting the crankback scheme, the requisite notation is introduced, leveraging from our earlier work in [17]. A multi-domain network is comprised of D domains, with the i-th domain having n^i nodes and b^i border/gateway nodes, $1 \le i \le D$. This network is modeled 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 domain nodes and $\boldsymbol{L}^i=\{l^{ii}_{jk}\}$ is the set of *intra-domain* links in domain i ($1 \le i \le D$, $1 \le j,k \le n^i$). The inter-domain link connecting border node v^i_k in domain i with border node v^j_m in domain j is further denoted as l^{ij}_{km}, $1 \le i,j \le D$, $1 \le k \le b^i$, $1 \le m \le b^j$. All connectivity is assumed to be bi-directional. Each node also maintains a list of traversing connections, $\underline{\boldsymbol{A}}^i_j$ for node v^i_j, where each entry in $\underline{\boldsymbol{A}}^i_j$ is a route vector. Finally, relevant RSVP-TE message fields support a path route vector, $\underline{\boldsymbol{R}}$, and other fields for crankback support [15]. The latter includes an exclude link vector, $\underline{\boldsymbol{X}}$, to track failed links and dual intra/inter-domain crankback counters, h_1 and h_2.

As all "per-domain" crankback schemes require some form of next-hop domain selection, a key saliency of the proposed framework here is the use of limited inter-domain routing state for "intelligent" next-hop domain selection, as in [17]. This is achieved by pre-computing *multi-entry* distance vector tables at all PCE nodes, listing up to K next-hop domains/egress links to each destination domain. Namely, at domain i, the k-th table entry to destination domain j, $T^i(j,k)$, is computed as the egress link on the k-th shortest "domain-level" hop-count path to domain j ($1 \le i,j \le D$, $i \ne j$, $1 \le k \le K$).

In the above, the actual computation of multi-entry distance vector tables is done by building a "skeleton" graph of the global network by extracting from the EGP path databases, $\boldsymbol{H}(\boldsymbol{U},\boldsymbol{E})$, where $\boldsymbol{U}$ is the set of domains $\{\boldsymbol{G}^i\}$ reduced to vertices and $\boldsymbol{E}$ is the set of physical inter-domain links $\{l^{ij}_{km}\}$, $i \ne j$. An iterative shortest-path algorithm is then used to compute multiple routes to all destination domains over $\boldsymbol{H}(\boldsymbol{U},\boldsymbol{E})$, see [17] for details. Overall, these tables are relatively static and will sequentially drive crankback recovery along inter-domain routes of increasing length. The joint intra/inter-domain crankback scheme is now detailed.

B. Multi-Domain Crankback: Working Mode

First consider regular per-domain provisioning for *non-crankback* operation, i.e., no link resource failures or

crankback retries, as detailed in [17]. Here a source node in domain i fielding a request for x units of capacity (to a destination in another domain) first queries its PCE to determine the egress link to the next-hop domain, i.e., first entry in table $T^i(\cdot)$. Upon receiving PCE response, the source uses its local OSPF-TE database to compute an *explicit route* (ER) to the specified egress border node. Granted that an ER path is found above, it is inserted into the path route vector, $\underline{\boldsymbol{R}}$, and RSVP-TE *PATH* messaging is then initiated (along the expanded route) to the ingress border node in the next-hop domain. Here each intermediate node checks for available bandwidth resources on its outbound link and propagates the message downstream. The above procedure is repeated at all next-hop domain border nodes until the destination domain. When the *PATH* message finally arrives at the destination domain, the border node (or PCE) expands the ER to the destination, which will then initiate upstream reservation by sending a *RESV* message. Note that all intermediate nodes also store the final path routes, i.e., list of traversing connection routes, for use in possible post-fault recovery operation.

Conversely, if there are insufficient resources on a link during *PATH* processing, crankback signaling is used to re-compute an alternate route. Current extensions to RSVP-TE [15] have outlined various alternatives for crankback operation, and two types are leveraged here, *intra-domain* (local) and *inter-domain* (intermediate). Along these lines, the proposed scheme makes use of the two crankback counters in the RSVP-TE *PATH* message, i.e., h_1 and h_2. These counters are initialized to pre-specified values at the start of the *working* path signaling phase, i.e., H_1 and H_2, respectively, and then decremented during crankback to avoid searching excessively-long paths. As such these values limit the number of intra and inter-domain crankbacks to $H_1 \cdot H_2$.

Overall, two key operations are defined in the joint crankback scheme, i.e., *notification* and *re-computation* [17]. The former refers to upstream signaling procedure executed upon link resource failure at intermediate nodes, whereas the latter refers to the re-routing procedure to select a new route. Namely, if a resource signaling failure occurs, i.e., bandwidth deficiency, upstream notification is sent via a *PATH_ERR* message to the domain's ingress border node, i.e., *notification*. In addition, the intra-domain counter h_1 is decremented and the failed link is noted in the exclude link vector, $\underline{\boldsymbol{X}}$, for future reference (note that for blocking in the source domain, a *PATH_ERR* is sent back to the source node). Meanwhile, upon receiving a notification from a downstream node, an ingress border node performs *re-computation*. Namely, if the intra-domain h_1 counter (in received *PATH_ERR* message) has not expired, another egress border node is chosen by the ingress border node from table $T^i(\cdot)$ and ER expansion attempted to it, i.e., intra-domain crankback. Otherwise, more aggressive inter-domain crankback is done by sending a *PATH_ERR* message to the ingress node in the upstream domain (with h_2 decremented and h_1 reset to H_1). Hence the connection request will be considered failed if both counters are zero.

General working mode crankback is shown in the top part of Figure 1 for H_1=3/H_2=3. Namely, intra-domain crankback is first illustrated in domain 3, where ingress border node v_1^3 makes H_1=3 unsuccessful ER expansion attempts to its border nodes. Subsequently, this node initiates inter-domain crankback to border node v_1^2 in its upstream domain, domain 2, which then resolves a new route via domain 5.

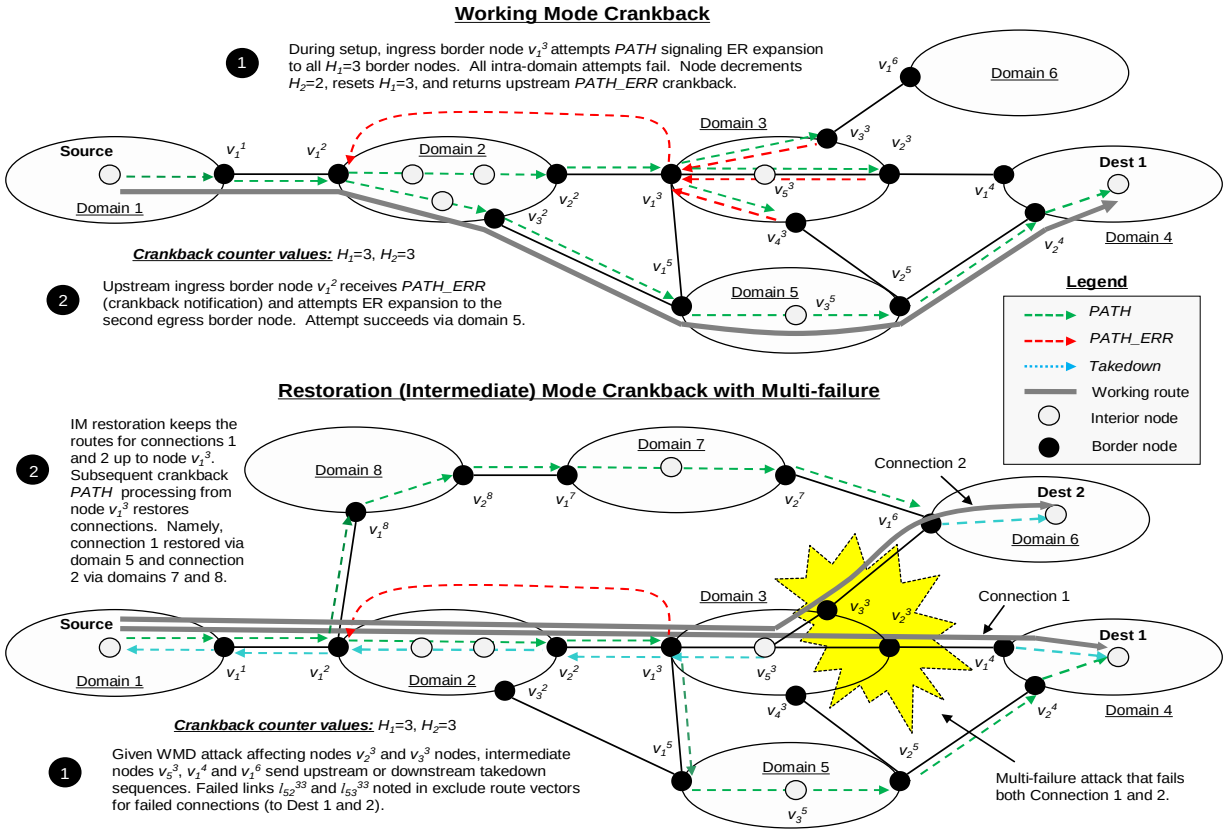

Figure 1: Joint intra/inter-domain crankback scheme for working and (intermediate) restoration mode

C. Multi-Domain Crankback: Restoration Mode

Now consider the case of crankback restoration under multi-failure scenarios with combination of node and link faults. Here it is assumed that a failed node will bring down all of its constituent links, i.e., lead to multiple link failures. As a result, any working node connecting to a failed node will quickly be able to resolve such occurrences via rapid lower layer detection mechanisms (which are out of the scope here). Hence with the failure notifications, appropriate crankback restoration procedures can be initiated by the working nodes bordering a failed region, detailed in the pseudo-code description in Figure 2. This extends upon [17] by also introducing RSVP-TE *PATH_ERR* message processing.

First, each failure-detecting node v^i_j loops through its active connection list, $\underline{\boldsymbol{A}}^i_j$, and searches for any routes traversing a failed link. All such connections are removed from $\underline{\boldsymbol{A}}^i_j$ and appropriate resource takedown messages issued along the failed upstream and/or downstream path segments. Specifically, this is done by either sending downstream *PATH_TEAR* or upstream *RESV_TEAR* messages (depending upon which side of link the destination node is). Next, if the node detecting the failure is on the *upstream* side of the failed link, it also notifies the source node of the connection via a *PATH_ERR* message with the failed link noted in exclude link vector, $\underline{\boldsymbol{X}}$. This notification message also sets a restoration flag to indicate that this is not a regular connection crankback.

```
/* Process all traversing connections at node v^i_j */
Loop through all traversing connections in list A^i_j at node v^i_j
    if (connection route uses failed link connected with v^i_j)
        Remove connection from list A^i_j
        if (failed link is incoming link to v^i_j)
            Send PATH_TEAR along downstream segment
        else /* failed link is outgoing link from v^i_j */
            Send RESV_TEAR along upstream segment;
            Generate PATH_ERR, set restoration flag,
            save route in R and failed link in X, send to source
```

Figure 2: Failure notification (at node with failed link)

Upon receiving the *PATH_ERR* failure notifications from a downstream endpoint node, the ingress/source node initiates crankback setup to re-compute new routes. Carefully note that since multiple crankback attempts can be initiated in close succession after a multi-failure event, signaling race condition can easily arise, particularly on inter-domain links. In order to resolve this concern, each notified source chooses to wait for a random back-off interval before initiating crankback re-tries, i.e., typically averaging from the 100's milliseconds to seconds range. In addition, the associated intra/inter-domain crankback counter limits are set to H_1 and H_2, respectively. Herein, two different restoration schemes are proposed:

End-to-end (E2E): In this case each source node simply generates and sends a new *PATH* message request to the destination, i.e., clear route vector $\underline{\boldsymbol{R}}$, reset crankback counters reset to $h_1=H_1$ and $h_2=H_2$, and copy failed links (from received *PATH_ERR*) to outgoing exclude route list, $\underline{\boldsymbol{X}}$. This request is then processed as per the regular working-mode setup procedures in Section III.B.

Intermediate (IM): In this case the source node tries to preserve as much of the original path as possible, i.e., even up to the domain in which the failure was detected. Namely, the failed connection route in the returning *PATH_ERR* message is copied up to the domain ingress border node of the failed domain. This partial route is then inserted into $\underline{\boldsymbol{R}}$ of a new *PATH* message with reset counter values and failed link noted in $\underline{\boldsymbol{X}}$. Hence this setup message will basically track along the original route sequence and then initiate crankback processing from the failed node/domain. Note that race conditions can occur if resources along the partial route are allocated between post-failure takedown (of original failed connection) and subsequent source-based crankback re-initiation. However, these conditions are rare at mid-light loads, and regardless, can be handled by the crankback process.

The case of IM post-fault crankback restoration is also shown in the lower part of Figure 1 for a multi-failure event affecting nodes v^3_2 and v^3_3 (connections 1, 2). Here, failure detection is first done by node v^3_5 (upstream endpoint of link l_{52}^{33}) as well as nodes v^4_1 and v^6_1 (downstream endpoints of link l_{21}^{34} and l_{31}^{36}, respectively). In response, these nodes issue RSVP-TE takedown sequences to free resources along the remaining links of the two failed connection routes. Besides, the upstream node detecting the failures, v^3_5, also sends a *PATH_ERR* notification to the source (in domain 1) for both failed connections. Upon receiving this notification, the source node extracts the partial (working) path route up to the border node v^3_1 and inserts it into route vector, $\underline{\boldsymbol{R}}$, for the two new downstream *PATH* messages. These two setup messages are then processed as per regular working mode crankback operation, restoring both connections via re-routed paths via domain 5 (connection 1) and domains 7 and 8 (connection 2).

Finally, connection bandwidth requests can also be re-sized, i.e., reduced, based upon attack severities and/or connection priorities. This helps improve post-fault restoration rates and also promotes a "tiered" restoration framework.

IV. PERFORMANCE EVALUATION

The performance of the proposed multi-domain post-fault restoration scheme is tested using specially-developed models in *OPNET Modeler*™, i.e., a modified NSFNET topology (with nodes replaced by domains) with 16 domains/25 bidirectional inter-domain links, as shown in Figure 3. In this topology, all intra and inter-domain link rates are set to 10 Gbps and the individual domain sizes average 7-10 nodes and correspond to metropolitan areas, i.e., cities. The NSFNET topology is chosen as it is very representative of a nationwide backbone and is widely used in many studies. Furthermore, all requests are generated between random nodes in random domains with mean holding times of 600 sec (exponential) and variable inter-arrival times (as per desired load). In addition, the actual request sizes are further varied uniformly between 200 Mbps–1 Gbps in increments of 200 Mbps, i.e., to model fractional Ethernet demands. Meanwhile, K=5 next-hop domain entries are computed in the distance vector table, although the number searched is limited by H_1 or H_2 values.

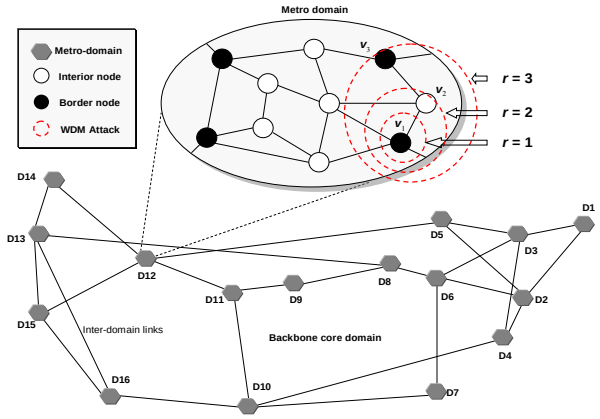

Figure 3: 16-domain NSFNET topology with multiple failures

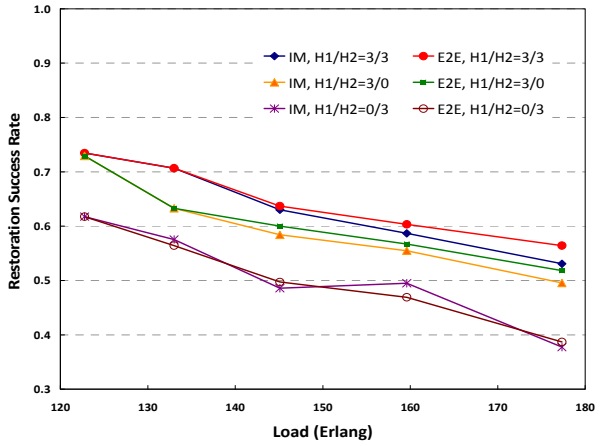

Figure 4: Restoration success with varying crankback counters (r=3)

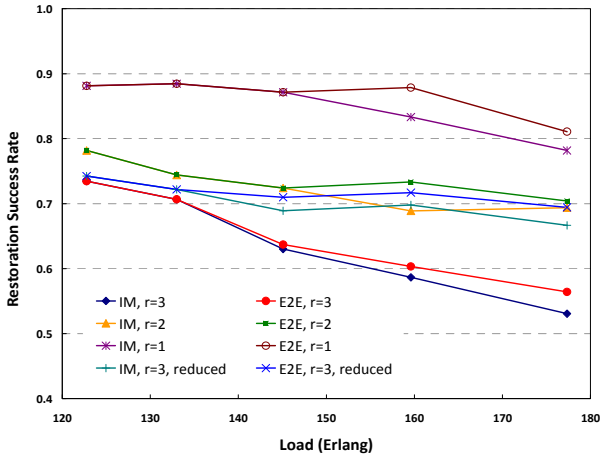

Figure 5: Restoration success rates of varying WMD attacks

Meanwhile, all multi-failure events are specified as a series of simultaneous node/link. To further incorporate different attack severities, a failure *region* is introduced to model WMD attacks. Namely, a region is circular area denoted by the radius r, Figure 3. Here, r=1 corresponds to the basic single-node failure case. Using the above failure region definition, the performance of the restoration scheme is gauged by measuring post-fault success rates after a large multi-failure event. Specifically, the network is first brought to steady state operation at a given input load, i.e., connection inter-arrival time, then a single multi-failure event triggered to randomly fail a cluster of nodes/links and subsequent recovery rates measured. This procedure is repeated 10 times at each input load point and the values averaged. Note that only transit connections on failed nodes are considered for restoration, i.e., failed source/destinations connections cannot be recovered.

First of all, restoration performance is tested for larger attack regions with r=3. The results are shown in Figure 4 for intra-domain only crankback (H_1=3/H_2=0), inter-domain only crankback (H_1=0/H_2=3), and also joint intra/inter-domain crankback (H_1=3/H_2=3). These results show that the E2E scheme gives slightly better recovery success rates versus IM restoration for all counter settings. This is due to the fact that the latter scheme tries to re-route multiple failed connections from the domain of (or immediately prior to) the attacked area. Such "localized" recovery exacerbates resource contention, especially on the domain egress links. This problem is particularly evident in the NSFNET topology, which has a relatively lower number of inter-domain links. By contrast, E2E scheme achieves better load distribution as it attempts path restoration over the wider network. It is also observed that the *joint* intra/inter-domain crankback yields notably higher restoration recovery versus intra- or inter-domain only crankback. Overall, these findings concur with similar results in non-failure crankback scenarios [17].

Next, post-fault restoration performance is evaluated by measuring the restoration success rates for varying input loads and three different attack region sizes, i.e., small (r=1), medium (r=2) and large (r=3), Figure 5. Here the intra/inter-domain crankback counter limits are set to H_1=3/H_2=3, as these values are shown to give better restoration performance in earlier runs (Figure 4). Overall, the results in Figure 5 show declining restoration success with increasing attack severities. This is expected as larger failure regions induce many more connection failures and result in higher resource contention during crankback. Akin to the findings in Figure 4, these results also show that E2E crankback gives slightly better restoration, particularly at higher loads. Moreover, even at very high traffic input loads, E2E crankback still gives very decent restoration for large attack regions, i.e., r=3. For example, post-fault recovery is almost 60% for a steady state *bandwidth blocking rate* (BBR) of 5% for r=3 (i.e., 178 Erlang, Figure 5). To further gauge crankback performance under such high-stress conditions, "tiered" recovery is also tested. Namely, the capacity of all post-fault connection re-tries are reduced by 50% and the results are also in Figure 5 for r=3. These findings show over 25% increase in recovery rates, pushing the aggregate restoration rates close to 70%.

In addition, resource efficiency and recovery delays are gauged for the two restoration schemes. Specifically, Figure 6 plots the average inter-domain path lengths whereas Figure 7 plots average restoration delays for successfully-restored connections (assuming distance-based link propagation delays and 0.5 ms node processing times). Again, it is clear that the E2E restoration consistently outperforms IM for both of these metrics. Namely, the average path utilization is about 9% lower at most loads and corresponding restoration delays are

about 8% lower. By contrast, IM restoration gives higher crankback re-tries, thereby leading to increased resource consumption and setup delay. Also, it is seen that more severe attacks, e.g., r=3, yield notably higher average path lengths and recovery times, e.g., path lengths about 25-30% higher and delays about 30% higher. Again, this is expected as longer routes will be required to bypass larger failed regions. Here it is noted that small-medium size attacks regions exhibit lower separation in terms of average path length and delay. This is due to the fact that such "smaller" attacks tend to affect fewer inter-domain nodes and hence are less likely to result in an expansion of "domain sequence" during restoration.

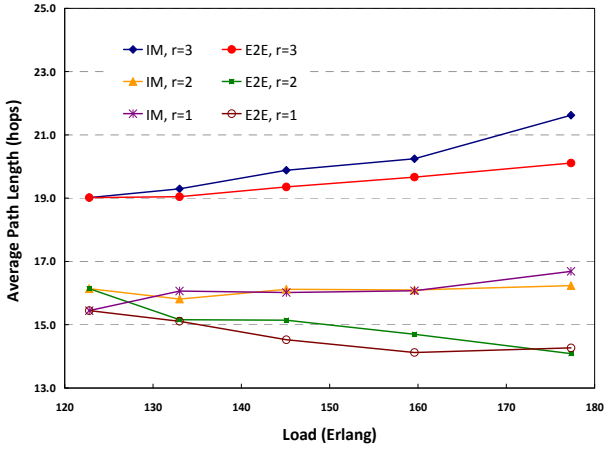

Figure 6: Average inter-domain restoration path lengths

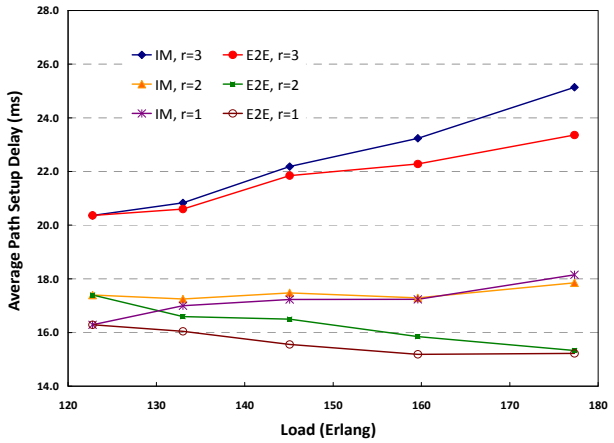

Figure 7: Average inter-domain restoration delays

V. CONCLUSIONS

This paper proposes novel crankback solution for post-fault restoration in large multi-domain backbone networks under multiple failures, i.e., WMD-type attacks. The scheme uses an enhanced next-hop domain selection strategy and incorporates full crankback history tracking to improve the effectiveness of the recovery process. In addition, a dual crankback counter approach is introduced to limit the number of intra/inter-domain retry attempts and both end-to-end and intermediate restoration modes are supported. The detailed performance results show very high post-fault recovery rates for correlated failure events with varying severity levels, particularly for E2E crankback. Future studies will look at extending this work with analytical modeling of restoration behaviors as well as integration with 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.

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

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

[4] T. Takeda, *et al*, "Analysis of Inter-Domain Label Switched Path (LSP) Recovery," *IETF Internet draft-ietf-ccamp-inter-domain-recovery-analysis-02.txt*, September 2007.

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

[6] D. Truong, B. Thiongane, "Dynamic Routing for Shared Path Protection in Multi-Domain Optical Mesh Networks," *OSA Journal of Optical Net.*, Vol. 5, Jan. 2006, pp. 58-74.

[7] 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, June 2004.

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

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

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

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

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

[13] 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.

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

[15] A. Farrel, *et al*, "Crankback Signaling Extensions for MPLS and GMPLS RSVP-TE," *IETF Request RFC 4920*, July 2007.

[16] S. Dasgupta, *et al*, "Path-Computation-Element-Based Architecture for Interdomain MPLS/GMPLS Traffic Engineering: Overview and Performance," *IEEE Network*, Vol. 21, No. 4, July/August 2007, pp. 38-45.

[17] F. Xu, *et al*, "Enhanced Crankback Signaling for Multi-Domain Traffic Engineering," *IEEE ICC 2010*, Cape Town, South Africa, May 2010.

[18] F. Aslam, *et al*, "Inter-Domain Path Computation Using Improved Crankback Signaling in Label Switched Networks," *IEEE ICC 2007*, Glasgow, Scotland, June 2007.

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