On the Characterization of the Structural Robustness of Data Center Networks

Kashif Bilal, *Student Member*, *IEEE*, Marc Manzano, Samee U. Khan, *Senior Member*, *IEEE*, Eusebi Calle, Keqin Li, *Senior Member*, *IEEE*, and Albert Y. Zomaya, *Fellow*, *IEEE*

Abstract—Data centers being an architectural and functional block of cloud computing are integral to the Information and Communication Technology (ICT) sector. Cloud computing is rigorously utilized by various domains, such as agriculture, nuclear science, smart grids, healthcare, and search engines for research, data storage, and analysis. A Data Center Network (DCN) constitutes the communicational backbone of a data center, ascertaining the performance boundaries for cloud infrastructure. The DCN needs to be robust to failures and uncertainties to deliver the required Quality-of-Service (QoS) level and satisfy service-level agreement (SLA). In this paper, we analyze robustness of the state-of-the-art DCNs. Our major contributions are: 1) we present multilayered graph modeling of various DCNs; 2) we study the classical robustness metrics considering various failure scenarios to perform a comparative analysis; 3) we present the inadequacy of the classical network robustness metrics to appropriately evaluate the DCN robustness; and 4) we propose new procedures to quantify the DCN robustness. Currently, there is no detailed study available centering the DCN robustness. Therefore, we believe that this study will lay a firm foundation for the future DCN robustness research.

Index Terms—Cloud computing, data center networks, multilayer graphs, network analysis, structural robustness

1 Introduction

Cloud computing has emerged as a promising paradigm in various domains of the information and communication technology (ICT). Recently, cloud computing has increasingly been employed to a wide range of applications in various research domains, such as agriculture, smart grids, e-commerce, scientific applications, healthcare, and nuclear science [1]. Data centers being an architectural and operational foundation of cloud, play a vital role in the economic and operational success of cloud computing. Cloud providers need to adhere and comply with the service-level agreement (SLA) and Quality of Service (QoS) for success. Any violation to the SLA may result in huge revenue and reputation loss. Cloud environment is dynamic and virtualized, with a shared pool of resources [2]. Therefore, the resources in the data center are prone to perturbations, faults, and failures. Cloud environment and data center networks (DCNs) need to function properly to deliver required QoS in presence of perturbations and failures [3].

- *K. Bilal and S.U. Khan are with the Department of Electrical and Computer Engineering, 1411 Centennial Blvd, North Dakota State University, Fargo, ND 58102. E-mail: {kashif.bilal, samee.khan}@ndsu.edu.*
- *M. Manzano and E. Calle are with the P-IV building, BCDS laboratory (204-205), Av. Lluís Santaló, s/n, 17071-Girona, University of Girona, Girona 17071, Spain. E-mail: mmanzano@eia.udg.edu, eusebi@eia.udg.edu.*
- *K. Li is with the Department of Computer Science, Faculty Office Building N-3 ,1 hawk Drive, State University of New York , New Paltz, NY 12561. E-mail: lik@newpaltz.edu.*
- *A.Y. Zomaya is with the School of Information Technologies, Building J12, The University of Sydney, Sydney, NSW 2006, Australia. E-mail: albert.zomaya@sydney.edu.au*

Manuscript received 18 Apr. 2013; revised 28 Aug. 2013; accepted 3 Sept. 2013; published online 18 Sept. 2013.
Recommended for acceptance by L. Lefevre.
For information on obtaining reprints of this article, please send e-mail to: tcc@computer.org, and reference IEEECS Log Number TCC-2013-04-0076.
Digital Object Identifier no. 10.1109/TCC.2013.6.

DCNs constitute the communicational backbone of a cloud, and hold a pivotal role to ascertain the data center performance and integrity [4]. A minor network performance degradation may result in enormous losses. Google reported 20 percent revenue loss, when an experiment caused an additional delay of 500 ms in the response time [5]. Moreover, Amazon reported 1 percent sales decrease for an additional delay of 100 ms in search results [5]. Currently, the network robustness quantification of the widely used DCN architectures is unavailable. Therefore, there is an immense need to carry out such a study to quantify the network behavior in the presence of perturbations. A minor failure in the O2 (leading cellular service provider in UK) network affected around seven million customers for three days [6]. Similarly, a core switch failure in the BlackBerry's network left millions of customers without Internet access for three days [6]. The significance of the interconnection networks is obvious from the aforementioned discussion, providing adequate evidences for the robustness requirement of the network. It can be inferred from the discussion that the network robustness holds a key role to ensure desired level of performance and QoS in cloud computing. In the said perspective, measuring the robustness of the DCN is crucial to identify the behavior and level of performance that a network can attain under perturbations and failure-prone circumstances. Therefore, DCN's robustness is a vital measure for proven performance and fault tolerance in cloud computing.

Network (or also referred to as topology) robustness is the ability of the network to deliver the expected level of performance when one or more components of the network fail [7]. Sydney et al. [8] defined robustness as the "ability of a network to maintain its total throughput under node and link removal". Ali et al. [3] consider a system robust, when the system is able to operate as expected in presence of

uncertainties and perturbations. System robustness, and network robustness in particular, has been widely discussed in the literature [7], [8], [9], [10], [11], [12], [13], [14], [15]. Network robustness metrics generally consider the graph theory-based topological features of the network [7]. Several metrics, such as the node connectivity, symmetry ratio, shortest path length, diameter, and assortativity coefficient are used to measure network robustness. However, DCNs exhibit various divergences from the conventional random networks and graph models, such as heterogeneity, multi-layered graph model, and connectivity pattern. DCNs follow a predefined complex architectural and topological pattern, and are generally composed of various layers, such as ThreeTier and FatTree DCNs. Therefore, proper modeling of DCNs is required to measure the robustness.

In this paper, we evaluate various topological features and robustness of the state-of-the-art DCNs namely: 1) ThreeTier [16], 2) FatTree [17], and 3) DCell [18]. Our major contributions include:

- modeling DCN topologies using multilayered graphs;
- developing a DCN graph topology generation tool;
- measuring several robustness metrics under various failure scenarios;
- comparative robustness analysis of the DCN topologies and indicating the inadequacy of the classical robustness metrics to evaluate DCNs;
- proposing new robustness metric for the DCN topologies.

The robustness analysis of the DCN topologies unveiled noteworthy observations. The results revealed that the classical robustness metrics, such as average nodal degree, algebraic connectivity, and spectral radius are unable to evaluate DCNs appropriately. Most of the metrics only consider the largest connected component for robustness evaluation. Consequently, the metrics are unable to depict the factual measurements and robustness of the network. Moreover, none of the DCNs can be declared as more robust based on the measurements taken: 1) without failure and 2) under various failure scenarios. Therefore, we present a new metric named deterioration to quantify the DCN robustness.

The remainder of the paper is organized as follows: Section 2 presents a brief overview of the various DCN architectures. The multilayered graph modeling for DCNs is presented in Section 3. Section 4 provides an overview of the various robustness metrics. The simulation scenarios and experimentation methodologies are detailed in Section 5. The set of networks considered in this work, as well as a discussion regarding their structural robustness, is presented in Section 6. Results are reported in Section 7. The deterioration metric for DCNs is detailed in Section 7.5. Finally, Section 8 concludes the work with possible future directions.

2 DCN Architectures

The network infrastructure holds a focal role in ascertaining the performance bounds of data centers. DCNs [16], [17], [18], [19], [20], [21], [22] can be classified broadly into: 1) switch-centric and 2) server-centric or hybrid models [4]. In the switch-centric models, network devices, such as switches and routers constitute the network. However, in the server-centric or hybrid models, computational devices also serve as network devices besides performing the computational tasks. Various DCN architectures require special routing protocols and possess deployment and management complexities. Details of the various DCN architectures, DCN routing protocols, and DCN comparative analysis are discussed in detail in [1], [4], [23].

The ThreeTier DCN is the most commonly used DCN architecture [16]. The ThreeTier DCN is a switch-centric architecture and the network devices are organized in three layers namely: 1) access, 2) aggregate, and 3) core network layer. The access layer switch connects various computational servers within a rack. Multiple access layer switches are connected to the aggregate layer switches. The core layer switches are used to connect all of the aggregate layer switches.

Al-Fares et al. [17] used commodity network switches to design the FatTree DCN architecture. The FatTree DCN is assembled in k modules (called pods). There are k network switches arranged in two distinct layers (access and aggregate) within each of the pods. Each of the pods contains $(\frac{k}{2})^2$ computational servers that are connected to $k/2$ access layer switches. The core network layer is comprised of $(\frac{k}{2})^2$ core switches. Every core level switch connects to one of the aggregate layer switches in each of the pods.

Guo et al. [18] proposed DCell, a hybrid DCN architecture. The DCell architecture is composed of a hierarchy of modules called $dcells$. In $dcell_0$, n servers are connected to a commodity network switch. $Dcell_0$ provides the foundation of the DCell architecture, and higher level $dcells$ are constructed using $dcell_0$. Each $dcell_{l-1}$ is connected to all of the other $dcells_{l-1}$ within the same $dcell_l$. The computational servers in the DCell architecture also perform packet routing and forwarding functionality besides performing computational tasks.

3 Graph Definitions for DCN Architectures

3.1 Previous Definitions

Kurant and Thiran proposed a general multi-layered graph model in [24]. The authors elaborated that although networks are deliberated as distinctive objects, these objects are usually fragments of complex network, where various topologies interdependently interact with each other. The authors defined two layers of the network: 1) *physical* $\mathcal{G}^{\phi} = \{V^{\phi}, E^{\phi}\}$ and 2) *logical* $\mathcal{G}^{\lambda} = \{V^{\lambda}, E^{\lambda}\}$. The physical graph represents the lower layer topology, and the logical graph represents the upper layer topology. Each logical edge e^{λ} exhibits mapping on the physical graph as a path $M(e^{\lambda}) \subset \mathcal{G}^{\phi}$. Because the number of layers is fixed, the proposed model is inapplicable to the DCN architectures. Moreover, none of the layers in DCNs are logical. Therefore, the idea of mapping one layer to the other is incapable to characterize DCNs.

Dong et al. in [25], defined a multilayered graph $\mathcal{G}$ composed of M layers $\mathcal{G}^{(i)}$, $i = 1, \ldots, M$. Each layer $\mathcal{G}^{(i)} = \{V, E^{(i)}, \omega^{(i)}\}$ represents an undirected weighted graph,

TABLE 1
Definition of the Variables Used in the DCN Models

Variable	Meaning
ν	set of vertices
ε	set of edges
P_i	a *pod* in the topology that is composed of servers and middle-layer switches.
k	number of pods/modules in the topology
n	number of nodes connected to a single access layer switch
m	number of access layer switches in each pod P_i
q	number of aggregate layer switches in each pod P_i
r	number of core switches
δ	servers
α	access layer switch
γ	aggregate layer switch
C	core layer switch

composed of a set of common vertices V and edges $E^{(i)}$ having $\omega^{(i)}$ associated weights. As the number of nodes (vertices) in each layer needs to be same, the proposal is inapplicable to DCNs. Moreover, the definition lacks the interconnection information between different layers in the proposal. Because none of the previously proposed graph models matches the DCN-based graph definition, we present a formal definition for each of the DCN architectures.

3.2 ThreeTier DCN Architecture

We define the ThreeTier architecture according to the definitions in Table 1 as

$$DCN_{TT} = (\nu, \varepsilon), \tag{1}$$

where ν are the vertices and ε represents the edges. Vertices are arranged in k pods P_i^k (servers, access switches, and aggregate switches), and a single layer of r core C_i^r switches:

$$\nu = C_i^r \cup P_i^k, \tag{2}$$

where C_i^r is a set composed of all of the core switches:

$$C_i^r = \{c_i, c_{i+1}, \ldots, c_r\}. \tag{3}$$

Each P_i is composed of three layers of nodes, namely: 1) servers layer (l^s), 2) access layer (l^a), and 3) aggregate layer (l^g). Nodes in each of the pods can be represented as

$$P_i = l^s_{m\alpha \times n\delta} \cup l^a_{m\alpha} \cup l^g_{q\gamma}, \tag{4}$$

where α represent access layer switches, γ represent aggregate layer switches, and δ represent servers. The total number of nodes in each of the pods can be calculated as

$$|P_i| = mn + m + q, \tag{5}$$

where $|P_i|$ stands for the cardinality of the set of nodes in each pod. The total number of vertices of a topology can be calculated as

$$|\nu| = k|P_i| + |C|. \tag{6}$$

There are generally three layers of edges, $\varepsilon = \{\S \cup \acute{\alpha} \cup \zeta\}$, where: 1) $\S$ are the edges that connect servers to the access layer, 2) $\acute{\alpha}$ edges connect the access layer to the aggregate layer, and 3) ζ edges connect the aggregate layer to the core layer. Beside the aforementioned, the ThreeTier topology also has a set of edges connecting the aggregate layer switches to each other within the pod, represented by Υ. Therefore, the set of edges of the ThreeTier DCN can be represented by

$$\varepsilon = \{\S_{(\forall\delta,\alpha)} \cup \acute{\alpha}_{(\forall\alpha,\forall\gamma)} \cup \Upsilon_{(\forall\gamma,\forall\gamma)} \cup \zeta_{(\forall\gamma,\forall C)}\}, \tag{7}$$

where $\S_{(\forall\delta,\alpha)}$, are the edges connecting each server to a single-access layer switch, $\acute{\alpha}_{(\forall\alpha,\forall\gamma)}$, connecting each access layer switch to all aggregate layer switches, $\Upsilon_{(\forall\gamma,\forall\gamma)}$, connecting each aggregate layer switch to all other aggregate layer switches within the pod, and $\zeta_{(\forall\gamma,\forall C)}$, connecting each aggregate layer switch to all of the core layer switches. The total edges ε in the topology can be calculated as

$$|\varepsilon| = k\left(mn + mq + \frac{q(q-1)}{2} + qr\right). \tag{8}$$

3.3 FatTree DCN Architecture

Similar to the ThreeTier DCN, the FatTree architecture is also composed of a single layer of computational servers and three layers of network switches arranged in k pods. However, the FatTree architecture follows a Clos-based topology [26], and the number of networking elements and interconnecting edges in the FatTree architecture are much higher than the ThreeTier architecture. We will use the conventions defined in Table 1 for the graph modeling of the FatTree architecture. The number of elements in each layer within each P_i is fixed based on the k. The number of vertices in n, m, q, and r can be calculated as

$$n = m = q = \left(\frac{k}{2}\right), \tag{9}$$

$$r = \left(\frac{k}{2}\right)^2. \tag{10}$$

The FatTree DCN can be modeled similar to that of the ThreeTier architecture as

$$DCN_{FT} = (\nu, \varepsilon), \tag{11}$$

where ν, C_i^r, P_i, $|P_i|$, and $|\nu|$ can be modeled by using (2) to (6), respectively. Contrary to the ThreeTier architecture, the aggregate layer switches in the FatTree architecture are not connected to each other. Moreover, every core layer switch C_i^r is connected only to a single-aggregate layer switch γ_i from each pod, which allow us to state:

$$\varepsilon = \{\S_{(\forall\delta,\alpha)} \cup \acute{\alpha}_{(\forall\alpha,\forall\gamma)} \cup \zeta_{(\forall C,\gamma_i)}\}, \tag{12}$$

and the total number of edges in the FatTree topology can be calculated as

$$|\varepsilon| = k(mn + mq) + kr. \tag{13}$$

3.4 DCell DCN Architecture

In contrast with the ThreeTier and FatTree DCN architectures, the DCell uses server-based routing architecture. Every $dcell_0$ within the DCell holds a switch to connect all of the computational servers within the $dcell_0$. The DCell uses a recursively built hierarchy, and $dcell_l$ is built of $x_i dcells_{l-1}$. The algorithm for the interconnections among the servers in various $dcells$ can be seen in [18]. The graph model of the DCell architecture can be represented as

$$DCN_{DC} = (\nu, \varepsilon), \tag{14}$$

$$\nu = \{\partial_i, \partial_{i+1}, \ldots, \partial_l\}, \tag{15}$$

where $0 \leq i \leq l$, and ∂_0 represents the $dcell_0$:

$$\partial_0 = \delta \cup \alpha, \tag{16}$$

where δ represents the set of servers within $dcell_0$, s is the number of servers within $dcell_0$, and α is the network switch connecting s servers within $dcell_0$:

$$\partial_l = x_l \partial_{l-1}, \tag{17}$$

where x_l is the total number of ∂_{l-1} in ∂_l.

$$\partial_1 = x_1 \partial_0, \tag{18}$$

$$x_1 = s + 1. \tag{19}$$

Similarly, for $l \geq 2$:

$$x_l = \left(\prod_{i=1}^{l-1} x_i \times s \right) + 1. \tag{20}$$

Let x_l denote the numbers of $dcell_{l-1}$ in a $dcell_l$. Let s_l denote the numbers of servers in a $dcell_l$. Then, we have:

$$\begin{aligned} &s_0 = 2, \\ &x_1 = 3, s_1 = 2 * 3 = 6, \\ &x_2 = 7, s_2 = 6 * 7 = 42, \\ &x_3 = 43, s_3 = 42 * 43, \\ &\ldots \\ &x_l = s_{l-1} + 1, s_l = x_l * s_{l-1} = (s_{l-1} + 1) * s_{l-1}. \end{aligned} \tag{21}$$

It can be verified by induction that (20) holds and

$$s_l = x_1 * x_2 * \cdots * x_l * s. \tag{22}$$

The DCell DCN is a highly scalable architecture and supports any level of $dcells$. However, a *3-level* DCell is sufficient to accommodate millions of servers. The total number of nodes in a *3-level* DCell can be computed as

$$|\nu_0^3| = \sum_1^{x_3} \sum_1^{x_2} \sum_1^{x_1} (s+1), \tag{23}$$

and the total number of edges in a *3-level* DCell are:

$$\begin{aligned} |\varepsilon_0^3| = \sum_1^{x_3} \Bigg[\sum_1^{x_2} \left(\left(\sum_1^{x_1} s \right) + \left(\frac{x_1(x_1-1)}{2} \right) \right) \\ + \left(\frac{x_2(x_2-1)}{2} \right) \Bigg] + \left(\frac{x_3(x_3-1)}{2} \right). \end{aligned} \tag{24}$$

The total number of nodes in the *l-level* DCell, ∂_0^l, can be computed as

$$|\nu| = \left(\frac{\prod_{i=1}^{l} \left(\sum_1^{x_i} (s+1) \right)}{(s+1)^{(l-1)}} \right), \tag{25}$$

and the total number of edges in the *l-level* DCell, ∂_0^l, can be computed as

$$|\varepsilon| = \left(\frac{\prod_{i=1}^{l} \left(\sum_1^{x_i} (s) \right)}{(s)^{(l-1)}} \right) + \frac{1}{2} \left[\sum_{j=1}^{l} \left(\left(\prod_{k=j}^{l} x_k \right) (x_j - 1) \right) \right]. \tag{26}$$

TABLE 2
Classical and Contemporary Robustness Metrics

	Characteristic	Reference
Classical	Average nodal degree ($\langle k \rangle$)	[13]
	Node connectivity (κ)	[9]
	Link connectivity (ρ)	[9]
	Heterogeneity ($\sqrt{\sigma_k^2}/\langle k \rangle$)	[10]
	Symmetry ratio ($\epsilon/(D+1)$)	[11]
	Diameter (D)	[32]
	Average shortest-path length ($\langle l \rangle$)	[12]
	Assortativity coefficient (r)	[13]
	Average neighbor connectivity ($\frac{k_{nn}}{\lvert\nu\rvert-1}$)	[13]
	Clustering coefficient ($\langle C \rangle$)	[13], [29]
	Betweenness centrality ($\langle b \rangle$)	[28]
	Largest eigenvalue or spectral radius (λ_1)	[13], [14]
	Second smallest Laplacian eigenvalue or algebraic connectivity ($\mu_{\lvert\nu\rvert-1}$)	[15]
	Average two-Terminal Reliability (A2TR)	[27]
Contemporary	Elasticity (E)	[8]
	Quantitative Robustness Metric (QNRM)	[7]
	Qualitative Robustness Metric (QLRM)	[7]
	R-value (R)	[30]
	Viral Conductance (VC)	[31]

4 Robustness Metrics

4.1 Background

This section briefly presents some of the well-known graph robustness metrics. Some of the metrics classified here (see Table 2) as *classical* are based on the concepts of the graph theory, while the *contemporary* metrics consider the services supported by the networks. In this paper, we consider the classical robustness metrics, leaving the dynamic aspects of the DCN robustness as future work. A brief description of the robustness metrics is presented in the following section.

4.2 Robustness Metrics Glossary

Assortativity coefficient (r): presents the tendency of a node to connect to other nodes having dissimilar degrees [13]. The value of r lies within the range $-1 \leq r \leq 1$. The value of $r < 0$ represents *dissassortative* network, having excess of links among nodes of dissimilar degrees.

Average neighbor connectivity ($\frac{k_{nn}}{|\nu|-1}$): delivers information about one-hop neighborhood of a node [9]. The value of $\frac{k_{nn}}{|\nu|-1}$ delivers *joint degree distribution* statistics, and is calculated as average neighbor degree of the average k-degree nodes.

Average nodal degree ($\langle k \rangle$): is one of the coarse robustness measures [13]. Networks having high $\langle k \rangle$ values are considered more robust and "better-connected" on average.

Average shortest path length ($\langle l \rangle$): is the average of all of the shortest paths among all of the node-pairs of the network [12]. Small $\langle l \rangle$ values exhibit better robustness, because such networks are likely to lose fewer connections in response to different types of failures (random or targeted).

Average two-terminal reliability (A2TR): delivers the probability of the connectivity between a randomly chosen node pair [27]. In a fully connected network, A2TR value is one. Otherwise, A2TR is the sum of total number of node pairs in each connected cluster divided by all of the node pairs in the network.

Betweenness centrality ($\langle b \rangle$): measures the number of shortest paths among nodes that pass through a node or

link. Betweenness centrality is used to estimate the prestige of node/link [28].

Clustering coefficient ($\langle C \rangle$): is the percentage of 3-cycles among all of the connected node triplets within the network [13], [29]. If two neighbors of a node are connected, then a triangle (3-cycle) is formed by these three nodes.

Diameter (D): is the longest path among all of the shortest paths of the network. Generally, low D represents higher robustness.

Elasticity (E): relates to the total throughput in response to the node removal [8]. The fundamental idea is to successively remove a certain fixed number of nodes r (in the original definition, $r = 1\%$) and measure the consequent throughput degradation. The more pronounced and abrupt is the throughput drop experienced by a given topology, the lower is the robustness.

Heterogeneity ($\frac{\sqrt{\sigma_k^2}}{\langle k \rangle}$): is the standard deviation of the average node degree divided by the average node degree [10]. The lower heterogeneity value translates to higher network robustness.

Largest eigenvalue or spectral radius (λ_1): is the largest eigenvalue of the adjacency matrix of a network [13], [14]. Generally, the networks with the higher eigenvalues have small diameter and higher node distinct paths.

Node connectivity (κ): represents the smallest number of nodes whose removal results in a disconnected graph [9]. The node connectivity is the least number of node-disjoint paths between any two nodes within the network, which provides a rough indication of network robustness in response to any kind of failures or attacks (random or targeted). The same definition can be applied to link connectivity ρ when considering links instead of nodes.

Quantitative Robustness Metric (QNRM): analyzes how multiple failures affect the number of connections established in a network [7]. The QNRM delivers the number of the *blocked connections* (that cannot be established because of failure).

Qualitative Robustness Metric (QLRM): analyzes the variation in the quality of service of a network under various types of failures [7]. The QLRM measures the variation of the average shortest path length of the *established connections*.

R-value (R): computes the robustness of a topology under one or more topological features [30]. The obtained value is normalized to [0, 1], where $R = 0$ represents minimum robustness, and $R = 1$ reflects the perfect robustness.

Second smallest Laplacian eigenvalue or algebraic connectivity ($\mu_{|\nu|-1}$): depicts how difficult it is to break the network into islands or individual components [15]. The higher the value of $\mu_{|\nu|-1}$, the better the robustness.

Symmetry ratio ($\frac{\epsilon}{D+1}$): is the quotient between the distinct eigenvalues of the network adjacency matrix and the network diameter [11]. The networks with low symmetry ratio are considered more robust to random failures or targeted attacks.

Viral Conductance (VC): measures the network robustness in case of epidemic scenarios (propagation/spreading of failures) [31]. The VC is measured by considering the area under the curve that provides the fraction of infected nodes in steady-state for a range of epidemic intensities.

5 Simulation Scenarios and Methodologies

This section details the simulation scenarios and methodologies used in this work. To generalize the robustness analysis of the state-of-the-art DCNs, we performed extensive simulations considering four *node* failure scenarios to measure the various robustness metrics, namely:

1. random failures,
2. targeted failures,
3. network-only (failures introduced only in the network devices), and
4. real DCN failures (using real DCN failure data collected over a period of one year).

To do so, we consider six DCN networks, which are presented in Section 6. For the first three failure scenarios, we analyzed the robustness of each DCN by introducing the failures within a range from 0.1 to 10 percent of the network size. With the purpose of providing a detailed robustness evaluation, we analyzed the robustness metrics by introducing 0.1 to 2.5 percent of failures with an increment of 0.1, whereas from 3 to 10 percent the increment was equal to 1.

In the real DCN failures case, we used the observations reported in [33]. Gill et al. analyzed the network failure logs collected over a period of around one year from tens of data centers. The authors derived the failure probability for various network components by dividing the number of failures observed in a specific network device type, such as access layer or aggregate layer switches, with the total population of the devices in the given device type. We used the frequentist probability to derive the number of failures in three DCN architectures. We analyzed the various robustness metrics under real failure scenario by instigating the derived failures at each layer. As the number of network elements in the FatTree is much higher than the ThreeTier architecture, the number of failed nodes is around five times in the FatTree as compared to the ThreeTier architecture.

We introduced random failures in data center nodes (including the computational servers) within a range of 0.1 to 10 percent of the network size, as discussed in the various studies, such as [7], [34], [35]. The node failures are distributed among the nodes at each layer and *dcell* level within a range of 31-3,266 nodes. Besides instigating failures randomly in the whole network, we also considered the scenario of the network-only node failure, as discussed in [33]. Another significant scenario to measure the system robustness is by introducing the targeted attacks [35], [36], [37]. In the targeted failures case, we considered the *betweenness centrality* of the nodes to introduce the node failures.

6 Network Topologies

In this section, we present six representative topologies of the DCN architectures. Moreover, robustness is discussed according to the characteristics of each of the DCN architectures. The selected topologies represent connected and symmetric DCN networks.

DCN architectures follow a complex interconnection topology that entails a detailed understanding of the architecture to generate the DCN topology. Therefore,

TABLE 3
30K DCN Topology Features

Topology	$\|\nu\|$	$\|\varepsilon\|$	$\langle k\rangle$	$\langle l\rangle$	$\langle d\rangle$	r	$\frac{2\cdot\|\varepsilon\|}{(\|\nu\|\cdot(\|\nu\|-1))}$
DCell30K	32656	61230	3.7500	11.1521	23	-0.25	0.00011
FatTree30K	30528	82994	5.4336	5.6200	6	-0.20	0.00017
ThreeTier30K	30676	31632	2.0620	5.9000	6	-0.95	0.00006

generating the representative DCN synthetic topologies is a difficult task. There is presumably no publically available DCN topology generation tool. We developed a DCN topology generator for custom and flexible creation of various DCN topologies. Based on various input parameters, such as number of pods for the FatTree, dcell levels, and number of nodes in $dcell_0$ for the DCell, and number of nodes and switches in various layers in the ThreeTier DCN architecture, the DCN topology generator engenders the network topology in various popular graph formats. We generated two representative network topologies for each of the DCN architectures:

- three large networks (*DCell30K*, *FatTree30K*, and *ThreeTier30K*),
- three smaller networks (*DCell2K*, *FatTree2K*, and *ThreeTier2K*).

Increasing a single server in the DCell topology exponentially expands the network. A *3-level* DCell with two servers in $dcell_0$ constitute a network of 2,709 nodes. An increase in the number of servers to three in $dcell_0$ results in a network of 32,656 nodes. Therefore, the considered topologies are 2K and 30K networks.

Table 3 depicts some of the features of the three large networks. As observed, all of the topologies have more than 30,000 nodes. The *FatTree30K* has the largest number of edges among the considered set of large networks. The density $\frac{2\cdot|\varepsilon|}{(|\nu|\cdot(|\nu|-1))}$ of the *FatTree30K* is around three times higher than the *ThreeTier30K*. The higher number of edges and density exhibit better resilience to failures. The value of the average shortest path length $\langle l\rangle$ for the *FatTree30K* and *ThreeTier30K* is less than six, whereas the *DCell30K* has a higher path length (11). A higher $\langle l\rangle$ means that the communication between the end hosts in the *DCell30K* is more susceptible to be affected by a failure than in the *FatTree30K* or *ThreeTier30K*. This is due to the fact that such a communication is going to be routed (in average) through a longer path. The higher the number of links and nodes involved in a path, the higher is the probability to be affected by failures. Similarly, the *DCell30K* diameter D presents a value four times higher than the *FatTree30K* and *ThreeTier30K* (both have $D = 6$). However, the *DCell30K* possesses high-average nodal degree $\langle k\rangle$ that depicts strong resilience against failures. Moreover, all of the three networks exhibit dissassortativity and have negative value of the assortativity coefficient. It means that all of the three networks have an excess of links among nodes with dissimilar degrees.

Tables 4 and 5 present features of the *DCell2K*, *FatTree2K*, and *ThreeTier2K* topologies. Each topology is composed of around 2,500 to 2,700 nodes. As observed previously in the 30K networks, the FatTree DCN architecture has the largest number of edges. Regarding the spectral radius λ_1 and algebraic connectivity $\mu_{|\nu|-1}$, the *FatTree2K* proves to be the most robust network. The higher the value of λ_1 and $\mu_{|\nu|-1}$, the higher the robustness. Although the *ThreeTier2K* also indicates better robustness when considering λ_1, the *ThreeTier2K* possess the highest maximum nodal degree $k_{\max}$. High $k_{\max}$ is an indicator of vulnerability, depicting that removal of such a node could seriously damage the network. Moreover, the minimum values of the node and link connectivity for all of the networks are $\kappa = 1$ and $\rho = 1$, respectively. Such values of κ and ρ indicate that a single node or link failure may cause the network fragmentation. Because of having the lowest symmetry ratio value $\frac{\epsilon}{D+1}$, the *DCell2K* exhibits a higher robustness.

It can also be observed that the *FatTree2K* and *ThreeTier2K* have a low average shortest path length $\langle l\rangle$ than the *DCell2K*, and consequently can be considered more robust with respect to $\langle l\rangle$. The average node betweenness centrality $\langle b\rangle$ depicts that although the *DCell2K* has the highest value of $\langle b\rangle$, the *DCell2K* exhibits least standard deviation in the individual node's $\langle b\rangle$ value. Therefore, it can be inferred that all of the nodes of the *DCell2K* have nearly similar value of the betweenness centrality. Alternatively, the value of $\langle b\rangle$ for the *FatTree2K* and *ThreeTier2K* is lower than the *DCell2K*, but they have higher standard deviation, which means that the *FatTree2K* and *ThreeTier2K* networks have an excess of centrality measures for some nodes, indicating the vulnerability of

TABLE 4
2K DCN Topology Features

Topology	$\|\nu\|$	$\|\varepsilon\|$	$\langle k\rangle \pm$ StDev	λ_1	$k_{\max}$	$\mu_{\|\nu\|-1}$	$\frac{k_{nn}}{\|\nu\|-1} \pm$ StDev	κ	ρ	$\frac{\epsilon}{D+1}$
DCell2K	2709	4515	3.3333 $\pm$0.9428	3.56155	4	0.12439	0.00066 $\pm$0.00052	1	1	169.31250
FatTree2K	2500	6000	4.8000 $\pm$7.6000	17.4186	20	0.31528	0.00337 $\pm$0.00330	1	1	318.28571
ThreeTier2K	2562	2740	2.1389 $\pm$4.6416	10.25044	40	0.02308	0.00119 $\pm$0.00170	1	1	318.71429

TABLE 5
2K DCN Topology Features

Topology	$\langle l\rangle \pm$ StDev	$\langle b\rangle \pm$ StDev	$\langle C\rangle \pm$ StDev	r	D	$\frac{2\cdot\|\varepsilon\|}{(\|\nu\|\cdot(\|\nu\|-1))}$	$\frac{\sqrt{\sigma_k^2}}{\langle k\rangle} \pm$ StDev
DCell2K	8.51062 $\pm$1.93990	0.00277 $\pm$0.00133	0 $\pm$ 0	-0.2500	15	0.00123	0.28284
FatTree2K	5.21063 $\pm$1.12342	0.00169 $\pm$0.00337	0.8000 $\pm$0.4000	-0.2000	6	0.00192	1.58333
ThreeTier2K	5.72473 $\pm$0.70278	0.00185 $\pm$0.01482	0.9404 $\pm$0.2303	-0.8961	6	0.00083	2.17007

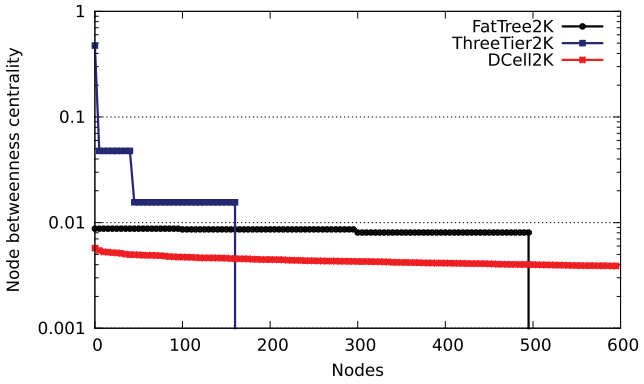


Fig. 1. Node betweenness centrality distribution in logarithmic scale of the 600 nodes with the highest value of *DCell2K*, *FatTree2K*, and *ThreeTier2K*.

networks under targeted failures. The node betweenness centrality distribution for the 600 highest values in the three networks is shown in Fig. 1. The *DCell2K* curve illustrates uniformly distributed values of $\langle b \rangle$ for all of the nodes.

The absence of 3-cycles in the clustering coefficient $\langle C \rangle$ measurements reveal that the *DCell2K* lacks two-hop paths to re-route the traffic in case of failure of one of its neighbors. On the contrary, the *FatTree2K* and *ThreeTier2K* exhibit better robustness by having high values of $\langle C \rangle$, which illustrate the existence of multiple alternative two-hop paths. Moreover, all of the three networks are dissassortative, $r < 0$. The density measurements show that the *FatTree2K* is the most dense and henceforth the most robust network. The low heterogeneity $\frac{\sqrt{\sigma_k^2}}{\langle k \rangle}$ value shows that the *DCell2K* can be considered as the most robust network when considering the heterogeneity.

The initial network analysis (for the whole network without failures) of the considered DCN topologies reveals that none of the three networks can be considered as the most robust architecture for all of the metrics. The robustness classification of the DCN networks for various metrics is reported in Table 6. The highest, average, and least values in the table depict the robustness level of the network. It can be observed that the FatTree architecture exhibits highest robustness for most of the metrics. Therefore, based on the initial network analysis without failures, it can be stated that the FatTree DCN exhibits better robustness than the DCell and ThreeTier architectures.

7 Results

This section presents a detailed analysis of the structural robustness of the DCN networks presented in Section 6. Initially, a comparison of the 30K with the 2K networks (7.1) is presented. Thereafter, the robustness analysis of the: 1) 30K networks (7.2) and 2) 2K networks (7.3) considering various failure scenarios is discussed. Although the study has been carried out within the range of 0.1 to 10 percent of the nodes affected by the failures, the results present a maximum of 6 percent of the affected nodes. This is due to the fact that the higher percentages in the targeted and network-only failures completely disconnect some of the networks. Therefore, the considered graph metrics do not deliver any useful information for higher failure percentages.

TABLE 6
Robustness Classification of the Three DCN Architectures

Metrics	FatTree	DCell	ThreeTier
$\lvert\varepsilon\rvert$	Highest	Average	Least
$\langle k \rangle$	Highest	Average	Least
$\langle l \rangle$	Highest	Least	Average
D	Highest	Least	Average
r	Highest	Average	Least
$\frac{2 \cdot \lvert\varepsilon\rvert}{(\lvert\nu\rvert \cdot (\lvert\nu\rvert - 1))}$	Highest	Average	Least
λ_1	Highest	Least	Average
$k_{\max}$	Average	Highest	Least
$\mu_{\lvert\nu\rvert-1}$	Highest	Average	Least
$\frac{\epsilon}{D+1}$	Average	Highest	Least
$\langle b \rangle$	Highest	Least	Average
$\langle C \rangle$	Average	Least	Highest
$\frac{\sqrt{\sigma_k^2}}{\langle k \rangle}$	Average	Highest	Least

Several graph metrics are computational intensive and require a large amount of CPU time. Therefore, the large (30K) networks are analyzed by their:

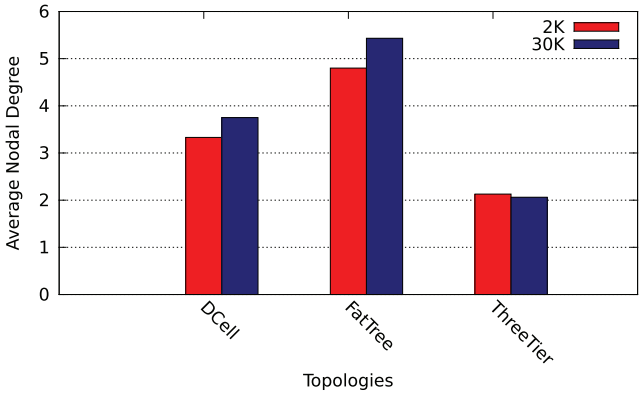
1. largest connected component,
2. average nodal degree,
3. node connectivity, and
4. number of clusters.

Whereas, the small networks are studied considering their: 1) algebraic connectivity, and 2) spectral radius or largest eigenvalue. It is noteworthy to consider that some of the metrics are applicable only to the largest connected component, as they require connected graph. Therefore, the result values of $\mu_{|\nu|-1}$, $\langle b \rangle$, and λ_1 are dependent on the largest connected component of the network.

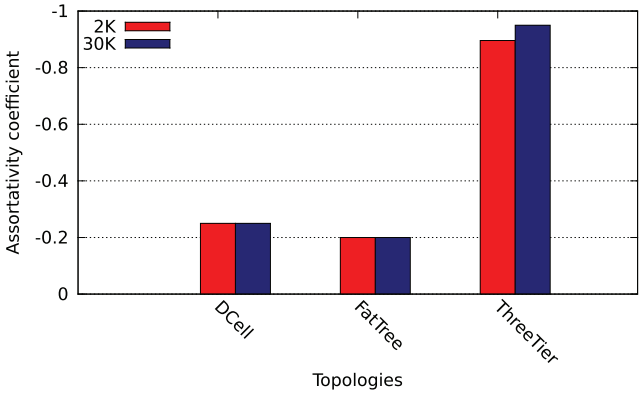
7.1 Network Size Comparison

The degree distribution of nodes in various DCNs exhibit homogeneous pattern, and the degree of each node is one among the few values in the degree set. For example, there are only two types of nodes (switches and servers) in the DCell. Therefore, the degree distribution follows two values: 1)having a similar value for all of the switches and 2) for all of the servers. Similarly, in case of the FatTree architecture, each node's degree is either one (for servers) or the k (for switches). In the ThreeTier architecture, the nodal degree falls within one of the four values, one each for the servers, access layer switches, aggregate layer switches, and core layer switches. The average nodal degree $\langle k \rangle$ of the DCNs does not strictly depend on the network size. A comparison of $\langle k \rangle$ for the 2K and 30K networks is illustrated in Fig. 2a. It can be observed that there is no significant difference in the values of $\langle k \rangle$ of the large and small DCN networks. Similarly, regarding the assortativity coefficient values for the large and small networks (see Fig. 2b), it can be observed that there is no remarkable difference between the assortativity coefficients of the large and small DCNs, and all of them remain dissassortative.

In essence, increasing the DCN size does not imply obtaining very different network topology characteristics. Therefore, we divide the robustness metrics analysis into

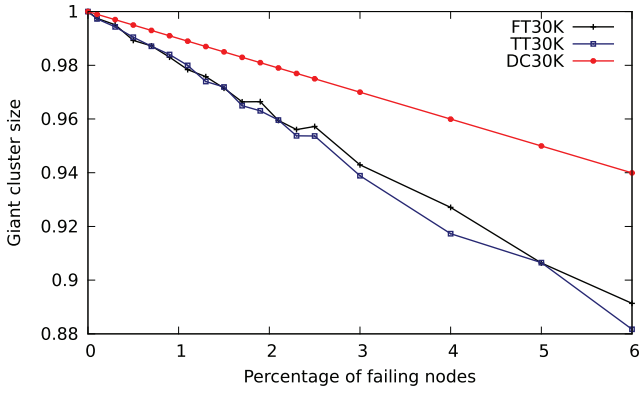


(a) Average nodal degree

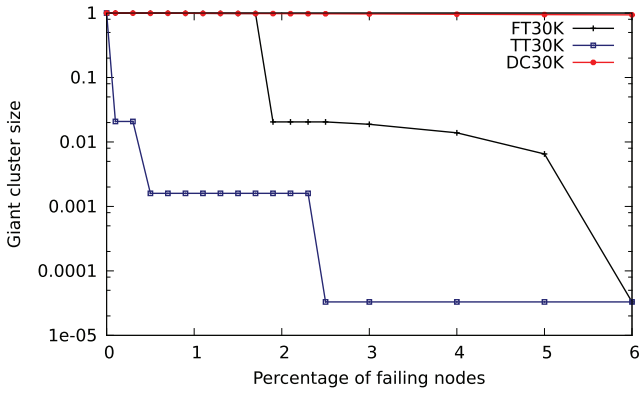


(b) Assortativity coefficient

Fig. 2. Average nodal degree and assortativity coefficient comparison of the 30K and 2K networks.



(a) Random failures



(b) Targeted failures

Fig. 3. Largest connected component size analysis under random and targeted failures of the 30K networks.

two parts: 1) the low CPU time consuming metrics are analyzed for the 30K network set and 2) the high CPU time consuming metrics are studied for the 2K network set.

7.2 30K Networks

Component structure of a network is one of the most important properties to be considered. Therefore, largest connected component is considered significant to measure the effectiveness of the network [38], [39], [40], [41], [42]. In case of random failures, all of the considered DCN architectures exhibit a robust behavior. As observed in Fig. 3a, the largest component size remains above 85 percent, even when 6 percent of the nodes fail randomly. However, in case of the targeted attack, the ThreeTier and FatTree topology behave contrary (see Fig. 3b). Removal of a very small fraction (<0.1 percent) of the nodes in the ThreeTier architecture results in segregation of the network. The network completely disconnects when around 2.5 percent of the nodes fail. This is due to the fact that the ThreeTier architecture have certain nodes (core and aggregate layer switches) with very high betweenness centrality values. Therefore, failure of such nodes segregate the network. However, the FatTree shows an altered behavior. Around 98 percent of the nodes reside in the connected component until a targeted failure of 1.8 percent of the nodes. An abrupt change is observed when the failure rate reaches 1.9 percent, resulting in the decline of largest component size from 98 to 2 percent, depicting a phase change. Therefore, the point at which 1.9 percent of the nodes fail can be considered as the critical point [43] for the FatTree architecture. Alternatively, the DCell confirms the resilience to the targeted attack. A smooth linear decline is observed in the largest component size decay. Around 94 percent of the network nodes reside in the largest component when 6 percent of the nodes fail in the DCell. The detailed values of the largest cluster size for the three networks can be found in Table 7.

The average nodal degree of the largest connected component are presented in Table 8 and Fig. 4 for the random, targeted, and network-only failures. It can be observed that in case of random failures, where the failure is introduced both in the network and computational nodes, the three DCNs behave similarly. However, when the failures are introduced in the network portion or in case of the targeted attack, the FatTree and ThreeTier exhibit a rapid decline in the nodal degree. The reason for such a rapid decay is the fact that the failure of a single access layer switch disconnects n nodes. Therefore, the average nodal degree decays rapidly. The failure analysis depicts that the *DCell30K* exhibits robustness in terms of the average nodal degree under all of the failures.

The node connectivity is an important measure to illustrate how many nodes need to fail to disconnect the network. The node connectivity can be measured by calculating the minimum node distinct paths between any two nodes within the network. As there is only a single

TABLE 7
Largest Connected Component Size of the 30K Networks

Percentage	Random			Targeted		
	FT30K	*TT30K*	*DC30K*	*FT30K*	*TT30K*	*DC30K*
0	1	1	1	1	1	1
0.1	0.997491	0.997268	0.998989	0.998985	0.020831	0.998928
0.5	0.989269	0.990452	0.994978	0.994988	0.001597	0.994488
0.9	0.983035	0.984007	0.990997	0.990992	0.001597	0.990323
1.1	0.978403	0.979955	0.988976	0.988994	0.001597	0.988180
1.5	0.971452	0.971903	0.984995	0.984997	0.001597	0.984015
1.9	0.966457	0.963040	0.980981	0.020440	0.001597	0.979912
2.3	0.956073	0.953742	0.976969	0.020440	0.001597	0.975778
3	0.942934	0.938871	0.969984	0.018802	0.000033	0.968030
4	0.927083	0.917294	0.959952	0.013889	0.000033	0.957343
5	0.906394	0.906516	0.949954	0.006519	0.000033	0.946564
6	0.891349	0.881696	0.939928	0.000033	0.000033	0.935755

TABLE 8
Average Nodal Degree ($\langle k \rangle$) of the 30K Networks

Percentage	Random			Targeted			Network-only failure		
	FT30K	*TT30K*	*DC30K*	*FT30K*	*TT30K*	*DC30K*	*FT30K*	*TT30K*	*DC30K*
0	5.433600	2.062000	3.750000	5.433600	2.06200000	3.750000	5.433600	2.06200000	3.750000
0.1	5.425747	2.058874	3.746289	5.341902	2.01481500	3.746437	5.342263	1.97187800	3.746289
0.5	5.401567	2.052683	3.731288	4.977778	1.79280500	3.733182	4.988352	1.61214200	3.731288
0.9	5.380802	2.046817	3.716587	4.610716	1.41162500	3.719697	4.644399	1.25845600	3.716587
1.3	5.359039	2.037837	3.701269	4.240682	1.03048500	3.706831	4.309270	0.91184070	3.701269
1.7	5.339445	2.026962	3.686467	3.867640	0.64309880	3.693558	3.988530	0.56251240	3.686467
2.1	5.317620	2.020499	3.671067	3.488456	0.25253900	3.681867	3.670923	0.21987950	3.671067
2.5	5.314319	2.016176	3.656359	3.109260	0	3.670571	3.372510	0	3.656359
3	5.270296	1.999509	3.637258	2.632446	0	3.656640	3.014096	0	3.637258
4	5.225032	1.972378	3.599930	1.657545	0	3.627364	2.333420	0	3.599930

edge that connects the node to the access switch, the node disjoint paths for the ThreeTier and FatTree DCNs are always one. However, from every access layer switch to every other switch, there are always, $k/2$ node distinct paths in the FatTree. Similarly, in the ThreeTier network, the maximum node distinct paths between any two access layer and the aggregate layer devices are equal to q and r, respectively. Because there is only a single edge between the network switch and servers within $dcell_0$, the minimum node disjoint paths of the network is one. However, the network switch only performs the packet forwarding within $dcell_0$ and the actual communication always occurs among the servers in the DCell architecture. The node distinct paths between any two servers in the DCell are equal to $l+1$, where l is the DCell level.

The number of the segregated clusters are depicted in Fig. 5 and the detailed values are presented in Table 9. It can be observed that the *FatTree30K* and *ThreeTier30K* behave similarly in random failures. The failure of a single access layer switch disconnects n servers, resulting in n segregated clusters. Therefore, the *FatTree30K* and *ThreeTier30K* networks disconnect into more than 45 clusters when 0.1 percent of the nodes fail. However, for the network-only failure case, the *ThreeTier30K* disconnects into more clusters than the *FatTree30K*. In the targeted failures case, the robustness difference is even more, where

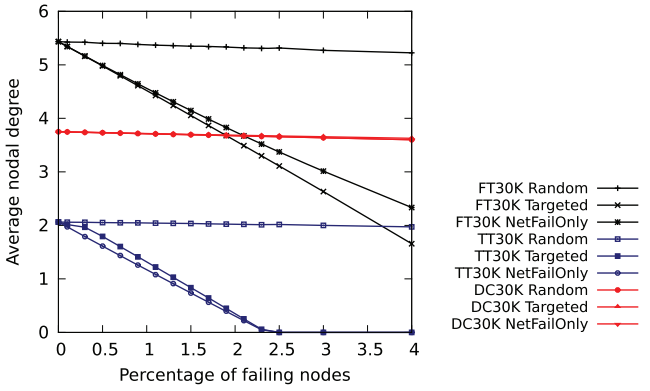


Fig. 4. Average nodal degree analysis of for the 30K networks.

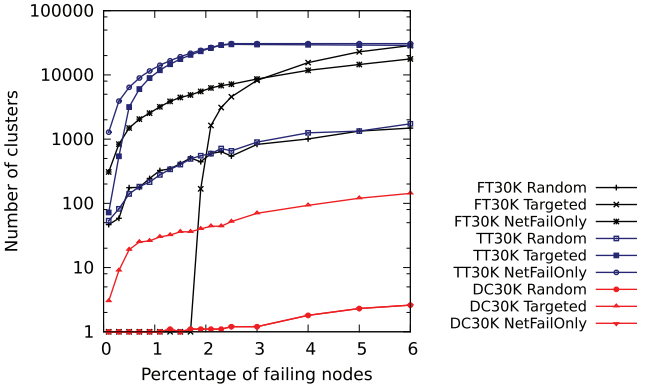


Fig. 5. Number of clusters analysis for the 30K networks.

TABLE 9
Number of Clusters of the 30K Networks

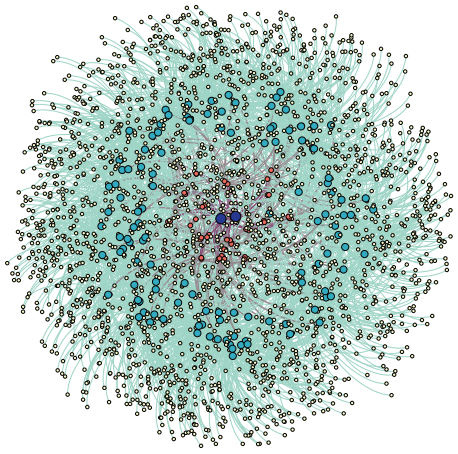
Percentage	Random			Targeted			Network-only failure		
	FT30K	*TT30K*	*DC30K*	*FT30K*	*TT30K*	*DC30K*	*FT30K*	*TT30K*	*DC30K*
0	1	1	1	1	1	1	1	1	1
0.1	46.6	53.8	1	1	72	3	308.2	1282.6	1
0.5	175.6	139.9	1	1	3162	19	1491.4	6405.7	1
0.9	243.9	214.6	1	1	8943	26	2540.2	11510.6	1
1.3	344	338.7	1.1	1	14677	32	3874.6	16580.6	1.1
1.7	507.9	492.1	1.1	1	20458	36	4865.8	21692.7	1.1
2.1	596.3	595	1.1	1632	26239	44	6269.8	26730.3	1.1
2.5	543.6	655.7	1.2	4560	29909	52	7165	30676	1.2
3	827.1	898.7	1.2	8208	29755	70	8626.6	30676	1.2
4	1005	1255.2	1.8	15552	29448	93	11792.2	30676	1.8
5	1331.6	1334.7	2.3	22872	29142	120	14504.2	30676	2.3
6	1485.9	1729.6	2.6	28696	28835	143	17710.6	30676	2.6

the *FatTree30K* remains connected until 1.9 percent of the nodes fail. This is due to the fact that the FatTree network has a considerable portion of the nodes ($(\frac{k}{2})^2$ core switches) with similar high betweenness centrality values (see the betweenness centrality distribution in Section 6). Therefore, the topology remains fully connected until $((\frac{k}{2})^2 - 1)$ nodes fail in the *FatTree30K*. It is noteworthy to mention that because of nearly similar betweenness centrality distribution among the nodes, the *DCell30K* outperforms the other two 30K networks. The DCell portrays high robustness in random or network-only failures, and remains connected until 4 percent of the nodes are affected. However, the network disconnects with only 0.1 percent of the nodes failure in case of the targeted attack. Nonetheless, the number of segregated cluster remains much less than the counterparts. In any of the failure cases, the *DCell30K* can be considered as the most robust network in terms of number of isolated clusters among all of the DCN architectures.

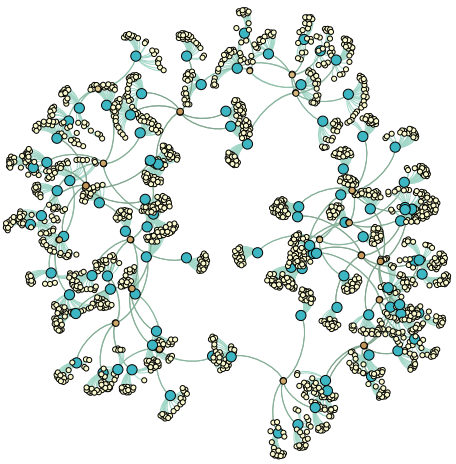
7.3 2K Networks

The robustness evaluation of the large (30k) networks is infeasible when considering the computational intensive metrics, such as $\mu_{|\nu|-1}$, $\langle b \rangle$, and λ_1. Therefore, we evaluate small (2k) networks for the computational intensive metrics. One of the most significant considerations in the evaluation of the 2k networks for the computational intensive metrics is that such metrics only consider the largest connected component of the network. In case of the targeted and network-only failures, the size of the largest connected component is typically very small, and it constitute very little portion of the network. Therefore, the resulting values are abrupt and unrealistic, and are unable to depict the factual robustness of the network. Fig. 6 illustrates the process of the propagation of targeted failures within a *ThreeTier2k* network. (Fig. 6a) depicts the initial network, and (Fig. 6b) shows the disconnected network with 1 percent targeted failures. The nodes at different layers of the network are shown in different colors and sizes.

The algebraic connectivity $\mu_{|\nu|-1}$ is an important measure to evaluate that how difficult it is to break the network into islands or individual components. The algebraic connectivity for the 2K networks is presented in Fig. 7 and the details can be observed in Table 10. It is noteworthy to consider that although the *DCell2K* does not possess the highest value of $\mu_{|\nu|-1}$, it exhibits a smooth and slow decline in the value of $\mu_{|\nu|-1}$ in random and network-only failures. However, in case of the targeted attack, the value of $\mu_{|\nu|-1}$ for the *DCell2K* drops significantly when 3 percent of the nodes fail. Such an abrupt decrease portrays that the *DCell2K* is vulnerable to the targeted failures when the percentage of node failure is increased. For the *FatTree2K* network, it can be observed that despite



(a) ThreeTier DCN without failures



(b) ThreeYier DCN after 1% targeted failures

Fig. 6. ThreeTier DCN before and after targeted failure.

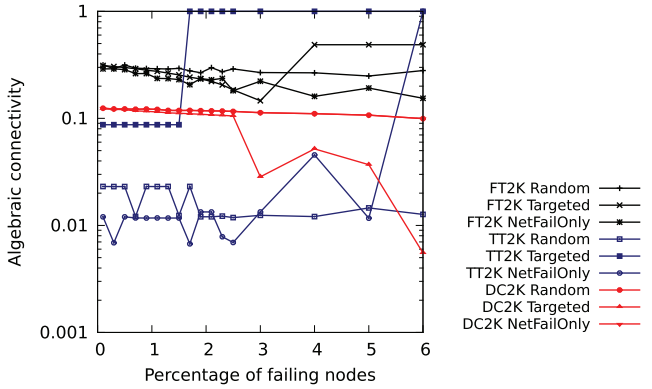


Fig. 7. Algebraic connectivity analysis of the 2K networks.

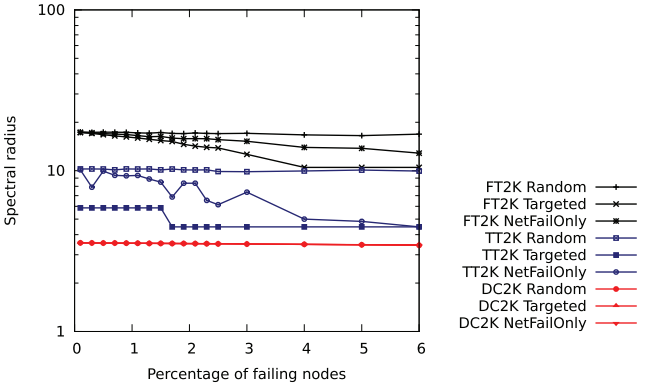


Fig. 8. Spectral radius analysis of the 2K networks.

showing a clearly descending curve under random and network-only failures, the value of $\mu_{|\nu|-1}$ increases in case of the targeted failure when more than 3 percent of the nodes fail. The *ThreeTier2k* also depicts a similar behavior as the *FatTree2k* network when percentage of the node failure increases in the targeted and network-only failures. However, such abrupt increase in the values of $\mu_{|\nu|-1}$ is due to the fact that $\mu_{|\nu|-1}$ is analyzed for a very small sized largest connected component. Therefore, the value of $\mu_{|\nu|-1}$ increases. It is noteworthy to consider that an increase in the value of $\mu_{|\nu|-1}$ for high percentages of node failures (4.5 or 6 percent) does not mean that the network is more robust. Because the value of $\mu_{|\nu|-1}$ is calculated only for the largest connected component, it cannot be inferred that the networks become more robust after failures.

The spectral radius or largest eigenvalue λ_1 analysis is presented in Fig. 8 and the detailed values are provided in Table 11. As observed in Fig. 8, the *DCell2K* has a smaller value of λ_1, but the value of λ_1 decrease slightly for all of the considered percentages and types of failures. The *FatTree2K* also exhibit slight decrease in the value of λ_1 in random failures case. However, the value of λ_1 decreases

TABLE 10
Algebraic Connectivity ($\mu_{|\nu|-1}$) of the 2K Networks

Percentage	Random			Targeted			Network-only failure		
	DC2K	*FT2K*	*TT2K*	*DC2K*	*FT2K*	*TT2K*	*DC2K*	*FT2K*	*TT2K*
0	0.12439	0.31528	0.02308	0.12439	0.31528	0.02308	0.12439	0.31528	0.02308
0.1	0.12431	0.31525	0.02308	0.12296	0.31164	0.08728	0.12431	0.28999	0.01202
0.5	0.12254	0.31528	0.02307	0.12056	0.29844	0.08729	0.12254	0.28539	0.01202
0.9	0.12229	0.29210	0.02308	0.11641	0.28220	0.08728	0.12229	0.26232	0.01172
1.3	0.11877	0.28940	0.02309	0.11267	0.26464	0.08729	0.11877	0.23556	0.01172
1.7	0.11881	0.27810	0.02308	0.11041	0.24262	1.00000	0.11881	0.20661	0.00672
2.1	0.11745	0.29799	0.01205	0.10776	0.22144	1.00000	0.11745	0.22910	0.01337
2.5	0.11621	0.28985	0.01182	0.10536	0.18455	1.00000	0.11621	0.18142	0.00693
3	0.11305	0.26777	0.01242	0.02845	0.14618	1.00000	0.11305	0.22228	0.01337
4	0.11070	0.26647	0.01208	0.05186	0.48750	1.00000	0.11070	0.16066	0.04555
5	0.10723	0.24866	0.01457	0.03693	0.48751	1.00000	0.10723	0.19194	0.01167
6	0.09943	0.27956	0.01267	0.00552	0.48750	1.00000	0.09943	0.15423	1.00000

FatTree2K is denoted as FT2K, ThreeTier2K as TT2K, and DCell2K as DC2K. Note that $\mu_{|\nu|-1}$ is computed for the largest connected component.

TABLE 11
Spectral radius (λ_1) of the 2K networks

Percentage	Random			Targeted			Network-only failure		
	DC2K	*FT2K*	*TT2K*	*DC2K*	*FT2K*	*TT2K*	*DC2K*	*FT2K*	*TT2K*
0	3.56155	17.4186	10.25044	3.56155	17.4186	10.25044	3.56155	17.4186	10.25044
0.1	3.55928	17.41844	10.25023	3.55873	17.26900	5.87059	3.55928	17.29531	10.11993
0.5	3.54998	17.37591	10.23774	3.55031	16.75502	5.87059	3.54998	17.01578	9.94468
0.9	3.54272	17.31246	10.24875	3.54182	16.23312	5.87059	3.54272	16.79820	9.29244
1.3	3.53132	17.12761	10.24804	3.53420	15.69121	5.87059	3.53132	16.26041	8.87328
1.7	3.52701	17.05352	10.24152	3.52819	15.19576	4.47214	3.52701	15.95128	6.86486
2.1	3.51929	17.17246	10.10359	3.52321	14.21964	4.47214	3.51929	15.82020	8.35723
2.5	3.50736	16.99966	9.87806	3.51720	13.85750	4.47214	3.50736	15.61130	6.15025
3	3.49963	17.08234	9.86205	3.50330	12.63656	4.47214	3.49963	15.24117	7.35323
4	3.48352	16.69618	9.96564	3.48346	10.48809	4.47214	3.48352	13.95714	5.00000
5	3.46261	16.53128	10.08541	3.46717	10.48809	4.47214	3.46261	13.77363	4.83971
6	3.43394	16.87975	9.94185	3.46481	10.48809	4.47214	3.43394	12.86074	4.47214

Note that λ_1 is computed for the largest connected component.

TABLE 12
DCN Features in Case of Real Failure

Feature	*FatTree30K*	*ThreeTier30K*	*DCell30K*	*FatTree2K*	*ThreeTier2K*	*DCell2K*
$max(\nu)$	0.91	0.94	0.99	0.98	0.95	0.9
$\langle k \rangle$	4.7453	1.9523	3.7155	4.4825 $\pm$ 7.07	2.0921 $\pm$ 4.51	3.2899 $\pm$ 0.95
$\langle l \rangle$	5.6417	5.901332	11.21	5.2863 $\pm$ 1.09	5.7204 $\pm$ 0.71	8.5909 $\pm$ 1.95
D	7	6	23	7	6	15
r	-0.288845	-0.974917	-0.215366	-0.21106	-0.95623	-0.23834
$\mu_{\lvert\nu\rvert-1}$	-	-	-	0.24157	0.00672	0.12017
$\langle b \rangle$	-	-	-	0.00181 $\pm$ 0.00372	0.00205 $\pm$ 0.02116	0.00284 $\pm$ 0.0014
λ_1	-	-	-	16.34671	7.68104	3.53358

almost linearly under network-only and targeted failures in the *FatTree2K*. The *ThreeTier2K* is significantly affected by the targeted failure, and the value of λ_1 divides almost to half with only 0.1 percent of the nodes failure.

The robustness analysis of the DCN architectures considering various failure types and percentages reveals the vulnerability of the ThreeTier and FatTree DCN architectures to the targeted and network-only failures. However, the DCell architecture exhibits graceful and little variation of the metric values in response to all of the failure types and percentages. Therefore, it can be inferred from the failure analysis that the DCell exhibits better robustness than the ThreeTier and FatTree architectures. Moreover, the results drawn from the initial robustness analysis of the DCN networks without failure (see Table 6) proves invalid. In contrary to the values reported in Table 6, the failure analysis reveals that the DCell architecture exhibits better robustness. Therefore, it is evident that the classical robustness metrics are inadequate to evaluate the DCN robustness.

7.4 Real Failures in DCNs

This section presents the robustness measurements obtained from the largest connected components of the six networks (three 30K and three 2K), when the real failures within the DCNs are produced. As defined in Section 5, a specific number of nodes from each layer of network topology (based on the failure logs of various data centers) have been selected to fail, and the graph metrics have been computed for the resulting largest connected component.

The Table 12 presents the results of the real failures. All of the networks possess more than 90 percent of the nodes in the largest connected component in response to the real failures, as indicated by the value of $max(\nu)$. The results obtained from the real failures illustrate that the average nodal degree decreases slightly in all of the networks. The average shortest path length and diameter exhibit minor increase in all of the three networks. The assortativity coefficient value also depicts minor change for all of the DCN architecture. The value of $\langle b \rangle$ increases for all of the networks, indicating the increased vulnerability of the networks. The value of $\mu_{|\nu|-1}$ exhibits comparatively higher decrease for the *FatTree2K* and *ThreeTier2K* than the *DCell2K* network. Similarly, the value of λ_1 also decreases significantly for the *FatTree2K* and *ThreeTier2K* as compared to the *DCell2K*. Therefore, the *DCell2K* network can be considered more robust network in case of the real failures while considering $\mu_{|\nu|-1}$ and λ_1.

All of the considered networks exhibit robust behavior in response to the real failures. However, the DCell architecture depicts graceful and minor variations in all of the observed metrics as compared to the ThreeTier and FatTree architectures. Therefore, the DCell DCN can be considered as the most robust architecture in case of the real failures.

7.5 Deterioration of DCNs

It has been observed that depending on the: 1) DCN architecture, 2) type of failure (whether it is random, targeted, network-only, or real), and (3) specific percentage of the nodes failed, the level of robustness according to a specific graph metric, computed from the largest connected component might be different. Moreover, the results for the various metrics exhibit strong dependence on the largest connected component, as observed in Section 7. Furthermore, the failure analysis depicts that the initial metric measurements are unable to quantify the DCN robustness appropriately (see Table 6 and Section 7.3). Therefore, we propose deterioration metric, a procedure for the quantification of the DCN robustness based on the percentage change in various graph metrics.

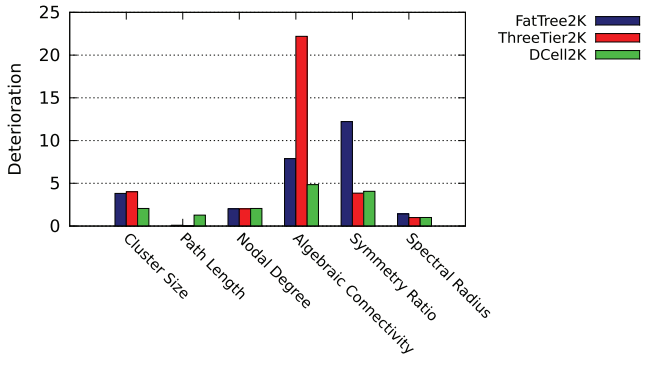
Deterioration σ_M, for any metric M can be calculated as the difference between the metric value for the whole network M_0, and the average of the metric values at various failure percentages M_i, divided by M_0.

$$\sigma_M = \left| \frac{1}{M_0} \left(\frac{\sum_{i=1}^{n} M_i}{n} - M_0 \right) \right|, \tag{27}$$

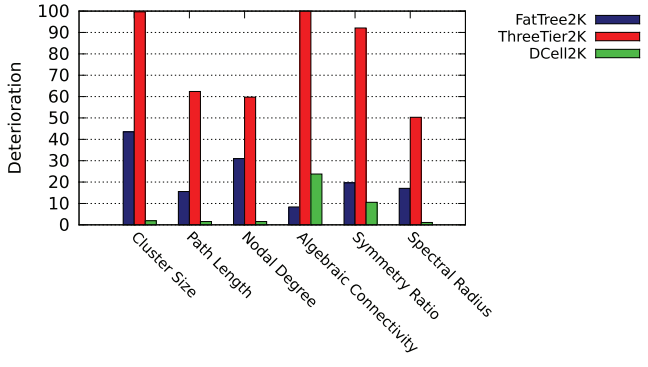
where M_i is measurement of the metric M at i percent of the nodes failure, and M_0 is the metric value for the whole of the network (without failure). To demonstrate that our proposed metric is able to quantify network robustness, we compute σ_M for:

1. six graph metrics namely:
   a. cluster size,
   b. average shortest-path length,
   c. nodal degree,
   d. algebraic connectivity,
   e. symmetry ratio, and
   f. spectral radius,
2. for the random, targeted, and real failures, taking into the account 1 to 6 percent of the nodes failure.

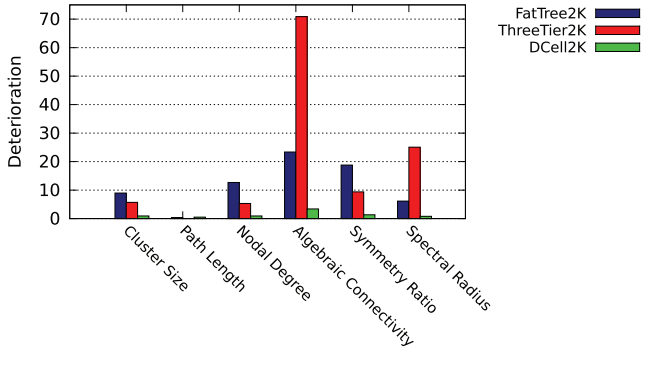
The results for the random, targeted, and real failures are presented in Figs. 9a, 9b, and 9c, respectively. The results depict that for almost all of the failure types, the σ_M for the



(a) Random failure



(b) Targeted failure



(c) Real failure

Fig. 9. Deterioration of the 2K networks in case of random, targeted, and real failures.

DCell is much less as compared to the ThreeTier and FatTree architectures. The ThreeTier DCN exhibits the highest deterioration in random and network-only failures. However, for the real failures, the FatTree DCN exhibits more deterioration than the ThreeTier network. It is noteworthy to consider that for the real failures, the number of the failed nodes for the FatTree is around five times higher than the ThreeTier architecture (see Section 5). Our proposed σ_M evaluates the network robustness, and also allows to compare the results among various DCN architectures. The lower the value of deterioration, the higher is the robustness of the network.

It can be observed that the robustness of the considered networks evaluated by the deterioration metric complies with the robustness of the networks observed in Sections 7.2 and 7.3. Therefore, it can be stated that the deterioration metric can be employed to evaluate the robustness of the networks where the classical robustness metrics are inapplicable, such as the DCNs.

8 Conclusions and Future Work

In this paper, we studied the structural robustness of the state-of-the-art data center network (DCN) architectures. Our results revealed that the DCell architecture degrades gracefully under all of the failure types as compared to the FatTree and ThreeTier architecture. Because of the connectivity pattern, layered architecture, and heterogeneous nature of the network, the results demonstrated that the classical robustness metrics are insufficient to quantify the DCN robustness appropriately. Henceforth, signifying and igniting the need for new robustness metrics for the DCN robustness quantification. We proposed deterioration metric to quantify the DCN robustness. The deterioration metric evaluates the network robustness based on the percentage change in the graph structure. The results of the deterioration metric illustrated that the DCell is the most robust architecture among all of the considered DCNs.

The DCN robustness analysis revealed the inadequacy of the classical robustness measures for the DCN architectures. Based on the hierarchical and pattern-based connectivity of the DCN architectures, new robustness metrics are necessary. Moreover, network traffic and throughput analysis in various DCN component failure scenarios also need to be performed. High level of network robustness leads to higher cost. New cost-effective robust DCNs are required to deliver the required level of robustness at minimum cost. Moreover, the performance aspects of the network, such as bisection bandwidth and bottleneck degree need to be considered for DCN performance quantification.

Acknowledgments

This work was partly supported by the projects TEC 2012-32336 and SGR-1202, and by the AGAUR FI-DGR 2012 Grant.

References

[1] K. Bilal, S.U.R. Malik, O. Khalid, A. Hameed, E. Alvarez, V. Wijaysekara, R. Irfan, S. Shrestha, D. Dwivedy, M. Ali, U.S. Khan, A. Abbas, N. Jalil, and S.U. Khan, "A Taxonomy and Survey on Green Data Center Networks," *Future Generation Computer Systems*, 2013.

[2] P. Mel and T. Grance, "The Nist Definition of Cloud Computing," http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf, 2011.

[3] S. Ali, A.A. Maciejewski, H.J. Siegel, and J. Kim, "Definition of a Robustness Metric for Resource Allocation," *Proc. Int'l Parallel and Distributed Processing Symp.*, p. 10, 2003.

[4] K. Bilal, S.U. Khan, L. Zhang, H. Li, K. Hayat, S.A. Madani, N. Min-Allah, L. Wang, D. Chen, M. Iqbal, C. Xu, and A.Y. Zomaya, "Quantitative Comparisons of the State-of-the-Art Data Center Architectures," *Concurrency and Computation: Practice and Experience*, vol. 25, pp. 1771-1783, 2012.

[5] A. Greenberg, J. Hamilton, D. Maltz, and P. Patel, "The Cost of a Cloud: Research Problems in Data Center Networks," *ACM SIGCOMM Computer Comm. Rev.*, vol. 39, no. 1, pp. 68-79, 2009.

[6] *ITProPortal*, http://www.itproportal.com/2012/07/12/o2-outage-latest-string-major-it- infrastructure-failures/, 2012.

[7] M. Manzano, E. Calle, and D. Harle, "Quantitative and Qualitative Network Robustness Analysis under Different Multiple Failure Scenarios," *Proc. Third Int'l Workshop Reliable Networks Design and Modeling (RNDM '11),* pp. 1-7, 2011.
[8] A. Sydney, C. Scoglio, P. Schumm, and R.E. Kooij, "Elasticity: Topological Characterization of Robustness in Complex Networks," *Proc. Third Int'l Conf. Bio-Inspired Models of Network,* pp. 19:1-19:8, 2008.
[9] A.H. Dekker and B.D. Colbert, "Network Robustness and Graph Topology," *Proc. 27th Australasian Conf. Computer Science,* pp. 359-368, 2004.
[10] J. Dong and S. Horvath, "Understanding Network Concepts in Modules," *BMC Systems Biology,* vol. 1, no. 1, pp. 1-24, 2007.
[11] A.H. Dekker and B.D. Colbert, "The Symmetry Ratio of a Network," *Proc. Australasian Symp. Theory of Computing,* pp. 13-20, 2005.
[12] C. Shannon and D. Moore, "The Spread of the Witty Worm," *IEEE Security and Privacy,* vol. 2, no. 4, pp. 46-50, July 2004.
[13] P. Mahadevan, D. Krioukov, M. Fomenkov, X. Dimitropoulos, K.C. Claffy, and A. Vahdat, "The Internet AS-Level Topology: Three Data Sources and One Definitive Metric," *SIGCOMM Computer Comm. Rev.,* vol. 36, pp. 17-26, Jan. 2006.
[14] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, "Epidemic Thresholds in Real Networks," *ACM Trans. Information and System Security,* vol. 10, no. 4, pp. 1-26, 2008.
[15] A. Jamakovic and S. Uhlig, "Influence of the Network Structure on Robustness," *Proc. 15th IEEE Int'l Conf. Networks (ICON '07),* pp. 278-283, 2007.
[16] *Cisco Data Center Infrastructure 2.5 Design Guide,* Cisco, 2010.
[17] M. Al-Fares, A. Loukissas, and A. Vahdat, "A Scalable, Commodity Data Center Network Architecture," *ACM SIGCOMM Computer Comm. Rev.,* vol. 38, no. 4, pp. 63-74, 2008.
[18] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu, "DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers," *SIGCOMM Computer Comm. Rev.,* vol. 38, no. 4, pp. 75-86, Aug. 2008.
[19] L. Gyarmati and T.A. Trinh, "Scafida: A Scale-Free Network Inspired Data Center Architecture," *ACM SIGCOMM Computer Comm. Rev.,* vol. 40, no. 5, pp. 4-12, 2010.
[20] L. Gyarmati, A. Gulyás, B. Sonkoly, T.A. Trinh, and G. Biczók, "Free-Scaling Your Data Center," *Computer Networks,* vol. 57, pp. 1758-1773, 2013.
[21] J. Kim, W.J. Dally, and D. Abts, "Flattened Butterfly: A Cost-Efficient Topology for High-Radix Networks," *ACM SIGARCH Computer Architecture News,* vol. 35, no. 2, pp. 126-137, 2007.
[22] D. Li, C. Guo, H. Wu, K. Tan, Y. Zhang, and S. Lu, "FiConn: Using Backup Port for Server Interconnection in Data Centers," *Proc. IEEE INFOCOM,* pp. 2276-2285, 2009.
[23] K. Bilal, S.U. Khan, J. Kolodziej, L. Zhang, K. Hayat, S. Madani, N. Min-Allah, L. Wang, and D. Chen, "A Comparative Study of Data Center Network Architectures," *Proc. 26th European Conf. Modeling and Simulation,* pp. 526-532, May 2012.
[24] M. Kurant and P. Thiran, "Layered Complex Networks," *Physical Rev. Letters,* vol. 96, p. 138701, 2006.
[25] X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov, "Clustering with Multi-Layer Graphs: A Spectral Perspective," *CoRR,* vol. abs/1106.2233, 2011.
[26] C. Clos, "A Study of Non-Blocking Switching Networks," *Bell System Technical J.,* vol. 32, no. 2, pp. 406-424, 1953.
[27] S. Neumayer and E. Modiano, "Network Reliability with Geographically Correlated Failures," *Proc. INFOCOM,* pp. 1658-1666, 2010.
[28] L.C. Freeman, "A Set of Measures of Centrality Based Upon Betweenness," *Sociometry,* vol. 40, no. 1, pp. 35-41, 1977.
[29] B. Bollobás, *Random Graphs.* vol. 73, Cambridge Univ. Press, 2001.
[30] P.V. Mieghem, C. Doerr, H. Wang, J.M. Hernandez, D. Hutchison, M. Karaliopoulos, and R.E. Kooij, "A Framework for Computing Topological Network Robustness," 2010.
[31] M. Youssef, R. Kooij, and C. Scoglio, "Viral Conductance: Quantifying the Robustness of Networks with Respect to Spread of Epidemics," *J. Computer Science,* vol. 2, no. 3, pp. 286-298, 2011.
[32] E. Weisstein, http://mathworld.wolfram.com/GraphDiameter.html, 2013.
[33] P. Gill, N. Jain, and N. Nagappan, "Understanding Network Failures in Data Centers: Measurement Analysis, and Implications," *Proc. ACM SIGCOMM,* 2011.
[34] R. Albert, H. Jeong, and A. Barabasi, "Error and Attack Tolerance of Complex Networks," *Letters to Nature,* vol. 406, pp. 378-382, 2000.
[35] J. Guillaume, M. Latapy, and C. Magnien, "Comparison of Failures and Attacks on Random and Scale-free Networks," *Proc. Eight Int'l Conf. Principles of Distributed Systems,* 2005.
[36] P. Holme, B. Kim, C. Yoon, and S. Han, "Attack Vulnerability of Complex Networks," *Physical Rev. E,* vol. 65, no. 5, p. 056109, 2002.
[37] M. Manzano, V. Torres-Padrosa, and E. Calle, "Vulnerability of Core Networks under Different Epidemic Attacks," *Proc. Fourth Int'l Workshop Reliable Networks Design and Modeling,* 2012.
[38] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, "Graph Structure in the Web," *Computer Networks,* vol. 33, nos. 1-6, pp. 309-320, 2000.
[39] D. Callaway, M. Newman, S. Strogatz, and D. Watts, "Network Robustness and Fragility: Percolation on Random Graphs," *Physical Rev. Letters,* vol. 85, no. 25, 2000.
[40] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, "Resilience of the Internet to Random Breakdowns," *Physical Rev. Letters,* vol. 85, no. 21, p. 3, 2000.
[41] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, "Giant Strongly Connected Component of Directed Networks," *Physical Rev. E,* vol. 64, no. 2 Pt 2, p. 4, 2001.
[42] M. Newman, S. Strogatz, and D. Watts, "Random Graphs with Arbitrary Degree Distributions and their Applications," *Physical Rev. E,* vol. 64, no. 2, p. 19, 2000.
[43] B. Luque and R. Solé, "Phase Transitions in Random Networks: Simple Analytic Determination of Critical Points," *Physical Rev. E,* vol. 55, no. 1, pp. 257-260, 1997.

Kashif Bilal is working toward the PhD degree in electrical and computer engineering at the North Dakota State University. His research interests include data center networks, distributed computing, and energy efficiency. He is a student member of the IEEE.

Marc Manzano is currently working toward the PhD degree at the University of Girona. His main research interests include the modeling and analysis of complex networks, the robustness of such networks under multiple failure scenarios.

Samee U. Khan is an assistant professor at the North Dakota State University. His research interest include encompasses topics, such as sustainable computing, social networking, and reliability. He has published more than 200 papers. He is a senior member of IEEE, and a fellow of IET and BCS.

Eusebi Calle is an associate professor at the University of Girona (UdG). He is a member of the research and teaching staff of the Broadband Communications and Distributed System Group at the UdG, where he develops his research in GMPLS fault management, routing, and network science.

Keqin Li is an SUNY distinguished professor. His research interests include mainly in the areas of design and analysis of algorithms, parallel and distributed computing, and computer networking. He has published more than 225 journal articles, book chapters, and research papers in refereed international conference proceedings. He is a senior memeber of the IEEE.

Albert Y. Zomaya is currently the chair professor of high performance computing and networking in the School of Information Technologies, The University of Sydney. He is the author/coauthor of seven books, more than 400 papers, and the editor of nine books and 11 conference proceedings. He is a fellow of IEEE, IET, and AAAS.