ε-Time Early Warning Data Backup in Disaster-Aware Optical Inter-Connected Data Center Networks

Lisheng Ma, Wei Su, Bin Wu, Tarik Taleb, Xiaohong Jiang, and Norio Shiratori

***Abstract*—Backup in data center networks (DCNs) against disasters is a critical task for avoiding huge data loss. In this paper, we optimize data backup for a particular DCN node threatened by a disaster by assuming geo-distributed optical interconnected DCNs, where the node can be aware of the disaster in an ε time before it is disrupted. We first formulate an integer linear program (ILP) to find the maximum amount of data in the threatened DCN node that can be protected. This helps to determine which data should be protected according to data importance. Then we formulate another ILP to achieve minimum-cost backup by properly selecting a set of safe backup DCN nodes and corresponding backup routes. To get real-time solutions for engineering practice, we also propose a heuristic to achieve cost-efficient backup in ε-time early-warning disasters. Extensive numerical results show that the proposed algorithms can automatically adapt to different early warning times ε for generating cost-efficient data backup solutions.**

***Index Terms*—Data backup; Data center networks; Disaster; ε-early warning time.**

I. Introduction

Many emerging data-intensive applications (e.g., e-commerce, social networking, etc.) are producing huge amounts of data, which rely on multiple worldwide deployed optical interconnected data center networks (DCNs) for data storage and processing. Generally, geo-distributed optical interconnected DCNs consist of massive computing and storage resources interconnected by broadband optical networks with low transmission latency [1,2]. In such networks, failures caused by disasters (e.g., hurricane, earthquake, electromagnetic pulse attack, etc.) can lead to a huge amount of data loss [3–11]. For example, the 2008 earthquake in Wenchuan, China disrupted over 60 enterprise DCNs [4,6], and the 2011 tsunami and earthquake in Japan disrupted tens of DCNs [8,9]. Accordingly, data protection against disasters is a major concern for DCN operators.

DCN data protection against disasters can be achieved through either proactive approaches [12–16] or post-disaster restoration schemes [17,18]. The former designs aim to prevent data loss by network planning before a disaster occurs. The latter utilize resources available at the time of disaster to recover data. Due to the uncertainty of disasters, proactive approaches require a relatively large amount of resources to achieve a desired level of protection. In contrast, post-disaster restoration saves cost, but the effect is generally poor due to its best-effort nature. Our work falls into the first category, with the objective to minimize the data protection cost.

In terms of a proactive approach, data backup is an important data protection method [15,16]. Based on the mutual backup model in [19], some periodic data backup schemes are proposed in [15] and [16] to jointly optimize backup site selection and data transmission paths. Such periodic backup schemes may not result in high data protection efficiency under the disaster scenario because a sudden disaster generally occurs in an unpredictable manner, and thus newly generated data may not be well or timely protected due to the fixed data backup period. On the other hand, [20] and [21] focus on real-time data replications in DCNs where data generated in a certain past period of time is not considered.

To solve these issues in the periodic backup and real-time data replication methods, early-warning-time backup against disasters is proposed in [22] and [23], based on the fact that many natural disasters (such as earthquakes, hurricanes, floods, and tsunamis) can be either observed or predicted with certain early warning information (e.g., affected region and time) [24], thus making it possible to carry out urgent backup in the warning period. For example, the real-time earthquake information system [25] is an earthquake early warning system deployed in Japan that can estimate the location and magnitude of an earthquake

Manuscript received October 13, 2016; revised February 27, 2017; accepted April 24, 2017; published June 1, 2017 (Doc. ID 278654).

L. Ma (e-mail: mls@chzu.edu.cn) is with the School of Systems Information Science, Future University Hakodate, 116-2, Kameda Nakano-Cho, Hakodate, Hokkaido 041-8655, Japan, and is also with the School of Computer and Information Engineering, Chuzhou University, Anhui 239000, China.

W. Su is with the School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044, China.

B. Wu is with the School of Computer Science and Technology, Tianjin University, Tianjin 300072, China.

T. Taleb is with Sejong University, Seoul, South Korea, and Aalto University, Espoo 02150, Finland.

X. Jiang is with the School of Systems Information Science, Future University Hakodate, 116-2, Kameda Nakano-Cho, Hakodate, Hokkaido 041-8655, Japan.

N. Shiratori is with the GITS, Waseda University, and RIEC, Tohoku University, Japan.

https://doi.org/10.1364/JOCN.9.000536

within 5 s after the P-waves arrive. The U.S. National Hurricane Center [26] can provide hurricane warnings from hours to days in advance. For different types of disasters, DCN operators can obtain different early warning times (from a few seconds to a few days) for urgent data backup. This method is more adaptive to sudden disasters by removing the fixed data backup period in [15] and [16]. Unlike the real-time data replications in [20] and [21], it can protect data generated in a certain past period of time as well.

Our study falls into the same category as [22] and [23]. In particular, we consider urgent data backup by properly exploring the ε early warning time, where ε denotes the time interval between the earliest moment that a DCN node is aware of the upcoming disaster and the latest moment, in which the disaster hits the DCN. Generally, ε can be obtained using an early warning system. We call the corresponding DCN node the threatened node for simplicity. Unlike [22], which evacuates as much content as possible from the threatened node to a single backup DCN node, we allow simultaneous backup to multiple safe DCN nodes in the disaster-disjoint zones, and thus can protect more data with a better resource utilization. Although our work considers a similar time-constrained urgent backup as in [23], it indeed differs from [23] on the optimization objective. The work in [23] carries out time-constrained urgent backup to maximize data owners' utility. In contrast, our work achieves cost-efficient backup within the given early warning time. Cost minimization is not considered in both [22] and [23]. As a major concern for DCN operators in selecting a protection strategy, we take backup cost minimization as our optimization objective.

In general, a DCN service provider (such as Google) needs to consider disaster scenarios at the network-planning stage. As a result, multiple DCNs are deployed in different geographical regions to avoid simultaneous failures [27]. Nowadays, such a geo-distributed DCN architecture is well supported by long-haul optical interconnects under wavelength division multiplexing technology. In this paper, we assume that multiple DCNs are affiliated with a single DCN provider and there is only one threatened DCN node. Of course, it is possible that multiple threatened DCN nodes from different DCN providers may need simultaneous data backup. This scenario is more complicated because it concerns service and resource competition among different service providers, and we leave it for future study.

Our design takes a given set of backup resources at different DCN nodes and fiber links as the inputs. Although dedicated backup resource reservation may be pre-planned at the network-planning stage, it is not a concern of this paper. Instead, our proposed scheme can use as much online available bandwidth as possible for urgent data backup against a disaster. Then, we focus on minimum-cost data backup by properly exploring the ε early warning time. To this end, DCN operators first need to know the maximum amount of data that can be protected under the given set of limited backup resources, and then determine which data should be backed up. Accordingly, we divide our design into two sub-problems: backup capacity evaluation (BCE) and backup cost minimization (BCM). BCE helps DCN operators find the maximum backup capacity and thus fully utilize the ε early warning time to back up as much data as possible. Since the maximum backup capacity may not be sufficient for backing up all data, priority can be given to the most important data. On the other hand, BCM minimizes backup cost by properly selecting a set of safe backup DCN nodes and routes for those important data. We propose both integer linear programs (ILPs) and a heuristic for the two sub-problems. Our solutions can be self-adaptive to different early warning times ε.

In this paper, data can be stored in either a distributed manner (across multiple DCNs) or a centralized manner (at a single DCN). Multiple replicas at different DCNs are allowed as well. Generally, a DCN backs up data or replicas periodically in its normal operation state when no disaster presents. At the earliest time when the DCN is aware of the disaster, some data may not have been successfully backed up in the current period or is still in an unsynchronized state. To this end, our work considers how to protect such unsafe data.

The rest of the paper is organized as follows. Section II introduces the network model. Section III formulates the ILPs. The heuristic is presented in Section IV. We carry out simulations in Section V and conclude the paper in Section VI.

II. Network Model

We assume multiple DCNs in an optical backbone network, each hosted by a distinct node. We also assume that data is transmitted in the network through all-optical paths where the optical cross-connects with wavelength converters (i.e., wavelength-conversion capabilities) are used at intermediate nodes for transparent optical connections. Network topology is denoted by a graph $G(V, E)$, where V is the set of all nodes and E is the set of all fiber links. There is a single threatened DCN node that will be affected by a disaster after ε early warning time. Other DCN nodes will not be affected by the disaster, and they can serve as candidate backup DCN nodes. Each candidate backup DCN node has a certain amount of backup storage, whereas the online bandwidth available on each link at the disaster time can be measured by the DCN operator. Figure 1 gives an example of the U.S. InternetMCI network [28], with five geo-distributed DCNs hosted at nodes 3, 8, 12, 14, and 16. Suppose that DCN node 3 will be affected by a disaster after ε early warning time, as shown by the shaded area in Fig. 1. Data hosted at the threatened DCN node 3 can be backed up to the backup DCN nodes 8, 12, 14, and 16. Backup cost consists of data storage and transmission costs. The former is the sum of costs of required storage (counted in data units) at all backup DCN nodes, where W_v denotes the storage cost per data unit at backup DCN node v. The latter counts for the cost of working wavelength capacity (including the cost of necessary wavelength converters) in all backup routing paths.

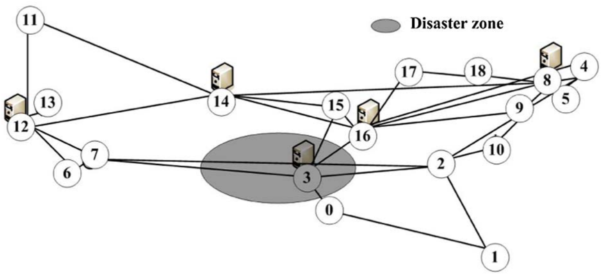


Fig. 1. Geo-distributed DCNs with DCN node 3 threatened by a disaster and other DCN nodes serving as candidate backup nodes.

Note that store-and-forward schemes (using safe DCNs as relays to forward data) are not considered in this paper. This is because we assume only a single threatened DCN, and other DCNs (including those that may possibly serve as intermediate DCNs in a store-and-forward scheme) are taken as safe DCNs. As a result, data will be safely protected as long as they can arrive at such a safe DCN.

III. ILP Formulations

In this section, we first provide an ILP to solve BCE under the ε early warning time constraint, which can be used to determine the amount of data that should be backed up in the threatened DCN node. For the determined amount of data that should be backed up, we also develop another ILP to solve BCM by identifying the optimal selection of backup DCN nodes and routes, such that the overall data backup cost is minimized.

A. Notation List

Inputs:

- V: The set of all nodes in network $G(V,E)$.
- E: The set of all fiber links in network $G(V,E)$.
- $V' \subset V$: The set of all backup DCN nodes in network $G(V,E)$.
- ε: The disaster early warning time for backing up data (quantified with the number of time units).
- $P = \{p | p = \langle Sc, De_p, L_p \rangle\}$: The set of paths between the threatened DCN node and the backup DCN nodes, where Sc, De_p, and L_p are the source DCN node (i.e., threatened DCN node), destination DCN node (i.e., backup DCN node), and the set of links on path p.
- $D = \langle Sc, VL, Vl \rangle$: The data in the threatened DCN node, where Sc is the threatened DCN node, VL is the amount of data D, and Vl is the amount of data that can be backed up, i.e., $Vl \le VL$ (VL and Vl are quantified with the number of data units).
- Re: The transmission rate of each wavelength (quantified with the number of data units that are transmitted by one wavelength per time unit).
- S_v: The available storage capacity in DCN node $v \in V'$ (quantified with the number of data units).
- B_e: The available bandwidth capacity on link $e \in E$ (counted in the number of wavelength channels).
- W_v: The cost of a data unit stored in the DCN node $v \in V'$.
- W_e: The cost of a wavelength on link $e \in E$.
- $A_p^e \in \{0,1\}$: Equals 1 if link $e \in L_p, p \in P$.
- PN: The maximum allowed number of paths between the threatened DCN node and a backup DCN node for backing up data.
- VN: The maximum allowed number of backup DCN nodes for backing up data.
- λ: A predefined constant that is larger than $\max\{B_p, Su_v | \forall v \in V', \forall p \in P\}$.

Variables:

- U_v: Binary variable. It takes 1 if the DCN node $v \in V'$ is used for backing up data and 0 otherwise.
- U_p: Binary variable. It takes 1 if the path $p \in P$ is used for backing up data and 0 otherwise.
- Su_v: Non-negative integer representing the used storage capacity in node $v \in V'$ for backing up data.
- B_p: Non-negative integer representing the used bandwidth capacity on path $p \in P$ for backing up data.
- M_ε: Non-negative integer representing the total amount of data that can be backed up in the threatened DCN node within time ε.

B. ILP for BCE

$$\text{Maximize}\{M_\varepsilon\}. \tag{1}$$

Subject to

$$Su_v \le S_v, \quad \forall v \in V'; \tag{2}$$

$$\sum_{v \in V'} Su_v = M_\varepsilon; \tag{3}$$

$$\sum_{p \in P} A_p^e B_p \le B_e, \quad \forall e \in E; \tag{4}$$

$$\sum_{p \in P} U_p \le PN, \quad \forall De_p \in V'; \tag{5}$$

$$M_\varepsilon \le \sum_{v \in V'} S_v; \tag{6}$$

$$\sum_{v \in V'} U_v \ge 1; \tag{7}$$

$$\sum_{v \in V'} U_v \le VN; \tag{8}$$

$$U_p \leq \frac{U_{De_p} + 1}{2}, \quad \forall\, p \in P; \tag{9}$$

$$\sum_{p \in P, De_p = v} U_p \geq U_v, \quad \forall\, v \in V'; \tag{10}$$

$$U_p \leq B_p, \quad \forall\, p \in P; \tag{11}$$

$$U_p \geq B_p/\lambda, \quad \forall\, p \in P; \tag{12}$$

$$U_v \leq Su_v, \quad \forall\, v \in V'; \tag{13}$$

$$U_v \geq Su_v/\lambda, \quad \forall\, v \in V'; \tag{14}$$

$$\frac{Su_v}{\sum_{p \in P, De_p = v} B_p} \leq \varepsilon \cdot Re, \quad \forall\, v \in V'. \tag{15}$$

Objective (1) maximizes the total amount of data that can be backed up. Constraint (2) ensures that the used storage capacity in a backup DCN node for backing up data does not exceed the available storage capacity of this DCN node. Constraint (3) guarantees that data with the amount M_ε can be backed up to the backup DCN nodes. Constraint (4) ensures that the used bandwidth capacity for backing up data on a link does not exceed the available capacity of this link. Constraint (5) indicates a bound on the number of paths between the threatened DCN node (i.e., Sc) and an arbitrary backup DCN node for backing up data. Constraint (6) guarantees that the amount of data that can be backed up does not exceed the total available storage capacity of all backup DCN nodes. Constraint (7) guarantees that data is backed up to at least one backup DCN node, whereas constraint (8) limits the number of backup DCN nodes to its maximum possible number. Constraint (9) implies that if a path is selected for backing up data, then the destination node of this path must be selected as the backup DCN node for storing data. Constraint (10) implies that if a DCN node is selected as the backup node for storing data, then at least one path destined to such a DCN node must be selected as the transmission path for backing up data. Constraints (11) and (12) define U_p while constraints (13) and (14) define U_v. Here, we use λ larger than $\max\{B_p, Su_v | \forall v \in V', \forall p \in P\}$ to ensure that constraints (12) and (14) can be properly established when $U_p = 1, B_p > 0$ and $U_v = 1, Su_v > 0$, respectively. Constraint (15) ensures that the time for backing up data does not exceed the time ε.

C. ILP for BCM

After we achieve the maximum amount of data M_ε^u that can be backed up within the given ε based on the preceding ILP, we can determine the amount of data D that should be backed up. For the determined amount of data D that should be backed up, we then develop an ILP shown as follows to generate optimal solutions of backup DCN nodes and routes under the time ε constraint, such that the overall data backup cost is minimized:

$$\text{Minimize}\left\{\sum_{v \in V'} W_v Su_v + \sum_{p \in P} \sum_{e \in L_p} W_e B_p\right\}. \tag{16}$$

Objective (16) minimizes the overall data backup cost, which consists of two terms. The first term is the cost of storing all data that should be backed up and the second is the total bandwidth cost for transmitting the data that should be backed up. The constraints in such an ILP are similar to those in Subsection III.B, in which M_ε is replaced by the amount of determined data that should be backed up $Vl, (Vl \leq \min(M_\varepsilon^u, VL))$ in constraints (3) and (6). Although the ILP for BCM has two terms, they can be integrated into a single objective for cost minimization, with storage and transmission costs counted into a total backup cost. Therefore, BCM can be taken as a single-objective optimization. Since we assume wavelength converters at intermediate nodes (if necessary) in the network for transparent optical connections, our ILP models can simply count the bandwidth capacity (i.e., number of available wavelengths) on each link to ensure non-overlapping wavelengths.

To simplify our analysis, we assume static network status within the early warning time. Nevertheless, this assumption can easily be extended to the scenario where network status changes within the early warning time. In this case, we can divide the early warning time into multiple time intervals within which network status can be taken as static for each. Then, our ILPs can be applied in each time interval for data backup. A similar technique is used in [23].

Algorithm 1 Data Backup (DBu):

Input:
$G(V, E)$, $V' \subset V$, Re, D, P, S_v, and W_v for $\forall v \in V'$, B_e and W_e for $\forall e \in E$, and the time for backing up data ε.

Output:
The backup scheme, i.e., the sets of backup DCN nodes (V_b) and backup transmission paths (T_p) for data D, the overall backup cost *Cost*, and the total amount of data that can be backed up within time ε, M_ε.

Set $V_b = \emptyset$, $T_p = \emptyset$, $Cost = 0$, $M_\varepsilon = 0$;
Set $S_v^I = \lfloor \frac{S_v}{\varepsilon \cdot Re} \rfloor \cdot \varepsilon \cdot Re$ for $\forall v \in V'$, $Vl_I = \lfloor \frac{Vl}{\varepsilon \cdot Re} \rfloor \cdot \varepsilon \cdot Re$ for data D;
Set $S_v^R = S_v - S_v^I$ for $\forall v \in V'$, $Vl_R = Vl - Vl_I$ for data D;
Call Procedure **Integer Data Backup**;
Set $Vl_R = Vl_R + Vl_I$ for data D;
Set $S_v^R = S_v^R + S_v^I$ for $\forall v \in V'$;
Call Procedure **Remainder Data Backup**.

Procedure 1 Integer Data Backup (IDBu):

while ($Vl_I > 0$) **do**
Select a path p, ($S_{De_p}^I > 0$) with nonzero available bandwidth $Min_{e \in p}\{B_e\}$ and the ability to back up the largest amount of data from the set P for BCE (Select a path p, ($S_{De_p}^I > 0$) with nonzero available bandwidth $Min_{e \in p}\{B_e\}$ and the smallest cost based on (17) from the set P for BCM);

3: **if** (p is found) **then**
4: Determine a bandwidth $B_p = Min\left(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, Min_{e \in p}\{B_e\}\right)$ on path p;
5: Set $Vl_I = Vl_I - B_p \cdot \varepsilon \cdot Re$, $S^I_{De_p} = S^I_{De_p} - B_p \cdot \varepsilon \cdot Re$;
6: Set $B_e = B_e - B_p$ for $\forall e \in p$;
7: Set $V_b = V_b \bigcup De_p$, $T_p = T_p \bigcup p$;
8: Set $Cost = Cost + W_{De_p} \cdot B_p \cdot \varepsilon \cdot Re + \sum_{e \in p} W_e \cdot B_p$;
9: Set $M_\varepsilon = M_\varepsilon + B_p \cdot \varepsilon \cdot Re$;
10: **else**
11: Exit procedure;
12: **end if**
13: **end while**

IV. Heuristic

Since solving the ILP for large-scale problems (e.g., a large amount of data to be backed up and a large number of backup DCN nodes deployed in a large-scale network) is intractable, it is generally hard to get an optimal ILP solution for data backup in real-time. To make our approach more scalable, in this section we propose a time-efficient heuristic for BCE and BCM to meet the practical engineering requirement. It is notable that, to solve BCE by the heuristic, we only need to set the amount of data to be backed up (Vl) as the total available capacity of all backup DCN nodes (i.e., $Vl = \sum_{v \in V'} S_v$), and then the amount of data (Vl) to be backed up for BCM is determined according to the result from BCE where $Vl \leq \min(M^u_\varepsilon, VL)$.

Procedure 2 Remainder Data Backup (RDBu):

1: **while** ($Vl_R > 0$) **do**
2: Select a path p, ($S^R_{De_p} > 0$) with nonzero available bandwidth $Min_{e \in p}\{B_e\}$ and the ability to back up the largest amount of data from the set P for BCE (Select a path p, ($S^R_{De_p} > 0$) with nonzero available bandwidth $Min_{e \in p}\{B_e\}$ and the smallest cost based on (17) from the set P for BCM);
3: **if** (p is found) **then**
4: Determine a bandwidth $B_p = 1$ on path p;
5: **if** ($S^R_{De_p} \geq Vl_R$) **then**
6: Set $Cost = Cost + W_{De_p} \cdot Vl_R + \sum_{e \in p} W_e \cdot B_p$;
7: Set $M_\varepsilon = M_\varepsilon + Vl_R$;
8: Set $S^R_{De_p} = S^R_{De_p} - Vl_R$, $Vl_R = 0$;
9: **else**
10: Set $Cost = Cost + W_{De_p} \cdot S^R_{De_p} + \sum_{e \in p} W_e \cdot B_p$;
11: Set $M_\varepsilon = M_\varepsilon + S^R_{De_p}$;
12: Set $Vl_R = Vl_R - S^R_{De_p}$, $S^R_{De_p} = 0$;
13: **end if**
14: Set $B_e = B_e - B_p$ for $\forall e \in p$;
15: Set $V_b = V_d \bigcup De_p$, $T_p = T_p \bigcup p$;
16: **else**
17: Exit procedure;
18: **end if**
19: **end while**

A. Algorithm Description

The proposed heuristic is illustrated in Algorithm 1, which includes two procedures, i.e., integer data backup and remainder data backup. Here, V', Re, D, P, S_v, W_v, B_e, W_e, ε, and M_ε are defined in Subsection III.A, and let $|A|$ denote the number of elements in an arbitrarily given set A. Note that in the proposed heuristic the bandwidth on the transmission path is assigned with the integer.

In Algorithm 1, the initialization is first shown in lines 1–3, where we set $V_b = \emptyset$, $T_p = \emptyset$ for data D, Cost = 0, and $M_\varepsilon = 0$, and the available capacity of each backup DCN node and the amount of data D that should be backed up are divided into two parts, respectively, i.e., S^I_v and S^R_v for $\forall v \in V'$, Vl_I and Vl_R for data D. Then we call Procedure 1 (i.e., integer data backup) by taking S^I_v, Vl_I, and the input of Algorithm 1 as its inputs. After executing Procedure 1, Procedure 2 (i.e., remainder data backup) is executed by taking S^R_v, Vl_R, and the input of Algorithm 1 that is updated by Procedure 1 as its inputs.

1) Integer Data Backup: In this procedure, for data D with the amount Vl that should be backed up, we consider only to back up the amount of data Vl_I. In line 2, for BCE, we select a path p ($S^I_{De_p} > 0$) from the set P which has nonzero available bandwidth and the ability to back up the largest amount of data. For BCM, we select a path p ($S^I_{De_p} > 0$) from the set P which has nonzero available bandwidth and the smallest cost. Here, the available bandwidth on path p is $\text{Min}_{e \in p}\{B_e\}$ and the cost is determined as

$$\frac{\sum_{e \in p} W_e + W_{De_p} \cdot \varepsilon \cdot Re \cdot \text{Min}\left(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, \text{Min}_{e \in p}\{B_e\}\right)}{\varepsilon \cdot Re \cdot \text{Min}\left(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, \text{Min}_{e \in p}\{B_e\}\right)}. \quad (17)$$

If we find an available path p, in line 4, the assigned bandwidth B_p on path p for backing up data D is determined as

$$\text{Min}\left(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, \text{Min}_{e \in p}\{B_e\}\right). \quad (18)$$

Expression (18) ensures that the assigned bandwidth on path p for backing up data satisfies the constraints of the available capacity of DCN node De_p, the amount of data that should be backed up, and the available bandwidth on path p. In lines 5–6, we update the values of Vl_I, $S^I_{De_p}$, and B_e for each $e \in p$, respectively. Node De_p is added into set V_b and the path p is also added into set T_p in line 7. The backup cost and total amount of data that can be backed up are obtained in lines 8–9, respectively. If we cannot find an available path p, the procedure exits in line 11.

2) Remainder Data Backup: In this procedure, for data D with the amount Vl that should be backed up, we consider to back up the amount of data Vl_R. In line 2, the path p is selected in the same manner as that in Procedure 1. If we find an available path p, we take a ceiling function of $\text{Min}(\frac{Vl_R}{\varepsilon \cdot Re}, \frac{S^R_{De_p}}{\varepsilon \cdot Re})$, where the value of $\text{Min}(\frac{Vl_R}{\varepsilon \cdot Re}, \frac{S^R_{De_p}}{\varepsilon \cdot Re})$ is less than 1. Then the assigned bandwidth B_p on path p for backing up data D is set as 1. From lines 5 to 13, we update the values of Vl_R, $S^R_{De_p}$, Cost, and M_ε for two cases (i.e., $S^R_{De_p} \geq Vl_R$ and $S^R_{De_p} < Vl_R$), respectively. In lines 14–15, the value of B_e for $\forall\ e \in p$, V_b, and T_p are updated, respectively. If we cannot find an available path p, the procedure exits in line 17.

B. Complexity Analysis of the Heuristic

In this subsection, we analyze the time complexity of the proposed heuristic. Before calculating the complexity of Algorithm 1, we first give the complexity of Procedure 1. In Procedure 1, for backing up data D with the amount Vl_I, the iteration from lines 1 to 13 is executed at most $|P|$ times, i.e., we traverse all paths in set P for backing up data. For line 2, since we need to traverse all available paths for backing up data D in set P, the complexity of this operation is at most $O(|P| \times N)$, where N denotes the maximum number of links on path p for $\forall p \in P$. Furthermore, since $N \leq |E|$, the complexity is no more than $O(|P| \times |E|)$. Besides, the complexity of the operations from lines 4 to 9 is $O(|E|)$. Thus, the complexity of Procedure 1 is no more than $O(|P|^2 \times |E|)$. From Procedure 2, we find that it has the same complexity of Procedure 1. Since both of the complexities of the operation from lines 1 to 3 and that from lines 5 to 6 in Algorithm 1 are $O(|V'|)$, the complexity of Algorithm 1 is $O(|V'| + |P|^2 \times |E|)$ and then the proposed heuristic runs in polynomial time.

V. Numerical Results

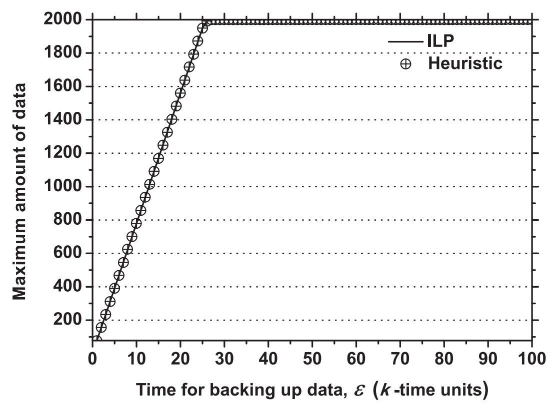
In this section, we carry out numerical experiments on the U.S. InternetMCI network with 19 nodes and 33 links to validate the proposed ILP models and heuristic. We assume that there is an ε-time early warning disaster which will affect DCN node 3 after ε time (i.e., DCN node 3 is the threatened DCN node). The number of available wavelength channels (i.e., available bandwidth capacity) on each link is set as a random integer between 10 and 30. The total available storage capacity in all backup DCN nodes is set as 2000 data units.

In our experiments, we set the cost of a wavelength on a link as the length of the link. In particular, the wavelength cost on each link in the U.S. InternetMCI network is shown in Table I. We set the cost of a data unit stored in a backup DCN node as a random value between 40 and 80. We also set $\lambda = 2000$, $\mathrm{Re} = 1$, and PN as the value of $|P|$. Here, we consider two scenarios, i.e., $|V'| = 4$ backup DCN nodes (i.e., backup DCNs host at nodes 8, 12, 14, 16, and $VN = 4$) and $|V'| = 10$ backup DCN nodes (i.e., backup DCNs host at nodes 2, 5, 7, 8, 9, 11, 12, 14, 15, 16, and $VN = 10$), respectively. Gurobi 6.0 is used to solve the ILPs in Section III. The experiments are run on a computer that has an Intel Core i3-4030U CPU at 1.90 GHz and 4 GB memory.

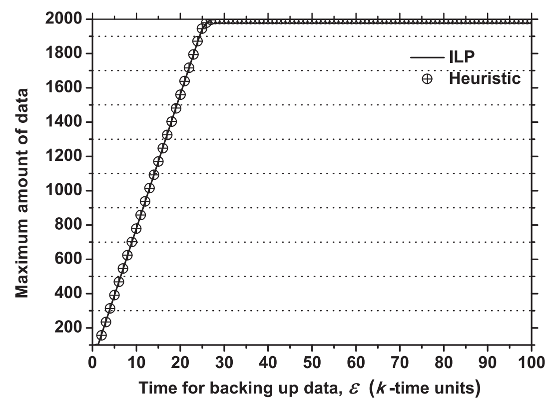
We first provide the comparisons on the maximum amount of data that can be backed up between the ILP and the heuristic for the cases $|V'| = 4$ and $|V'| = 10$, respectively, when ε ranges from 1 to 100 time units, as shown in Fig. 2. From Fig. 2, we can observe that the maximum amount of data that can be backed up by the proposed heuristic is the same as that backed up by the ILP. Note that the ILP gives a mathematical formulation for the BCE sub-problem, whereas its optimal solution can indeed be found by the corresponding heuristic. In other words, our heuristic for BCE is an exact algorithm for generating an optimal solution. This is because we assume only a single threatened DCN node; its maximum amount of protectable data is determined by the available storage capacity

TABLE I
Costs of Wavelength on Each Link in the U.S. InternetMCI Network

Link	Cost	Link	Cost	Link	Cost
(0,1)	625	(4,8)	105	(9,10)	157
(0,3)	133	(4,9)	240	(9,16)	602
(1,2)	352	(4,16)	826	(11,12)	393
(2,3)	488	(5,8)	9	(11,14)	761
(2,7)	1309	(6,7)	35	(12,13)	49
(2,9)	365	(6,12)	223	(12,14)	701
(2,10)	213	(7,12)	249	(14,15)	423
(3,7)	824	(8,9)	135	(14,16)	532
(3,15)	269	(8,14)	1230	(15,16)	128
(3,16)	256	(8,16)	725	(16,17)	249
(4,5)	99	(8,18)	300	(17,18)	252



(a) $|V'| = 4$



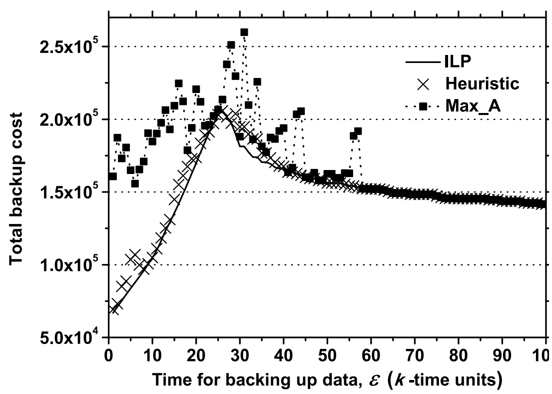
(b) $|V'| = 10$

Fig. 2. Comparison between the ILP and the heuristic of the maximum amount of data that can be backed up for different times ϵ.

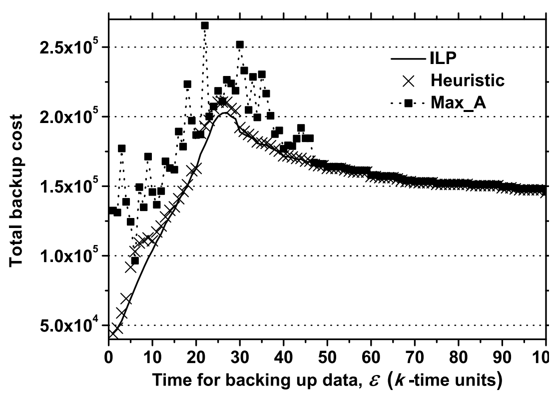
of each backup DCN, as well as the available bandwidth on the paths to those backup DCNs. Since our heuristic fully utilizes the available bandwidth on all paths to those backup DCNs, it can exactly achieve an optimal solution as the ILP. Also note that this is only for BCE, and the situation is different for BCM. We also find that the maximum amount of data that can be backed up increases as the time ε increases and the amount of data that can be backed up reaches the maximum value of 2000 data units for the case $|V'| = 4$, when ε is equal to 27 time units, and for the case $|V'| = 10$ when ε is equal to 28 time units.

In Fig. 3, we then show the total backup cost for each maximum amount of data achieved in Fig. 2. For comparison, inspired by Refs. [15] and [22], we also show the results from the backup scheme with the objective of maximizing the amount of data that can be backed up (referred to as Max_A), which is defined here as a benchmark of our proposed scheme in terms of backup cost. From the results in Fig. 3, we observe that Max_A involves a large cost and our proposed scheme is effective in reducing the backup cost. We also use our proposed ILP as a benchmark to evaluate the performance of the proposed heuristic in Fig. 3. We find that the maximum gap between the ILP and the heuristic is 23.9% for the case $|V'| = 4$, when ε is equal to 5 time units, and 34.9% for the case $|V'| = 10$ when ε is equal to 6 time units. The average gap between the ILP and the heuristic is 5.2% for the case $|V'| = 4$, when ε ranges from 1 time unit to 27 time units, and 8.2% for the case $|V'| = 10$ when ε ranges from 1 to 28 time units. The results in Fig. 3 also show that, after the maximum amount of data that can be backed up reaches the maximum value of 2000 data units, the total backup cost decreases as the time ε increases. This is because more time is available for data backup, and thus less bandwidth is consumed. These results indicate that the proposed heuristic has very good performance.

To further validate the performance of the proposed heuristic, we also give the following comparisons of the ILP and the heuristic. Figure 4 shows the total backup costs from the ILP and the heuristic for the cases $|V'| = 4$ and $|V'| = 10$, when Vl ranges from 1000 to 2000 data units at a fixed $\varepsilon = 28$ time units. The results in Fig. 4 indicate that the total backup cost increases with the increase of Vl. Although the gap of the backup cost between the ILP and the heuristic varies as Vl and $|V'|$ increase, the gap is always less than 5%. In Fig. 5, we show the total backup costs from the ILP and the heuristic for the cases $|V'| = 4$ and $|V'| = 10$ when we increase ε from 10 to 100 time units at a fixed $Vl = 700$ data units. The results in Fig. 5 indicate

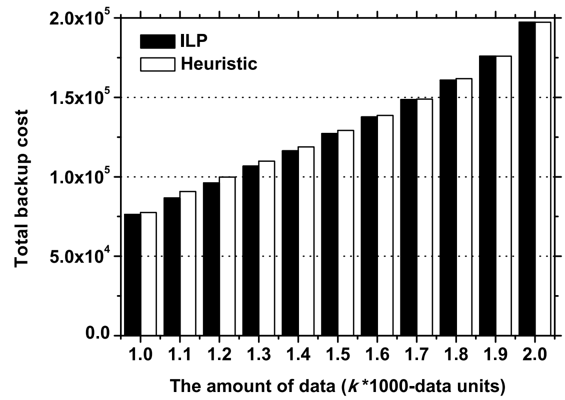


(a) $|V'| = 4$

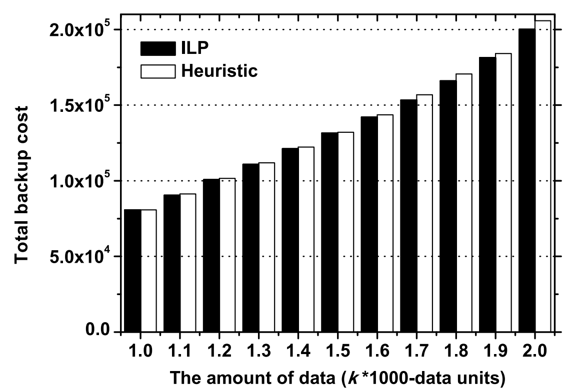


(b) $|V'| = 10$

Fig. 3. Comparison of the total backup cost of the maximum amount of data that can be backed up based on the ILP, the heuristic, and Max_A for different times ε.

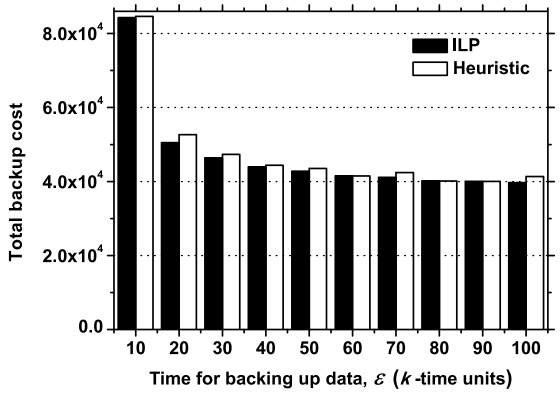


(a) $|V'| = 4$

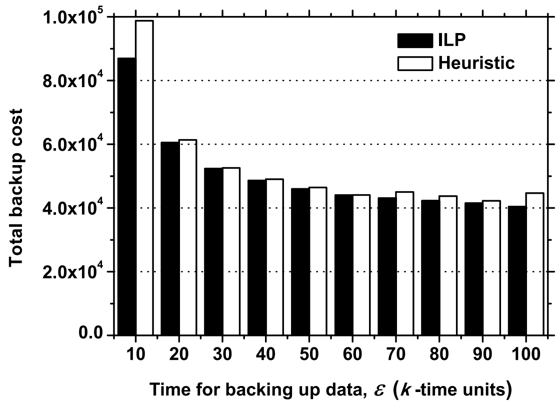


(b) $|V'| = 10$

Fig. 4. Total backup cost comparison between the ILP and the heuristic for different amounts of data with $\varepsilon = 28$ time units.



(a) $|V'| = 4$



(b) $|V'| = 10$

Fig. 5. Total backup cost comparison between the ILP and the heuristic for different times ϵ with $Vl = 700$ data units.

that the total backup cost decreases as ϵ increases. We also find that the gap of the backup cost between the ILP and the heuristic is less than 14%. These results also indicate that the proposed heuristic has very good performance. It is notable that the results from Figs. 2, 3, and 5 show that our scheme can automatically adapt to disasters with different values of ϵ for generating efficient data backup solutions.

TABLE II
RUNNING TIME (IN SECONDS) FOR THE ILP AND HEURISTIC AT $\epsilon = 28$ TIME UNITS

	ILP		Heuristic	
Vl	$\lvert V'\rvert = 4$	$\lvert V'\rvert = 10$	$\lvert V'\rvert = 4$	$\lvert V'\rvert = 10$
1000	3.072	11.587	0.168	0.281
1100	4.066	8.178	0.049	0.984
1200	5.123	7.546	0.056	0.13
1300	5.649	7.391	0.062	0.078
1400	3.041	4.75	0.037	0.21
1500	3.541	6.03	0.903	0.103
1600	2.695	5.822	0.033	0.057
1700	3.485	8.429	0.036	0.08
1800	4.254	9.845	0.034	0.065
1900	3.369	72.342	0.047	0.093
2000	1.717	3.723	0.036	0.063

TABLE III
RUNNING TIME (IN SECONDS) FOR THE ILP AND HEURISTIC WITH $Vl = 700$ DATA UNITS

	ILP		Heuristic	
ϵ	$\lvert V'\rvert = 4$	$\lvert V'\rvert = 10$	$\lvert V'\rvert = 4$	$\lvert V'\rvert = 10$
10	5.224	5.969	0.197	0.312
20	1.411	3.683	0.056	1.027
30	3.525	5.138	0.048	0.087
40	1.917	2.635	0.061	0.082
50	2.159	2.867	1.077	0.083
60	1.505	3.069	0.039	0.081
70	1.442	2.502	0.04	0.052
80	2.942	4.414	0.036	0.18
90	1.519	3.088	0.036	0.058
100	1.633	2.458	0.03	0.126

Tables II and III show the running times for the ILP and the heuristic in Figs. 4 and 5, respectively. We observe that the running time ILP increases with the increase of $|V'|$. In particular, the running time for the ILP reaches the maximum value of more than 72 s when $Vl = 1900$ and $|V'| = 10$. However, the running time for the heuristic increases slowly with the increase of $|V'|$ and Vl, and thus the proposed heuristic is more scalable. Since the running time for the heuristic is small for large-scale backup problems, we can achieve a real-time solution based on the proposed heuristic to meet the practical engineering requirement for an ϵ-time early-warning disaster. For example, under the aforementioned hardware settings (i.e., Intel Core i3-4030U CPU at 1.90 GHz and 4 GB memory), the proposed heuristic can provide backup schemes for all the scenarios in Fig. 4 against a disaster with $\epsilon = 29$ time units early warning time.

VI. CONCLUSION

We studied minimum-cost data backup in geo-distributed optical interconnected DCNs against an ϵ-time early warning disaster under a given set of backup resources. Two sets of algorithms were proposed, each consisting of an optimal ILP and a corresponding heuristic. With the ϵ early warning time constraint, the first set of algorithms can help DCN operators to evaluate the maximum backup capacity under a limited amount of backup resources, and the second set of algorithms can minimize backup cost by properly selecting a set of backup DCN nodes and corresponding backup routes. By properly exploring the ϵ early warning time, the proposed scheme can be more flexible and adaptive to disasters compared with existing periodic backup and real-time replication schemes. Our scheme allows simultaneous data backup from the threatened DCN node to multiple safe DCN nodes in the disaster-disjoint zones. It was shown that the optimal solution changes with different early warning times ϵ, indicating that the proposed scheme is disaster-adaptive under different values of ϵ.

It is notable that we assume only one threatened DCN node in this paper. The backup scheme for the case of multiple threatened DCN nodes is an interesting future research topic. Notice also that we divide the backup problem into two sub-problems in this paper; another future research topic is how to jointly design an optimization problem to maximize the backup capacity while keeping the backup cost minimized.

Acknowledgment

This work was supported by the National Basic Research Program of China (973 Program) (2013CB329101), the National Natural Science Foundation of China (NSFC) (61372085), the Natural Science Foundation of Anhui Province (1708085QF152), the Key Project of Anhui University Science Research (KJ2015A285, KJ2017A425, KJ2017A418), and the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (A), 2015 (26240012).

References

[1] X. Zhao, V. Vusirikala, B. Koley, V. Kamalov, and T. Hofmeister, "The prospect of inter-data-center optical networks," *IEEE Commun. Mag.*, vol. 51, no. 9, pp. 32–38, Sept. 2013.

[2] C. Kachris, K. Kanonakis, and I. Tomkos, "Optical interconnection networks in data centers: Recent trends and future challenges," *IEEE Commun. Mag.*, vol. 51, no. 9, pp. 39–45, Sept. 2013.

[3] K. Tanaka, Y. Yamazaki, T. Okazawa, T. Suzuki, T. Kishimoto, and K. Iwata, "Experiment on seismic disaster characteristics of underground cable," in *The 14th World Conference on Earthquake Engineering*, 2008.

[4] "2008 Sichuan earthquake" [Online]. Available: http://en.wikipedia.org/wiki/2008_Sichuan_earthquake.

[5] A. Kwasinski, W. W. Weaver, P. L. Chapman, and P. T. Krein, "Telecommunications power plant damage assessment for hurricane Katrina-site survey and follow-up results," *IEEE Syst. J.*, vol. 3, no. 3, pp. 277–287, Sept. 2009.

[6] Y. Ran, "Considerations and suggestions on improvement of communication network disaster countermeasures after the Wenchuan earthquake," *IEEE Commun. Mag.*, vol. 49, no. 1, pp. 44–47, Jan. 2011.

[7] K. Morrison, "Rapidly recovering from the catastrophic loss of a major telecommunications office," *IEEE Commun. Mag.*, vol. 49, no. 1, pp. 28–35, Jan. 2011.

[8] "2011 Tohoku earthquake and tsunami" [Online]. Available: http://en.wikipedia.org/wiki/2011_Tohoku_earthquake_and_tsunami.

[9] T. Adachi, Y. Ishiyama, Y. Asakura, and K. Nakamura, "The restoration of telecom power damages by the Great East Japan Earthquake," in *IEEE 33rd Int. Telecommunications Energy Conf.*, 2011, pp. 1–5.

[10] "Flooding, power outages from Hurricane Sandy lead to Internet, phone service disruptions," 2012 [Online]. Available: http://nypost.com/2012/10/30/flooding-power-outages-from-hurricane-sandy-lead-to-internet-phone-service-disruptions/.

[11] A. Kwasinski, "Lessons from field damage assessments about communication networks power supply and infrastructure performance during natural disasters with a focus on Hurricane Sandy," in *FCC Workshop Network Resiliency*, 2013.

[12] M. F. Habib, M. Tornatore, M. D. Leenheer, F. Dikbiyik, and B. Mukherjee, "Design of disaster-resilient optical datacenter networks," *J. Lightwave Technol.*, vol. 30, no. 16, pp. 2563–2573, Aug. 2012.

[13] S. S. Savas, M. F. Habib, F. Dikbiyik, M. Tornatore, and B. Mukherjee, "Disaster-aware service provisioning with manycasting in could networks," *Photon. Netw. Commun.*, vol. 28, no. 2, pp. 123–134, Oct. 2014.

[14] F. Dikbiyik, M. Tornatore, and B. Mukherjee, "Minimizing the risk from disaster failures in optical backbone networks," *J. Lightwave Technol.*, vol. 32, no. 18, pp. 3175–3183, Sept. 2014.

[15] J. Yao, P. Lu, L. Gong, and Z. Zhu, "On fast and coordinated data backup in geo-distributed optical inter-datacenter networks," *J. Lightwave Technol.*, vol. 33, no. 14, pp. 3005–3015, July 2015.

[16] P. Lu, L. Zhang, X. Liu, J. Yao, and Z. Zhu, "Highly efficient data migration and backup for big data applications in elastic optical inter-data-center networks," *IEEE Network*, vol. 29, no. 5, pp. 36–42, Sept.–Oct. 2015.

[17] N. H. Bao, M. F. Habib, M. Tornatore, C. U. Martel, and B. Mukherjee, "Global versus essential post-disaster re-provisioning in telecom mesh networks," *J. Opt. Commun. Netw.*, vol. 7, no. 5, pp. 392–400, May 2015.

[18] C. N. da Silva, L. Wosinska, S. Spadaro, J. C. W. A. Costa, C. R. L. Francês, and P. Monti, "Restoration in optical cloud networks with relocation and services differentiation," *J. Opt. Commun. Netw.*, vol. 8, no. 2, pp. 100–111, Feb. 2016.

[19] J. Yao, P. Lu, and Z. Zhu, "Minimizing disaster backup window for geo-distributed multi-datacenter cloud systems," in *IEEE Int. Conf. on Communications (ICC)*, 2014, pp. 3631–3635.

[20] D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Zomaya, "Energy-efficient data replication in cloud computing datacenters," *Cluster Comput.*, vol. 18, pp. 385–402, Jan. 2015.

[21] R. S. Couto, S. Secci, M. E. M. Campista, and L. H. M. Costa, "Server placement with shared backups for disaster-resilient clouds," *Comput. Netw.*, vol. 93, pp. 423–434, Dec. 2015.

[22] S. Ferdousi, M. Tornatore, M. F. Habib, and B. Mukherjee, "Rapid data evacuation for large-scale disasters in optical cloud networks," *J. Opt. Commun. Netw.*, vol. 7, no. 12, pp. B163–B172, Dec. 2015.

[23] P. Lu, Q. Ling, and Z. Zhu, "Maximizing utility of time-constrained emergency backup in inter-datacenter networks," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 890–893, May 2016.

[24] B. Mukherjee, M. F. Habib, and F. Dikbiyik, "Network adaptability from disaster disruptions and cascading failures," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 230–238, May 2014.

[25] H. Nakamura, S. Horiuchi, C. Wu, S. Yamamoto, and P. A. Rydelek, "Evaluation of the real-time earthquake information system in Japan," *Geophys. Res. Lett.*, vol. 36, L00B01, Jan. 2009.

[26] "National Hurricane Center" [Online]. Available: http://www.nhc.noaa.gov/.

[27] "Google data center locations" [Online]. Available: https://www.google.com/about/datacenters/inside/locations/index.html.

[28] "InternetMCI network," 2011 [Online]. Available: http://www.topology-zoo.org/dataset.html.

Lisheng Ma received the B.S. and M.S. degrees in computer science and technology from Taiyuan Normal University, China, in 2004 and from Southwest University, China, in 2007, respectively. He is currently pursuing the Ph.D. degree at the School of Systems Information Science, Future University Hakodate, Japan, and is also a faculty member at the School of Computer and Information Engineering, Chuzhou University, China. His research interests include switching networks and data center networks.

Wei Su received the Ph.D. degree in communication and information systems from Beijing Jiaotong University in January 2008. He is currently a teacher in the School of Electronics and Information Engineering, Beijing Jiaotong University. He was granted the title of Professor in November 2015. Dr. Su is mainly engaged in researching key theories and technologies for the next-generation Internet and has taken part in many national projects such as the National Basic Research Program (also called the 973 Program), the Projects of Development Plan of the State High Technology Research, and the National Natural Science Foundation of China. He currently presides over the research project Fundamental Research on Cognitive Services and Routing of Future Internet, a project funded by the National Natural Science Foundation of China.

Bin Wu (S'04-M'07) received the Ph.D. degree in electrical and electronic engineering from the University of Hong Kong (Pokfulam, Hong Kong) in 2007. He worked as a Postdoctoral Research Fellow from 2007–2012 in the ECE Department at University of Waterloo (Waterloo, Canada). He is now a Professor in the School of Computer Science and Technology at Tianjin University (Tianjin, China). His research interests include computer systems and networking as well as communication system design.

Tarik Taleb is currently a Professor at the School of Electrical Engineering, Aalto University, Finland. Previously, he was a Senior Researcher and 3GPP Standards Expert at NEC Europe Ltd., Germany. Prior to his work at NEC and until March 2009, he worked as an Assistant Professor at Tohoku University, Japan, in a lab fully funded by KDDI, the second largest network operator in Japan. He received the B.E. degree in information engineering with distinction and the M.Sc. and Ph.D. degrees in information sciences from Tohoku University in 2001, 2003, and 2005, respectively.

Prof. Taleb's research interests lie in the field of architectural enhancements to mobile core networks, cloud-based mobile networking, mobile multimedia streaming, inter-vehicular communications, and social media networking. Prof. Taleb has also been directly engaged in the development and standardization of the Evolved Packet System as a member of the 3GPP's System Architecture Working Group. Prof. Taleb is a board member of the IEEE Communications Society Standardization Program Development Board. In an attempt to bridge the gap between academia and industry, Prof. Taleb has founded and has been the Steering Committee chair of the IEEE Conference on Standards for Communications and Networking (IEEE CSCN).

Prof. Taleb is/was on the editorial board of the *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Vehicular Technology*, *IEEE Communications Surveys and Tutorials*, and a number of Wiley journals. He is serving as Chair of the Wireless Communications Technical Committee, the largest in IEEE ComSoC. He also served as Secretary and then as Vice Chair of the Satellite and Space Communications Technical Committee of IEEE ComSoc (2006–2010). He has been on the Technical Program Committee of several IEEE conferences, including GLOBECOM, ICC, and WCNC, and chaired some of their symposia.

Prof. Taleb is the recipient of the 2009 IEEE ComSoc Asia-Pacific Best Young Researcher award, the 2008 TELECOM System Technology Award from the Telecommunications Advancement Foundation, the 2007 Funai Foundation Science Promotion Award, the 2006 IEEE Computer Society Japan Chapter Young Author Award, the Niwa Yasujirou Memorial Award, and the Young Researchers Encouragement Award from the Japan chapter of the IEEE Vehicular Technology Society (VTS). Some of Prof. Taleb's research work has also been awarded best paper awards at prestigious conferences.

Xiaohong Jiang received the B.S., M.S., and Ph.D. degrees from Xidian University, China. He is currently a Full Professor at Future University Hakodate, Japan. Dr. Jiang was an Associate Professor at Tohoku University, Japan, from February 2005 to March 2010 and an Assistant Professor at the Japan Advanced Institute of Science and Technology (JAIST) from October 2001 to January 2005. Dr. Jiang was a JSPS Research Fellow at JAIST from October 1999 to October 2001. He was a Research Associate at the University of Edinburgh from March 1999 to October 1999. Dr. Jiang's research interests include computer communications networks, mainly wireless networks, optical networks, etc. He has published over 260 technical papers in premium international journals and conferences, which include over 50 papers published in IEEE journals like the *IEEE/ACM Transactions on Networking*, the *IEEE Journal of Selected Areas on Communications*, etc. Dr. Jiang was the winner of the Best Paper Award and Outstanding Paper Award at IEEE HPCC 2014, IEEE WCNC 2012, IEEE WCNC 2008, IEEE ICC 2005 Optical Networking Symposium, and IEEE/IEICE HPSR 2002. He is a Senior Member of IEEE and a member of ACM and IEICE.

Norio Shuratori is currently an Emeritus and Research Professor at the Research Institute of Electrical Communication (RIEC), Tohoku University, Japan. He is also a board member of Future University of Hakodate and a Visiting Professor at Chuo University, Japan. He is a Fellow of the Institute of Electrical and Electronic Engineers, the Information Processing Society of Japan (IPSJ), and the Institute of Electronics, Information and Communication Engineers. He was the President of the IPSJ from 2009 to 2011. He has published more than 15 books and over 400 refereed papers in computer science and related fields. He was the recipient of the IPSJ Memorial Prize Winning Paper Award in 1985, the Telecommunication Advancement Foundation Incorporation Award in 1991, the Best Paper Award of ICOIN-9 in 1994, the IPSJ Best Paper Award in 1997, and many others, including the most recent Outstanding Paper Award of UIC-07 in 2007.