Probabilistic region failure-aware data center network and content placement

Lisheng Ma , Xiaohong Jiang, Bin Wu, Achille Pattavina, Norio Shiratori
Computer Networks
July 2016

Sifat Ferdousi

June 3, 2016

UCDAVIS

Introduction

Given a network for DCN placement, a general probabilistic region failure model is adopted to capture the key features of a region failure and to determine the failure probability of a node/link in the network under the region failure.

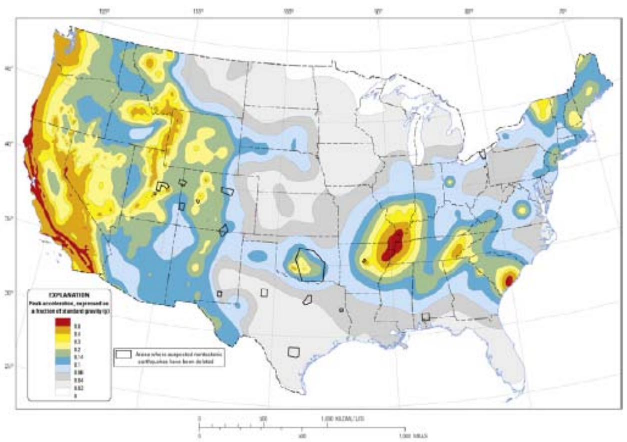
They propose a general grid partition-based scheme to flexibly define the global non-uniform distribution of potential region failure in terms of its occurring probability and intensity. Such grid partition scheme also helps to evaluate the vulnerability of a given network under a region failure and thus to create a "vulnerability map" for DCN and content placement in the network.

UCDAVIS
UNIVERSITY OF CALIFORNIA

- Given a network, the placement of DCN and content in the network with the consideration of potential region failure usually concerns with the following two aspects:

  (1) to assess the network vulnerability due to a region failure;

  (2) based on the network vulnerability information, to properly place the DCN and content in the network such that the DCN failure probability due to region failure is minimized.



- Other works on network vulnerability assessment all assume that both occurring probability and intensity of region failure(s) follow the uniform distribution in the network area. As seen on U.S. national seismic hazard map, we can observe that in the real world, however, a disaster may happen in different areas with different probabilities and different intensities.

UCDAVIS
UNIVERSITY OF CALIFORNIA

They combine the probabilistic region failure model and grid partition scheme (in which the probabilistic region failure model is applied to determine the failure probability of a node/link) to capture the key features of the general non-uniform distribution of a potential disaster in terms of its location and intensity, and then apply them to conduct the network vulnerability assessment.

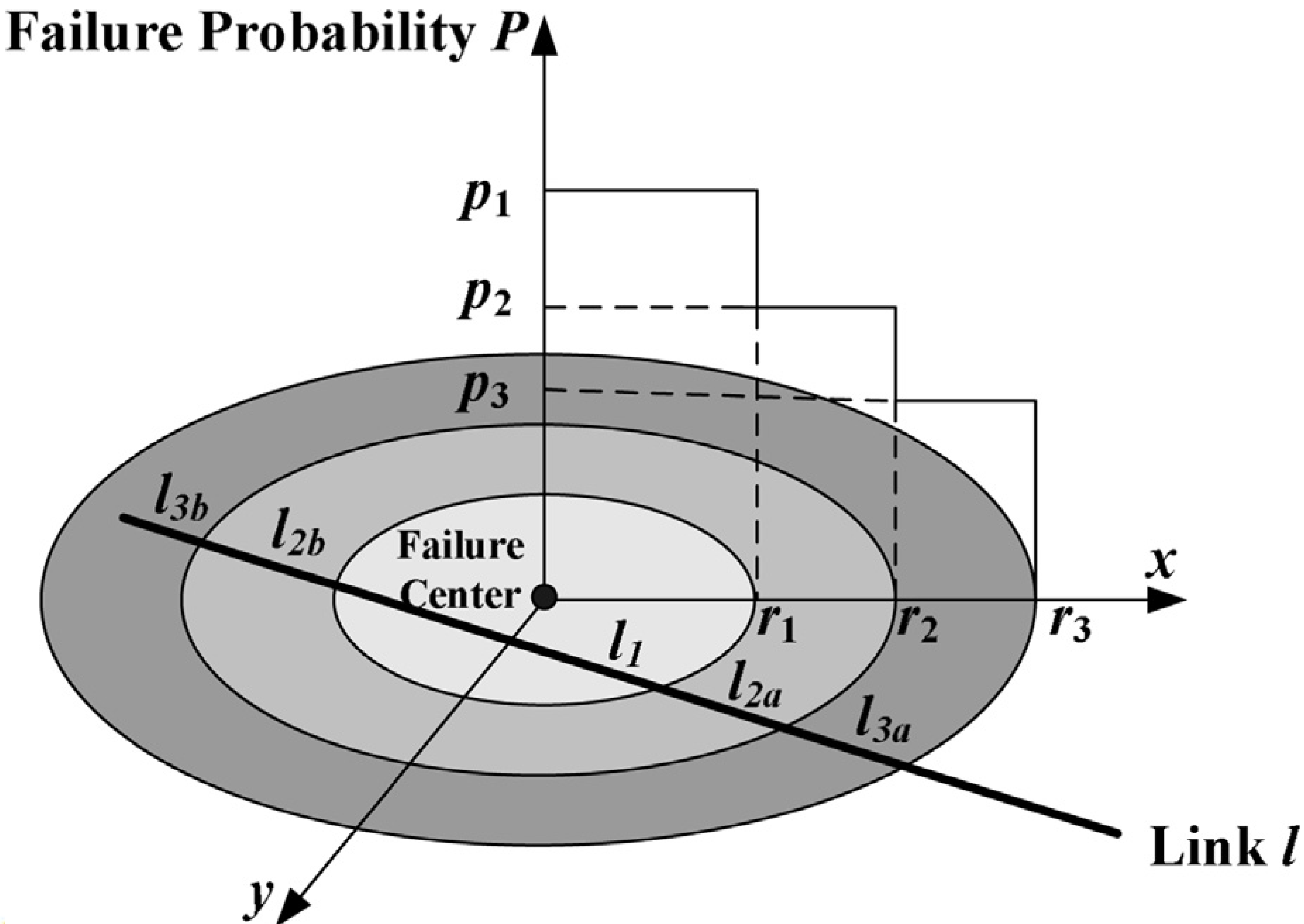
Based on the vulnerability information of a given network for DCN and content placement, an optimal DCN and content placement scheme is proposed with the consideration of the tradeoff among failure probabilities of DCN hosting nodes, failure probabilities of requesting paths and traffic transmission delay.

Network vulnerability evaluation

Probabilistic region failure model

- A real-world disaster (or attack) is usually confined in a specific geographical region. A network component (like a link or node) in this disaster region will fail with certain probability, and such a failure probability depends on the intensity of failure, the distance to failure center and also the dimension of the component (such as the length of a link).
- Under a probabilistic region failure, multiple network components (e.g. nodes and links) may simultaneously fail, but with a certain probability for each. In this paper we evaluate failure probabilities of node and link separately without any dependency between the two.

UCDAVIS
UNIVERSITY OF CALIFORNIA



UCDAVIS
UNIVERSITY OF CALIFORNIA

- Since failure probability evaluations of nodes and links are different from each other as follows, the proposed approaches can properly handle various scenarios. Based on the PRF model, the failure probability *Pv* for a node *v* in the *i*th annulus can be formulated as $P_v = p_i$

- In general, a link spans multiple annuluses of a region failure, and each annulus contains a segment of the link. Then, failure probability of the link is determined by that of all those segments. Therefore, the failure probability *Pl* for a link *l* can be formulated as

$$P_l = 1 - \prod_{i=1}^{m} (1 - P_{l_i})$$

UCDAVIS
UNIVERSITY OF CALIFORNIA

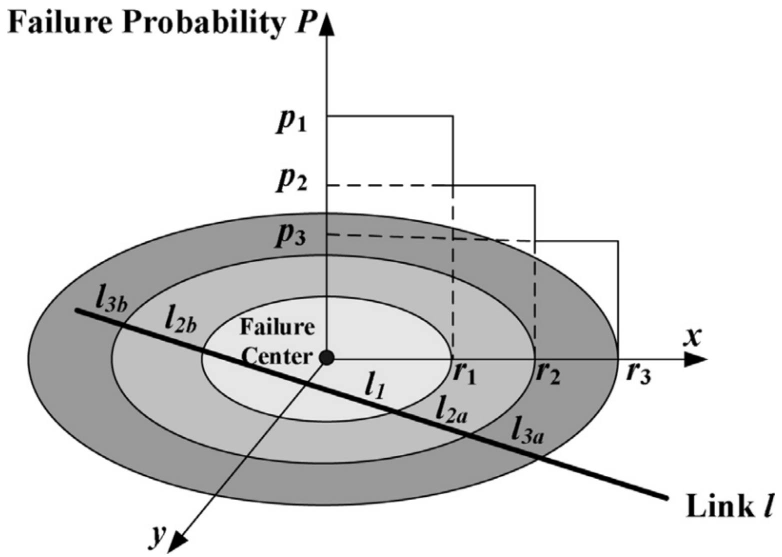
- Consider a segment *li* on link *l* that falls into the *i*th annulus. They first divide such a segment into multiple shorter segments, and each of them is approximated as a node to evaluate the failure probability of *li*. Then, the failure probability *Pli* for *li* can be formulated as

$$P_{l_i} = 1 - (1 - p_i)^{\frac{|l_i|}{\xi}}$$

where ξ is a pre-defined factor representing the length of the shorter segment and |*li*| represents the length of segment *li*.

- Each fiber link has a set of amplifiers. Generally, a link failure is mainly caused by failures of those amplifiers. They treat a segment on a particular link as a sequence of amplifiers, with each approximated as a node to evaluate its failure probability.

UC DAVIS
UNIVERSITY OF CALIFORNIA



$$P_{l_1} = 1 - (1 - p_1)^{\frac{|l_1|}{\xi}},$$
$$P_{l_2} = 1 - (1 - p_2)^{\frac{|l_2|}{\xi}},$$
$$P_{l_3} = 1 - (1 - p_3)^{\frac{|l_3|}{\xi}},$$

where

$$|l_2| = |l_{2a}| + |l_{2b}|, |l_3| = |l_{3a}| + |l_{3b}|$$

Based on link failure probability, failure probability *Pr* for a path *r* can be formulated as

$$P_r = 1 - \prod_{l \in r} (1 - P_l)$$

where *Pl* is the failure probability of a link *l* on path *r* .

Vulnerability metrics

- To evaluate the vulnerability of a network, they consider the following two vulnerability metrics:
  - **NFP (node failure probability):** The probability that a node fails due to a PRF.
  - **LFP (link failure probability):** The probability that a link fails due to a PRF.

- For a given network, one straight-forward approach to assessing the vulnerability of a metric is to first partition the overall network area into some disjoint region failure location (RFL) zones.
  - ***RFL zone definition*** **:** An RFL zone for a specified metric (e.g. NFP or LFP) is a network subarea that any PRF with center in it will always induce the same value of to the network.

UCDAVIS
UNIVERSITY OF CALIFORNIA

Δ NFP or LFP

$$\Delta = \sum_{Z_n} P_{Z_n} \cdot \Delta_{Z_n}$$

Here, *PZn* denotes the probability that a PRF falls within the RFL zone *Zn*

Network deployment area *Z* into a set of disjoint RFL zones *Zn*, where a PRF in *Zn* induces the value *Zn* Δ to the network.

Grid partition-based vulnerability estimation

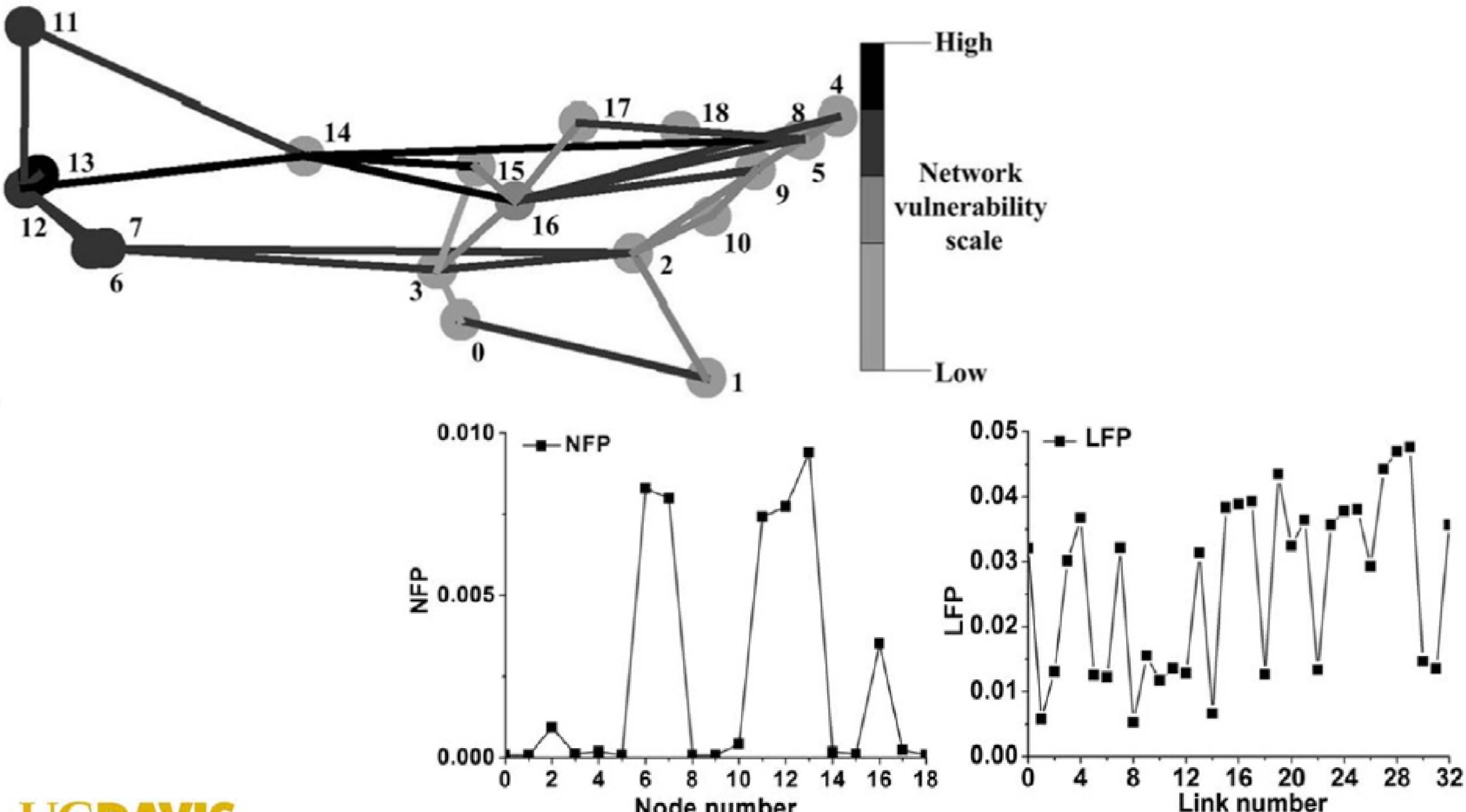
- General grid partition-based scheme helps to flexibly define the non-uniform distribution of PRF and to efficiently estimate the vulnerability of a network.

- Apply a grid partition scheme to evenly divide the network area Z into M small square cells. Based on this grid partition scheme, regard each cell as an "RFL" zone and take the center point of the cell as the failure center to get an estimation of metric.

- Since the intensity of a disaster may be different in different regions, a PRF with center falling within different cells may have different parameters of *ri and pi .*

UCDAVIS
UNIVERSITY OF CALIFORNIA

Vulnerability map



UCDAVIS
UNIVERSITY OF CALIFORNIA

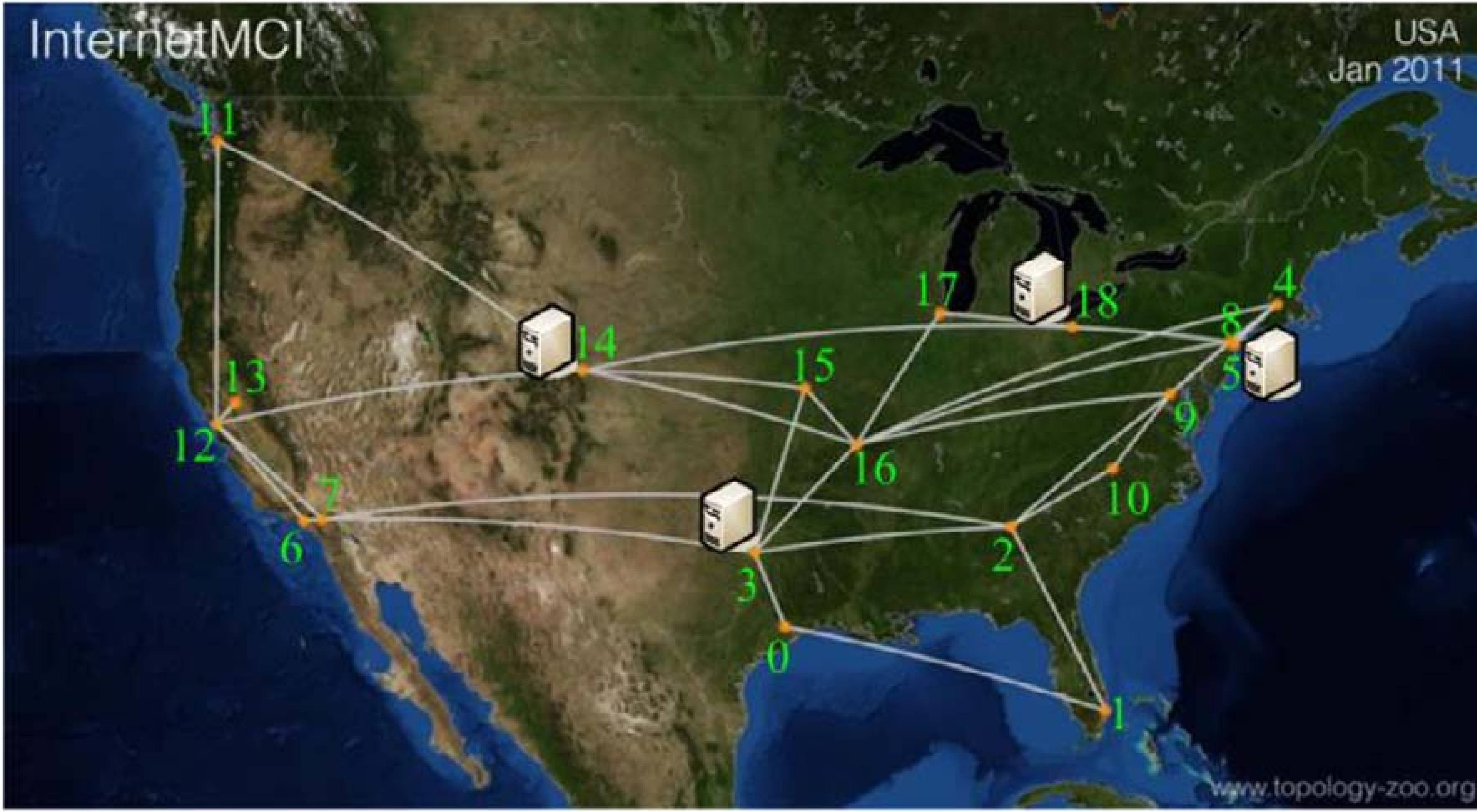
- V: The set of all nodes in network G(V, E).
- E: The set of all links in network G(V, E).
- V': The set of DCN candidate hosting nodes, $V' \subseteq V$.
- C: The set of contents provided by DCNs.
- δ: The scaling factor for adjusting the weight among total failure probability of DCN hosting nodes, total failure probability of requesting paths and total traffic transmission delay.
- S: The set of requesting nodes, $S \subseteq V$.
- R_{sv}: The set of paths between requesting node s and DCN hosting node v.
- N_d: The number of DCNs to be placed.
- N_c: The maximum number of replicas of content c.
- N_{sv}: The number of paths between requesting node s and DCN
- β: Predefined constant greater than the number of contents $|C|$.
- PF_v: The failure probability of DCN candidate hosting node v (Δ_{NFP_v}) obtained by "vulnerability map".
- PF_{rsv}: The failure probability of path r between requesting node s and DCN hosting node v obtained by $P_r = 1 - \prod_{l \in r}(1 - P_l)$.
- PF_{sv}: The average failure probability of paths between requesting node s and DCN hosting node v.
- L_{rsv}: The length of path r between requesting node s and DCN hosting node v.
- L_{sv}: The average length of paths between requesting node s and DCN hosting node v.

- U_v: Binary variable. It takes 1 if a DCN is placed at node v and 0 otherwise.
- U_v^c: Binary variable. It takes 1 if content c is hosted at DCN hosting node v and 0 otherwise.
- U_v^{sc}: Binary variable. It takes 1 if requesting node s requests content c provided by DCN hosting node v and 0 otherwise.

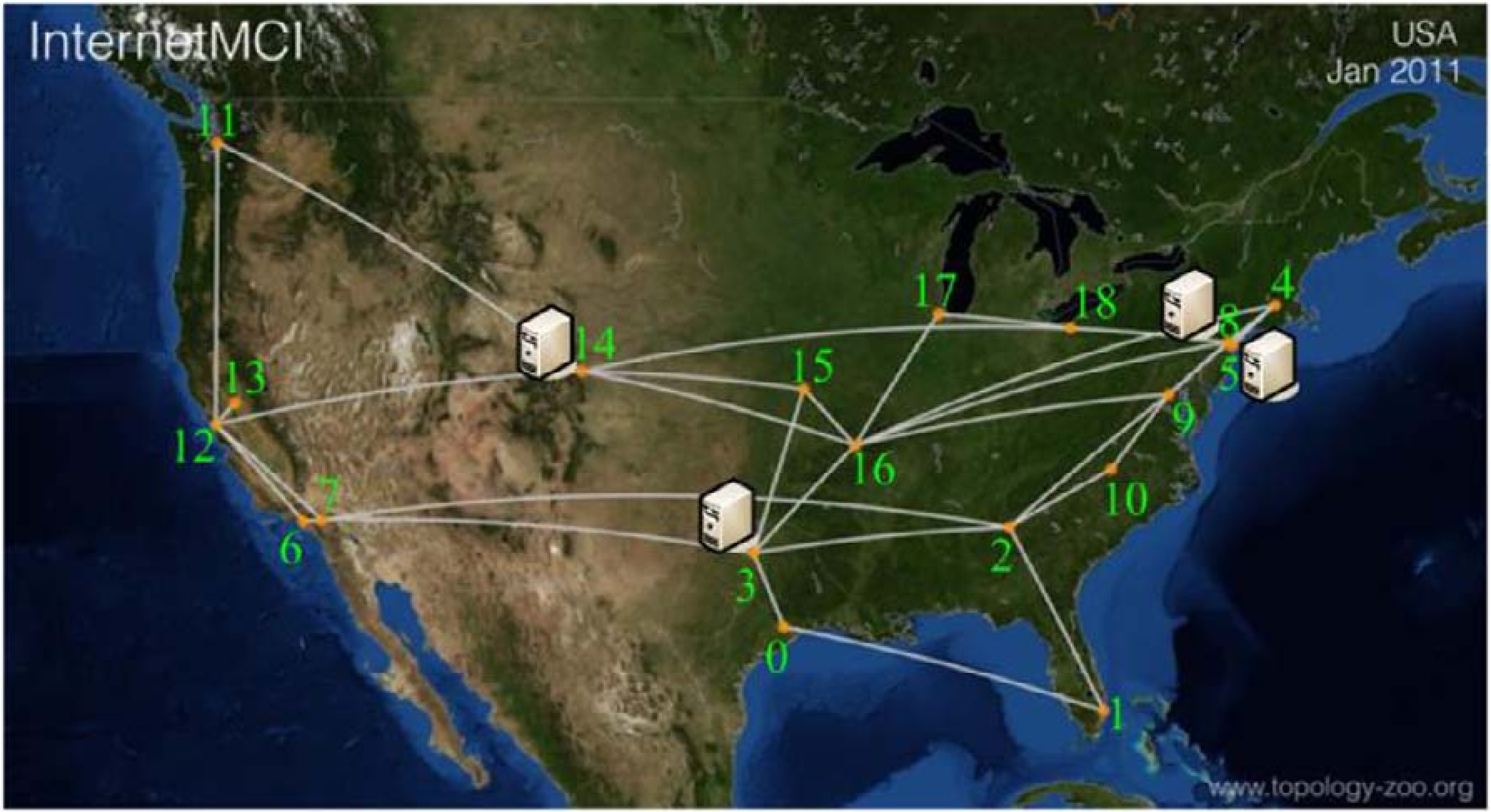
$$\text{Minimize} \quad \left\{ \delta \sum_{v \in V'} U_v PF_v + \sum_{v \in V'} \sum_{s \in S} \sum_{c \in C} U_v^{sc} (PF_{sv} + L_{sv}) \right\}$$

Objective (abbreviated as failure risk) minimizes the total failure probability of DCN hosting nodes and requesting paths, as well as the total traffic transmission delay. The scaling factor δ is used to control the weight among total failure probability of DCN hosting nodes, total failure probability of requesting paths and to- tal traffic transmission delay.

UCDAVIS
UNIVERSITY OF CALIFORNIA



UCDAVIS
UNIVERSITY OF CALIFORNIA


InternetMCI
USA
Jan 2011
11
14
13
12
7
6
15
16
17
18
8
4
5
9
10
3
2
0
1
www.topology-zoo.org