Green Data Center Networks: Challenges and Opportunities

Kashif Bilal, Samee U. Khan
Department of Electrical and Computer Engineering
North Dakota State University
Fargo ND, USA
{kashif.bilal, samee.khan}@ndsu.edu

Albert Y. Zomaya
School of Information Technologies
University of Sydney
Sydney, Australia
albert.zomaya@sydney.edu.au

Abstract— **The concerns about environmental impacts, electricity cost, and energy needs of data centers are mounting. These concerns have given significant impetus to consider energy efficiency as one of the major data center design parameters for researchers and IT executives. Tremendous growth in the size and computing demands of data centers mandate scalable, fault tolerant, and energy efficient network infrastructure within data centers. In the recent years, new Data Center Network (DCN) architectures have aptly proposed a viable solution to the issues posed by the legacy DCNs. New DCN architectures are required to be energy efficient and energy proportional, besides dealing with other identified design drawbacks, such as scalability, fault tolerance, end-to-end bandwidth, and non-agility. In this article, we will deliberate on the various DCN architectures, their inherent problems, and the potentials to achieve energy efficiency. We will also shed light on myriad techniques that may be suitable to make the DCN architectures more energy efficient and energy proportional. Finally, we will outline future directions.**

Keywords- data center networks; energy efficiency; energy efficicent networks; cloud computing

I. INTRODUCTION

For the last several decades, Quality of Service (QoS), reliability, and performance have been given preference over all other design issues in the Information and Communication Technology (ICT) sector [1]. The customers and government agencies are concerned about Green House Gases (GHG) emissions and environmental impacts of energy-thirsty data centers. Therefore, today's ICT executives are appropriately concerned about the energy efficiency within data centers, besides providing highly available and high performance ICT infrastructures. Nowadays, companies are expected to uphold the service level with much lower GHG emissions and carbon footprints. A scheme named "Carbon Reduction Commitment" (CRC) requires big companies to purchase annual "carbon allowance" for companies estimated energy usage, and deliver the required services within the CRC limits. Around 10,000 business firms received warning letters from the UK government that the companies could be affected by the CRC. Gartner now lists "Green IT" as the leading issue of the "top ten strategic trends and technologies" for ICT organizations [2].

The networking equipment accounts for around 15% of a data center's overall cyber energy budget [5]. In 2010, the communication infrastructure accounted for approximately 15.6 billion kWh of energy consumption within all data centers worldwide. The servers and cooling infrastructures are the major energy consumers within data centers. Therefore, energy efficiency of the servers and cooling equipment has received more consideration by the research community. As the servers and cooling infrastructures are becoming more energy efficient, it is expected that the power consumption share of the networking devices within a data center will increase. It has been projected that the energy consumption of the network component will increase by up to 50% within data centers [7]. Moreover, the bandwidth demands of new network applications are doubling every 18 months [4]. Furthermore, in the near future, one should expect major network upgrades to support emerging technologies, such as network virtualization and Virtual Desktop Infrastructure (VDI). Table 1 summarizes the electricity usage within data centers [5].

As our reliance on fossil fuels is expected to rise substantially in the coming decades [2], factors, such as: **(a)** environmental aspects, **(b)** high energy costs, **(c)** increased demand, and **(d)** economic consequences are driving the need to consider energy efficiency as one of the foremost data center design concerns [7]. The heat dissipation from the networking devices makes a significant contribution to the overall thermal signature of the data center. Therefore, energy efficient networking will also result in the reduced data center cooling costs. Energy efficient techniques are emerging rapidly for the networking devices, such as IEEE 802.3az [4]. However, a large base of legacy networking equipment will be continually used for many years. Therefore, there is a need for energy efficiency of the legacy networking equipment [10].

We can categorize the existing green networking techniques into: **(a)** consolidation, **(b)** selective connectedness, and **(c)** proportional computing [4]. The consolidation techniques exploit the network overprovisioning and path diversity to consolidate the network traffic on a subset of devices. For instance, idle networking devices may be turned off to save energy. The idle devices are transitioned to sleep modes transparently in the selective connectedness using the "interface proxying" technique. Proportional computing refers to the idea of the system consuming energy in proportion to its utilization.

TABLE I. ELECTRICITY USAGE WITHIN DATA CENTERS

Location	Year	Communication Infrastructure Electricity Usage (billion KWh)	Total Data Center Electricity Usage (billion KWh)	Total Data Center Electricity Usage (% of total worldwide electricity usage)
Worldwide	2000	3.8	70.8	0.53%
US	2000	1.4	28.2	0.82%
Worldwide	2005	7.3	152.5	0.97%
US	2005	2.7	56.0	1.53%
Worldwide	2010	15.6	271.8	1.50%
US	2010	4.9	85.6	2.20%

Such techniques can be mainly classified into: **(a)** Dynamic Voltage and Frequency Scaling (DVFS) and **(b)** Adaptive Link Rate (ALR). Typically, the DVFS techniques are applied to processors for energy savings. The ALR techniques are aimed to scale down the network link data rates to conserve energy. For example, the IEEE 802.3az Energy Efficient Ethernet (EEE) standard provides a mechanism for energy efficiency in the Ethernet using the ALR [4].

II. DCN ARCHITECTURES

The network infrastructure of a data center plays a pivotal role in ascertaining the performance factors and initial capital investment. The legacy DCN infrastructure inherently lacks the capability to meet the current growth trend and bandwidth demands. The legacy DCN architectures mainly suffer from: **(a)** energy-inefficiency, **(b)** poor scalability, **(c)** high cost, **(d)** low cross-section bandwidth, and **(e)** non-agility [6]. Many new DCN architectures, such as the fat-tree, DCell, VL2, BCube, flattened butterfly, and FiConn are proposed in response to the limitations posed by the legacy DCN architectures [6]. Moreover, various hybrid DCN architectures using amalgamation of optical and wireless technologies have been proposed in recent years [14]. The state of the art DCN architectures mainly focus on: **(a)** performance, **(b)** reliability, **(c)** fault tolerance, **(d)** high end-to-end bandwidth, and **(e)** agility. As a consequence, the new DCN architectures are often highly overprovisioned and underutilized [6], which triggers the inquest to design newer energy-efficient architectures and techniques.

Based on the network traffic routing model, we can classify the DCN architectures into two major categories: **(a)** switch-centric model, **(b)** hybrid model, and **(c)** server-centric model. The switch-centric models rely on the network switches to perform network communication and traffic routing. The hybrid models use an amalgamation of electrical network switches and optical or wireless devices to accomplish the network communication and traffic routing. The server-centric models place the routing and switching capability within the computational servers. Computational servers are used for packet forwarding and routing besides performing computational tasks [6].

The three-tier, fat-tree, VL2, Portland, and flattened butterfly DCN architectures are switch-centric. The three-tier DCN architecture is by far the most commonly used network model, in which the switches are arranged in three layers, namely: **(a)** access, **(b)** aggregation, and **(c)** core. Expensive and energy-inefficient high-end equipment is used at the higher layers (aggregation and core) of the topology. However, a high oversubscription ratio in the network leads to low end-to-end bandwidth and high network latency in the three-tier DCN architecture. Al-Fares *et al*. proposed the fat-tree DCN architecture using commodity network switches. To deliver 1:1 oversubscription ratio, the commodity network switches are arranged in a Clos topology within the fat-tree. The fat-tree architecture delivers high end-to-end bandwidth at a much lower cost and energy consumption [6]. Subsequently proposed successors of the fat-tree, such as VL2 and Portland, also use the Clos structure and commodity network equipment [7]. Kim *et al*. proposed the flattened butterfly DCN architecture using high radix routers [8]. The flattened butterfly delivers better end-to-end bandwidth and path diversity than the baseline butterfly network.

Recursively defined DCN architectures, such as DCell, BCube, and FiConn use the server-centric routing model [6]. In the DCell architecture, the whole network is composed of a hierarchy of cells called *dcells*. Higher level *dcells* are built from lower level *dcells*. The $dcell_0$ is composed of *n* servers and a commodity network switch. $dcell_0$ provides the foundation of the DCell architecture. Each server in the DCell architecture has multiple network interfaces, which are used to connect severs from one *dcell* with corresponding servers in other *dcells*. Each $dcell_{l-1}$ is connected to all of the other $dcells_{l-1}$ within the same $dcell_l$. The DCell delivers a highly scalable architecture [6]. A simulation of the DCell and fat-tree DCN is shown in Fig. 1. We will deliberate on various inherent problems and potentials to achieve energy efficiency in the state of the art DCNs in the ensuing section.

Optical and hybrid (optical + electrical) DCN architectures are proposed recently to overcome numerous challenges posed by the legacy electrical-only DCNs. Various all-optical interconnects, such as Data center Optical Switch (DOS), Petabit optical switch, E-Rapid, IRIS, and Data vortex, have been proposed recently [14]. Optical DCN architectures, such as Proteus, wield Optical Wavelength Switching (OWS) to reconfigure network topology on the fly. Hybrid interconnects, such as c-Through, Helios, and

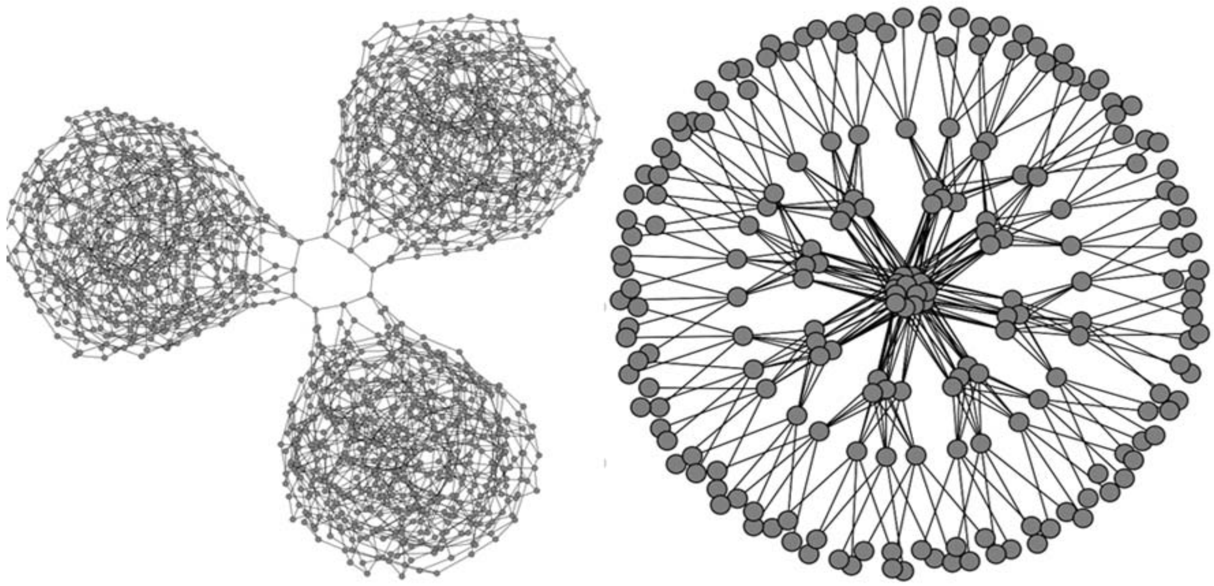

Figure 1: Simulation of DCell (left) and fat-tree (right) DCN architectures in ns-3

HyPac have been proposed to supplement the DCN using optical interconnects [14].

With the advent of 60 GHz wireless technology, numerous wireless and hybrid-wireless (wireless + electrical) DCNs are proposed recently. Wireless connectivity offers cost-effective enrichments in existing DCN infrastructure with little manual intervention. 60 GHz technology exhibits promising opportunities for DCNs by exploiting 2D and 3D beamforming, directional antennas, signal reflection, and frequency reuse. Shin *et al*. and Vardhan *et al*. proposed completely wireless DCN using 60 GHz technology. Various hybrid DCNs are proposed to augment the existing DCN by placing one or more wireless devices on servers and Top of Rack (ToR) switches. The IEEE P802.11 task group is working to standardize 60 GHz wireless technology with IEEE 802.11ad standard [15].

III. DCN ARCHITECTURES: ISSUES, SOLUTIONS, AND POTENTIALS

Servers are interconnected to each other in the recursively defined server-centric DCN architectures. Such DCN architectures do not use multiple layers of network switches. Therefore, the network links interconnecting the cells experience high oversubscription ratio (up to 256:1 for 4096 nodes) [6]. The recursively defined DCNs exhibit strong reliance on the network size. The results of our simulation analysis presented in Fig. 2 show that the network throughput and inter-node bandwidth are inversely proportional to the network size in the DCell. Moreover, the servers also perform the additional task of traffic processing and routing within the server-centric routing model, which usually requires a dedicated processor/core. Furthermore, the routing schemes used in the recursive DCN architectures are usually not based on the shortest path routing. The path between the source and destination may possess additional intermediate hops, which results in higher packet delays and increased link utilization. We have simulated the DCell architecture with the DCell's customized routing and shortest path routing schemes. Our simulation results, presented in Fig. 3, demonstrate that the shortest path routing outperforms the DCell's routing in terms of network throughput and average packet delay. Because servers are used to inter-connect cells, the idle servers cannot be placed into sleep state. Therefore, Dynamic Power Management (DPM) techniques for energy savings are not feasible for the recursive DCNs, resulting in continuously full energy consumption despite being idle. Details of the simulation analysis can be seen in [6].

The server-centric DCNs inherently save energy that is used in the switch-centric DCNs by network switches. The network traffic flows may be managed by applying load-balancing techniques to overcome the network congestion problem. The DCell architecture also exhibits path diversity from source to destination. Network flow based adaptive routing protocols may exploit the path diversity to select the best and most conducive path from the source to destination.

The switch-centered DCNs exhibit better candidature for energy efficiency, because of high path overprovisioning. One of the major drawbacks of the contemporary switch-centric networks is the use of a large number of network switches to ensure 1:1 oversubscription ratio. For example, the 8-pod (128 nodes) fat-tree network requires 80 network switches [6]. There is a great deal of overprovisioning and path diversity in the switch-centric networks. Moreover, the average link utilization of network links is reported around 5% - 25% [9]. Therefore, underutilization of the links and path diversity may be exploited for energy efficiency. For instance, ALR based techniques can be very conveniently applied to save energy. Moreover, end-to-end path diversity

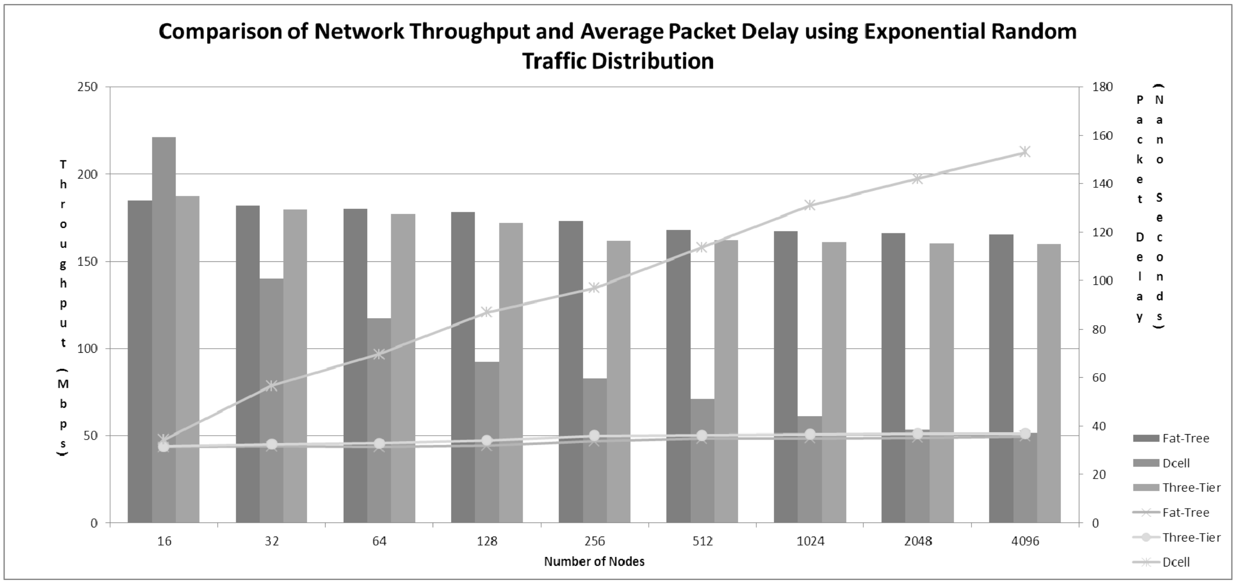

Figure 2: Throughput and average packet delay of DCNs using exponential random traffic distribution

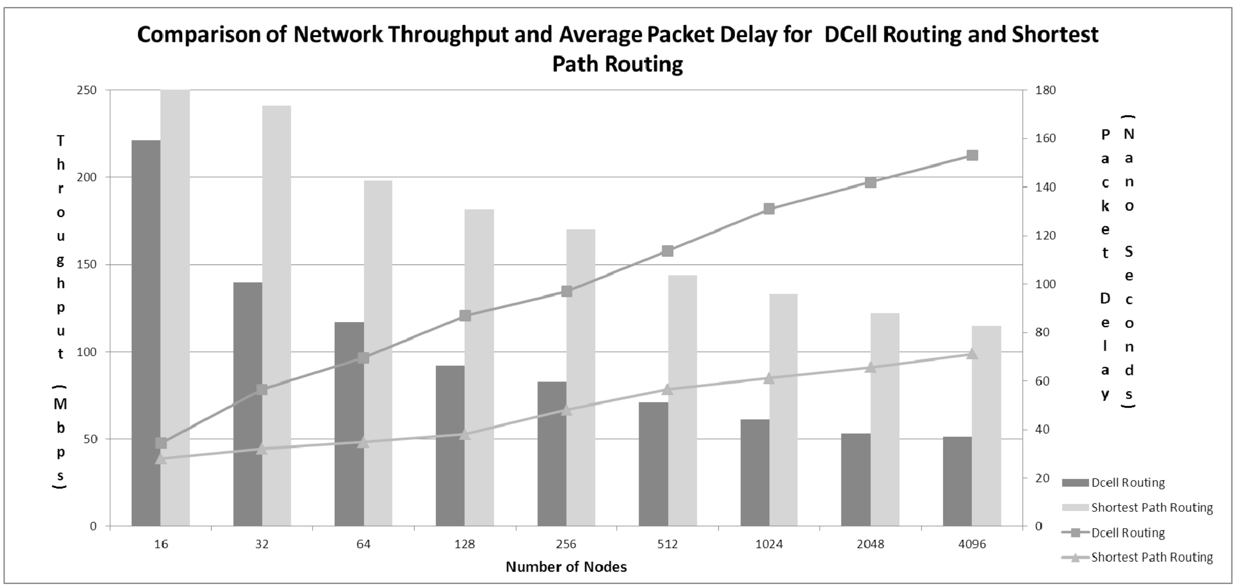

Figure 3: Throughput and average packet delay comparison of DCell routing and shortest path routing using exponential random traffic distribution

offers an opportunity for network traffic consolidation and re-routing on a subset of links and devices. Therefore, the remaining idle devices may be transitioned to low power sleep mode. Furthermore, IEEE 802.3az EEE may be employed to increase energy savings in amalgamation with other energy efficiency techniques.

Hybrid (electrical + optical/wireless) DCNs offer solutions to various DCN problems. Wireless connectivity offers a feasible solution to extend existing DCN infrastructure eliminating cabling cost, complexity, and installation. Wireless links can be created among server and racks on the fly, reducing the network load on the core network. Energy efficiency is one of the foremost design requirements for 60 GHz technology, resulting in energy efficient 60 GHz devices and technology. Wireless interconnects can be exploited to migrate the traffic from underutilized network devices to wireless links to place the idle devices in sleep mode for energy saving. Optical interconnects offer higher port density and bandwidth at considerably low energy consumption. Estimated energy consumption of an optical transmitter is around 0.5nJ/bit, whereas the energy consumed by a high-end electrical router is around 20nJ/bit. Consequently, overall energy consumption can be reduced significantly. A complete optical DCN is estimated to deliver around 75% energy savings. Hybrid DCNs can be used to relieve the hotspots (congested links) in the DCN. Elephant flows and high fan in/out traffic are considered the chief reasons for the network congestion and performance deprivation. 60 GHz wireless flyways and optical circuit switched paths offer load balancing opportunities by offloading elephant flows from electrical switches to reduce network load and congestion.

IV. Green Networking in Data Centers

Green networking research follows two major trajectories: **(a)** workload consolidation and **(b)** scaling down the communication link data rate (based on ALR). Workload consolidation and traffic redirection are used to utilize a subset of network links and devices to serve all of the network traffic flows. Consequently, a subset of idle network links and devices may be transitioned to the sleep or low power idle mode to save energy. Transitioning from sleep to active mode is time consuming and may result in performance degradation. Therefore, another class of network energy optimization based on ALR is used to scale down the data rate of network equipment for underutilized devices. The ALR scales down the data rate of a part of the network equipment, so the energy savings are less as compared to turning off the equipment. The ALR delivers energy savings without affecting the network state and operations. A detailed discussion of energy savings techniques using the ALR is presented in [4].

Mahadevan *et al.* optimized network energy consumption by consolidating: **(a)** server workload and **(b)** network traffic, and reported around 74% energy savings [10]. Abts *et al.* used network traffic prediction to scale down the network link data rate to obtain energy savings of 30% - 45% at 25% link load [9]. Carrega *et al*. discussed the energy saving in DCNs using the network traffic aggregation [9]. The authors proposed to merge the traffic from multiple network links so that fewer links are utilized and idle links may be transitioned to sleep. The authors reported 22% energy savings for link loads of 50% [9]. ElasticTree, a network level energy optimizer used the overprovisioning and idleness of network links and devices in the fat-tree topology. The optimizer module finds a subset of links and devices to serve traffic, and powers down the remaining idle devices for energy savings. The authors estimated up to 50% energy savings using the ElasticTree. Zhang *et al.* proposed an energy-aware traffic engineering scheme to redirect network traffic to underutilized links. The authors reported energy savings of 27% - 42% when the maximum link load is less than 50%. Thu *et al*. optimized network energy consumption by designing a network control system named Ecodane. The authors implemented topology-aware heuristics and demonstrated energy savings up to 30%. Bolla *et al*. exploited: **(a)** resource virtualization and **(b)** modular architecture of network devices to report 51% energy savings. The aforementioned is achieved by transitioning the underutilized network components into sleep mode and migrate the virtual processes to the active components of the network. Chiaraviglio et al. discussed heuristics to turn-off links and devices while preserving the QoS and connectivity constraints. Simulation results showed up to 24% of energy savings [4]. The detailed discussion of energy efficiency in data centers can be seen in [16].

V. Green DCN Challenges and Future Directions

New DCN architectures are required to handle the increasing GHG emissions produced by the ICT sector. DCNs typically experience an average load of not more than 25% of the peak load. Moreover, around 70% of the time, a considerable number of the links remain idle within data centers [13]. However, the links do not remain idle constantly for long periods of time. Benson *et al*. analyzed data center traffic over a period of ten days and observed that the set of idle links continuously varied for the entire time period. Moreover, it also was observed that 80% of the links remained idle only for 0.002% of the time [13]. Therefore, it is important to consider the traffic characteristic within a data center prior to applying ALR or other energy saving techniques.

Overprovisioning and underutilization of links and devices enable opportunity for energy efficiency techniques. However, due consideration is required to be given to the QoS and performance constraints. Performance degradation and increased latency may result in substantial revenue loss. Google reported 20% revenue loss because of an experiment that added an extra delay of 500ms in displaying the search results. Amazon experienced 1% sales decrease because of 100ms additional delay [11]. Therefore, green networking initiatives must be reliable and ensure required performance and QoS constraints.

Hybrid DCNs offer promising opportunities to DCNs. However, hybrids DCNs are in their infancy and are facing numerous challenges. 60 GHz wireless technology is limited by line of sight, short range, propagation loss, and signal attenuation. 60 GHz technology poses serious challenges in transceiver positioning, beam forming, interference due to power leaks, and signal reflection in densely populated data centers. Similarly, optical interconnects also experience numerous challenges, such as cost, scalability, link setup, switching time, and insertion loss. The wavelength switching time for commercially available optical switches is around 10 to 25ms. Moreover, hybrid networks lack sufficient efficacy for DCNs multi-tenant based mixed and heterogeneous workloads. Various hybrid DCN architectures make stringent overlay assumptions, such as that: **(a)** flows are independent, **(b)** flows do not have priority, and **(c)** random hashing for flow distribution is effective. However, in practice such assumptions do not hold true for the DCN traffic patterns. Hybrid interconnects promise significant network upgrades. Aforementioned are some of the numerous unresolved challenges that pose a barrier in adopting hybrid technologies in data centers.

Network protocols may be optimized or re-designed for enhanced performance and energy efficiency. Network-aware and energy-aware adaptive routing protocols are needed for better performance, high link utilization, and traffic consolidation and redirection. Moreover, many network services remain active to ratify their availability to periodic heartbeat messages or network chatter. The "interface proxying" techniques may be used to transparently transition such services to sleep without affecting the network operations [4]. The EEE is a promising energy efficient technology but is still in its infancy. Efficient and reliable ALR policies are required to be designed for the IEEE 802.3az EEE.

There are very little details available on characteristics of the data center traffic [11]. There is presumably no network workload generator at hand, which may generate data center traffic for various scenarios, such as one-to-one, all-to-all, and one-to-all, and for data intensive, computational intensive, and mixed workloads. A realistic data center traffic generator will substantially help the research community to analyze DCN under various scenarios, and tune the DCN for energy efficiency and reliability.

Energy efficiency of network equipment has not increased following the Dennard's law and current network equipment is not energy-proportional. Energy consumed by network devices in idle state is around 80% - 90% of energy consumed in peak load [12]. Energy proportional network devices are required to be designed, and can save enormous amount of wasted energy.

The DCN is one of the most significant data center components wielding a marked impact on initial capital investment and performance parameters. The state of the art DCN are implemented at a very small scale and tested under non-realistic data center traffic [6]. There are very less comparative studies for DCNs [6], and presumably no comparative study of different DCN architectures under realistic traffic conditions. Different DCN comparative studies under realistic workloads are required to highlight the DCN drawbacks and future research for enhancement.

VI. Conclusions

This paper presented an overview of the major challenges faced by DCNs. We discussed the need and potentials to achieve energy efficiency within the network portion of a data center. Moreover, we discussed the inherent problems and challenges related to energy efficiency of DCNs. As the energy efficiency efforts for data centers are predominantly targeted towards energy-efficient servers and data center cooling, the networking component in data centers has not yet received the sufficient attention. Therefore, the energy consumption share of the network portion is expected to increase. The ever-increasing share of the network energy consumption in data centers needs concerted and dedicated efforts to conserve significant amount of energy by exploiting overprovisioning and idleness of the network equipment. Several streams of research thrusts may further emerge by resolving the issues highlighted in this paper.

References

[1] S. Murugesan, "Harnessing Green IT: Principles and Practices," *IEEE IT Professional*, Vol. 48, No. 1, 2008, pp. 24–33.

[2] Energy efficiency in the data center, Hewlett-Packard, Sep. 2008.

[3] B. Raghavan, D. Irwin, J. Albrecht, J. Ma, and A. Streed, "An Intermittent Energy Internet Architecture," In *proceedings of the Third International Conference on Future Energy Systems (e-Energy '12)*, Madrid, Spain, May 2012.

[4] K. Bilal, S.U. Khan, J. Kolodziej, L. Zhang, K. Hayat, S.A. Madani, N. Min-Allah, L. Wang, and D. Chen, "A Survey on Green Communications using Adaptive Link Rate," *Cluster Computing*, (DOI 10.1007/s10586-012-0225-8).

[5] J. Koomey, Growth in Data Center Electricity Use 2005 to 2010, Analytics Press, 2011.

[6] K. Bilal, S.U. Khan, L. Zhang, H. Li, K. Hayat, S.A. Madani, N. Min-Allah, L. Wang, and D. Chen, "Quantitative Comparisons of the State of the Art Data Center Architectures," *Concurrency and Computation: Practice and Experience,* (DOI: 10.1002/cpe.2963).

[7] D. Abts, M. Marty, P. Wells, P. Klausler, and H. Liu, "Energy Proportional Datacenter Networks," In *Proceedings of the 37th International Symposium on Computer Architecture,* Saint Malo, France, June 2010, pp. 338–347.

[8] J. Kim, W. Dally, and D. Abts, "Flattened Butterfly: a Cost-Efficient Topology for High-Radix Networks," In *Proceedings of the 34th International Symposium on Computer Architecture*, San Diego, CA, June 2007, pp. 126–137.

[9] A. Carrega, S. Singh, R. Bruschi, and R. Bolla, "Traffic Merging for Energy-Efficient Datacenter Networks," *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS '12)*, Genoa, Italy, July 8–11, 2012.

[10] P. Mahadevan, S. Banerjee, P. Sharma, A. Shah, and P. Ranganathan, "On energy efficiency for enterprise and data center networks," *IEEE Communication Magazine*, Vol. 49 No. 8, 2011, pp. 94–100.

[11] A. Greenberg, J. Hamilton, D. Maltz, and P. Patel, "The Cost of a Cloud: Research Problems in Data Center Networks," *ACM SIGCOMM Computer Communication Review,* Vol. 39, No. 1, Jan. 2009, pp. 68–79.

[12] L. Niccolini, G. Iannaccone, S. Ratnasamy, J. Chandrashekar, and L. Rizzo, "Building a Power Proportional Router," *Usenix ATC '12*, Boston, June 2012.

[13] T. Benson, A. Anand, A. Akella, and M. Zhang, "Understanding Data Center Traffic Characteristics," *SIGCOMM Computer Communication Review*, Vol. 40, No. 1, 2010, pp. 92–99.

[14] C. Kachris and L. Tomkos, "A Survey on Optical Interconnects for Data Centers," *Communications Surveys & Tutorials, IEEE* , Vol. 14, No.4, 2012, pp.1021-1036

[15] X. Zhou, Z. Zhang, Y. Zhu, Y. Li, S. Kumar, and A. Vahdat, "Mirror Mirror On The Ceiling: Flexible Wireless Links For Data Centers," *ACM SIGCOMM Computer Communication Review*, Vol. 42, No. 4, pp. 443-454, 2012.

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