Datacenter-Carrier Cooperation over Optical Networks during Disaster Recovery

Subhadeep Sahoo[1], Sugang Xu[2], Sifat Ferdousi[1], Yusuke Hirota[2], Massimo Tornatore[1, 3], Yoshinari Awaji[2], Biswanath Mukherjee[1, 4]
[1]University of California, Davis, USA, [2]National Institute of Information and Communications Technology (NICT), Japan, [3]Politecnico di Milano, Italy, [4]Soochow University, China
subsahoo@ucdavis.edu

Abstract: A novel cooperation strategy among DC service provider and carriers (operating optical networks) is proposed for disaster recovery. This cooperation improves service restoration by 70% w.r.t. benchmark methods for typical scenarios, with reduced cost.

1. Introduction

Internet users and applications are increasingly migrating to the cloud [1]. To improve the quality and availability of cloud services, datacenter service providers (DCP) replicate contents among their geographically-distributed DCs through DC interconnect (DCI) networks, established by using connection services (e.g., IP-over-WDM or optical lightpaths) over telecom carrier networks [2]. In such an ecosystem, resilience to failures is essential. In particular, when a large-scale disaster occurs, DC services can be severely disrupted due to failures in both carrier network infrastructure and DC infrastructure [3]; thus, resiliency strategies for swift and efficient restoration of both infrastructures are very desirable. Post-disaster recovery strategies for joint DC and network ecosystems have been studied [4], showing the *potential benefit of DC-carrier joint restoration when complete information of DC/network is owned by the same entity*. However, when the networks and the DCs are owned by different entities, some confidential information, such as content locations in DCs, network topologies of carriers, detailed damage situations, etc., is not disclosed, making it challenging to achieve restoration optimization. In these cases, cooperation among DCP and carriers, in the form of sharing network resources (based on incentives), brings greater benefit (in terms of quick service recovery) compared to relying on individual network recovery. Moreover, *to facilitate cooperative restoration and avoid confidential information leakage, a third-party mediator called emergency shareable exchange (ESE) and a public reference ESE network (ESEN) topology were introduced in* [5].

Considering DC-carrier cooperation and the presence of an ESE, in this work we investigate how to maximize the restoration efficiency as well minimize the restoration cost from DCP's perspective. To the best of our knowledge, for the first time, *we propose a novel DC-carrier cooperation-based recovery strategy with the aid of ESE without sharing of confidential information of individual entities*. A *five-phase* DC-carrier cooperative restoration planning process, comprising different subtasks performed by DCP and carriers, is proposed. Our method, on the considered case studies, improves restoration efficiency by at least 70%, compared to conventional non-cooperation-based recovery, with minimum additional cost.

2. Problem Statement and Information Exchange in Cooperative Restoration Strategy

We consider the problem of maximizing post-disaster service restoration in a DCI network. A DCP prepares its own DCI topology (top layer in Fig. 1), to interconnect its DCs and serve content requests of its customers, by leasing connection services, i.e., DCI links (as IP-over-WDM connections with specified bandwidth or optical lightpaths) from multiple carriers. As illustrated in Fig. 1, blue and pink links (without arrows) in the DCI topology represent links that are mapped on underlying carrier networks of Carrier-A and Carrier-B, respectively. A third-party mediator (the ESE) builds a common public ESEN reference topology, where each ESEN node is an actual electronic switch employed to interconnect different carriers' nodes (e.g., that are in the same proximity) via optical-electrical-optical (OEO) conversion [5]. In case of a disaster, the carriers abstract their topologies to the subset of the ESEN reference topology within the disaster area to hide their actual damage information. As shown in Fig. 1, abstracted *node 0* in both carriers' networks represents the portion of their networks outside of the disaster area and the remaining portion (*node 1* to *node 11*) represents the abstracted disaster-affected area. The dashed and solid links in the carrier networks represent the damaged and survived links, respectively, after a disaster. Due to disaster, DCI network also gets affected and several DCI links (dashed blue and pink links) and a few DCs (*10, 11*, *1*) are damaged. We assume that, in a damaged DC, switching elements can be restored immediately for traffic exchange, but full restoration of content service requires a longer time. Hence, content requests towards damaged DCs should be redirected to survived backup DCs. For example, in Fig. 1, a content k is placed in a primary DC (*D10*) and replicated in a backup DC (*D2*). Since

D10 is damaged, requests for content *k* need to be redirected to *D2*, for which DCP first tries to do IP-layer recovery, then DCI optimization with content redistribution if needed (discussed in Sec. 3). After that, DCP requests connection services for the failed DCI links from the carriers. The carriers jointly provision the DCP requests using carrier-cooperation scheme, facilitated by ESE (as in [5]). In Fig. 1, DCI links *D6-D2* and *D10-D6* are provisioned by Carriers A and B, respectively (shown as thick pink and blue lines with arrows).

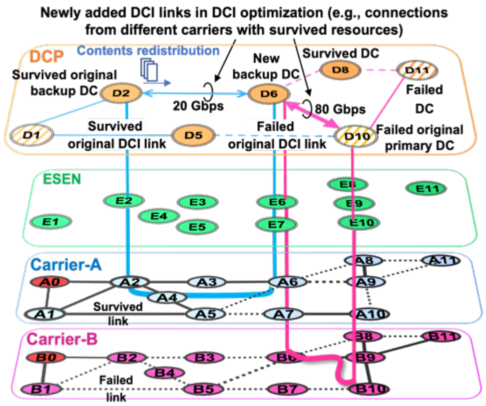

Fig. 1. Network model.

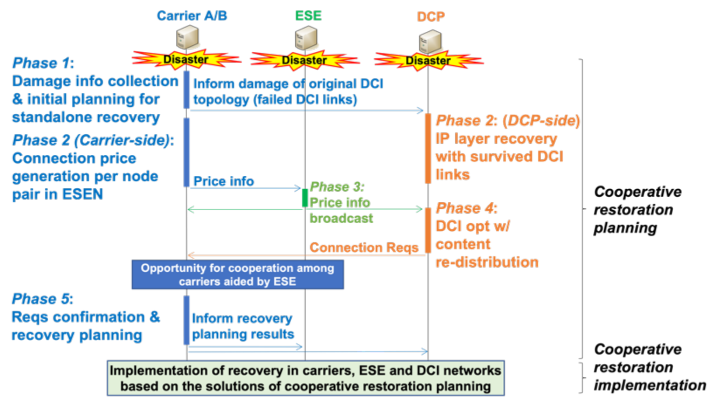

Fig. 2. A five-phase planning for cooperative restoration.

In post-disaster restoration, some services can be restored immediately owing to survived resources and some may not be restored readily (due to delay in recovery of damaged resources). To facilitate the recovery optimization in both DC and carriers, during information exchange among different entities, each carrier first advertises an abstracted *pseudo* price (public information) of connection services for each ESE node pair in ESEN. For immediate service restoration based on survived resources, regular (pre-disaster) service price is adopted. For non-immediate service restoration, which requires recovery of damaged resources involving repair expenses, an extra high-value dummy price is added. With this *pseudo* price, the carriers can hide their actual damage information and guide the DCP to utilize survived resources first and avoid carriers' damaged network resources (if possible). For the final payment of all connections, we assume carriers charge regular price to retain their trust from their customers.

3. Decomposed Optimizations in Cooperative Restoration Strategy

The optimization of DC-carrier coordinated restoration is decomposed into DCP-side and carrier-side subtasks. From DCP's perspective, the objective is to maximize service restoration for its users and identify the required DCI links for restoration with minimum cost (payment to carriers). A carrier's goal is to restore requests -- of its own customers and the DCP -- and minimize recovery tasks and costs. Our *five-phase* restoration plan is based on cooperation by DCP, carriers, and ESE as illustrated in Fig. 2.

In *phase 1*, carriers analyze their damage situation, perform initial planning for standalone recovery, and inform DCP of the failed DCI links. In *phase 2*, DCP performs **DCI IP-layer recovery** by redirecting the affected content requests to backup DCs based on its survived link resources. Simultaneously, carriers generate the aforementioned *pseudo* price for each node pair in ESEN. In *phase 3*, ESE broadcasts the price information to all carriers and DCP. In *phase 4*, DCP jointly performs **DCI-topology optimization** (based on price information) and **content redistribution** with the objective to minimize restoration cost. As illustrated in Fig. 1, since the original backup DC of content *k* (*D2*) is far away from the damaged primary DC (*D10*), DCP replicates content *k* from *D2* to a survived DC (*D6*) which is relatively nearer to *D10*, and designates a corresponding new DCI link (*D6-D2*) with fewer resources (e.g., 20 Gbps). In addition, DCP designates a new DCI link, *D10-D6,* with more resources (e.g., 80 Gbps) to redirect the requests from *D10* to the new backup DC (*D6*) with minimal hops. After that, DCP **requests connection services** (IP-over-WDM connection or lightpath support) for the DCI links from the carriers. In *phase 5*, each carrier performs carrier-side subtasks using carrier-cooperation scheme [5] to optimize its recovery plan while satisfying the requests and minimizing recovery tasks and costs. If DCP's requests cannot be satisfied, the aforementioned planning process can be performed again to adapt to new conditions. After the planning process, recovery implementation is carried out by all the entities.

4. Simulation Setup and Numerical Analysis

We consider the ESEN reference topology, which is a subset of Japan photonic network model [6] with 11 nodes as shown in Fig. 1. For simplicity, we assume that the abstracted post-disaster topologies of Carrier-A and Carrier-

B are identical to the reference topology with 15 links, and all ESEN nodes and carriers' nodes are colocated. The pre-disaster DCI topology has 7 DC nodes colocated with ESEN nodes. Each lightpath's capacity is 100 Gbps. Three damage scenarios are generated based on strong correlation (e.g., with damage similarity = 0.8) among the carrier links: (a) *heavy damage* with 5 survived links in both carrier networks, (b) *light damage* with 10 survived links of both carriers, and (c) *mixed damage* with 5 and 10 survived links of Carrier A and B, respectively. For each damage scenario, we assume three ranks of carrier link recovery costs (i.e., cost of physical repair) as: {1, 4}, {1, 7}, or {1, 10} units [5]. For each damage scenario/recovery cost, 50 instances are generated to observe the consistency of the findings. We also consider that 3 to 4 DCs are damaged, and in total 10000 requests are assigned to damaged DCs that need to be restored. Bandwidth requirement of each user is uniform over 30–50 Mbps. Each carrier has its own IP traffic demands (10 Gbps) e.g., for urgent post-disaster victim relief between *node 0* and all other nodes. We set the *pseudo* price of a carrier's lightpath as 2 units per survived link, and the extra dummy price as 50 units per damaged link in ESEN. The *pseudo* price of IP-over-WDM connection is proportional to the bandwidth demand. Final payment of the connection services is based on regular price (2 units per link) for both immediate and non-immediate service restoration. Average simulation runtime is in the order of minutes.

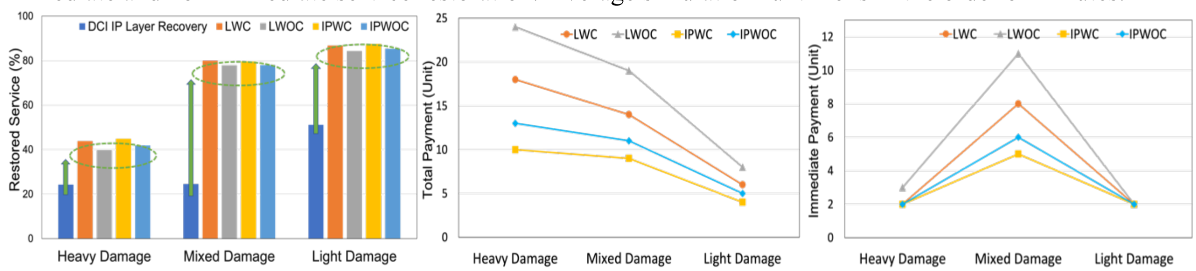

Fig. 3 a) Service restoration efficiency b) Total Payment c) Payments for immediate service restoration, for (i) LWC: Lightpath w/ content-redist (ii) LWOC: Lightpath w/o content-redist (iii) IPWC: IP-over-WDM w/ content-redist (iv) IPWOC: IP-over-WDM w/o content-redist.

To analyze restoration efficiency, we compare four strategies of DCI restoration, namely: IP-over-WDM connection service w/ content redistribution (IPWC) and w/o content redistribution (IPWOC); and lightpath support w/ content redistribution (LWC) and w/o content redistribution (LWOC), with a baseline scheme, i.e., DCI IP-layer recovery without topology optimization and carrier cooperation. With carrier link recovery cost of {*1*, *4*}, Fig. 3a reports that, with the baseline scheme, only a limited number of DC user requests could be restored immediately, whereas with DCI restoration (all four strategies), amount of immediately restored services was more than 70% compared to baseline. Similar trend is observed for other carrier link recovery costs (results are not reported to conserve space).

Fig. 3b compares total payments for connection services (for immediate and non-immediate service restoration) to carriers from DCP. With IPWC, the total payment for establishing new DCI links is significantly reduced in all damage scenarios. The reason is that, compared to lightpath support service which requires payment for the entire lightpath, with IP-over-WDM connection service, DCP only pays for the desired bandwidth to carriers; hence, the cost is comparatively low. By introducing joint content-redistribution scheme in DCI optimization, the DCI link requests could be further optimized. Fig. 3c reports the payment for immediate service restoration in all damage scenarios. For heavy damage, both carriers were lacking survived resources; hence the amount of immediate service restoration is low. For light damage, DCP has limited requirements for connection service from carriers. Hence, both damage scenarios lead to less payment by DCP. However, in mixed damage scenario, the payment was increased since a large number of connection service requests from DCP could be satisfied immediately by Carrier B as it comparatively has more survived resources. Besides DC-carrier cooperation, which is the focus of this paper, further carrier-cooperation can be performed to reduce recovery costs (results are omitted due to space limitation).

5. Conclusion

We proposed a novel DC-carrier cooperation-based strategy for quick post-disaster restoration. It involves abstraction through ESE to avoid leakage of confidential information, DCP-side and carrier-side optimization tasks. The results show significant improvement in restored services with reduced cost facilitated by cooperation.

Acknowledgement: This work is supported by US-Japan JUNO2 project: NSF Grant no. 1818972.

References

[1] T. Fukai, et al., "Live migration in bare-metal clouds," IEEE TCC, 9 (1), 2021.
[2] A. Muhammad et al., "Manycast, anycast, and replica placement in optical inter-datacenter networks," IEEE/OSA JOCN, 9 (12), 2017.
[3] Y. Liu, et al., "Disaster protection in inter-datacenter networks leveraging cooperative storage," IEEE TNSM, 18 (3), 2021.
[4] S. Ferdousi, et al., "Joint progressive network and datacenter recovery after large-scale disasters," IEEE TNSM, 17 (3), 2020.
[5] S. Xu et al., "A novel carrier-cooperation scheme with an incentive to offer emergency lightpath support during disaster recovery," PNET, 2020.
[6] T. Sakano et al., "A study on a photonic network model based on the regional characteristic of Japan," IEICE Tech. Rpt. PN2013-01, 2013.