Reliable Provisioning for Dynamic Content Requests in Optical Metro Networks

G. Le[1], S. Ferdousi[1], A. Marotta[2], S. Xu[5], Y. Hirota[5], Y. Awaji[5], M. Tornatore[1,3], B. Mukherjee[1,4]
[1]University of California, Davis, USA, [2]University of L'Aquila, Italy, [3]Politecnico di Milano, Italy, [4]Soochow University, China
[5]National Institute of Information and Communications Technology (NICT), Japan
dgle@ucdavis.edu

Abstract: We investigate new methods for reliable provisioning of dynamic content requests in optical metro networks. Our methods leverage content replication across multiple edge datacenters and multipath routing.

1. Introduction

Optical metro networks are currently attracting significant investments to evolve from a rigid ring-based aggregation infrastructure to a composite network-and-computing ecosystem where emerging applications and services can be implemented and supported. With adoption of edge computing, several distributed datacenters (DCs) are now available within the footprint of a metro network to reduce latency [1]. Typically, micro DCs are available in metro-access nodes and medium-size DCs in metro core nodes, while hyper-scale DCs are available in core network locations, and they communicate with metro networks via metro core backbone nodes as gateways [2, 3].

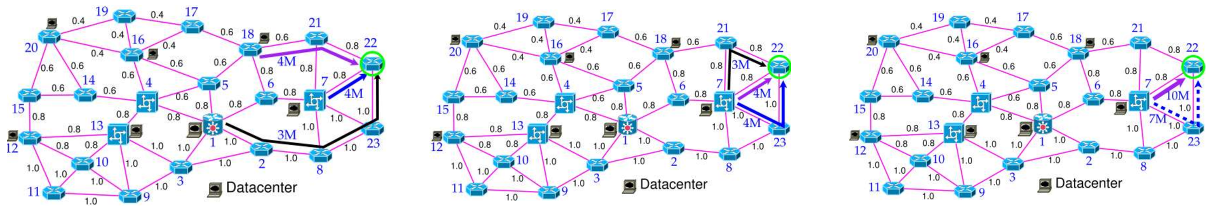
Such redundancy of DC locations offers new opportunities to provide reliable service to users, especially for content-based services, e.g., video delivery which, by 2022, will be 79% of world's mobile data traffic and 82% of all consumer Internet traffic [4]. Other content-based services are emerging, such as augmented reality (AR) and virtual reality (VR), which allow users to interact intuitively with the environment through six degrees of freedom (6DoF) [5]. These services require high bandwidth, low latency, reliable connections, and are classified as mobile broadband reliable low-latency communication (MBRLLC) in the vision of 6G communications [6]. While some of these services require full protection, others can continue to operate with a reduced, i.e., degraded, quality in case of failures and can be served with partial protection (e.g., video streaming can switch to lower resolution depending on available bandwidth).

This study considers the problem of fulfilling a request for a content that is replicated at multiple DCs in a network. This problem has been addressed in [7-10]; and is referred to as anycast (i.e., the user's request can be served by any DC hosting the content) and manycast (i.e., the user's request can be served by any subset of DCs hosting the content). Refs. [7, 8] investigated various content-replication strategies and proposed solutions for anycast provisioning. Refs. [9, 10] focused on manycast provisioning. The above solutions were proposed for a static environment. In this work, we propose, for the first time to the best of our knowledge, a flexible provisioning strategy to fulfill *dynamic* content requests while guaranteeing strict requirements on bandwidth, latency, and survivability. We consider multipath routing where a user request is simultaneously served by multiple DCs. The algorithm must select the optimal subset of the DCs hosting the desired content, find appropriate paths from the selected DCs to the user, and assign enough bandwidth on each path. This solution combines the benefits of multipath routing, provides protection against link and destination (i.e., DC) failures, and uses minimal additional network resources due to the nature of multipath routing.

2. Reliable Provisioning for Dynamic Content Requests

Our reliable provisioning strategy inversely multiplexes a dynamic content request over multiple paths from multiple DCs. Due to the asymmetric traffic of content retrieval, in this work, we only consider the downstream traffic.

A dynamic content request Q is characterized by a tuple $Q =< t, N, C, B, M >$ where t is arrival time, N is requesting node, C is desired content size, B is desired bandwidth, and M is minimum degraded service (which is ratio of survivable bandwidth to desired bandwidth in case of a single failure either on a path or a DC). Without loss of generality, we consider three levels of degraded service, M = 0.5, 0.7, or 1.0, where M = 1.0 denotes full protection. Total bandwidth is sum of bandwidth over all paths while occupied network bandwidth is sum of bandwidth on each path weighted by number of hops. For illustration, in Fig. 1.a, consider a content request from node 22 with desired bandwidth 10 Mbps which is split into three paths, with bandwidth on each path being 4 Mbps, 4 Mbps, and 3 Mbps. Here, total provisioned bandwidth is 11 Mbps while total occupied network bandwidth is 24 Mbps (i.e., 1×4+2×4+4×3), by taking into account the number of hops on each path. For any single failure on either a path or a DC, there is at least 7 Mbps survivable bandwidth which is 70% of desired bandwidth (i.e., M = 0.7).



a. Multiple paths from Multiple DCs (MMD).

b. Multiple paths from a Single DC (MSD).

c. Dedicated-Path Protection (DPP).

Fig. 1. Provisioning strategies: Requesting node 22.

Our problem can be formally stated as follows. Given a content request arriving at node N at time t, **find** optimal subset of DCs that host the content, **compute** link-disjoint paths from these optimal DCs to the requesting node, and **assign** enough bandwidth on each path. The **objective** (1), in descending priority order, is to minimize total bandwidth over all paths (first term), total occupied network bandwidth (second term), total distance from requesting node to the DCs (third term), and total normalized risk factor of all physical links on which the paths are mapped (fourth term) (i.e., $\alpha \gg \beta \gg \gamma \gg \delta$). Here, $x_{d,N}$ is bandwidth on path from DC d to node N, $y_{d,N}^{i,j}$ (integer) and its binarization $z_{d,N}^{i,j}$ are mapping variables of path (d, N) on link (i, j), and $L^{i,j}$ and $R^{i,j}$ are length and normalized risk factor of link (i, j).

$$\min\left(\alpha \sum_d x_{d,N} + \beta \sum_{d,i,j} y_{d,N}^{i,j} + \gamma \sum_{d,i,j} L^{i,j} z_{d,N}^{i,j} + \delta \sum_{d,i,j} R^{i,j} z_{d,N}^{i,j}\right). \quad (1)$$

The solution must fulfill **constraints** on minimum total bandwidth, minimum degraded service in case of a single failure (either on a path or a selected DC), differential delay between paths, survivable routing, available capacity on each link, and computing capacity in each DC at time of request arrival. We formulate this optimization formulation as an integer linear program (ILP) and its compact description is reported in Table 1. This ILP can find optimal solution in less than 0.5 second, on average, in a metro-size network with 23 nodes and 86 links (as reported in next section). We are working on heuristic version for this optimization model. Below, we refer this model as *multiple paths from multiple DCs* (MMD).

Table 1: Optimization model for MMD.

Inputs:	**Outputs:**	**Objective:** Minimize:	**Constraints:**
– $Q = < t, N, C, B, M >$	– Optimal subset of DCs	– Total BW over all paths	– Minimum BW
– Physical network graph	– Number of paths	– Occupied network BW	– Degraded service
– Content replication locations	– BW on each path	– Total distance to DCs	– Differential delay
– Link capacity at t	– How to route each path	– Total norm. risk of links	– Survivable routing
– Comp. capacity in DCs at t	– Content transfer time		– Link, comp. capacity

We use two reference models to benchmark our proposed model, viz. *multiple paths from single DC* (MSD) (Fig. 1.b) and *dedicated-path protection* (DDP) (Fig. 1.c) [9, 10]. It is worth noting that, in terms of total bandwidth, the content request is 10% overprovisioned in case of MMD and MSD while we need 70% overprovisioning (i.e., 10 Mbps on primary path and 7 Mbps on backup path) in case of DPP. As mentioned before, compared to [9, 10], MMD is designed for content requests in a dynamic environment; and depending on resource availability, a request can flexibly select either fewer paths with higher bandwidth on each path or more paths with lower bandwidth on each path.

3. Case Study and Numerical Results

We consider the Tokyo23 optical metro network in Fig. 1 [11], which consists of 43 100-Gbps-bidirectional links and 23 nodes, with each node located at each ward office building in Tokyo metropolitan area. As marked on the figure, node 1 is the metro backbone node which operates as a gateway to core networks; nodes 4, 7, and 13 are metro core nodes; and the remaining nodes are metro access nodes. We assume hyper-scale DCs are accessible via node 1, while medium-size DCs are available at nodes 7 and 13. Edge DCs are deployed at metro access nodes 12, 16, 18, and 20. For this network, we also consider natural disasters, such as earthquakes and tsunamis, as the major cause of fiber cuts, hence we assume the links along the Tokyo Bay coast (bottom right) are at higher risk. Each physical link is characterized by a normalized risk factor (attached to each link) dependent on its location. When multiple lightpaths equally utilize network resources between a pair of virtual nodes, the optimization model selects the lowest-risk path.

Our first experiment considers a catalog of 10000 contents with popularity rank in descending order from 10000 to 1 according to a Zipf distribution. Per [2], top 10% most popular contents are cached in edge DCs, next 10% popular contents are cached in DCs at metro core nodes, and remaining contents are only available in remote DCs through node 1. We use an event-driven simulator to generate 20000 requests where a request's arrival time t follows a Poisson distribution. Requesting node N is randomly selected among metro access nodes; content size C is in range 7 GB (e.g., a short HD video) to 4 TB (e.g., a long live AR-VR video); bandwidth B is randomly selected from discrete values in the range 8 Mbps (e.g., a HD video stream) to 600 Mbps (e.g., a 6DoF video stream); and degraded service M is randomly selected from $\{0.5, 0.7, 1.0\}$. Maximum differential delay is 25 μs. Hit rates of edge DCs, DCs at metro core nodes, and remote DCs are 0.7, 0.15, and 0.15, respectively. Our goal is to compute the acceptance ratio of MMD at different arrival rates and compare it with those of the reference models (MSD and DPP). Fig. 2.a shows that MMD outperforms MSD and DPP in terms of acceptance rate: MMD is 20% and 8% better than MSD and DPP, respectively.

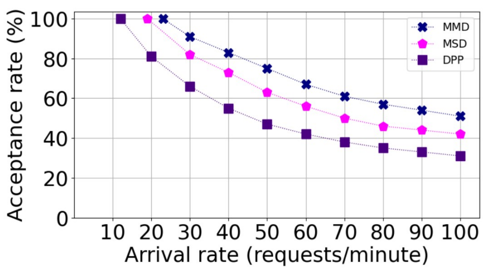


Fig. 2.a. Acceptance rate vs. arrival rate.

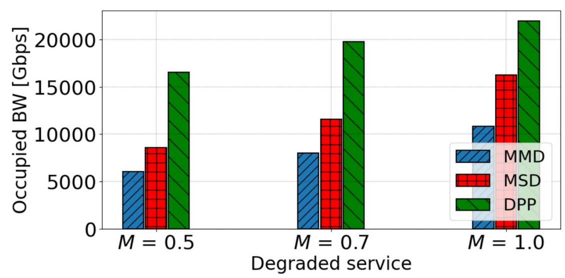


Fig. 2.b. Network resource utilization vs. degraded service.

In the second experiment, we run simulations for MMD, MSD, and DPP for 5000, 10000, 15000, and 20000 requests at 10 requests/minute (i.e., 100% acceptance rate). Average bandwidth per request and average latency per request are reported for MMD, MSD, and DPP in Table 2. On average, MMD needs 50% less bandwidth per request than DPP. MMD and MSD use comparable average bandwidth per request because they both use multipath routing. However, acceptance rate of MSD is lower than that of MMD (Fig. 2.a) because MSD must find enough computing capacity (e.g., for content inspection) in a single DC while MMD can share this computational task among multiple DCs. For each request, latency includes fiber propagation delay (5 μs/km) and processing delay in nodes (e.g., 5 μs/node) on the longest path to DCs. As expected, MMD has highest latency among the three models since MMD requires multiple paths to multiple DCs, probably including a remote DC. Nevertheless, even after adding the latency in the wireless segment (e.g., 1 ms), this total latency is still acceptable for the considered services (e.g., 5 ms) [5].

Table 2: Average bandwidth per request (Mbps) and average latency per request (μs).

# of requests	5000		10000		15000		20000	
	Bandwidth	Latency	Bandwidth	Latency	Bandwidth	Latency	Bandwidth	Latency
MMD	237.5	152.2	236.8	153.6	237.7	151.9	237.3	152.1
MSD	238.1	137.7	237.2	138.2	236.8	136.8	237.4	137.5
DPP	501.7	87.3	502.9	85.1	500.8	88.2	501.7	87.6

In the third experiment, we run simulation for 20000 requests at 10 requests/minute and compare in Fig. 2.b the total occupied network bandwidth of MMD, MSD, and DPP for different levels of degraded service (i.e., M = 0.5, 0.7, 1.0). We find that, at all levels of degraded service, MMD uses much less network resources than MSD and DPP; MMD reduces the total occupied bandwidth in the network by 30% and 64% compared to MSD and DPP, respectively.

4. Conclusion

We proposed and investigated the properties of a reliable provisioning strategy for dynamic content requests in optical metro networks. Numerical results show that this provisioning strategy, leveraging the new redundancy opportunities provide by DC replication in metro area, has higher acceptance rate and minimizes network resource utilization.

Acknowledgement: This work is supported by NSF: Japan-US Network Opportunity 2 (JUNO2), Grant No. 1818972.

References

[1] J. Yao et al., "On mobile edge caching," *IEEE Comm. Surveys & Tutorials*, 2019.
[2] Metro-Haul, "Deliverable D2.1: Definition of use cases, service requirements and KPIs," 2018.
[3] O. Ayoub et al., "Energy-efficient video-on-demand content caching and distribution in metro area networks." *IEEE Trans. on Green Commun. & Networking (TGCN)*, 2018.
[4] Cisco, "Cisco VNI: Global mobile data traffic forecast update, 2017-2022," 2019.
[5] Qualcomm, "Augmented and virtual reality: The first wave of 5G killer apps," 2017.
[6] W. Saad et al., "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, 2020.
[7] C. Natalino et al., "Content placement in 5G-enabled edge/core data center networks resilient to link cut attacks," 2020.
[8] A. Muhammad et al., "Manycast, anycast, and replica placement in optical inter-datacenter networks," *IEEE/OSA JOCN*, 2017.
[9] S. Savas et al., "Disaster-aware service provisioning with manycasting in cloud networks," *Photonic Net. Commun.*, 2014.
[10] R. Goścień et al., "Survivable multipath routing of anycast and unicast traffic in elastic optical networks," *IEEE/OSA JOCN*, 2016.
[11] T. Tachibana et al., "Construction algorithm of metropolitan area network based on regional characteristics in Japan: A case of Tokyo metropolitan area," *IEICE Tech. Rep.*, 2019.