Data Center Architectures: Challenges and Opportunities

Rana W. Alaskar, and Imtiaz Ahmad
Computer Engineering Department
College of Computing Sciences and Engineering
Kuwait University, P. O. Box 5969, Safat 13060, Kuwait
rana.alaskar@hotmail.com, imtiaz.ahmad@ku.edu.kw

ABSTRACT

Recently, data centers became a main part in information and communication technology. Data centers are facing many challenges because of the rapid and continuous growth of applications' size and complexity. In this paper, we present an overview of data center architectures challenges and their proposed solutions related to networking equipment, information technology equipment, power conservation equipment as well as cooling systems. We also mentioned how software defined network, a promising technology, can enhance the performance of data centers, and what challenges are faced while applying such technology to existing data centers.

KEYWORDS

Data centers; challenges; network equipment; routing; security; servers; software defined networks; power consumption; cooling systems.

1 INTRODUCTION

A data center is a centralized protected data repository which contains access networks, storage systems, cooling systems, applications, and power distributing infrastructure. It also contains web, application, database, and streaming servers. Providing scalable, flexible, available, reliable and secure cloud computing services and maintaining organizations' data are the major goals of data centers. Data centers' size and complexity are increasing enormously, and they continue to grow day after day. Many conventional approaches, design methodologies and tools which are applied in current data centers are not applicable to the dynamic environment, which raises many problems related to the scalability, availability as well as the security, leading lowering the throughput of data centers. Thus, improving existing techniques that are applied in data centers became an active research topic in the information technology field.

Significant research work has been done to address various data center challenges, and to improve the current techniques in order to meet customers' requirements. Kant [1] discusses networking, management, and power challenges that are faced by the telecommunication and data center industries. The author considers, in his study, the layered model of virtualized data centers. Garimella et al. [2] present particularly the thermal management challenges, and focused on the liquid cooling technology which can improve chip performance. Optical interconnection networks technology is considered the promising solution to handle the increase of network traffic, because commodity switches, which are prevalent in data centers and consume an enormous amount of power, could not deal with that increase efficiently. Kachris et al. [3] outlined how optical interconnection networks technology can provide a high throughput, with minimum power consumption. In addition, they mentioned the challenges related to applying such technology to the current data centers. Although virtualization can reduce the consumed power significantly, it leads to new challenges related to scalability, failure tolerance, and addressing schemes (i.e. networking challenge) [4]. Since software defined network is considered as the future technology of networks, many researchers pinpointed issues that may be faced while applying it in data centers. Lara et al. [5] presented how Openflow, which has been deployed by Google, can enhance data centers virtualization. Our main motivation is to present an overview of data centers' challenges

from different perspectives, and research efforts which handle those challenges efficiently, and improve data centers' throughput, availability, reliability, and security.

In order to overcome issues and challenges of the existing techniques and architectures, the following requirements should be met. First, a new technique, topology, or architecture for data centers should be independent of any factor such as the number of servers, the protocols, the location or the applications. Having a dependent techniques will affect the overall performance of the data center and will restrict dealing with any new technologies. Second, it is important to have a technique that provides powerful security mechanisms to deal with all kinds of attacks (e.g., eavesdropping, unauthorized access, denial of service, etc.) and protect data centers from them. Failing to have such a secure technique will negatively affect the availability of data centers. Third, data center location, the applied topology, and the scheduling approaches must be taken into consideration while proposing a new technique for data centers, because the mentioned factors highly affect the operation cost of the data center. Without the slightest doubt, energy consumed by information technology equipment, networking equipment, as well as power conservation equipment costs more than the aforementioned elements. It is essential to have different techniques that can intelligently minimize the total consumed energy.

The rest of the paper is structured as follows: Section II describes different data center challenges, and it includes subsections that discuss; traffic engineering, multicasting, security, servers, network and routing, software defined network, power and consumed energy, cooling systems, construction, as well as simulation challenges. Section III concludes the paper.

2 DATA CENTERS CHALLENGES

Since the number of servers that are placed in data centers, such as those owned by Google, Amazon, Facebook and Microsoft, keep increasing exponentially, connecting those servers becomes a big challenge, and raises many other challenges related to:

- Network equipment: It is used simply for communication purposes. Because of the dynamic nature of new communication technologies, conventional techniques for traffic engineering, multicasting, security and routing are not applicable anymore for large scale data centers. Software defined network also raises multiple challenges for existing data centers.

- Servers: They are used mainly for data processing, and to run workload. The best technology that can enhance servers' performance and reduce the power consumed by them is virtualization. It provides efficient utilization, and facilitates applications sharing.

- Power conservation equipment: It maintains the desirable high quality power and the proper temperature. The quality of that equipment is very important to ensure that all servers contained in the data center are working well.

- Cooling systems: Heat-generating computer equipment need to be cooled using efficient cooling systems which operate different thermal control schemes. Electricity price is affected by the geographical location of the data center. So, choosing an area with a suitable climate will decrease cooling systems cost.

In this section, various challenges of data centers from different perspectives are discussed; the causes and effects of these challenges, as well as the schemes and techniques that have been proposed by researchers to address them.

2.1 Traffic Engineering Challenges

Traffic engineering is a method of optimizing reliability, throughput and latency that are required for complex routing schemes in modern data centers [6]. It plays a major role in decreasing the congestion in data centers. As mentioned in [6], multiple mechanisms are proposed for such a kind

of optimization. One of these mechanisms is described in [7]. In order to utilize network resources efficiently, traffic in multi-stage switch topologies must be scheduled. Thus, AL-Fars *et al.* [7] present a scalable dynamic flow scheduling system named Hedera. This system is designed for multi-stage switch topologies like VL2 [8] and PortLand [9]. The key idea of the proposed system is based on global knowledge of active flows to provide an efficient distribution of massive flows on different paths.

Another mechanism to provide the required optimization is presented by Xi *et al.* [10]. Equal-Cost Multi-Path (ECMP) routing scheme is recently considered poor and ineffective, because of its inability to ensure optimal resource utilization and traffic distribution between various paths. As a result, an ECMP complement scheme named Probe and RerOute based on ECMP (PROBE) is proposed [10]. It works as follows: setting ECMP as a default routing scheme, and in case of congestion in any link, big flows are rerouted to other alternative paths. Through this way, the congestion that usually occurs because of the absence of load checking before path assignment in ECMP will be handled in an efficient way. The most interesting thing in this scheme is that it can be applied to current data centers because it is designed for commodity switches and routers.

Chen *et al.* [6] determine a set of principles that should be considered when trying to solve traffic engineering problems in data centers. The first one is reliability. It is essential for both providers and subscribers because data providers depend on information stored in data centers to run different operations and to deliver services to subscribers; thus a redundant power lines and emergency power generators should be available permanently [11]. The second one is load balancing. Traffic should be distributed in an optimal and efficient way between links in data centers, and the proposed traffic engineering design should improve link utilization greatly. The last one is energy efficiency. Idle links should be turned off intelligently in order to save energy. This will decrease consumed energy, especially for green data centers [6].

2.2 Multicast Routing Challenges

Multicasting is the process of sending packets simultaneously to multiple receivers, so it provides a group communication for data centers. Group communication is highly required in data centers because it helps increasing applications' throughput, saving networks' bandwidth and reducing traffic loads. Low-end switches in data centers are facing difficulties in holding routing entries of huge multicast groups. As mentioned in [12], ideal access switches are not able to handle more than 1500 multicast group states. Also, link density, in most data centers topologies, is very high, meaning that the conventional multicast routing schemes and current multicast protocols are not appropriate for modern data center networks (DCNs), because they may lead to wasting link resources. The mentioned challenges motivated Li *et al.* [13] to propose Efficient and Scalable Multicast (ESM) routing scheme. The key idea is combining both in-packet Bloom Filters and in-switch entries in order to get scalable multicast routing in data centers. This new network-level multicast routing scheme helps to decrease the number of links used in the multicast tree and enhance the throughput greatly.

Li *et al.* pinpoint challenges that should be addressed in order to achieve a reliable multicast approach for data centers [14]. First, both node and link failure must be handled, because if the broken node or link is involved in the multicast tree, packets transmission will be paused. Second, because of the unpredictable nature of traffic in data centers, congestion can occur frequently in multicast tree. Third, low-end commodity switches which have small buffer space are not suitable for modern data centers because they only can perform basic multicast packet forwarding that offer limited reliability. As a result, the authors designed a reliable multicast approach for data centers named Reliable Data Center Multicast (RDCM). Its idea is based on using rich link resources to reduce the effect of packet loss on multicast performance [14].

Limited forwarding table memory space in commodity switches, which are used widely in

data centers, has become a challenge for group communication improvement in data centers. To handle such an issue, Multi-class Bloom Filter (MBF) is proposed in [15]. The idea of this new approach which is an extension for the standard Bloom Filters is based on determining the number of hash functions in a pre-element level and this number is calculated by a low complexity algorithm. MFB decreases multicast traffic leakage more than the standard Bloom Filter.

2.3 Security Challenges

Security is very essential in data centers which house tens of thousands of servers. It is very important to avoid different types of threats such as DoS, unauthorized use of resources and data theft or alteration. Different approaches are used to provide a high level of security in data centers, but conventional approaches are complicated and have different limitations. Lack of scalability, flexibility and resilience are the limitations of these conventional approaches. To handle the mentioned issues, a new scalable, flexible and cost-effective scheme named Hybrid Security Architecture (HSA) is proposed [16]. In HSA, middleboxes (e.g., firewalls and Intrusion Detection Systems (IDS)) can be placed anywhere in data centers, in contrast with conventional approaches which restrict the distribution of middleboxes in data centers. Additionally, HSA decouples routing and security services to increase the simplicity of operations, thus topology and routing changes have minimal effects on security services. In order to save costs, HSA combines hardware and software security devices; and because of this combination, the name of this approach is Hybrid SA. Moreover, this new approach does not require modifications to existing switches and routers when applied in current data centers [16].

Xue *et al.* present Policy Aware Data center network (PAD) strategy [17] to handle middleboxes deployment problem in data center networks and to make middleboxes traversal flexible and reliable. Middleboxes, which are responsible for applications' protection, are usually deployed in two fashions, either deployed in-path

or one-arm. In order to achieve flexible and scalable deployment, PAD uses two principles:

- Packet determines the path that it will follow because system policy may require a packet to pass through many middleboxes in its way to the destination server, which requires complicated configurations.

- Policies of packets should be separated from routing. This makes switch configuration easier, because only edge switch will concern the policy and all other switches will try to deliver packets without considering policies. Because PAD works with various types of DCNs (i.e. tree architectures), it is considered as flexible strategy [17].

Defense mechanisms in data centers are either cyber (firewalls) or physical (sensors), but usually not both. Szefer *et al.* present cyber-physical security defense for data centers against physical attacks [18]. The new security system aims to combine both cyber defenses and physical attack-detection sensors in order to provide protection for customers' data stored in data centers. Its idea is based on the time window between attack detection and attack occurrence. The system consists of three components: the physical sensors and monitors which collect different information from sensors, management infrastructure software which receives input from physical sensors and monitors, and the actual computing and networking nodes which perform defensive measures. The authors stipulate that this combination provides better dynamic security in data centers [18].

2.4 Servers' Challenges

The primary component in the data center is a server. A collection of servers that are connected to each other through an Ethernet switch is called a rack. Multiple racks are connected by a data center switch within a cluster. This architecture is called cluster architecture and it's prevalent in data centers. Because of the rapid growth of the size of applications run on data centers' servers, conventional cluster architecture which is used in

most data centers is considered as a challenge. Cluster architecture does not support resource sharing among various nodes, therefore it does not meet the previously mentioned requirements, and it downgrades the overall performance. Using traditional approaches for resources sharing is costly because of the massive number of servers existing in data centers. This encouraged Hou *et al.* to propose a flexible and cost-effective system [19] that handles the problem of memory capacity in big applications by allowing a node (i.e. server) to use other nodes resources (e.g., memory, NIC, I/O devices) as long as these resources are idle. For example, the memory can be shared among different nodes by either a simple load/store instructions, or Direct Memory Access (DMA) engine. NIC can be shared via a PCIe switch, and the tests show a great performance of this approach. It provides average 95% bandwidth [19].

Distributing loads among servers is one of the most important issues in data centers, and optimizing load distribution can lead to a great improvement in data center performance. Using multi-core processors affect positively the consumed energy, multicasting, as well as networking applications, and can provide the required load distribution balance. Cao *et al.* discussed load distribution for multiple heterogeneous multi-core server processors which are used widely in large-scale data centers [20]. They provide optimal server speed setting for two different core speed models; the idle speed model (i.e. speed = zero) and the constant speed model.

Since two-thirds of the total IT energy is consumed by servers [21], it's strongly desirable to have virtualized servers in data centers. Applying virtualization to servers means that collecting different applications running on different physical servers into one server virtually. This technology will increase the number of idle servers that can be turned off, and accordingly, the consumed energy will be reduced. Besides, virtualization can decrease both capital and operational costs. A study [22] shows that 20% of consumed energy can be saved by applying virtualization. Y. Jen et al. [23] provides a brief comparison between two different virtualization models applied in data centers, Xen and KVM, which are used particularly for CPU virtualization. The scheduler

used in the former allocates credits to each virtualized CPU, while the latter treat every guest as a normal thread. Both of aforementioned schedulers are used to balance global load on multi-cores and to get a good resource allocation.

2.5 Network and Routing Challenges

Data centers contain a huge number of servers and Ethernet switches which need to communicate with each other in order to provide the required services. Servers and switches can be connected in different ways (i.e. different topologies). As mentioned in [24], network topologies used in DCNs can be either fixed or flexible. The former means the connections between devices cannot be modified after the network is deployed, while the latter means the connections can be changed and modified. Examples of the fixed architecture include BCube [25] and DCell [26]. OSA [27] and Helios [28] are examples of the flexible architecture.

Each device in the data center should have a unique name or IP address. The process of assigning IP address to each device is called address configuration, and it is considered a challenge in the data centers because of two reasons. First, locality and topology information should be embedded in the addresses, and this requirement cannot be provided by DHCP, which is used in the IP network to configure devices. Second, a manual configuration increases the probability of errors. Research shows that the major reason behind data centers' failure is human errors. Thus, automatic address configuration is highly required in data centers.

Data centers can suffer from three different types of malfunctions. The first is node failure which occurs when the device (server or switch) is disconnected from the network because of hardware failure. The second is link failure which occurs when the connection between two devices (servers or switches) is lost because of network card failure. The third is mis-wiring which occurs when a particular wire is not connected to the proper device. Node failure causes a halt in the address autoconfiguration process, because there is

no ability to reach a non-working device and autoconfigure it. On the other hand, link failures and mis-wirings in Server-centric Structures do not affect the availability of devices that are connected to a failed or mis-wired link, so there is an ability to configure any device if and only if it is still connected through at least one link to the network.

Server-centric Structures' properties will be helpful to handle address autoconfiguration in the data centers with the existence of link failures and mis-wirings. In Server-centric Structures, such as BCube and Dcell, the degree of all nodes is greater than one, meaning that if there is a link failure or mis-wiring, we can still reach devices by other links. By using the mentioned property, the required goal which is maximizing the number of autoconfigured devices with malfunctions existence will be achieved. Data center Address Configuration system (DAC) is a generic and automatic address configuration system proposed by Chen *et al.* [29], which uses DCN's blueprint and physical network topology graphs to autoconfigure DCN devices. Basically, this system performs two main functions: device-to-logical ID mapping and malfunction detection.

The inputs of DAC are two graphs: the blueprint which provides the system with the connections between all servers and switches in the data center (Fig. 1), and physical network topology graph which is generated by following interconnections present in the blueprint (Fig. 2). Because of the iterative or recursive nature of data centers, the blueprint can be simply generated by software.
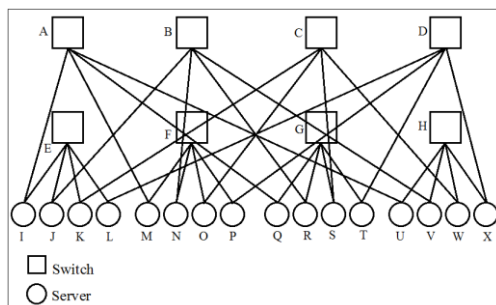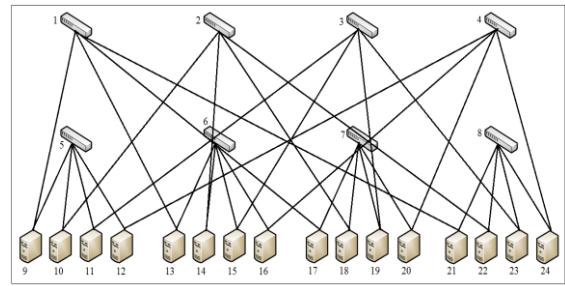

Figure 1. The blueprint

Figure 2. The physical network topology graph

DAC is one of the first successful attempts at aotuconfiguring generic data center networks. In spite of the great performance of DAC in address autoconfiguration, it has multiple limitations. First, DAC halts all its operations when malfunctions are detected, and cannot resume until all malfunctions are resolved manually by the administrator which can lead to substantial delay. Second, DAC is designed to be generic and suitable for all well-known structures, it may not be applied for new structures of data centers, which means we need a system that can work with arbitrary structures.

Error Tolerant Address Configuration (ETAC) is an optimized version of DAC, proposed by Ma *et al.* in [30]. Unlike DAC, ETAC can intelligently autoconfigure DCNs if malfunctions exist. The ETAC's idea is based on generating device graph by removing devices that are involved in the malfunctions from physical network topology graph, and then performing the mapping between the blueprint and device graphs. The inputs of ETAC are two graphs: the blueprint and the physical network topology graph. If a malfunction is detected during the mapping between two graphs, ETAC automatically generates a new graph called device graph which only contains devices that are not included in the malfunction. Once the device graph is generated, mapping between non-malfunction devices and logical IDs in the blueprint can proceed.

Although ETAC is intelligent in autoconfiguring data centers when malfunctions are present, it has some weaknesses. First, ETAC removes all devices included in the malfunction regardless of malfunction type when mapping is performed. There is no need to remove devices that are connected to a failed or mis-wired link. Second, there is no indication that ETAC can

locate malfunctions, unlike DAC which provides the administrator with all malfunctions and their locations (i.e. MAC address) in the data center. This means that ETAC needs human intervention to locate malfunctions, and this contradicts with the requirements of a good autoconfiguration system (i.e. minimal human intervention).
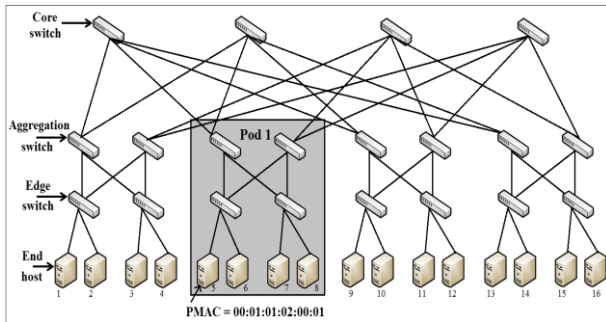

Figure 3. The PortLand structure

PortLand [9] is a data center network virtualization architecture, and a switch-centric structure which provides forwarding and routing for a data center that is organized as a multi-rooted tree [24]. It contains core, aggregate and edge switches as illustrated in Fig. 3. End hosts are connected directly to the edge switches. Each host has an Actual MAC (AMAC) address and a Pseudo MAC (PMAC) address. PMAC address encodes the topology information of the end host. The form of a PMAC is *pod.position.port.vmid*; where pod is the pod number of an edge switch, position is the pod's position, port is the port number of the switch that the end host is connected to, and vmid presents the virtual machines in the same physical machine. For example, end host # 5 is the first host in pod 1, so its PMAC address is 00:01:01:02:00:01 (see Fig. 3).

Although PortLand addresses multiple issues related to the VMs' population scalability, migration and management, it has limitations such as lack of bandwidth guarantees and the requirement of a multi-rooted tree physical topology [31].

2.6 Software Defined Network Challenges

Software defined network (SDN) is a new networking approach that is based on decoupling hardware forwarding from controller decisions.

The two main components of SDN are: 1) Software-based controller which controls network elements remotely. 2) Packet forwarding devices which are physical network equipment (e.g., switches, routers, wireless access points, etc.), and behave as a basic packets forwarding hardware, according to a set of rules that reside in the controller.

Multiple SDN architectures have been proposed based on the concept of forwarding all incoming packets according to controller decisions. Each architecture has its own design, forwarding model and protocol interface. Openflow [32] is one of the well-known SDN architectures that was developed to provide a separation between the control plane and the data plane. The Openflow architecture consists mainly of an Openflow switch, a controller, and a secure channel. The Openflow switch includes two parts, the first is Openflow client which is an abstraction layer that communicates with the controller by Openflow protocols via the secure channel, and the second is flow tables. A flow table contains multiple flow entries that consist of three fields:

- Matching rules: compared with the received packets to determine what kind of actions should be taken.

- Actions: defines a set of instructions that should be applied to the received packets after matching.

- Counters: records statistics about packets, such as the number of received packets.

The controller manages and controls the entries of the flow table. It can add, delete, or update flow entries through the secure channel. The secure channel is the communication link between the controller and all switches. Fig. 4 represents the components of the Openflow architecture.
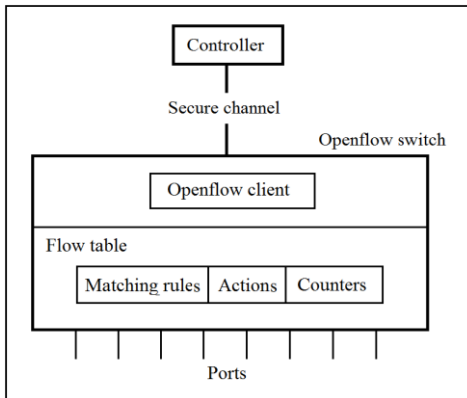
Figure 4. The componants of OpenFlow

Because of the scalability limitations in Openflow architecture which lead to limitations in the memory of forwarding devices (Openflow switches); Curtis *et al.* have proposed an extension to the Openflow called Devoflow [33]. Devoflow handles the problem of Openflow rules which are more complex in comparison with the traditional switches' rules. The main idea of Devoflow is based on allowing the Openflow switch to deal only with mice flows, and let the controller to deal with elephant flows. Having one centralized controller that is connected to all switches in the network negatively affects the reliability and availability, and raises the problem of a single point of failure. HyperFlow [34] which is presented by Tootoonchian *et al.* deals with the problem by creating a logically centralized but physically distributed controller [35].

SDN has been applied to different networking environments with the aim to facilitate the development of new services, features, and protocols. DCN is one of the environments where SDN is applied. Energy consumption is a critical issue in the large scale DCNs. Turning off unneeded devices is desirable to decrease the consumed energy, especially for green data centers. This motivated Heller *et al.* to propose a new power manager called ElasticTree [36]. Its key idea is based on using SDN to find the optimal and minimum power network subset that is appropriate to the current state of the traffic. As a result, the number of unused devices and the amount of saved energy will increase.

Although applying SDN to data centers enhances the performance and gives an opportunity to improve existing services, it has some weaknesses related to resources allocation (i.e. distributing available resource in an efficient way among end users). As mentioned above, the task of forwarding decisions is assigned to a single controller; accordingly, resources allocation in the DCNs is also assigned to that controller. This leads to multiple design challenges that have been discussed by Abu Sharkh *et al.* [37]. Since data centers house thousands of servers, the controller can be the bottleneck of the network if the number of requests exceeds the allowable load, or the number of switches in the DC increases.

Controller placement is another challenge in the SDN. It depends on multiple factors such as the applied topology, distribution options, bandwidth options, and port speed. The trade-offs between these factors should be analyzed carefully, and the location of the controller should be chosen accurately in order to minimize the downsides of the chosen location [37].

The security issue also must be taken into consideration while applying SDN technology for the DCN. Since the controller contains all critical information of the network, it should be protected strongly against attackers. Besides, the channel where all data exchanged between the controller and switches exist should be protected against eavesdropping [5]. The aforementioned issues must be handled in an efficient manner to get the maximum benefits of applying SDN.

2.7 Power and Consumed Energy Challenges

The electricity used in the data center keeps increasing because of the continues growing of applications, users' requests, and equipment. A study done by New York Times founded that the electricity consumed by data centers increased by more than 50% from 2005 to 2010 [38]. 40% of data centers energy is consumed by information technology equipment (i.e. computing servers) [39]. Many methods have been recently proposed in order to reduce the amount of consumed power in data centers. One of these methods is presented by Warkozek *et al.* [40]. Their method is based on layering the energy consumed by servers,

depending on whether the server is working or in idle mode, and whether the number of virtual machine in a server is small or large.

Another report written by a U.S. Environmental Protection Agency states that networking devices in data centers in the United States consume 3 billion kWh/year in 2006 [41], and the number is growing more and more. As a result, many research efforts have been done to make network components more energy efficient. One of these efforts is based on using Wok-on-Lan (WoL) mechanisms which wake host only in case of arrival of packets, and use a proxy to deal with idle-time traffic on behalf of a sleeping host [42]. Another algorithm called Dynamic Ethernet Link Shutdown (DELS) [43] is presented by Gupta *et al.* to address the same problem. DELS algorithm takes sleeping decision based on the arrival time of the previous packet, buffer occupancy, and configurable maximum bounded delay. The results show that DELS can shut down links and save power without delaying or losing packets. Shang *et al.* [44] solve the energy-saving problem from routing perspective in DCNs. They propose a model of energy aware routing and a heuristic algorithm which aims to: minimize the number of used routers and switches without affecting the network throughput, and put idle switches in the sleep mode to save energy. Results showed that their algorithm was efficient for saving energy in the networking elements, especially when the workload is low.

The massive number of servers which are placed in geographically distributed data centers consumes an enormous amount of power to meet the huge customer demands. Power management could be achieved by distributing workload across data centers efficiently. An optimized algorithm called Green-Fair has been proposed by Alawnah *et al.* [45] for workload distribution in geographically distributed data centers based on the standard gradient method. Results show that Green-Fair distributes the work between different regions in such a way that no region will be idle. Besides, it reduces latency, electrical and cooling cost.

Liu *et al.* [46] studied whether the use of brown energy can be reduced by geographic load balancing. They proposed a general power cost model which is a linear combination of the electricity prices and the lost revenue due to the network propagation delay and the load-dependent queuing delay. Based on that model, the authors introduced a distributed algorithm which aims to determine dynamically: the optimal way of distributing requests across data centers and the number of servers that should be on/off in each data center (i.e. sleep to save energy). Trace-based numerical simulations which have been used for evaluation showed that over 40% of the cost is saved during light-traffic periods [47].

Most of existing scheduling approaches in data centers do not consider the energy, but rather consider jobs distribution. A new methodology called Data center Energy-efficient Network-aware Scheduling (DENS) [21] is proposed by Kliazovich to provide a balance between the individual performance, traffic demands, and consumed energy. It reduces the number of servers needed for a particular job, and accordingly, the amount of consumed energy will be reduced.

It's worth mentioning that Power Usage effectiveness (PUE) is one of the primary approaches that is used to measure data centers energy efficiency. The two factors that affect PUE are the total power used in the data center and the power used for IT equipment. It is obvious that many factors are affecting the amount of consumed energy in data centers. The tradeoffs between energy and all mentioned factors should be analyzed and handled carefully for financial purposes and to achieve the desirable results.

2.8 Cooling Challenges

Large-scale data centers contain thousands of servers that need be managed through efficient thermal management systems with least amount of power. 45% of the total power of the data centers is consumed by cooling systems [21]. Various thermal control schemes have been proposed to manage data centers' cooling systems. Having a cooling system that saves power is highly desirable in huge data centers. As a result, a free cooling

economizer system is proposed by Sujatha *et al.* [48]. The proposed system cooperates with traditional cooling system to reduce consumed energy. Its idea is based on using outside air only when the climate is suitable. Using such a system will eliminate the work done by the chiller which consumes the greater part of cooling system's power.

One of the most important requirements of thermal management schemes is to detect the location of hot areas and send information to the controller as quick as possible. Wang *et al.* proposed an intelligent sensor placement scheme [49] based on Computational Fluid Dynamics (CFD) which takes servers and cooling systems information as inputs in order to analysis the current thermal condition of the data center. CFD is simply a method uses algorithms and numerical techniques to analyze issues that include fluid flows. Once the hot area is detected, an intensive cooling should be provided to this area. According to simulations, this scheme outperforms the conventional placement schemes. Another thermal control scheme [50] proposed by Bash *et al.* is also using thermal sensors that are placed in a distributed network and attached to racks in order to provide measurements of the environment. Results show 50% improvement in the performance, beside a better utilization of different spaces in the data center according to the provided measurements.

All aforementioned schemes are proposed with a goal of reducing cooling systems' cost and having efficient thermal management systems for the current and future data centers.

2.9 Construction Challenges

Multiple factors should be taken into consideration while choosing the physical location of a data center. First, the chosen location should be less prone to the natural disasters such as earthquakes, floods, ice storms and hurricanes. Second, the data center location should be proximity to the fiber backbone and has a good telecommunications and power infrastructure, because the availability of the mentioned elements will affect positively the performance and the throughput of the data center. The third factor is

the geographical location which plays a key role in a site selection. Consumed power and electricity price are affected by the temperature [51]. Choosing an area with a cool or moderate climate will reduce the power consumed by cooling systems, and then the total cost of the data center will be reduced consequently. Additionally, flight paths, construction costs and tax rates must also be considered since they may affect the services' availability and response time. According to the Forbes report [52], Province of Ontario in central Canada has been chosen as one of the ideal locations where data centers can be built, because it meets almost all the before mentioned requirements.

2.10 Simulation Challenges

Simulation is defined as the process of providing a flexible, efficient and an operative working model of a system. The workload moved in data centers' servers is increasing due to the rapid growth of online services. Cost-effective and energy-efficient simulation tools need to be developed in order to handle the huge amount of data stored and transmitted through data centers. Since power consumption is a critical issue in data centers as mentioned previously, it is highly recommended to have an energy efficient simulator in order to reduce operation cost. Banerjee *et al.* proposed an error-bound hybrid simulator for cyber-physical energy systems which captures the necessary cyber and physical events and processes [53]. Lim *et al.* proposed another simulation platform called Multi-tier Data Center Simulation Platform (MDCSim) [54]. It uses steady-state queuing models to provide a comprehensive, flexible and scalable simulation for data centers servicing multi-tier applications. Conventional detailed micro-architectural simulators can be replaced with BigHouse [55] which has been introduced by Meisner *et al*. BigHouse is a simulation infrastructure for data centers which is based on using a combination of queuing theory and stochastic modeling. This combination allows BigHouse to simulate server systems in minutes instead of seconds. Results show that BigHouse can deal efficiently with large cluster systems.

3 DISCUSSION AND CONCLUSIONS

In this paper, we discussed data center challenges from different perspectives. Each mentioned challenge can affect negatively the performance of data centers if not handled efficiently, especially with the rapid change in the communication and data centers industries. First, we discussed challenges related to network equipment such as traffic engineering, multicasting, security, and routing. Designing a reliable and flexible network architecture is required to provide an automatic naming for each device placed inside the data center. Although two different solutions have been mentioned above to a handle the naming issue, each has multiple disadvantages. We also presented some issues related to servers residing in the data centers and how virtualization can enhance servers' performance. A software defined network, which is the promising technology, has been discussed with its challenges. Virtualizations techniques, resource allocation schemes and energy consumption methods must be improved to keep up with software defined network technology. Then, we presented power and consumed energy challenges faced by data centers. The power needed to operate data centers keeps increasing, so new approaches are required to tackle downsides of the traditional power conservation equipment and cooling systems. Additionally, we summarized the most critical factors that play a key role in selecting the optimal data center's location. Finally, we briefly mentioned different energy-efficient simulation tools that can be used for data centers.

4 REFERENCES

1. Kant, K.: Data Center Evolution: A Tutorial on State of the Art, Issues, and Challenges. Computer Networks 53 (17), 2939-2965 (2009).
2. Garimella, S., Yeh, L., Persoons, T.: Thermal Management Challenges in Telecommunication Systems and Data Centers. IEEE Transactions on Components, Packaging and Manufacturing Technology 2 (8), 1307-1316 (2012).
3. Kachris, C., Kanonakis, K., Tomkos, I.: Optical Interconnection Networks in Data Centers: Recent Trends and Future Challenges. IEEE Communications Magazine 51 (9), 39-45 (2013).
4. Bari, Md., Boutaba, R., Esteves, R., Granville, L., Podlesny, M., Rabbani, Md., Zhang, Q., Zhani, F.: Data Center Network Virtualization: A Survey. IEEE Communications Surveys and Tutorials 15 (2), 909-928 (2013).
5. Lara, A., Kolasani, A., Ramamurthy, B.: Network Innovation Using OpenFlow: A Survey. IEEE Communications Surveys and Tutorials 16 (1), 1-20 (2013).
6. Chen, K., Hu, C., Zhang, X., Zheng, K., Chen, Y., Vasilakos, A.: Survey on Routing in Data Centers: Insights and Future Directions. IEEE Network 25 (4), 6-10 (2011).
7. Al-Fares, M., Radhakrishnan, S., Raghavan, B., Huang, N., Vahdat, A.: Hedera: Dynamic Flow Scheduling for Data Center Networks. In: Proc. 2010 USENIX Conference on Networked Systems Design and Implementation, pp. 19-19, USENIX Association Berkeley, California (2010).
8. Greenberg, A., Hamilton, J., Jain, N., Kandula, S., Kim, C., Lahiri, P., Maltz, D., Patel, P., Sengupta, S.: VL2: A Scalable and Flexible Data Center Network. In: Proc. 2009 ACM SIGCOMM Conference on Data Communication, pp. 95-104, ACM, New York (2008).
9. Mysore, R., Pamboris, A., Farrington, N., Huang, N., Miri, P., Radhakrishnan, S., Subramanya, A., Vahdat, A.: PortLand: A Scalable Fault-Tolerant Layer 2 Data Center Network Fabric. In: Proc. 2009 ACM SIGCOMM Conference on Data Communication, pp. 39-50, ACM, New York (2008).
10. Xi, K., Liu, Y., Chao, H.: Enabling Flow-based Routing Control in Data Center Networks using Probe and ECMP. In: Proc. 2011 IEEE Conference on Computer Communications Workshops, pp. 608-613, IEEE, Shanghai (2011).
11. 8 Important Datacenter Considerations, http://www.itmanagement.com/articles/it-blog/8-important-datacenter-considerations-54593
12. Newman, D.: 10 Gig Access Switches: Not Just Packet-Pushers Anymore. Network World 25 (12), 1-6 (2008).
13. Li, D., Li, Y., Wu, J., Su, S., Yu, J.: ESM: Efficient and Scalable Data Center Multicast Routing. IEEE/ACM Transactions on Networking 20 (3), 944-955 (2012).
14. Li, D., Xu, M., Zhao, M., Guo, C., Zhang, Y., Wu, M.: RDCM: Reliable Data Center Multicast. In: Proc. 2011 IEEE INFOCOM, pp. 56-60, IEEE, Sanghai (2011).
15. Li, D., Cui, H., Hu, Y., Xia, Y., Wang, X.: Scalable Data Center Multicast using Multi-Class Bloom Filter. In: Proc. 2011 19th IEEE International Conference on Network Protocols, pp. 266-275, IEEE, Vancouver (2011).
16. Lam, H., Zhao, S., Xi, K., Chao, H.: Hybrid Security Architecture for Data Center Networks. In: Proc. 2011 IEEE International Conference on Communications, pp. 2939-2934, IEEE, Ottawa (2011).
17. Xue, H., Liu, X., Feng, X., Guo, Z., Dai, Y.: PAD: Policy Aware Data Center Network. In: Proc. 2011 IEEE 3rd International Conference on Communication

Software and Networks, pp. 410-414, IEEE, Xi'an (2011).

18. Szefer, J., Jamkhedkar, P., Chen, Y., Lee, R.: Physical Attack Protection with Human-Secure Virtualization in Data Centers. In: Proc. 2012 IEEE/IFIP 42nd International Conference on Dependable Systems and Networks Workshops, pp. 1-6, IEEE, Boston (2012).

19. Hou, R., Jiang, T., Zhang, L., Qi, P., Dong, J., Wang, H., Gu, X., Zhang, S.: Cost Effective Data Center Servers. In: Proc. 2013 IEEE 19th International Symposium on High Performance Computer Architecture, pp. 179-189, IEEE, Shenzhen (2013).

20. Cao, J., Li, K., Stomenovic, I.: Optimal Power Allocation and Load Distribution for Multiple Heterogeneous Multi-core Server Processors across Clouds and Data Centers. IEEE Transactions on Computers 63 (1), 45-58 (2013).

21. Kliazovich, D., Bouvry, P., Khan, S.: DENS: Data Center Energy-Efficient Network-Aware Scheduling. In: Proc. 2010 IEEE/ACM International Conference on Green Computing and Communications & 2010 IEEE/ACM International Conference on Cyber, Physical and Social Computing, pp. 69-75., IEEE, Hangzhou (2010).

22. Data Center Energy Forecast, http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Data-Center-Energy-Forecast.pdf

23. Energy Efficiency and Server Virtualization in Data Centers: An Empirical Investigation, http://www.ntu.edu.sg/home/ygwen/Paper/JWC-Infocom-12.pdf

24. A Survey of Data Center Network Architecture, http://www.ece.virginia.edu/mv/pubs/recent-samples/Data-center-Survey.pdf

25. Guo, C., Lu. G., Li., D, Wu, H., Zhang, X., Shi, Y., Tian, C., Zhang, Y., Lu, S.: BCube: A High Performance, Server-Centric Network Architecture for Modular Data Centers. In: Proc. ACM SIGCOMM 2009 conference on Data Communication, pp. 63-74, ACM, New York (2009).

26. Guo, C., Wu, H., Tan, K., Shi, L., Zhang, Y., Lu, S.: Dcell: A Scalable and Fault-Tolerant Network Structure for Data Centers. In: Proc. ACM SIGCOMM 2008 Conference on Data Communication, pp. 75-86, ACM, New York (2008).

27. Chen, K., Singla, A., Singla, A., Ramachandram, K., Xu, L., Zhang, Y., Wen, X., Chen, Y.: OSA: An Optical Switching Architecture for Data Center Networks with Unprecedented Flexibility. IEEE/ACM Transactions on Networking 22 (2), 498-511 (2014).

28. Farrington, N., Porter, G., Radhakrishnan, S., Bazzaz, H., Subramanya, V., Fainman, Y., Papen, G., Vahdat, A.: Helios: A Hybrid Electrical/Optical Switch Architecture for Modular Data Centers. In: Proc. ACM SIGCOMM 2010 Conference, pp. 339-350, ACM, New York (2010).

29. Chen, K., Guo, C., Wu, H., Yuan, J., Feng, Z., Chen, Y., Lu, S.: DAC: Generic and Automatic Address Configuration for Data Center Networks. IEEE/ACM Transactions on Networking 20 (1), 84-99 (2012).

30. Ma, X., Hu, C., Chen, K., Zhang, C., Zhang, H., Zheng, K., Chen, Y., Sun, X.: Error Tolerant Address Configuration for Data Center Networks with Malfunctioning Devices. In: Proc. 2012 IEEE 32nd International Conference on Distribute Distributed Computing Systems, pp. 708-717, IEEE, Macau (2012).

31. Rygielski, P., Samuel, K.: Network Virtualization for QoS-Aware Resource Management in Cloud Data Centers: A Survey. PIK - Practice of Information Processing and Communication 36 (1), 55-64 (2013).

32. Mckeown, N., Anderson, T., Balakrishnan, H., Parullkar, G., Peterson, L., Rexford, J., Shenker, S., Turner, J.: OpenFlow: Enabling Innovation in Campus Networks. ACM SIGCOMM Computer Communication Review 38, pp. 69-74 (2008).

33. Curtis, A., Mogul, J., Tourrilhes, J., Yalagandula, P., Sharma, P., Banerjee, S.: Devoflow: Scaling flow Management for High-performance Networks. ACM SIGCOMM Computer Communication Review 41 (4), 254-264 (2011).

34. Tootoonchian, A., Ganjali, Y.: HyperFlow: A Distributed Control Plane for OpenFlow. In: Proc. 2010 Internet Network Management Conference on Research on Enterprise Networking, pp. 3-3, USENIX Association, California (2010).

35. A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks, http://hal.inria.fr/docs/00/87/76/27/PDF/hal-survey-v3.pdf

36. Heller, B., Seetharaman, S., Mahadevan, P., Yiakoumis, Y., Sharma, P., Banerjee, S., Mckeown, N.: ElasticTree: Saving Energy in Data Center Networks. In: Proc. 7th USENIX Conference on Networked Systems Design and Implementation, pp. 17-17, USENIX Association, California (2010).

37. Abu Sharkh, M., Jammal, M., Shami, A., Ouda, A.: Resource Allocation in a Network-based Cloud Computing Environment: Design Challenges. IEEE Communications Magazine 51 (11), 46-52, (2013).

38. Growth in Data Center Electricity Use 2005 to 2010, http://www.missioncriticalmagazine.com/ext/resources/MC/Home/Files/PDFs/Koomey_Data_Center.pdf

39. Report to Congress on Server and Data Center Energy Efficiency: Public Law 109-431, http://hightech.lbl.gov/documents/data_centers/epa-datacenters.pdf

40. Warkozek, G., Drayer, E., Debusschere, V., Bacha, S.: A New Approach to Model Energy Consumption of Servers in Data Centers. In: Proc. 2012 IEEE International Conference on Industrial Technology, pp. 211-216, IEEE, Athens (2012).

41. Mahadevan, P., Banerjee, S., Sharma, P., Shah, A., Ranganathan, P.: On Energy Efficiency for Enterprise and Data Center Networks. IEEE Communications Magazine 49 (8), 94-100 (2011).

42. Nedevschi, S., Chandrashekar, J., Liu, J., Nordman, B., Ratnasamy, S., Taft, N.: Skilled in the Art of being Idle: Reducing Energy Waste in Networked Systems. In: Proc. 6th USENIX Symposium on Networked Systems Design and Implementation, pp. 381-394, USENIX Association, California (2009).

43. Gupta, M., Singh, S.: Dynamic Ethernet Link Shutdown for Energy Conservation on Ethernet Links. In: Proc. IEEE International Conference on Communications, pp. 6156-6161, IEEE, Glasgow (2007).

44. Shang, Y., Li, D., Xu, M.: Energy-Aware Routing in Data Center Network. In: Proc. first ACM SIGCOMM Workshop on Green Networking, pp. 1-8, ACM, New York (2010).

45. Alawanah, R., Ahmad, I., Alrashid, E.: Green and Fair Workload Distribution in Geographically Distributed Datacenters. Journal of Green Engineering 4, 69-98 (2014).

46. Liu, Z., Lin, M., Wierman, A., Low, S., Andrew, L.: Greening Geographical Load Balancing. In: Proc. ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems, pp. 233-344, ACM, New York (2011).

47. Rahman, A., Liu, X., Kong, F.: A Survey on Geographical Load Balancing Based Data Center Power Management in the Smart Grid Environment. IEEE Communications Surveys and Tutorials 16 (1), 214-233 (2014).

48. Sujatha, D., Abimannan, S.: Energy Efficient Free Cooling System for Data Centers. In: Proc. 2011 IEEE Third International Conference on Cloud Computing Technology and Science, pp. 646-651, IEEE, Athens (2011).

49. Wang, X., Wang, X., Xing, G., Chen, J., Lin, C., Chen, Y.: Intelligent Sensor Placement for Hot Server Detection in Data Centers. IEEE Transactions on Parallel and Distributed Systems 24 (8), 1577-1588 (2013).

50. Bash, C., Patel, C., Sharma, R.: Dynamic thermal management of air cooled data centers. In. Proc. The Tenth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronics Systems, pp. 445-452, IEEE, California (2006).

51. Malkamaki, T., Ovaska, S.: Data Center and Energy Balance in Finland. In: Proc. 2012 International Green Computing Conference, pp. 1-6, IEEE, California (2012).

52. The Best Countries For Business, http://www.forbes.com/sites/kurtbadenhausen/2011/10/03/the-best-countries-for-business/

53. Banerjee, A., Banerjee, J., Varsamopoulos, G., Abbasi, Z., Gupta, S.: Hybrid Simulator for Cyber-Physical Energy Systems. In: Proc. 2013 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems, pp. 1-6, IEEE, California (2008).

54. Lim, S., Sharma, B., Nam, G., Kim, E., Das, C.: MDCSim: A Multi-tier Data Center Simulation Platform. In: Proc. 2009 IEEE International Conference on Cluster Computing and Workshops, pp. 1-9, IEEE, New Orleans (2009).

55. Meisner, D., Wu, J., Wenisch, T.: BigHouse: A Simulation Infrastructure for Data Center Systems. In: Proc. 2012 IEEE International Symposium on Performance Analysis of Systems and Software, pp. 35-45, IEEE, DC (2012).