2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS): 2016 IEEE Infocom Green and Sustainable Networking and Computing (GSNC) Workshop

Crash Me If You Can: Rethinking Sustainable Data Center Networking From a Topological Perspective

Ritchie Qi[1], William Liu[1], Jairo Gutierrez[1] and Asad W. Malik[2]
[1] School of Engineering, Computer and Mathematical Sciences,
Auckland University of Technology, Auckland, New Zealand
[2] National University of Sciences and Technology
NUST - SEECS, Islamabad, Pakistan
{yyj7588, william.liu and jairo.gutierrez} @aut.ac.nz and asad.malik@seecs.edu.pk

Abstract— a data center is a pool of resources such as computational, storage facilities, and servers interconnected using a communications network. The Data Center Networking (DCN) holds a pivotal role, and it needs to be scalable and efficient to connect the growing number of servers so as to handle the intensive demands in cloud computing. Recently there has been a rapidly growing literature on DCNs which focuses on studying how to model and evaluate the resource provisioning and allocation algorithms for more effective and efficient resource management of cloud systems. Unfortunately there is not much work to reveal how the underlying network topological connectivity can affect the overlay DCNs performance in aspects such as energy consumption and service resilience. There is a saying that, it is not what you know i.e., the algorithm itself but who you know i.e., connectivity, which argues that people get ahead in life based on their connections, not on their skills or knowledge, and every day offers evidence of this aphorism. This case also applies to DCNs. DCNs performance is not merely a function of resource provisioning and allocation, but also it is a network-wide activity. The structure and ties that links a data center to other data centers are also critical factors. In this paper, we are paving a new developing line of research on revealing the roles of network topological characteristics to foster more sustainable DCNs in terms of resource efficiency and service resilience. We believe this research is thought-provoking, and opens a new conversation for researchers working in the area of sustainable DCNs communications and networking.

Keywords—Data Center Networking (DCN); QoS; Topological connectivity; Energy efficiency; Service Resiliency; Sustainable network structure; CloudNetSim++

I BACKGROUD

The rise in information service provisioning is a product of user demands, and the rapid growth of users demands in obtaining data has pushed Information Communications Technologies (ICT) services to a whole new level, which has led to a new evolution of cloud computing services. The prominent characteristic for a cloud is its centralized distribution which applies high performance computing to the task it receives, and delivers services by sharing resources which are stored inside a cloud rather than having local servers to handle applications traditionally [1, 2]. This "pay-as-you-go" model has brought huge benefits to customers but there is an enormous cost associated with maintaining its infrastructure. The cost has to be considered seriously, as this leads to a threshold level for those organizations intending to provide such services.

At present, cloud services are still evolving and, as a consequence, are far from perfect. Although the cloud services are still running with the aid of network infrastructure, the cloud performance is still being impacted and varied by different layers. Numerous cloud competitors had dedicated efforts to alleviate the limitations of clouds such as the work linked to bandwidth and internal management. Data Center Networking (DCN) is a vital component in the cloud infrastructure and it is being increasingly adopted by organizations to handle the core business and operational data that interconnects all the components in the cloud, while delivering main cloud services such as data storage and protection. It follows that the maintenance work turns to be extremely significant to the cloud service providers. However, with the sustained growth in computing capacity, the cost and operational expenses (OPEX) are showing a sharp increase. Energy consumption has been a great concern for data centers' operators, according to the survey conducted by Gartner Group, approximately 40% of data center OPEX comes from the energy consumed by ICT equipment, which is composed of computing servers (2/3) and communication links (1/3) [3]. The remainder 60% energy consumption comes from cooling and power distribution, which are 45% and 15% respectively. On the other hand, the cooling cost of heat generated by data center infrastructure ranges between \$2 to \$5 million per year [4, 5]. Therefore the optimized data center architecture plays an important role in OPEX reduction [6]. A robust network topology is essentially required to comprehensively address malicious attacks or nodes and link failures. While the current studies on DCNs are mainly focused on how to develop more effective and efficient resource provision and allocation algorithms among data centers, there is not much work to reveal how the underlying topology changes can affect the overlay DCNs` performance.

The rest of the paper is organized as follows. Section II surveys the state of the art of DCNs architectures. Section III presents the models of network performance and energy consumption. In section IV, we conduct extensive simulation studies to reveal how the underlying network structure changes can impact on the overlay DCNs performance. The conclusion and layout of future work are drawn in Section V.

II RELATED WORK

A three-tier DCN is the most commonly known configuration for the current cloud data centers. This architecture is based on three layers of switching including core, aggregate and access switches from top to bottom [7]. The core layer connects the layer 2 aggregate switches with the network outside the DCN, and the aggregate switches can be easily added due to its inexpensive character that plays a transitive role. This setup is capable of supporting an increasing number of servers up to 10,000 [7].

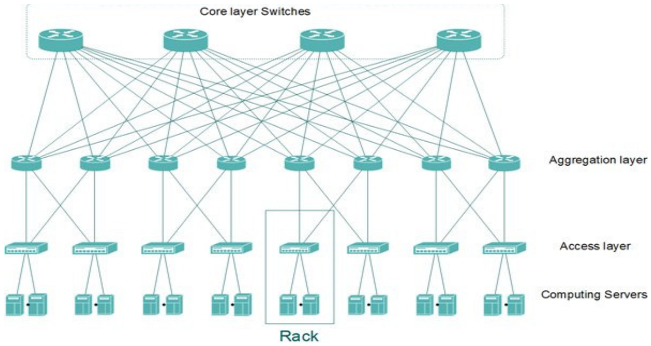

Fig. 1. Conventional Three-tier DC architecture

The Three-tier network topology, as shown in Fig.1, is easily setup and it uses fewer network components and this lowers the hardware costs significantly [8]. However, the Three-tier DCN shows intense capacity because of the cost involved which is reflected in link oversubscriptions, and growing demands on services which is the reality noticed during recent years. The architecture then lacks scalability, has energy efficiency problems, and has issues with the use of cross-sectional bandwidth.

The Fat-tree oriented DCN as shown in Fig.2 (and proposed by [8]) is the most widely adopted DCN configuration. It follows the hierarchy architecture which contains core, aggregate and access layers, and this structure is composed of *k* pods, where in each pod, there are *(k/2) ^2* servers, *k/2* access layer switches, and *k/2* aggregate layer switches. The core layers contain *(k/2) ^2* core switches where each of the core switches is connected to one aggregate layer switch in each of the pods. The Fat-tree DCN presents advantages in strengthening the ability of resolving oversubscription and cross section bandwidth by contrast to Three-tier DCNs. The other DCNs architecture includes BCube [9] and HyperFlatNet [10] etc. which are different from the Fat-tree and the Three-tier architectures as they adopt a server-centric methodology that relies on mini-switches for the interconnections.

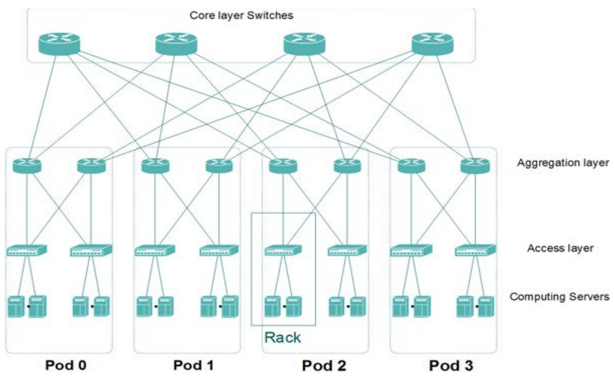

Fig. 2 Fat-tree DC architecture

The BCube topology [9] is a datacenter structure built inside shipping containers that introduce a brand-new DCN configuration and proposed to be used as a Modular Data Center (MDC), which simplifies the installation procedure and implements physical migration as compared with conventional data centers. Datacenter migration facilitates energy saving, because shipping datacenters to regions promotes strategic positioning, allowing the placement close to regions with high service demands. As MDCs are built in sealed containers with a high equipment density, they need to be highly reliable. Furthermore, the equipment has to be moved under careful control because the failures increase when the hardware is not protected well during the shipping process.

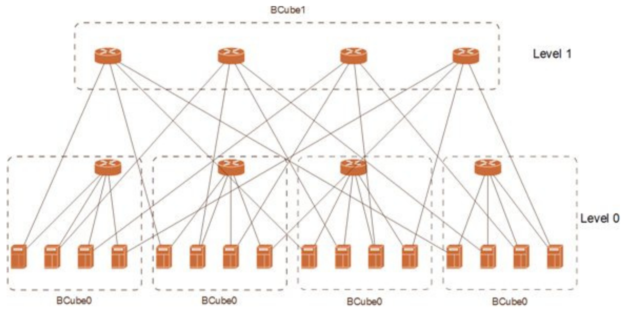

Fig. 3 BCube DC architecture

HyperFlatNet is a recursive DCN topology which was proposed by [10], HyperFlatnet is formed by two layers with the first layer containing n servers connected by one n-port switch; the second layer consists of n^2 first layers. Hence, the total n^3 servers can be taken as n^2 clusters of n servers. Moreover, different servers can be represented as a n^2 ∗ n matrix where the row and column indexes correspond to the cluster number (*i*) and the index in the cluster (*j*). The authors propose a connection algorithm called Linked Clusters Maximization (LCM) algorithm which increases the number of directly connected clusters and reduces the number of intermediate hops used to transmit the packet to the destination. A 64-server HyperFlatNet DC architecture is shown in Fig. 4.

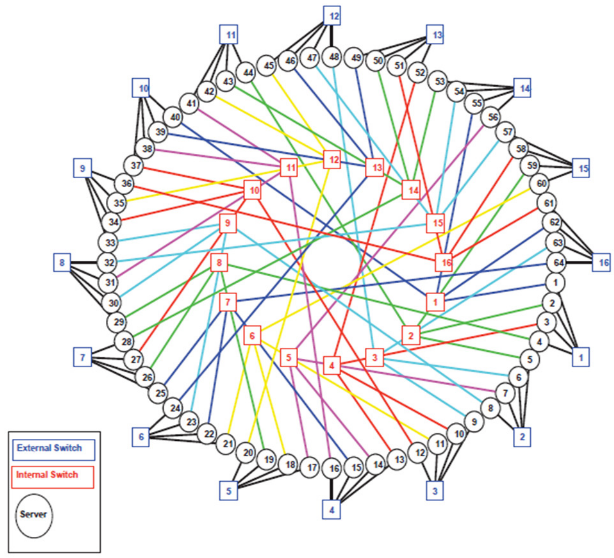

Fig. 4 HyperFlatNet DC architecture [10]

The trend of growing reliance on online services and the increasing demands from mobile devices has already converted many client based applications into cloud services [11], which has motivated the data center networking to be geo-distributed. The distributed data centers are highly adopted in modern cloud infrastructure that consists of a number of data centers combined by being interconnected to each other. The most used connection alternatives are based on mesh and star topologies. Mesh networks relay data by using either a flooding technique or a routing technique. It utilizes the shortest path bridging (SPB) algorithm for the OSPF routing protocol particularly in wired Ethernets [12]. The message is propagated along a path by hopping from node to node until it reaches its destination. The advantages of a mesh network are stability and reliability; a fault on the cable attached between two nodes will not affect others [13]. The star network, as its shape indicates, uses a centralization arrangement, this topology has benefits but the central hub holds the vantage point so that an optimal configuration for the central point can generate a higher performance. However, the failure of the centralization node will crash the whole network.

III DCN Performance and Energy Models

Based on the existing DCNs infrastructure, there is scope to conduct topological changes in order to improve the DCNs performance. The communication components for networking, such as the switches and links, directly influence the network performance of the cloud infrastructure. Normally, due to the geographical distribution characteristics, the links connecting the data centers cover a long distance, which means link delays cannot be neglected as a performance metric. According to [14], the latency or packet delivery time indicates the time spent from the first bit sent from the transmitter to the last bit received by the receiver. Furthermore, latency is composed of packet transmission time and link propagation delay. The average packet delay can be calculated as follows: where D_{avg} refers to the average packet delay and n represents the number of packets received, d_i is the delay of the packet i.

$$D_{avg} = \frac{1}{n}\sum_{i=1}^{n} d_i \qquad (1)$$

On the other hand, when a data stream is transmitted over a communication channel, there is a possibility that the number of received bits be altered due to the link noise, interference, etc. [15], and this leads to a probability of packet dropping. In addition, the packets also have the probability of being dropped when a large data stream traffic pass the switch especially when the switch load burdens beyond its frame capacity. In those cases, the packets are dropped automatically by the switch queuing mechanism. The average network throughput can be calculated as the following, where T_{avg} refers to the average throughput in the network, $p_i \in [0,1]$ while 0 reflects the loss of the packet and 1 indicates the receipt of the packet, δ_i refers to the size of packet in bits and d_i represents the delay of the packet, and n the number of packets received.

$$T_{avg} = \frac{\sum_{i=1}^{n}(p_i \times \delta_i)}{\sum_{i=1}^{n} d_i} \qquad (2)$$

A. Power Management Model

1. Dynamic Voltage/Frequency Scaling

Power management can be achieved by using the Dynamic Voltage/Frequency Scaling (DVFS) technique [16], In DVFS, chip switching power decreases proportionally to $V^2 * f$, where V is voltage and f is switching frequency. The core principle is that the average power consumed has a cubic relationship with the CPU frequency, moreover, the power consumption for the components not related to f remains fixed, such as bus, memory, and disk. Therefore, the server power consumption can be stated as follows, where P_{fix} is power consumption of hardware components not linked with frequency, P_f is CPU power consumption linked with frequency.

$$P = P_{fix} + P_f \times f^3 \qquad (3)$$

On the other hand, according to [6], the total energy consumption for the DCN can be divided into three main portions in general: computing by servers, communication supported by links and network equipment operations; and the power consumed by infrastructure for supporting data centers (e.g. cooling/air conditioning system), which is to say that only a fraction of energy consumption has been delivered to the computing server directly, another considerable portion of the energy is consumed for maintaining interconnection links and network equipment operations. The power consumption for switches makes up a great proportion to the overall DCN power consumption. As stated in [6] the energy consumed by a switch can be expressed as:

$$P_{switch} = P_{chassis} + n_{linecard} * P_{linecard} + \sum_{i=0}^{configs} nports_{configs_i} * P_{configs_i} \qquad (4)$$

Where P_{switch} is the power consumption for the switch, $P_{chassis}$ is the power consumed by switch hardware, $P_{linecard}$ indicates the line card power consumption with no ports turned on, $n_{linecard}$ represents the number of cards plugged into a switch, $P_{configs_i}$ is related to the power consumed for a port running at rate i. $P_{chassis}$ & $P_{linecard}$ are fixed due to the operation of a switch, so in this equation, only $P_{configs_i}$ is dependent on transmission rate i which is proportional to overall power consumption of the switch.

2. Dynamic Power Management

As reported in [17, 18], although a server stays in an idle state, it still consumes around 66% of energy compared to its full load energy consumption, which comes from the fixed components that are not related to the frequency but which also consume power. According to [19], the minimum length of time for a server staying in an idle period is referred to as the break-even time (Tbe), the state transition delay (To) which consists of shutdown delay (Tsd) and wake-up delay (Twu), the energy consumed during this period is Eo. Considering the power consumed in working and sleeping states is Pw and Ps. Fig. 5 (a) represents the working state of the server; Fig. 5 (b) demonstrates the shutdown state of the server. The break-even time makes energy consumption in both cases equal. The total energy consumed by a server that is in a working state going through the minimum time length to save power is $\mathrm{Pw} \times \mathrm{Tbe}$, While the total energy consumed by a server that is in sleep mode with the minimum time length to save power is $\mathrm{Eo} + \mathrm{Ps} \times (\mathrm{Tbe} - \mathrm{T})$ therefore, $\mathrm{Pw} \times \mathrm{Tbe} = \mathrm{Eo} + \mathrm{Ps} \times (\mathrm{Tbe} - \mathrm{To})$, so that $\mathrm{Tbe} = (\mathrm{Eo} - \mathrm{Ps} \times \mathrm{To})/ (\mathrm{Pw} - \mathrm{Ps})$, The break-even time has to be larger than the transition delay; therefore, $\mathrm{Tbe} = \max[(\mathrm{Eo} - \mathrm{Ps} \times \mathrm{To})/(\mathrm{Pw} - \mathrm{Ps}), \mathrm{To}]$.

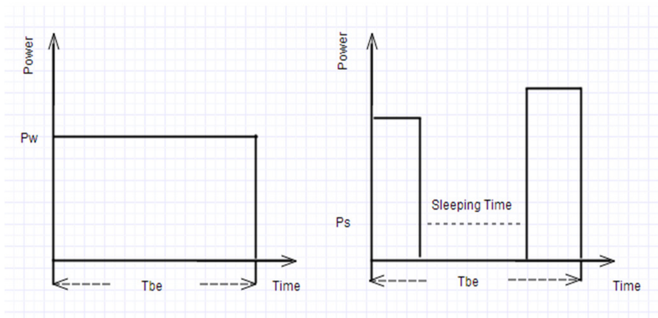

Fig. 5 (a) Server with working state, (b) Server with shutdown and wake-up state

Fixed timeout is the most conventional policy that is used in Dynamic Power Management (DPM) and it implements a timeout value τ. The device is put into sleep mode if it is kept in idle for more than τ. The basic mechanism is that if the device remains idle up to τ, then it should further stay idle for at least Tbe. However, this policy wastes energy within the value τ.

In this paper, we deploy the DVFS technique according to Eq.3 and Eq.4, combined with a DPM fixed timeout policy to calculate the total energy consumption and the overall energy consumption includes server consumption and switch consumption (it omits the link consumption).

B. Metrics of Robustness on Fault tolerence

1. Average Nodal Degree (k)

This is the coarsest connectivity feature of any topology [20]. Networks with higher k are "better-connected" on average, and consequently, are likely to be more robust. However, if a node with a higher nodal degree fails, potential higher numbers of connections are also prone to be affected. Thus, this metric by itself provides only a limited measure of the robustness of a network which is likely to vary depending on how the nodal degree is actually distributed over the network graph. Fig. 6 shows a calculation example of the average nodal degree.

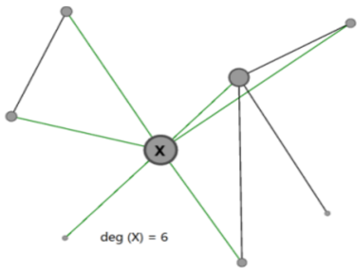

Fig. 6 – Calculation of Avg. nodal degree

$$deg_{avg} = \frac{node\ count \times value}{total\ number\ of\ node} = \frac{2\times1+4\times2+1\times4+1\times6}{8} = 2.5$$

2. Network Diameter

The network diameter is known as the maximum length among all the calculated shortest paths in a network [20], under the All-to-All server traffic scenario, is the longest path from one server to another reachable server. From a topological view, a smaller network diameter indicates relative lower network latency because of fewer opportunities of producing transmission delays and less queuing time taken so that generates more effective routing ideally.

3. Average Shortest Path Length (ASPL)

Average shortest path length (ASPL) is calculated as an average of all the shortest paths between all the possible origin-destination node pairs of the network [20]. Generally, networks with smaller ASPL are more robust in terms of network latency performance, but they are prone to lose connections due to having a smaller link cardinal number. When comprehensively comparing the network performance, ASPL cannot tell all and it is usually a network performance metric with regard to the network latency.

IV Simulation Studies

We have built an energy-aware Data Center Networking through simulation in order to reveal the impact of the underlying network structure and connectivity on overlay DCNs performances. We modeled four identical data centers that are geographically distributed by using the CloudNetSim++ simulator [21] and the data centers are interconnected by a mesh topology.

TABLE I SCENARIO SETUP

Inter-DCN	Mesh Network
Intra-DCN	Fat-tree; BCube-2 level; BCube-3 level; Three-tier; HyperFlatnet
Traffic Protocol	UDP
Traffic Type	(all-to-all) All servers to all servers
Packet Size	2500 bytes
Packet Send Interval	0.8s
Connections	Gigabit Ethernet links
Link quality	Packet Error Rate = 0.8%
Link Bandwidth	intra-DCN: 1Gbps - 10Gbps; inter-DCN: 100Gbps
Simulation Time	2000 simulation seconds

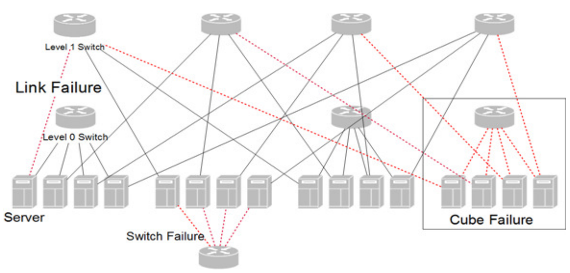

Fig.7 – DCN Failure classification

In each scenario, the network performance of the Fat-tree, Three-tier, BCube with 2 switch layers, BCube with 3 switch layers and HyperFlatNet DCN topologies are compared against an increasing Link Failure Ratio (LFR). The traffic we simulated is servers-to-servers, where each server transmits UDP packets to other servers, following with an All-to-All traffic pattern. In order to seek the prominent feature of various DCN topologies, the maximum fault level is adopted as the failure mechanism; failed links are selected as representative, connections lying on important routes such as rack failures, Cube failures are chosen to be disconnected as in Fig. 7 so that the characteristics of each DCN topology could be enlarged.

A. Scenarios Setup

In order to make network performance comparably between each data center topology, the number of servers and the inter-DCN configuration are maintained fixed, and as a consequence, the performance will vary based on the differences of DCN type, the internal architecture, the number of switches & links, the server port number, and the switch port number.

TABLE II. DCN TOPOLOGY SETUP DETAILS

	Fat Tree	**Three Tier**	**BCube**	**BCube2**	**HyperFlat Net**
Switch degrees	3 layer	3 layer	2 layer	3 layer	2 layer
Switches Breakdown	Core=16 Aggr=32 Edge=32	Core=16 Aggr=32 Edge=32	Level0=32 Level1=32	Level0=48 Level1=48 Level2=48	Internal=64 External=64
No. of Switches	80	80	64	192	128
No. of Servers	***256***	***256***	***256***	***256***	***256***
No. of Links (4 DCs)	384	448	512	512	976
No. of Links (mesh)	***16***	***16***	***16***	***16***	***16***
Total Links	400	464	528	528	992
Core Switch ports	4	8	-	-	-
Aggr Switch ports	4	6	-	-	-
Edge Switch ports	10	10	-	-	-
Level 0 Switch ports	-	-	8	4	-
Level 1 Switch ports	-	-	8	4	-
Level 2 Switch ports	-	-	-	4	-
Internal Switch ports	-	-	-	-	5
External Switch ports	-	-	-	-	4
Max. Server ports	1	1	2	3	2

When the number of servers is maintained, there are 3 layers of switches in the Fat-tree, Three-tier and BCube-3 layer configurations; as a result these three DCNs possess a higher number of switches than the other 2-layer DCNs (BCube-2 layer & HyperFlatNet). However, due to the fact that the connecting patterns varied diversely, there is a resulting difference on the number of links used. The HyperFlatNet consumed the maximum number of links (976) while Fat-tree obtained the least number of links, which are 384. On the other side, although two BCubes are diverse from the switch layers and the switch numbers, the links used remained as the same because in both cases the DCNs` switches and servers differentiated between port numbers: BCube-2 layer employs 8-port switches and only 2-port servers while BCube-3 layer applied comparatively less costly 4-port switches and 3-port servers. Fat-tree and Three-tier DCNs have similar architectures, which only varied in terms of the number of links consumed and switch types; Fat-tree has fewer link oversubscriptions than the mainstream the Three-tier DCN so that fewer links are used when compared to Three-tier case. In addition, from the view of DCN arrangement, although Fat Tree and Three Tier obtain the same number of racks (32 racks in this scenario), the cost of establishing a Fat-tree DCN can be reduced not only in relation to the number of links used, but also because the number of maximum-port switches used can be rationalized. In other words, the core and aggregation layer switches can be substituted for commodity 4-port switches.

B. Result Analysis

TABLE III THE GRADIENT OF AVERAGE NODAL DEGREE

Fat Tree	Three Tier	BCube	BCube2	HyperFlatNet
0.058	0.069	0.079	0.061	0.112

From TABLE III, we can see that by comparing Fat-tree/Three-tier/BCube-2 layers and 3 layers set-ups, HyperFlatNet has a higher gradient of decreasing its average nodal degree, which implies that with the LFR increased, the HyperFlatNet mitigates the probability of connections being affected. While in this case, even the Fat-tree gains the lowest average nodal degree, it lacks the ability of "stopping" the connections from being affected.

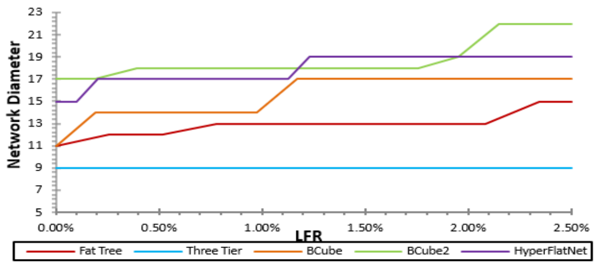

Fig. 8 Network Diameter against LFR

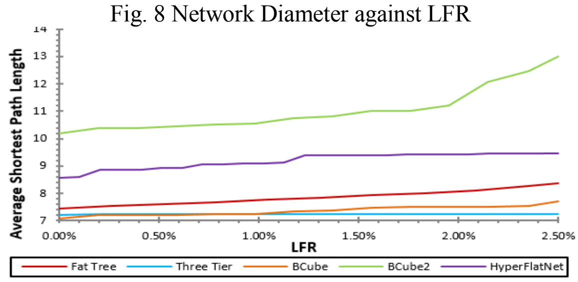

Fig. 9 ASPL against LFR

As shown above, the Three Tier DCN topology maintains the lowest network diameter and ASPL with regard to the increase of LFR, which implies that Three-tier obtains the lowest network latency than the others. On the other side, BCube-3 layer is suspected to gain the highest network latency.

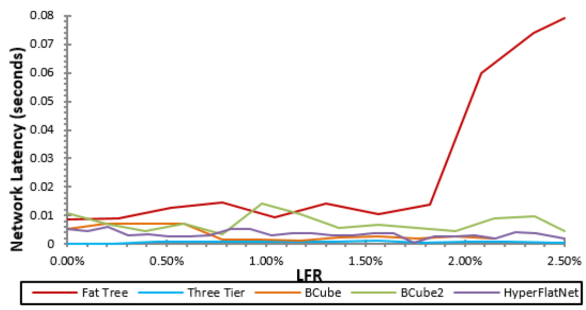

Fig. 10 Network Latency against LFR

The previous assumption could have identified by the latency curves; i.e., the Three-tier gains the predicted result with the lowest network latency with the increase of LFR which verified the hypothesis that it could take more chances to go through less transmission delay with small ASPL and network diameter. However, BCube-3 layers did not achieve the highest latency as predicted, but Fat-tree did. The discrepancy occurred mainly within the range 2.50% > LFR > 2.00%, as the traffic generated is unbalanced. It is a coincidence that within that LFR range, the specific racks which generated a large number of data flows failed to transport packets to other racks on short paths so that more packets were necessarily taken on longer paths. Furthermore, the transmission time spent for a large proportion of total packets has been increased as a result of the paths becoming longer while the packets received decreased due to longer queueing times required within the fixed UDP application transmitting time as shown in the following figure which has led to a more steep gradient after 2.00% LFR of avg. network latency for the Fat-tree DCN.

The joggle of latency for HyperFlatNet is created by random traffic load as the packet received stays smoothly with a little decrease, which implies that the HyperFlatNet DCN topology is relatively stable when compared with others. And also, from the topological point, HyperFlatNet possesses 1.9 times more links than BCube and 2.54 times than the Fat-tree topology, which is supposed to have much more powerful stable transmission ability despite that it has the second larger ASPL and network diameter out of the five topologies.

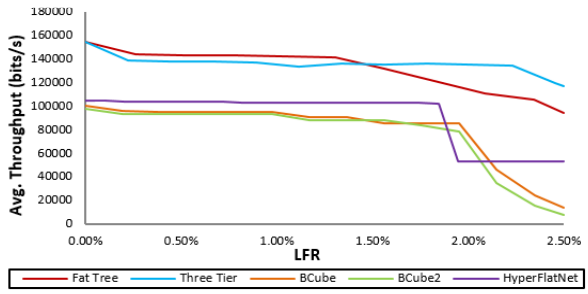

Fig. 11 Avg. Network Throughput against LFR

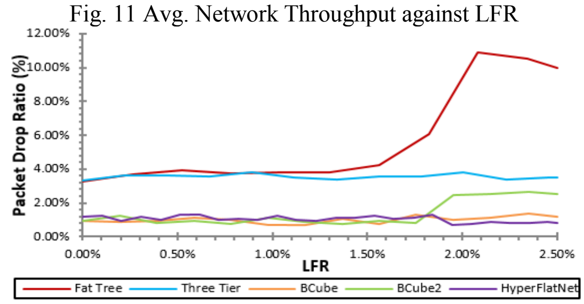

Fig. 12 Packet Drop Ratio against LFR

The above two figures show that the result curves could be clustered by different types of topologies; the tree-based conventional Three-Tier and Fat-tree have very closed initial points on avg. throughput which is around 150,000 bits/s. Even though the Fat-tree suffer less from link oversubscriptions than the Three-tier topology; With the link failure ratio increased, the Fat-tree decreases more urgently than Three-Tier because Three-tier has more effective connections for maintaining the number of ASPL as well as the network robustness. The server-centric topologies all appeared steady until the LFR reached around 2.00%. Relatively, the HyperFlatNet is more robust within limited LFR, but a fall that large beyond the failure range indicates specific effective connections failed, which may imply a substantial decrease on network performance. However, the performance could be still steady at a accepted level by compared to other topologies. The BCube-2 layers and BCube-3 layers represent highly reliable capacity when LFR is less than 2.00%, the avg. throughput decreases approximate 20% which compares to Fat-tree, with 31.25% and HyperFlatNet, with 46.8%.

In a similar fashion, with the avg. throughput, the result of PDR against LFR can also be classified into two clusters, tree-based DCN topologies and server-centric topologies. Obviously, the PDRs for tree-based topologies are higher than for the server-centric topologies. The Fat-tree still presents an unstable trend, owing to the connections failed to reliably relay messages due to fewer effective links utilized on Shortest Path. BCube-3 layers has more switches used than the other two server-centric topologies, this generates higher probabilities of the packet dropping when the messages are getting through the queuing process.

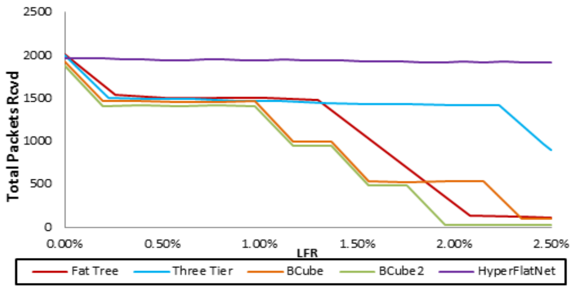

Fig. 13 Total Packets Received against LFR

There are a total of 2100 packets generated for all topologies. The packets received for all topologies present ladder-like trends except for the HyperFlatNet, while the BCube-2 layers and 3 layers produced several steps downward after 1.00% LFR, and the Fat-tree appeared to show lower reliability on higher LFR than the other tree-based topologies. The HyperFlatNet with the largest number of links remains the arrangement with the highest reliability against the other 4 topologies.

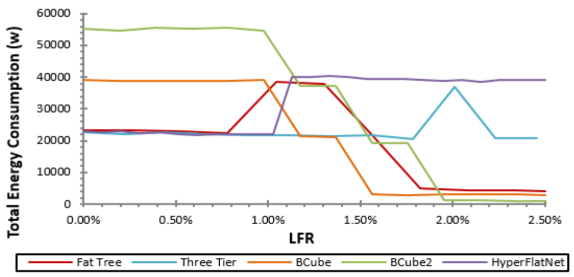

Fig.14 - Energy Consumption against LFR

In addition, we deploy DVFS and the DPM energy model together to maximize the total energy savings. The energy consumption is adjusted according to the node voltage / CPU frequency by using the DVFS technique and idle nodes are shut down automatically based on the DPM technique so that energy consumption can be reduced to the lowest level. Within the limited simulation time, the energy consumption varied intensely as the LFR varied; this is shown in the above figure. In Fat-tree/Three-tier architectures, any failures of a top-of-rack (ToR) switch will cause the disconnection of a whole rack of servers; so that these servers will not receive any packets outside that rack. Therefore, the idle servers decreased generate a decrease pattern on the total energy consumption. Otherwise, any failures of an end-of-rack (EoR) switch (aggregation or core switch) will transfer the switching load to alternative switches thus the distributed load on switches increases the energy consumption for the network. On the other hand, in order to capture the prominent topological features of the different configurations, connections lie on important routes that have been selected as failure scenarios, so that the results are amplified as representative. The energy consumption decreases in both two BCube topologies due to the Cube failure as shown in Fig.7. The BCube-3 layers consumed more energy than 2 layers due to the fact that a large number of switches supported the workload concurrently. Comparatively, the HyperFlatNet is a more robust DCN topology so that there is less probability of link failures leading to server failure; only a few switches are disconnected during the LFR increasing process so that the workloads are efficiently transferred to the existing operating switches which results in the energy efficiency increases.

V CONCLUSION AND FUTURE WORK

DCNs performance is not merely a function of resource provisioning and allocation, but also it is a network-wide activity. We have revealed how the DCN QoS performance and robustness can be impacted by the underlying network structure in a cloud environment. Some topological metrics have been studied to reveal their impact on DCN performance and robustness. Then the robustness in the presence of an increase on LFR for different DCN architectures are evaluated which also shows their correlations to DCN performance. Under the same network settings, the topological metrics such as network diameter and ASPL can be rough indicators on network latency because they cannot comprehensively explain the complicated DCN structure. The conventional Three-tier DCN in this case generated the lowest avg. latency due to the amount of links deployed and because its tree architecture provides the highest efficiency on network transmission. Additionally, with the adoption of DVFS and a fixed timeout DPM policy, the energy consumption by different DCN architectures illustrates irregular patterns given that the failed links were selected as prominent to enlarge the consumption effects. Considering various network scenarios will further develop more realistic and complex DCN energy aware performance models and also better topological indicators should be sought.

REFERENCES

[1] S. Khan and A. Zomaya, *Handbook on data centers*.

[2] Y. Liu, J. Muppla, M. Veeraghavan, D. Lin and M. Hamdi, *Data Center Networks*. Springer International Publishing, 2013, pp. 1-5.

[3] Gartner.com, 'Technology Research | Gartner Inc., 2015. [Online].

[4] A. Banerjee, T. Mukherjee, G. Varsamopoulos and S. Gupta, 'Integrating cooling awareness with thermal aware workload placement for HPC data centers', *Sustainable Computing: Informatics and Systems*,

[5] J. Carter and K. Rajamani, 'Designing Energy-Efficient Servers and Data Centers', *Computer*, vol. 43, no. 7, pp. 76-78, 2010.

[6] D. Kliazovich, P. Bouvry and S. Khan, 'GreenCloud: a packet-level simulator of energy-aware cloud computing data centers', *The Journal of Supercomputing*, vol. 62, no. 3, pp. 1263-1283, 2010.

[7] Cisco System, Inc, 'Cisco Data Center Infrastructure 2.5 Design Guide', 2011.

[8] M. Al-Fares, A. Loukissas and A. Vahdat, 'A scalable, commodity data center network architecture', *SIGCOMM Comput. Commun.* vol. 38, pp. 63-74, 2008.

[9] C. Guo, G. Lu, D. Li, H. Wu, X. Zhang, Y. Shi, C. Tian, Y. Zhang and S. Lu, 'BCube', *SIGCOMM Rev.*, vol. 39, no. 4, p. 63, 2009.

[10] Z. Chkirbene, S. Foufou, M. Hamdi and R. Hamila, "HyperFlatnet: A Novel Network Architecture For Data Centers", in *Communication Workshop (ICCW), 2015 IEEE International Conference, London*, 2015, pp. 1877 - 1882.

[11] K. Choo, 'Mobile Cloud Storage Users', *IEEE Cloud Computing*, vol. 1, no. 3, pp. 20-23, 2014.

[12] D. Allan, P. Ashwood-Smith, N. Bragg and J. Farkas, 'Shortest path bridging: Efficient control of larger ethernet networks', *IEEE Communications Magazine*, vol. 48, no. 10, pp. 128-135, 2010.

[13] B. Li, Q. Zhang, J. Liu, C. Wang and K. Farkas, 'Advances in Wireless Mesh Networks', *Mobile Networks and Applications*, vol. 13, pp. 1-5, 2008.

[14] R. Goonatilake and R. Bachnak, 'Modeling Latency in a Network Distribution', *Network and Communication Technologies*, vol.1, 2012.

[15] Jayanthi, 'Improving the performance and reducing bit error rate on wireless deep fading environment receivers', *Journal of Computer Science*, vol. 10, no. 3, pp. 458-468, 2014.

[16] Li Shang, L. Peh and N. Jha, 'Power-efficient Interconnection Networks: Dynamic Voltage Scaling with Links', *IEEE Computer Architecture Letters*, vol. 1, no. 1, pp. 6-6, 2002.

[17] G. Chen, W. He, J. Liu, S. Nath, L. Rigas, L. Xiao and F. Zhao, "Energy-aware server provisioning and load dispatching for connection-intensive internet services.", 2008, pp. 337–350.

[18] X. Fan, W. Weber and L. Barroso, "Power provisioning for a warehouse-sized computer", *ACM SIGARCH Computer Architecture* vol. 35, p. 13, 2007.

[19] Yung-Hsiang Lu and G. De Micheli, "Comparing system level power management policies", *IEEE Design & Test of Computers*, vol. 18, pp. 10-19, 2001.

[20] R. Diestel, *Graph theory*. New York: Springer, 1997.

[21] A. W. Malik, K. Bilal, K. Aziz, D. Kliazovich, N. Ghani, S. U. Khan, R. Buyya, 'CloudNetSim++: A toolkit for Data Center Simulations in OMNeT++', *Proceedings of the 11th International Conference on High-capacity Optical Networks and Enabling/ Emerging Technologies, Dec*, 2014, Charlotte United States.