On the Connectivity of Data Center Networks

Marc Manzano, Kashif Bilal, Eusebi Calle, and Samee U. Khan, *Senior Member, IEEE*

***Abstract*—Data Center Networks (DCNs) constitute the communication backbone for the cloud computing paradigm. Recently, network connectivity analysis in terms of reliability has received attention from the network research community. The traditional network features are useful; however, they are insufficient to determine how *well-connected* or *well-designed* a DCN is against the node or link removals. In this letter, we present a connectivity analysis of three well-known DCN architectures, namely: (a) *ThreeTier*, (b) *FatTree*, and (c) *DCell*. Our analysis reveals that the classic connectivity measures are inadequate for evaluating DCN connectivity. Therefore, we propose μ-A2TR, a novel metric to characterize network connectivity in the case of node or link failures. Experimental results reveal that the DCNs exhibit a moderate level of connectivity in the case of random node removals. However, connectivity decays abruptly when considering the targeted nodes removal. Moreover, the connectivity analysis depicts significant differences among the considered DCNs.**

***Index Terms*—Cloud computing; data center networks; connectivity analysis; distributed systems.**

I. Introduction

CLOUD computing is an emerging paradigm that in the forthcoming years is expected to play a pivotal role in the Information and Communication Technology (ICT) sector. Data centers are the foundations of the cloud computing paradigm and are crucial for its operational and economic success. Data centers are composed of tens of thousands of hosts that are organized in clusters. Services are sourced from multiple clusters within the data center, and each cluster may host multiple services to increase system utilization. Most of the network communication, such as indexing, search, or other Map-Reduce tasks [1], take place within the data center [2]. For example, to process a single search query, thousands of servers within the data center are contacted in parallel [2]. The expected response time to the user is generally in tens of milliseconds [1], and a minor performance degradation or network congestion may result in a Quality of Service (QoS) violation.

Data Center Networks (DCNs) that constitute the communicational backbone of the cloud computing paradigm are of paramount importance to guarantee the system integrity [3]. The DCNs can be broadly classified into: **(a)** switch-centric and **(b)** server-centric or hybrid models [3]. The *ThreeTier* DCN is the most commonly used switch-centric architecture [4]. Al-Fares *et al.* used commodity network switches to design the *FatTree* switch-centric DCN architecture [5]. Guo *et al.* proposed the *DCell*; a hybrid DCN architecture [6] composed of recursive building units called $dcells$. The aforementioned are the three most common DCNs [3].

Various network robustness and connectivity metrics have been proposed, such as the Average Two-Terminal Reliability (A2TR) [7], which take into consideration the physical topology and node interconnection of the network. To operate successfully, the DCNs are expected to possess high tolerance to network failures [3]. However, the networks may behave diversely when exposed to various types of node or link failures.

The A2TR is used to evaluate network connectivity in response to random failures [8], [9]. In this work we extend and customize the A2TR procedure to evaluate targeted failures. Our analysis reveals that the DCNs exhibit diverse connectivity features and robustness in response to the targeted and random failures. As a consequence, we propose a new connectivity metric called μ-A2TR, which evaluates how difficult it is to break a network into components according to a specific type of failure. We believe that our proposal will aid network engineers and the research community in designing more robust and better-connected DCNs.

Our major contributions include: (a) comparing the traditional network features of the state of the art DCNs namely: *ThreeTier*, *FatTree*, and *DCell*; (b) studying the DCNs architectural network connectivity in response to random and targeted node removals; and (c) proposing μ-A2TR, a metric to characterize the underlying connectivity of the DCNs.

The rest of the letter is organized as follows. Section II presents the details of the connectivity analysis. The results of our study are reported in Section III. Our proposed metric, μ-A2TR, is presented in Section IV. Finally, Section V concludes the work by outlining future directions.

II. Connectivity Analysis

In this section, we detail the various scenarios of our connectivity study. Firstly, we present the various robustness features of the considered DCN topologies. Secondly, we introduce the methodology for the connectivity analysis of the DCNs.

A. Robustness of the DCN Topologies

In this work, three DCN topologies have been taken into account: (a) *ThreeTier*, (b) *FatTree*, and (c) *DCell*. We developed a DCN topology generator tool to generate the network topologies. The main characteristics of the DCNs are presented in Table I. Each of the topologies is composed of around 2,500 to 2,700 nodes ($|N|$). The reason for a differing number of nodes $|N|$ in each DCN architecture is that each DCN follows

Manuscript received May 21, 2013. The associate editor coordinating the review of this letter and approving it for publication was G. Lazarou.

M. Manzano and E. Calle are with the Department of Architecture and Computer's Technology, University of Girona, Spain. M. Manzano is the corresponding author (e-mail: mmanzano@eia.udg.edu).

K. Bilal and S. U. Khan are with the Department of Electrical and Computer Engineering, North Dakota State University, USA.

Digital Object Identifier 10.1109/LCOMM.2013.13.131176

TABLE I
MAIN DCN FEATURES.

Feature	*DCell*	*FatTree*	*ThreeTier*
$\|N\|$	2709	2500	2562
$\|E\|$	4515	6000	2740
d	0.00123	0.00192	0.00083
$\langle k\rangle\pm\sigma$	3.33 $\pm$0.94	4.8 $\pm$7.6	2.13 $\pm$4.64
λ_1	3.56155	17.4186	10.25044
$\mu_{\|N\|-1}$	0.12439	0.31528	0.02308
$k_{\max}$	4	20	40
κ	1	1	1
ρ	1	1	1
$\langle l\rangle\pm\sigma$	8.51 $\pm$1.93	5.21 $\pm$1.12	5.72 $\pm$0.71
D	15	6	6
$\langle b\rangle\pm\sigma$	0.003 $\pm$0.001	0.002 $\pm$0.003	0.002 $\pm$0.014
r	-0.25	-0.2	-0.8961

a predefined complex topology and connectivity pattern [3]. We have configured the servers and switches to obtain the closest possible number of nodes for each architecture. As can be observed, the *FatTree* DCN architecture has the largest number of edges ($|E|$) and the highest density value (d). Consequently, the *FatTree* has the highest average nodal degree ($\langle k\rangle$), which means that it is better-connected on average than the other DCNs. Moreover, when regarding spectral radius (the largest eigenvalue of the adjacency matrix, λ_1), and algebraic connectivity (the second smallest Laplacian eigenvalue, $\mu_{|N|-1}$), the *FatTree* proves to be the most robust network. This is because the higher the value of λ_1 and $\mu_{|N|-1}$, the higher the difficulty to segregate the network is [10]. Although the *ThreeTier* architecture indicates better robustness than the *DCell* architecture when considering λ_1, it possesses the highest maximum nodal degree ($k_{\max}$) that indicates a high vulnerability. This is because the removal of such a node can seriously affect the network. The values of the node (κ) and link (ρ) connectivity for all of the networks are $\kappa = 1$ and $\rho = 1$, respectively. These values imply that a single node or link failure may cause network fragmentation [11]. This behavior is not expected from DCN architectures, as they are dependent on the element connectivity for successful operation.

The high values of average shortest path length ($\langle l\rangle$) and the diameter (D) indicate that the inter-node communications within the *DCell* architecture are more susceptible to being affected by a failure. The average node betweenness centrality ($\langle b\rangle$) depicts that, although the *DCell* has the highest average value of $\langle b\rangle$, the *ThreeTier* is the most vulnerable network. This is due to the fact that it presents the highest standard deviation in the individual node's $\langle b\rangle$ values. Therefore, it can be inferred that the *ThreeTier* network has an excess of centrality measures for some of the nodes, indicating the vulnerability under targeted failures [10]. Finally, although all of the networks are dissassortative ($r < 0$), the *ThreeTier* exhibits a value near to -1 which implies that this network has an excess of links connecting nodes of dissimilar degrees [10].

B. DCN Connectivity Analysis

According to the characteristics of the DCNs presented in the previous section, it can been inferred that the *FatTree* is the least vulnerable network, followed by the *DCell* and *ThreeTier* architectures, respectively. To examine the connectivity of the DCNs in detail, we evaluate the A2TR [7] value of each network in the case of three different types of node removals. The nodes to be removed are selected: **(a)** randomly, as discussed in various studies, such as [8], [9], [12]; **(b)** by their nodal degree; and **(c)** by betweenness centrality. The nodes with high betweenness centrality and nodal degree are selected for removal to demonstrate the system connectivity under targeted attacks [12], [13], [14].

The $A2TR(p)$ is the probability that a randomly chosen pair of the nodes is connected when p nodes are removed from the network. If the network is fully connected, the value of A2TR is equal to 1. Otherwise, when p nodes are removed, the A2TR value is calculated as the sum of the number of the node pairs in every strongly connected component (SCC) divided by the total number of node pairs in the network:

$$A2TR(p) = \frac{\sum_{i=1}^{|SCC|} |C_i| \cdot (|C_i| - 1)}{|N'| \cdot (|N'| - 1)}, \tag{1}$$

where $|C_i|$ is the number of nodes of the SCC number i, and $|N'|$ is the vertex size of the residual graph $|N| - p$. This ratio indicates the fraction of node pairs that are connected to each other. Therefore, the higher the A2TR value (for a given number of removed nodes), the more connected the DCN is.

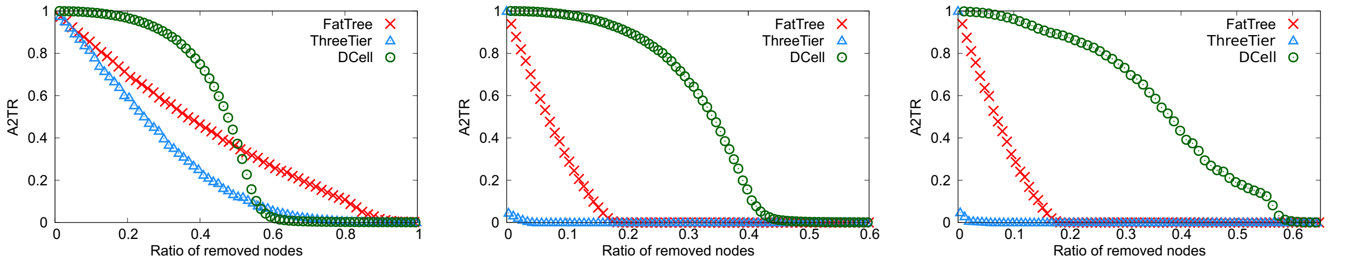
We compute the A2TR value from $p = 0$ to $p = |N| - 2$, where $|N|$ is the total number of nodes in a DCN. In the procedure described in this section, the most expensive computation is in obtaining, for each p, the strongly connected components of the network. For that step, we use Tarjan's algorithm [15], whose running time complexity is in $O(|N| + |E|)$. The simulation was performed on a Linux system with an 8-core 64-bit Intel Xeon processor of 2GHz and 16 GB of RAM. We employed a discrete-event simulation tool called PHISON [16].

III. RESULTS

The results of the connectivity analysis are presented in Fig. 1, which depicts the A2TR evolution according to the three type of node removals. The depicted values are the average of 1,000 runs with different random seeds, this being a widely used value in the bootstrap literature to carry out replications because it guarantees low variance [17].

In Fig. 1a it can be observed that for a lower percentage of randomly removed nodes (up to 40%), the *DCell* exhibits highly connected network, because of the high A2TR values as compared to the *ThreeTier* and *FatTree* architectures. The *ThreeTier* network is more affected by the random removal of the nodes than the *ThreeTier* network. However, it is interesting to note that the connectivity of the *DCell* decreases extremely rapidly within the interval of 40% to 60% of removed nodes. Nevertheless, *FatTree* maintains a smooth linear decline for any percentage of removed nodes. Consequently, as discussed in Section II, of all three architectures considered and in response to high percentages of random node failures, the *FatTree* proves to be the most connected network.

The connectivity analysis of the DCNs observed in the case of high nodal degree and betweenness centrality based



(a) Random (b) Average Nodal Degree (c) Betweenness Centrality

Fig. 1. A2TR of DCNs.

node removal differs significantly from the random nodes removal. The results presented in Fig. 1b and Fig. 1c depict the targeted removal of the nodes. It can be observed that the *ThreeTier* architecture is the most vulnerable network. Less than 10% of the node pairs remain connected to each other when removing only four nodes (core layer nodes) in the *ThreeTier* architecture. Contrary to the random nodes removal case, the *FatTree* network is significantly affected by the targeted failures. However, the A2TR value curves of the *FatTree* exhibit a smoother decline than the *ThreeTier* A2TR value curves. Finally, of the three architectures considered, the *DCell* is the most connected network for targeted nodes removal cases.

To conclude, it is worth noting that the network features analyzed in Section II do not accurately translate when evaluating the connectivity of the networks in various failure/node removal scenarios. Despite the fact that the *FatTree* exhibits better robustness features ($\langle k \rangle$, λ_1, $\mu_{|N|-1}$, $\langle l \rangle$ and D) than the *DCell* (k_{max}) architecture, the connectivity analysis demonstrates that the *DCell* architecture exhibits less vulnerability than the *FatTree* architecture. Therefore, it necessitates defining a new metric, which can accurately evaluate the connectivity of the DCNs.

IV. μ-A2TR

In this section we present μ-A2TR as our third contribution to this letter, a novel metric to evaluate the connectivity of DCNs. We compute μ-A2TR for a given network and a given type of failure, from the A2TR values which are obtained by conducting the analysis defined previously in this letter. The idea of considering the performance curve for increasing network damage was initially proposed in [18]. As a result, our proposal characterizes how difficult it is to break a network into components when considering an incremental node failure scenario. Therefore, μ-A2TR can be defined as:

$$\mu\text{-A2TR} = \frac{\sum_{p=0}^{|N|-2} A2TR(p)}{|N|-1}, \quad (2)$$

where p is the number of nodes that have been removed from the network, and $A2TR(p)$ is the A2TR value of the network for p removed nodes. μ-A2TR takes values over the interval [0, 1]. The higher the value of μ-A2TR, the more robust the DCN, in terms of connectivity, and more difficult to segregate the DCN into smaller clusters is.

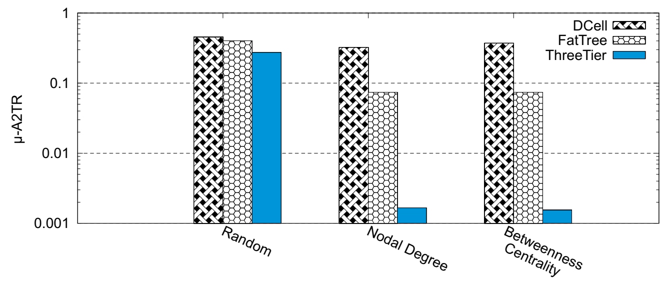


Fig. 2. μ-A2TR in logarithmic scale of the DCNs.

Fig. 2 presents the μ-A2TR values for the three DCNs, and for the three types of node removals. As can be observed, the *DCell* architecture exhibits the highest μ-A2TR: 0.45, 0.32 and 0.37 for the random, nodal degree, and betweenness centrality, respectively. The *DCell* architecture follows a recursively built topology, where each $dcell$ connects to l other $dcells$ (l is the level of the *DCell* architecture [3]). Moreover, the nodes within the *DCell* architecture exhibit low standard deviation in the nodal degree and betweenness centrality. Therefore, the *DCell* architecture exhibits high resilience to the node failures. On the contrary, the *FatTree* and *ThreeTier* architectures follow a hierarchical topology, where some of the nodes (core layer nodes) possess high nodal degree and betweenness centrality. In case of the targeted failures, these nodes are chosen for removal, resulting in network segregation and low connectivity. However, the number of nodes in the core layer of the *FatTree* are higher than in the *ThreeTier* architecture, the latter only having 4 nodes in our case. Therefore, the *FatTree* architecture exhibits better connectivity in terms of μ-A2TR (0.41, 0.07 and 0.07 for the three types of node removals) than the *ThreeTier* (0.27 in the case of random removals and close to 0 in both of the targeted cases).

Unlike traditional graph features (see Section II-A), our proposal describes how robust a network is in the case of specific failure scenarios. The μ-A2TR metric is able to denote significant differences between the three DCN topologies considered in this work. For instance, according to the algebraic connectivity ($\mu_{|N|-1}$) or the spectral radius (λ_1), the *FatTree* is the most robust network. However, μ-A2TR demonstrates that, although the *DCell* and the *FatTree* perform similarly in the case of random node removals, the former is more robust under targeted failures. In conclusion, μ-A2TR provides further insight into the connectivity of the DCNs, this being highly beneficial for the network research community given the critical role played by such networks currently.

V. Conclusions and Future Work

This letter presented a comparison of the network features of the three well-known DCN architectures namely: **(a)** *ThreeTier*, **(b)** *FatTree*, and **(c)** *DCell*. Moreover, we conducted a connectivity analysis of the considered DCNs. Finally, we proposed μ-A2TR, a novel robustness metric, which is able to characterize network connectivity.

It has been observed that, based on several classical robustness features such as density, average nodal degree, spectral radius, algebraic connectivity, average shortest path length, and diameter, the *FatTree* architecture is the most robust and connected network. However, the connectivity analysis of the DCNs based on the A2TR values in response to three types of node removals (random, nodal degree, and betweenness centrality) demonstrated that the *DCell* and *FatTree* are similar in terms of network connectivity in the case of random removals. Nevertheless, as regards to the targeted removals, the *FatTree* and *ThreeTier* depicted low network connectivity. From the connectivity analysis, it can be inferred that, although the traditional network features are useful in determining network robustness and connectivity, there is a need for an appropriate connectivity metric.

We presented μ-A2TR and demonstrated its ability to characterize network connectivity. We hope that the μ-A2TR metric will help the engineers and research community to design more robust DCNs.

For future work, a wide range of the node or link removal scenarios can be considered to compute μ-A2TR of the DCNs. Moreover, network optimization procedures could be defined to enhance the connectivity of the DCNs. Finally, we will mathematically formalize the μ-A2TR metric by providing analytical proofs for different failure scenarios.

Acknowledgements

This work is partly supported by the Spanish Ministerio de Ciencia e Innovacion through project RoGeR (TEC 2012-32336) and by the Generalitat de Catalunya through the research support program projects SGR-1202 and AGAUR FI-DGR 2012 grant.

References

[1] D. Abts and B. Felderman, "A guided tour of datacenter networking," *Commun. ACM – ACM Queue*, vol. 55, no. 6, pp. 44–51, 2012.

[2] A. Vahdat, H. Liu, X. Zhao, and C. Johnson, "The emerging optical data center," in *Proc. 2011 Optical Fiber Communication Conference*, pp. 8–10.

[3] 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*, 2012.

[4] Cisco, *Cisco Data Center Infrastructure 2.5 Design Guide*, 2010.

[5] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 4, pp. 63–74, 2008.

[6] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu, "Dcell: a scalable and fault-tolerant network structure for data centers," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 4, pp. 75–86, 2008.

[7] S. Neumayer and E. Modiano, "Network reliability with geographically correlated failures," in *Proc. 2010 Conference on Information Communications*, pp. 1658–1666.

[8] M. Manzano, E. Calle, and D. Harle, "Quantitative and qualitative network robustness analysis under different multiple failure scenarios," in *Proc. 2011 International Workshop on Reliable Networks Design and Modeling*, pp. 1–7.

[9] M. Manzano, J.-L. Marzo, E. Calle, and A. Manolova, "Robustness analysis of real network topologies under multiple failure scenarios," in *Proc. 2012 European Conference on Networks and Optical Communications*.

[10] 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," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 36, pp. 17–26, 2006.

[11] A. H. Dekker and B. D. Colbert, "Network robustness and graph topology," in *Proc. 2004 Australasian Conference on Computer Science*, pp. 359–368.

[12] J. Guillaume, M. Latapy, and C. Magnien, "Comparison of failures and attacks on random and scale-free networks," in *Proc. 2005 International Conference on Principles of Distributed Systems*, pp. 186–196.

[13] P. Holme, B. Kim, C. Yoon, and S. Han, "Attack vulnerability of complex networks," *Physical Review E*, vol. 65, no. 5, p. 056109, 2002.

[14] M. Manzano, V. Torres-Padrosa, and E. Calle, "Vulnerability of core networks under different epidemic attacks," in *Proc. 2012 International Workshop on Reliable Networks Design and Modeling*.

[15] R. E. Tarjan, "Depth-first search and linear graph algorithms," *SIAM J. Computing*, vol. 1, no. 2, pp. 146–160, 1972.

[16] M. Manzano, J. Segovia, E. Calle, and J.-L. Marzo, "Phison: playground for high-level simulations on networks," in *Proc. 2012 International Symposium on Performance Evaluation of Computer and Telecommunication Systems*.

[17] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Chapman & Hall, 1993.

[18] A. Sydney, C. Scoglio, M. Youssef, and P. Schumm, "Characterising the robustness of complex networks," *International J. Internet Technol. and Secured Transactions*, vol. 2, no. 3/4, pp. 291–320, 2010.