DC-Carrier Cooperation for Rapid Restoration against PNE-Node Failure in Optical Networks

Subhadeep Sahoo,[1,*] Sifat Ferdousi,[1] Sugang Xu,[2] Yusuke Hirota,[2] Massimo Tornatore,[1,3] Yoshinari Awaji,[2] and Biswanath Mukherjee[1,4]

[1] University of California, Davis, USA; [2]National Institute of Information and Communications Technology, Japan; [3]Politecnico di Milano, Italy;[4]Soochow University, China

[*]Corresponding author: subsahoo@ucdavis.edu

Abstract: We propose a rapid restoration strategy against PNE-node failure during post-disaster cooperation among DC providers and optical-network carriers. Our strategy reduces disruption and improves DC-service restoration by 35% in 20% less time compared to baseline.

1. Introduction

Today's Internet applications increasingly rely on cloud datacenters (DCs) for ubiquitous computing [1]. DC providers (DCPs) support these cloud applications by interconnecting their geo-distributed DCs through DC interconnects (DCIs), which are established using connection services (e.g., IP-over-WDM connections or optical lightpaths) provided by carrier networks [2]. In such a network-cloud ecosystem, failures from a large-scale disaster can severely disrupt DC services due to damages in carrier/DC infrastructures. Restoration of the impacted services can also be largely constrained by post-disaster resource crunch in the ecosystem. In such case, cooperation among DCPs and carriers is important for efficient service restoration, e.g., DCPs can cooperate by re-configuring their DCIs to adapt to the failures, while carriers can cooperate by sharing their available bandwidth resources to satisfy the modified DCIs [3]. However, this type of cooperation is challenging due to regulatory policies that limit sharing of information (e.g., DC/content location, detailed topology, resource availability, etc.) among DCPs and carriers. Ref. [2] introduced the concept of a third-party mediator, named ***Provider Neutral Exchange*** (PNE) [e.g., a consortium of distributed co-location centers or Internet Exchange Points (IXPs)], which facilitates cooperation among DCPs and carriers while maintaining information privacy through abstraction. PNE can build a public reference topology, a PNE network (PNEN), for the region of interest (e.g., disaster-affected area). PNEN consists of several electronic switches (PNE nodes) that interconnect different carrier nodes and DCs, located in close proximity (e.g., within same city), and enable traffic exchange via optical-electrical-optical (OEO) conversion.

Ref. [3] presented a DC-carrier cooperation strategy aided by PNE, in which a DCP adapts to different failure scenarios by adjusting its DCIs and carriers cooperate through resource-sharing to satisfy the DCP requests. Ref. [4] illustrated how to address the imbalance in service restoration when multiple DCPs are present in the ecosystem. Both these works considered that the PNE nodes seamlessly execute the required functionalities; however, the PNE nodes can also be vulnerable to failures (e.g., switch malfunctions). A PNE-node failure during post-disaster service restoration can severely impact the DC-carrier cooperation plan and disrupt the ongoing restoration process. In this work, we investigate how to address such failures and propose a novel PNE-failure-aware Rapid Restoration (PRR) strategy to reduce disruption during post-disaster service restoration of multiple DCPs. Results under different PNE-node failures and disaster scenarios show that our strategy can significantly reduce the disruption by restoring the impacted services efficiently while incurring less restoration time.

2. Problem Statement and Network Model

We study the problem of rapidly restoring the impacted DC services by reducing disruption caused by PNE-node failure during post-disaster cloud-service restoration though DC-carrier cooperation. Fig. 1 illustrates the network model used to evaluate the cooperation framework. Let us consider a subset of *Japan photonic network* as a disaster-affected area (marked in red oval) [5]. PNE creates a public reference PNEN topology for that area, represented by *P1-P11*. Carrier-A and Carrier-B abstract their detailed topologies to the PNEN topology (to avoid sharing infrastructure details), which are represented by *A1-A11* and *B1-B11*, respectively. The solid and dashed lines in both carriers represent the survived and failed links (due to disaster), respectively. Before disaster, DCP-X and DCP-Y create their own DCI topologies by leasing connection services from Carrier-A and Carrier-B, represented by solid blue and pink lines, respectively. These DCIs are also impacted by the disaster; the dashed lines and the grey nodes represent failed DCI links and failed DCs, respectively. The carriers evaluate failures in their networks and advertise abstracted connection prices for each PNE-node pair accordingly (please refer to [3] for details) to the DCPs through PNE. Based on the connection prices, DCPs generate new connection requests for the carriers to restore the DC services. For example, DCP-Y requests new connection services *A6-A8* and *B8-B10* from Carrier-A and Carrier-B, respectively. Then, the carriers cooperate by sharing their survived bandwidth

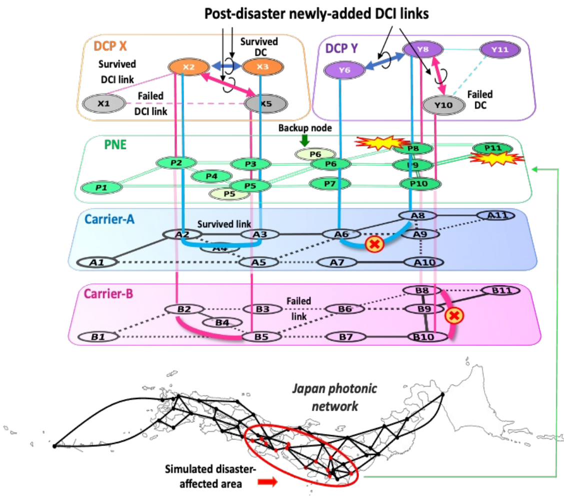

Fig. 1: Network model.

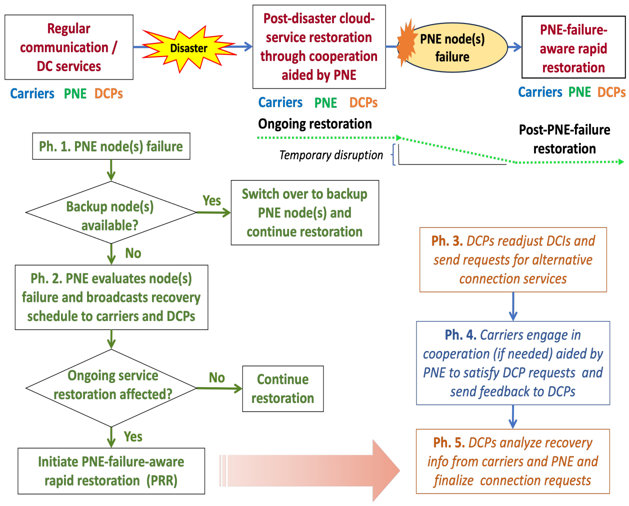

Fig. 2: PNE-failure-aware rapid restoration (PRR) framework.

resources (if sufficient resources are not available in their own networks) to satisfy the DCP requests, while PNE interconnects the DCs and the carrier nodes (e.g., through *P6*, *P8*, and *P10*) for such traffic exchange.

During service restoration, if any PNE node fails, it can severely disrupt the restoration process. To protect against PNE-node failure, backup switches can be deployed at the PNE nodes. Since deploying backup at all PNE nodes is a cost-intensive solution, in this study, we assume that backup switches are deployed at PNE nodes selectively, based on critical parameters, such as, nodal degree, traffic volume, number of interconnected nodes, etc. For example, in Fig. 1, PNE nodes *P5* and *P6* have backup switches and hence their failures do not disrupt the ongoing restoration. However, deploying backup only at selective PNE nodes is not sufficient to reduce disruption; hence, we propose PRR strategy that does not rely on availability of backup switches. We consider a combination of different PNE-node failures to analyze the varying impact across different connections, e.g., if *P11* fails, ongoing restoration is not impacted as *P11* does not exchange traffic among any DCP and carrier; if *P8* fails, restoration of DCP-Y is severely impacted due to disruption of *A6-A8* and *B8-B10* (denoted by red-cross in Fig. 1) as *P8* interconnects *Y8* with *A8* and *B8*. The impacted services can be restored after scheduled recovery of the failed PNE node(s), which may prolong the disruption. To rapidly restore the disrupted services, it is crucial to adjust the cooperation plan, e.g., DCP-Y can generate alternative connection requests, such as *B6-B9* and *B9-B10* for Carrier-B, which does not involve exchange through *P8* and Carrier-B can allocate resources accordingly to satisfy the request. The above described problem can be summarized as follows. Given a PNEN reference topology (with backup switches deployed at selective PNE nodes), a set of failed PNE nodes, PNE-node recovery schedule, and connection prices for each PNE-node pair advertised by the carriers, our objective is to restore the disrupted DC services rapidly by reducing disruption from PNE-node failure during post-disaster service restoration.

3. Proposed PNE-Failure-Aware Rapid Restoration (PRR) Framework

In a network-cloud ecosystem, regular communication/DC services can be impacted from disaster failures in carrier/DC infrastructures. As shown in Fig. 2, carriers and DCPs can cooperate for post-disaster DC-service restoration with the aid of PNE [3]. Such ongoing restoration can be disrupted due to PNE-node failure. To reduce disruption and restore the services rapidly, we propose a PRR framework as illustrated in Fig. 2. In ***phase 1***, after PNE node(s) failure, PNE checks if backup switch is available for the failed PNE node(s). If backup is available, the impacted traffic of ongoing restoration is immediately switched over to the backup node(s); otherwise, PNE evaluates the failure and prepares a recovery schedule for its failed node(s) (we assume recovery time of one PNE node as l units). In ***phase 2***, PNE informs DCPs and carriers about the disrupted interconnections (if any) of ongoing restoration and broadcasts the PNE-node recovery schedule. For the disrupted interconnections, PRR strategy is initiated which involves *phases 3*, *4*, and *5*. In ***phase 3***, DCPs first perform IP-layer restoration (based on available resources in the adjusted DCI topologies) of the disrupted services in a best-effort manner. For the unrestored services, DCPs generate an alternative set of connection requests for the carriers by re-adjusting their DCIs while avoiding the failed PNE nodes. For example, in Fig. 1, newly-added links of DCP-Y (i.e., *Y6-Y8* and *Y8-Y10*) are disrupted by *P8* failure. Thus, DCP-Y can generate alternative connection requests, such as *B6-B9* and *B9-B10* for Carrier-B to directly interconnect DCs *Y6* and *Y10*. In ***phase 4***, carriers evaluate the new set of connection requests and create a recovery plan. In this phase, multi-carrier cooperation aided by PNE can be performed [6]. Then, carriers send feedback to the DCPs on whether a connection request is: (a) immediately acceptable, (b) acceptable but requires provisioning time of k units, or (c) rejected. In ***phase 5***, DCPs analyze the feedback from carriers, compare it with the PNE-node recovery schedule, and finalize the connection requests.

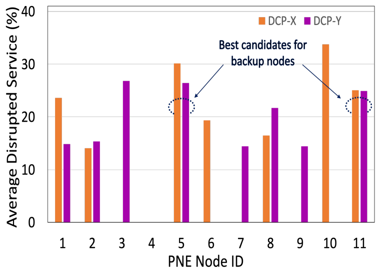

Fig. 3: Impact of PNE-node failure.

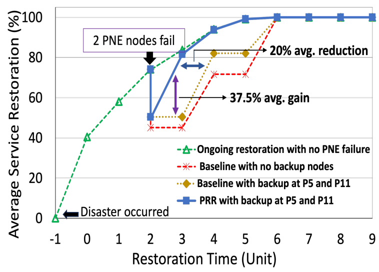

Fig. 4: Concurrent PNE-node failure.

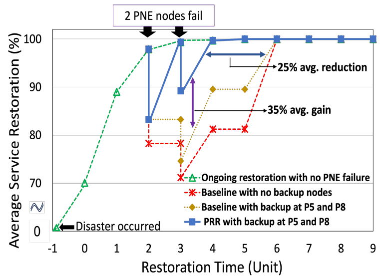

Fig. 5: Sequential PNE-node failure.

For example, Carrier-B can provision a connection request of DCP-Y (i.e., *B6-B9*) in k = 1 unit of time and PNE can recover *P8* in l = 2 units of time. Hence, DCP-Y finalizes the connection request *B6-B9* for Carrier-B instead of waiting for *P8* to recover. Finally, carriers implement the recovery plan and restore the disrupted DC services.

4. Simulation Setup and Illustrative Numerical Analysis

For performance evaluation, we consider the network model as shown in Fig. 1. The PNEN is modeled as a subset of *Japan photonic network* [5] and Carrier-A and Carrier-B abstract their topologies to the PNEN with 11 nodes and 15 links. For DCP-X and DCP-Y, the pre-disaster DCI topologies have 7 and 8 DCs with 11 and 12 DCI links, respectively. Two damage scenarios, each with 50 instances, are generated based on a strong correlation value of 0.8 among the link failures in both carrier networks: (i) *heavy damage*, where 10 links are failed in both carriers, and (ii) *mixed damage*, where, 10 links are failed in Carrier-A and 5 links are failed in Carrier-B or vice-versa. To simulate post-disaster resource crunch in the network, we assume that only 4 lightpath channels (each with 100 Gbps capacity) are available in each optical fiber link. We also consider that 3 to 4 DCs are failed in different disaster instances and a total of 15000 user requests (each with bandwidth requirement uniformly distributed between 70 to 100 Mbps) are affected. During post-disaster DC-service restoration, two random PNE nodes fail concurrently and sequentially for *heavy* and *mixed damage*, respectively. We assume backup switches are deployed at two PNE nodes selectively and that one failed PNE node can be recovered in $l = 2$ units of time. We consider a *baseline* strategy in which DCPs wait for the failed PNE nodes to be recovered to restore the disrupted services.

In Fig. 3, we illustrate the impact of different PNE-node failures during post-disaster DC-service restoration in case of *heavy damage*. Based on DC-node location, different DCPs experience disruption in service restoration differently. For example, we see that failure of *P4* does not impact service restoration of any DCP. Failure of *P3*, *P7*, and *P9* does not impact DCP-X, and failure of *P6* and *P10* does not impact DCP-Y. However, about 25-30% services of both DCPs are disrupted by failure of either *P5* or *P11* and hence, we consider deploying backup switches at these two PNE nodes. Similarly, for *mixed damage*, *P5* and *P8* are selected for deploying the backup switches. Figs. 4 and 5 illustrate the average service restoration of DCPs w.r.t. restoration time. We show that disaster occurred at time unit "-1" and post-disaster DC-service restoration is initiated through DC-carrier cooperation. Average DC-service restoration reaches 40% for *heavy damage* (Fig. 4) and 70% for *mixed damage* (Fig. 5) at time unit 0. The ongoing restoration is then disrupted due to PNE-node failure. We observe that, even with two backup PNE nodes, the disruption is about 25% in case of concurrent node failures (*heavy damage*) at time unit 2 and about 15% and 9% in case of sequential node failure (*mixed damage*) at time units 2 and 3, respectively. Note that, without any backup PNE node, the amount of disruption is at least 5% higher in all cases. Now, let us focus on the restoration gain. We first see that the *baseline* with backup PNE nodes provides 10% restoration gain (on average) compared to the *baseline* with no backup nodes. We further compare PRR strategy and the *baseline*, both considering two backup PNE nodes. We see that PRR ensures at least 35% restoration gain (on average) and at least 20% less restoration time (on average) compared to the *baseline* in both damage scenarios. This is because, while the *baseline* only waits for recovery of the failed PNE nodes to restore the disrupted services, PRR employs efficient DC-carrier cooperation for rapid service restoration.

5. Conclusion

We presented the impact of PNE-node failure during post-disaster cloud-service restoration through DC-carrier cooperation. Our proposed PNE-failure-aware rapid restoration strategy significantly reduces disruption and improves DC-service restoration by 35% while incurring 20% less restoration time compared to baseline approach.

Acknowledgement: This work is supported by US-Japan JUNO3 project: NSF Grant no. 2210384.

References

1. Forbes Technology Council, "Why migrate to the cloud: the basics, benefits and real-life examples," Mar. 2021.
2. S. Xu et al., "Enhancement of network-cloud ecosystem resilience with openness disaggregation and cooperation," *Proc. OFC*, Mar. 2023.
3. S. Sahoo et al., "Datacenter-carrier cooperation over optical networks during disaster recovery," *Proc. OFC*, Mar. 2022.
4. S. Sahoo et al., "Strategic cooperation among datacenter providers and optical-network carriers for disaster recovery," *Proc. Globecom*, Dec. 2022.
5. T. Sakano et al., "A study on a photonic network model based on the regional characteristic of japan," *IEICE Rpt.*, 2013.
6. S. Xu et al., "A novel strategy of carrier cooperation with coordinated scheduling for swift failure/disaster recovery," *Proc. ONDM*, May 2023.