Computer Networks and Communications
https://ojs.wiserpub.com/index.php/CNC

Review

Optical Switching Data Center Networks: Understanding Techniques and Challenges

Yisong Zhao[1], Xuwei Xue[1,*], Xiongfei Ren[1], Wenzhe Li[2], Yuanzhi Guo[1], Changsheng Yang[1], Daohang Dang[1], Shicheng Zhang[1], Bingli Guo[1] and Shanguo Huang[1]

[1] State Key Laboratory of Information Photonics and Optical Communications (IPOC), Beijing University of Posts and Telecommunications, Beijing, 100876, China
[2] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 100190, China
E-mail: x.xue@bupt.edu.cn

Received: 3 June 2023; **Revised:** 8 August 2023; **Accepted:** 18 August 2023

Abstract: Relying on the flexible-access interconnects to the scalable storage and compute resources, data centers deliver critical communications connectivity among numerous servers to support the housed applications and services. To provide the high-speeds and long-distance communications, the data centers have turned to fiber interconnections. With the stringently increased traffic volume, the data centers are then expected to further deploy the optical switches into the systems infrastructure to implement the full optical switching. This paper first summarizes the topologies and traffic characteristics in data centers and analyzes the reasons and importance of moving to optical switching. Recent techniques related to the optical switching, and main challenges limiting the practical deployments of optical switches in data centers are also summarized and reported.

Keywords: optical interconnects, data center network, optical switches, clock and data recovery, packet contention, switch control

1. Introduction

Dater centers (DCs), consisting of tens thousands of servers connected by large switching networks, provide the infrastructure for online applications and services such as cloud computing, social networks, file storage, and web search [1]. The topology of data center networks (DCNs) plays significant roles in determining the communication bandwidth between servers, the flow completion time and fault tolerance [2]. The design of the DCN topology is thus to build a robust network that provides the high bandwidth links and low (typically hundreds of microseconds) flow completion time across servers with low building cost and power consumption. Due to the various hosted application and services, the DC traffic, consisting of the tenant-generated interactive traffic, deterministic traffic and traffic with deadlines, is a mix of several classes with differentiate characteristics [3]. Thus, the DCN infrastructure should be effectively and efficiently utilized to provide high performance and quality access to the variety of services and applications deployed in DCs.

With the escalation of traffic-increasing applications, such as high-definition streaming, Internet of Things and cloud computing, traffic growth in DCs exceeds the bandwidth growth rate of application-specific integrated circuits (ASICs) based electrical switch [4]. The ASIC switches are expected to hit the bandwidth bottleneck in two generations from now, because the Ball Grid Array (BGA) packaging technique is hard to increase the pin density [5]. Moreover, the hierarchical network topologies based on electrical switches further deteriorates the bandwidth provisioning, especially at the top-layer switch, stringent increasing the flow

Copyright ©2023 Yisong Zhao, et al.
DOI: https://doi.org/10.37256/cnc.1220233159
This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)
https://creativecommons.org/licenses/by/4.0/

completion time. As future-proof solutions supplying high bandwidth, switching traffic in the optical domain has been considerably investigated to overcome the bandwidth bottleneck of electrical switches. The optical switches with high bandwidth, because of the optical transparency, are independent of the data-format and data-rate of the traffic [6]. Moreover, switching the traffic in the optical domain eliminates the power-consuming optical-electrical-optical (O-E-O) conversions. Migration of the traffic switching from electrical domain to the optical domain also removes the electronics circuits for dedicated various-format modulation at transceivers, thereby, significantly reducing cost expenses and data processing delay [7]. To date, the optics and networking communities have proposed many solutions on optical switches with milliseconds to nanoseconds switching configuration time, and variety switches based DCN topologies.

To practically deploy optical switches in DCNs, there are still several challenges that need to be addressed. First, to fully utilize the nanoseconds-level hardware switching time, a corresponding switch control mechanism is required to manage the optical switches in nanoseconds time scale to fast switch the traffic data [4]. Second, the conflicted packets at optical switches would be dropped, as no optical buffer existing to store the conflicted packets. This would result in high packet loss when packet contention happens at the optical switch nodes. Thus, packet contention resolution is another unsolved challenge to practical deploy the optical switches in DCNs [8]. Third, in optically switched network, new optical connections are generated every time as the switch reconfiguring. This requires that the receivers having to continuously adjust the local clock to properly sample the incoming packets and then recover the data [9]. The longer this recovery process takes, the lower the network throughput will be, particularly for the intra data center scenarios where many applications produce short traffic packets. Four, the multi-tenant services and applications with various data flows impose their own set of heterogeneous traffic requirements to the DC infrastructure [10]. Thus, a reconfigurable and highly flexible connectivity for DCN is required to provide the customized network frameworks to the various applications.

In this paper, we present a review of optical switching techniques capable of meeting the requirements of the next generation of large-scale data center networks. We start with a summarization of current data center traffic characteristics and topologies that revels the requirements of improving network performance in terms of both switch nodes and network topologies. To overcome the bandwidth limitation and multi-tier architecture of electrically switched networks, optical switching techniques have been proposed and investigated to replace the current electrical switches. We then review the technologies involved in the optical switch fabrics and the switch based optical topologies. The challenges of limiting the practical deployment of optical switching data centers have also been proposed to inspire researchers to propose more solutions. Finally, we summarize our conclusions.

2. Data Center

A data center is a physical facility, dedicated space within a group of buildings that operators use to compute, store and forward large amounts of traffic data. A data center's design is based on a computing network and associated components, such as storage and telecommunications that enable the operations of housed applications and services [1]. DCs consist of servers, switches, storage systems, routers, firewalls and application delivery controllers. Because these components are operated to process and store the business-critical data, stability is critical in the design and build of data centers. Together, they provide [11]: 1) Computing resources. The servers provide the computing, local memory, storage and network connectivity to drive the services and applications that are the engines of the network; 2) Storing infrastructure. Applications and services' data is the fuel of data centers. Storing systems are utilized to reliably hold this valuable information; 3) Network infrastructure. This system connects internal servers, storages, and external connectivity to end-users. Based on these infrastructures, data centers of worldwide enterprise IT are built to operate business applications and services that include but not only limited to [12]: 1) Productivity applications; 2) Email and file-sharing; 3) Database and analytics; 4) enterprise resource planning (ERP) and Customer relationship management (CRM); 5) Communications and collaboration services; 6) Artificial intelligence, big data and machine learning.

In the 5G era, data centers are evolving to provide high-speed and quality access, triggered by the new emerging cloud computing, 5G smart services and Internet of Things (IoT) applications. In this era, the digital ecosystem chain covers centralized data center network, edge computing, and terminal-device and everything in between [13]. This requires the DC infrastructure migrating from conventional on-premises physical servers to virtualized infrastructure that supports workloads across pools of physical infrastructure and into a multi-cloud environment [14]. Moreover, the escalation of traffic-boosting applications and the scale out of powerful servers

have heavily increased the traffic volume in DCs. As shown in Figure 1, by the end of the year 2021, annul global data center IP traffic is projected to reach 19.5 Zettabytes, which represents almost four-fold increase from the year 2016 [15]. About three-quarters of the business and consumer traffic flowing in data centers resides within the data center network. As the core hub of the digital ecosystem, data centers are phenomenally growing in complexity and size to evolve infrastructure with high network performance and to satisfy the keep-increasing traffic, playing the pivotal role in this innovative era and cross-generation evolutionary.

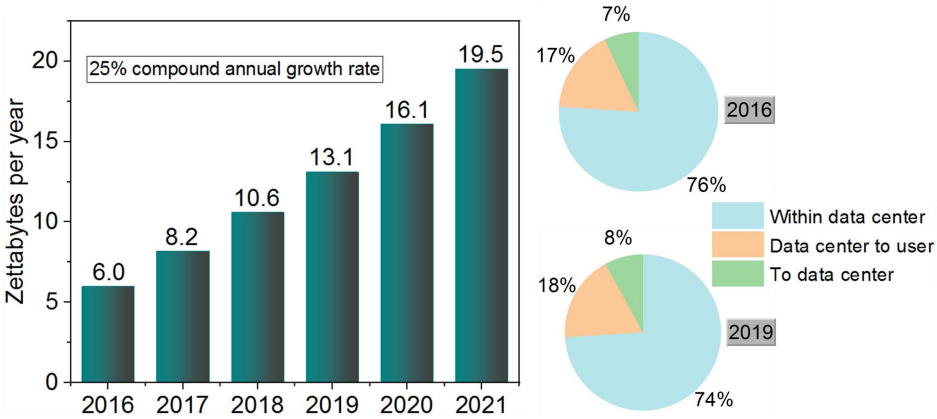

Figure 1. Data center traffic growth and distributions.

2.1 *Data Center Network Topology*

Data center networks (DCNs) establish and connect the entire network-based equipment and devices within the data center facility to enable a reliable interconnection. The DCNs thus ensure that the inside facility nodes can communicate and transfer data between each other and to the external users [16]. Existing electrical switched based DCN architectures are classified into two categories: server-centric and switch-centric architectures [17]. In server-centric networks, switch nodes are utilized as cross-connects, and routing intelligence should be placed on servers, where multiple Network Interface Card (NIC) ports are used per server. In switch-centric networks, routing intelligence is placed on switch nodes and each server usually uses only one NIC port to connect to the network.

The major advantage of switch-centric networks with the separation of communication and computation is that they are based on proven traffic forwarding and routing technologies available in commodity switches (e.g., Ethernet switches), such as IP broadcasting, link-state routing, and equal-cost multi-path forwarding [18]. Although a number of architectures in server-centric design have been proposed exploiting low-cost switches, the switch-centric based architectures are the mainstream scheme for the DCNs [19]. For instance, the multi-tier tree-like architectures continues to be the most widely deployed, and the fat-tree, leaf-spine and expander graph topology are the most promising architectures in terms of robustness, scalability and cost. All these architectures are switch-centric design.

The multi-tier design of DCNs comprises of hierarchy of switches layers as depicted in Figure 2. The leaves of the network tree form the access switching layer. Switches in this layer are usually top of rack (ToR) switches at low-cost, connecting servers (typically 20-40 servers) locating in the same rack. The middle tier of the network tree forms the aggregation switching layer, and the root of the network tree forms the core switching layer. In the multi-tier network, multiple racks are grouped together into one cluster. The intra-cluster and inter-cluster communication are handled by layers of aggregation switches and core switches, respectively. As illustrated in Figure 2, these access switches are connected through optical links to the aggregation switches to forward intra-cluster traffic. The inter-cluster traffic data is forwarded by the aggregation switches connected to the core switches. One or more border core layer switches provide connectivity between network infrastructure and users.

This multi-tier tree-like architecture features high fault tolerance enabled by the extensive path diversity even under failures. The main drawback of this tree-like design is the less than 1:1 oversubscription because of prohibitive costs. The 1:1 oversubscription means that any server can communicate at full bandwidth of their network interface with other arbitrary servers [20]. Due to the linear increasing costs associated with the scaling of link bandwidth and the port density of conventional electrical switches, building a tree-like scheme with 1:1 oversubscription would be prohibitively expensive for large-scale DCNs [21]. Therefore, practical oversubscription in this topology typically ranges from 8:1 to 3:1. Under the high oversubscription, if more

traffic is generated at the certain time on the active link, the large-stocked traffic exceeding the routing table entries will significantly increase the transmission latency [22].

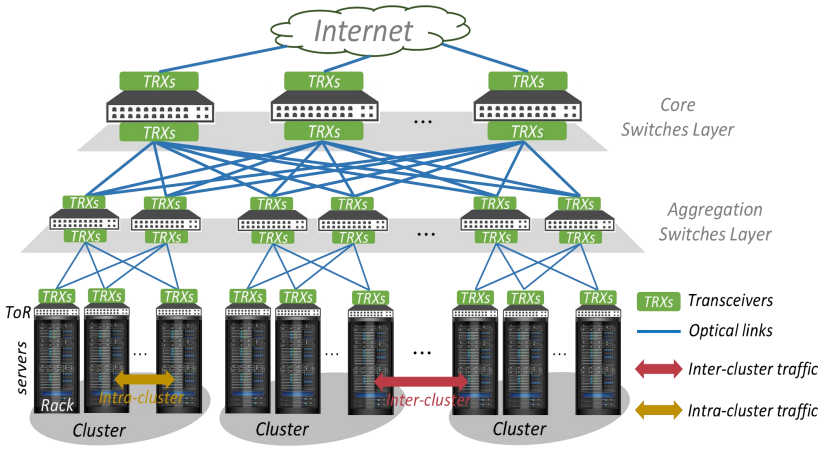

Figure 2. Multi-tier tree-like data center network.

As schematically illustrated in Figure 3, the fat-tree, leaf-spine and xpander data center network topologies have been evolved as the typical DCN architectures. As compared to multi-tier tree-like networks, the fat-tree architecture supports the use of commodity, identical switches in all switching layers, thereby decreasing cost in multiple-times. As depicted in Figure 3a, servers in the same rack are connected to a ToR switch, and this ToR switch is connected to a set of aggregation switches. At the root of fat-tree topology, a set of core switches are connected to aggregation switches in each cluster. In fat-tree network, aggregation switches and ToR switches of every cluster provide sufficient bandwidth for intra-cluster traffic forwarding, such that servers in the same cluster can communicate with each other at full speed [23].

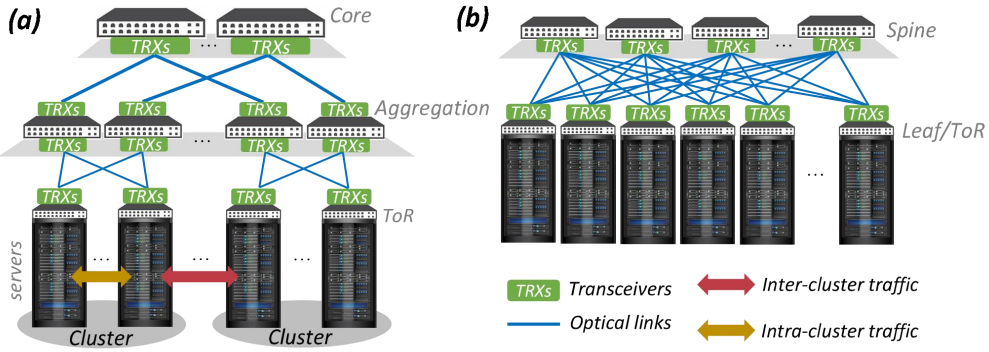

Figure 3. (a) Fat-tree and (b) leaf-spine DCN topology.

As a typical DCN topology, especial for the big data industry, the leaf-spine design has only two layers as shown in Figure 3b, the spine switching layer and leaf switching layer. The spine layer is the backbone of the network, where every leaf switch is interconnected with each-and-every spine switches. The leaf layer consists of ToR switches that connect to rack servers. With such a leaf-spine network topology, no matter which server is connected to which leaf switch, its traffic always crosses the same number of devices to arrive at the other server. Because a packet only needs to hop to a spine switch and another leaf switch to reach its destination, this guarantees the packet completion time at a predictable level [24]. Another benefit of leaf-spine topology is the capability of adding additional hardware and bandwidth. When oversubscription occurs on a certain link, an additional spine switch can be added and bandwidth can be thus extended to every leaf switch to reduce the oversubscription [25].

The xpander graph topology has been broadly investigated in schematic and industry [26]. Compared to leaf-spine architecture, the spine switches are fully connected to other spine switches in xpander topology. And ToRs are divided in different sets to connect only one spine switch. To select the best performance of throughput, there are some algorithms to solve such problem.

2.2 *Data Center Traffic*

Variety of applications and services are deployed on data centers benefiting from the flexible and cost-effective access to scalable storage and compute resources [12]. Data center traffic is often a mix of several classes with different service priorities and application requirements, which intimately determines the topology, scale, and even the technology selection of the DCNs [27]. To this end, a full understanding of the data center traffic characteristics is extremely necessary before putting efforts into the topology identifying and technology selection. Traffic characteristics, such as flow types, size and arrivals are highly correlated with applications. It is thus relatively hard to conclude a general pattern about traffic characteristics, given such strong dependency on data center applications. Some interesting findings include the following.

Mix of Flow Types: Data center traffic is a mix of various flow types and sizes [28]. A traffic flow here is specified as an established link between any two servers. User interactions like soft real-time applications such as Web search can create interactive traffic which are latency-sensitive flows that are usually short and should be forwarded with high priority [29]. Examples include short messages (100 KB to 1 MB) and queries (2 to 20 KB) [30]. Throughput-oriented flows require consistent bandwidth, but they are not sensitive to delay. These flows length ranges from moderate transfers (1 MB to 100 MB), such as one created by data computing applications (e.g., MapReduce), to long-running background flows, such as delivering large volumes of data across network sites for data storage [31]. Deadline flow has to complete the flow prior to the certain deadlines [32]. The size of deadline flow is either pre-known or a good estimate can be typically drawn [33]. The flow deadline could be either hard or soft which implies how value of flow completion decrease as time passed [34]. The hard deadline means zero value once the deadline has passed, while the soft deadline implies that it is still profitable to complete the flow and the value drops as the time passes away from the software deadline.

Traffic Burstiness: Burst is a common characteristic for data center traffic. Hardware offloading features, such as Interrupt Moderation and Large Send Offload that release the traffic workload of CPU, can lead to high burstiness [35]. Transport control in software, such as TCP slow start, can also create burst traffic when sends together a large window of flows [36]. In a burst environment, packet loss due to higher burstiness has been found more frequent at the network edge nodes like ToRs [37]. Burstiness can also lead to higher average buffer queue occupancy and thereby increasing packet drops probability and flow completion time due to buffers overflowed [38,39]. Moreover, high-burst traffic can deteriorate the buffer utilization in the memory shared switches when a certain switch port exhausts the shared buffer resulting from the long received bursty traffic [40].

Unpredictable Traffic Matrix: The various applications and services running on a data center create variety flows with different properties. Most bytes are delivered by large flows and most flows are short (less than 1 MB). The data center monitored in [41] shows that 99% of flows are shorted than 100MB and more than 90% of bytes are forwarded in flows between 100MB and 1GB. The flow arrival rates and distributions are also determined by the data center applications. The inter-arrival time of median flows in Facebook's DCs are between 2 ms to 10 ms for a single server (100 to 500 flows per second) while Microsoft finds the median arrival rate for the intra-cluster traffic to be 100 flows per millisecond [29,42]. Reference [43] reports between 100 to 10000 flow arriving at the switch per seconds in different private and educational data centers. Such variety of flow length and flow arrival rate can create a fluctuating and unpredictable traffic matrix which makes it hard to operate capacity planning and perform traffic engineering in a long-term scale to improve the network performance.

Data center infrastructures can be shared by multi-tenants and hosted applications offering various services to the network end-users. Traffic analyzation is then a necessary task in DCNs to guide the efficient use of the network resources. Knowledge of various traffic requirements and characteristics, as can be seen from the above empirical results, can help us design transport protocols and even network topologies to more efficiently use network resources.

2.3 *Design Requirements for Data Center Networks*

So much more than just a warehouse for servers, the data center is a sophisticated data networking environment that offers reliable and high-quality services to its users and customers, empowering them to design new infrastructures to enable their business forward. When design a new data center infrastructure, it's important to comprehensively consider the following summarized requirements that have a direct impact on performance.

Capacity: Recently, the escalation of traffic-boosting applications and the scale-out of powerful servers have significantly increased the traffic volume within the data centers. As reported in [15], the annual global data center traffic has reached 10.4 Zettabytes by the end of the year 2019. Consequently, each aggregation

switching node in the data center networks has to handle multiple to hundreds of Tb/s traffic. In addition, the traffic inside the modern data centers is expected to increase in the 5G era with a very high compound annual growth rate. Consequently, future data centers require ultra-high capacity networks to interconnect the infrastructure resources [44].

Scalability: Network scalability is the ability to easily subtract or add more storage and compute resources. For 'the old days' of on-premise DCNs, scalability was incredibly slow, costly, and difficult to manage. In 'the new days' of 5G era, the majority of network routing functions, physical network access, and innovative applications such as cloud computing and machine learning are typically deployed in centralized large-scale cloud data centers, while the Internet of Things device access and corresponding services platform, as well as diversified third-party applications are distributed and located at small-scale data centers [45]. Thus, the new scheme used to scale (add or subtract) the present built data centers to suit the new 5G fashion should be designed in a time-and-cost efficient manner.

Flexibility: Flexibility is increasingly important because many applications and services are dynamically deployed that require elastic network resource to efficiently accommodate them in the data center. The promising solution is to flexibly slice the network infrastructure following the virtualization strategy in a fully manageable and operable way to map the various applications' requirements. For the virtualized data center network, each infrastructure component (such as computing, storage and network) and its connections can be virtually recreated to slice the network [46]. Supervisory software creates these virtual components as various applications are required. In addition, virtualization enabled flexible data center often requires less power and space than a traditional data center. It can also be simpler to automate and update than the traditional data center [47].

Intelligence: The intelligent engines that are easier to use and configure with high efficiency, combined with network topologies and diverse domain knowledge of traffic characteristics, are required for modern data centers to quickly learn valuable information and execute targeted strategies from the massive amounts of data traffic generated by various applications. These intelligent engines should enable the data centers to provide rich platform services and application programming interfaces (APIs) with pre-integrated artificial intelligence (AI) services, search capabilities and graphics engine, as well as APIs in common fields such as voice, visual and language processing. These intelligent platforms and general pre-integrated services should work closely with the heterogeneous computing hardware such as field-programmable gate arrays (FPGAs) and graphics processing units (GPUs) to implement the application performance in-depth optimization [48].

Effectiveness and Efficiency: The traffic-boosting applications hosted in modern data centers impose stringent requirements in terms of packet loss, latency/jitter and throughput to the network infrastructure. For instance, the professional audio/video services and automatic driving applications require zero packet loss, low and bounded latency performance which is named deterministic quality of service (QoS) [49]. In addition, a data center represents a significant investment with the huge costs of hardware and software installation. Moreover, the cost of a DC's cooling and power typically is higher than the cost of IT devices inside it [50]. This is because components of modern data centers, including powerful servers, electrical switches and related network equipment, packages more transistors on the chip and more power-hungry chips in a smaller footprint. With the aim to satisfy these stringent network performance, cost/power-efficient data center networks featuring high bandwidth, providing extremely lower packet loss and latency performance should be investigated to host a broad range of mission-critical and latency-sensitive applications.

2.4 *Challenges for Data Center Networks*

Driven by the emerging of traffic-boosting applications and the scaling-out of powerful servers, more stringent requirements as abovementioned are imposed on the data centers with variety traffic characteristics. Current data center networks based on electrical switches are organized in a hierarchical topology, which is challenged by the bandwidth bottleneck and poor power efficiency to deliver the necessary and high quality of services [51].

Electrical Switch: The electrical switches double their bandwidth roughly for every two years at the same cost accordingly to Moore's law [52]. This allows data centers to keep up with the network bandwidth demands while maintaining the relatively steady and low network cost over the passing years [53]. However, the move towards traffic-boosting applications and powerful servers will greatly boost the demands of network bandwidth. Meeting the requirements of higher network bandwidth especial for the aggregation switch nodes, would greatly inflate costs.

Due to the limited number of high-speeds pins available on the switch chip and the limited number of connectors on the front panel of the rack unit, the bandwidth of electrical switches is expected to hit the

bandwidth bottleneck soon [5]. Furthermore, the electrical switches consume the power proportional to the data rate, as the switch dissipates energy with every bit transition [54]. With the speed scaling-up requirements, the electrical switches based DCNs face the stringent pressures on the power-consumption. Despite new technologies based on multi-tier packaging, monolithic integration and Silicon Photonics (SiPh) are being investigated, several challenges, however, still have to be solved before these technologies become viable [55,56]. For instance, the high manufacturing (including both packaging and testing) costs, not to mention the complexity of packaging a large number of fiber coupling and external laser sources. Even if these issues were able to be solved, these technologies will ultimately hard to keep increasing the transistor density limited by the CMOS scaling [57].

Hierarchical Network Topology: One of the performance deteriorations appeared with hierarchical data center topology is oversubscription that could dramatically disrupt the network performance [21]. Oversubscription is the ratio between the total uplink bandwidth to the servers' bandwidth at the ToR switch layer. Thus, due to the hierarchical network topology, as moving up to aggregation and core layers, the number of servers (and thus the bandwidth) sharing the uplink bandwidth increases and, hence, increases the oversubscription ratio, resulting in bandwidth bottlenecks at aggregation/core layers. Oversubscription limits the server to server capacity, especial for servers locating in different clusters/pods, where the ratio exceeds 1:1. The bandwidth contesting leads to switch buffers overloading, which then in turn start losing packets [23]. In addition, the buffer queuing and processing delay at the multi-tier switches bring large latency for the inter-cluster traffic that a packet needs to traverse the aggregation and core switches to reach its destination. Another challenge introduced with hierarchical network topology is the lack of network fault tolerance, especially at the core switching layer resulting from the lower physical connectivity. Hardware (switch) failures in aggregation of core switching layers will significantly deteriorate overall network performance.

To interconnect the multi-tier switching nodes, the electrical-to-optical-to-electrical (O/E/O) conversions exist between switching layers, which thereby significantly increases the number of transceivers and, hence, cost and power consumption [58]. Instead of the low-rate on-off keying (OOK) modulation, multi-level modulation like pulse-amplitude modulation (PAM)-4 schemes are gradually employed in the data centers fueled by the demand for higher data-rate [59]. To process these format-dependent signals, dedicated parallel optics and electronic circuits are required at the front-end of the electronic switches. This introduces extra cost and power consumption in the hierarchical network topology.

3. Optical Data Center Networks

Optical switching, as a future-proof solution to overcome the bandwidth bottleneck of electrical switches, has attracted the widespread attention to researchers. Due to the optical transparency, switching the data in the optical domain is independent of the bit-rate and data-format of the traffic. Thus, optical switching supports much higher bandwidth than electrical switching and at much lower packet completion time due to the removing of electronic circuits for switching. In addition, WDM technology can be employed to boost the optical network capacity at a superior power-per-unit bandwidth performance. Combining with the WDM, optical switching is a viable solution to overcome the count limitation of high-speeds pins and front panel connectors at electrical switches. Moreover, the optical switches do not require any power-consuming E/O and O/E conversions, which significantly reduces the number of expensive and power-hungry transceivers. All these benefits can be exploited to flatten the network topology and thus sidestepping the scaling wall of the hierarchical data center topology.

3.1 *Optical Switching Technologies*

To date, three main optical switching technologies have been investigated which resulted in increasing data transfer capabilities for the data center networks.

Optical Circuit Switching (OCS): OCS has three distinct steps: links set-up, data transmission and links tear-down. One of the main features of OCS is its two-way reservation process in the phase of link circuit set-up, where a source sends a request for setting up a circuit and then receives an acknowledgement back from the corresponding destination [60]. The overall transfers suffer from long set-up times relative to connection holding time, seriously deteriorating the network throughput. In addition, all data transmission of a connection in OCS network follow the same path, no statistical multiplexing of the client packet can be achieved at any intermediate node. More specifically, bandwidth allocation is a coarse granularity, which is allocated by one wavelength at a time. However, most modern applications in practice require the sub-wavelength connectivity and these high-bit-rate applications often involve "traffic bursts" that last only a few milliseconds or less.

Furthermore, since no optical buffer existing, the capacity of the circuit link must equal the peak data rate, which can be orders of magnitudes higher than the average data rate, for bursty sources [61]. This results in the low bandwidth utilization for OCS-based data centers, especial for networks with many communication pairs with bursty traffic patterns.

Optical Burst Switching (OBS): In OBS-based data center schemes, the source nodes first send a burst header (control packet) on a separate control link (similar to the link set-up step of OCS) to reserve the optical bandwidth with the configuration of switches along an optical path for the burst forwarding of optical payload. The OBS scheme, unlike OCS, can send out the payload on a data channel without having to receive the response signal first. After sending out the burst payload, another control signal (similar to the link tear-down of OCS) is sent out to release the reserved optical bandwidth [62]. This implies that the offset time T between the burst header and the burst packets can be much less than the circuit set-up time, improving the network throughput. Benefitting from the one-pass bandwidth reservation for the duration of actual data transfer, the OBS paradigm provides the sub-wavelength switching granularity. However, due to OBS schemes generally do not need optical buffer, a big issue related to the one-way reservation OBS is how to deal with packet contention and prevent the contention caused packet dropping [63]. Another challenge of OBS related to the long-time bandwidth reservation and the using of a non-zero offset time is the high payload completion time encountered by each burst communication, not fully supporting the bursty traffic patterns as well [64].

Optical Packet Switching (OPS): OPS paradigm uses in-band control information where the header or label follows the rest of the packet payload closely, so there is no reservation possible (thus decreasing the end-to-end latency) and the bandwidth can be utilized in the most flexible way [65]. Due to OPS scheme allows statistical sharing of the optical bandwidth among packets belonging to different source and destination pairs, OPS scheme is thus suitable for supporting burst traffic scenarios. The packet payloads in OPS-based data centers remain in the optical domain, while the header or label may be electronically or optically processed (though the optical logic is very primitive). The generating and processing of fast header rely heavily on optical labeling techniques. To keep the percentage of the control overhead down, OPS-based data centers normally employ fast (nanoseconds reconfiguration time) optical switches based on semiconductor optical amplifiers (SOAs) or arrayed waveguide grating routers (AWGRs). Contention resolution is typically achieved by a combination of wavelength conversion, fiber-optic delay lines (FDLs) and, in rare cases, deflection routing [66]. Exploiting the WDM technique, the OPS paradigms significantly improve the network capacity where multiple streams of packets are multiplexed in the wavelength domain [67]. Benefitting from these features, OPS technologies offer a suitable solution for data center applications which requires on-demand transmitting the bursty and small data sets.

Table 1 summarizes the characteristics of these three optical switching technologies. The OCS has coarse switching granularity and thus resulting in high packet completion time and low bandwidth utilization, but it benefits the low implementation complexity that is easier to be deployed. As a comparison, the OPS, benefitting from the fine switching granularity, can fast complete the traffic flow with high bandwidth utilization. However, the implementation complexity and high control overhead limit its practical application in large-scale optical DCNs. As a compromise, the characteristics of OBS lies between OCS and OPS technologies.

Table 1. Comparison of optical switching technology.

Switching Technology	**OCS**	**OBS**	**OPS**
Switching Granularity	coarse	medium	fine
Flow Completion Delay	high	medium	low
Bandwidth Utilization	low	medium	high
Control Overhead	low	low	high
Complexity	low	medium	high
Applicability	medium	low	high

3.2 *Optical Switches: State-of-Art*

To date, exploiting various building blocks, many solutions employing optical switches have been investigated [68], such as 3D micro-electrical mechanical switches (MEMS), Liquid Crystal on Silicon (LCoS) display matrices, micro-ring resonators (MRRs), Mach-Zehnder interferometers (MZIs), tunable lasers and arrayed waveguide grating routers (AWGRs), and semiconductor optical amplifiers (SOAs). Determined by the exploited building blocks, switch reconfiguration time, as the main switch investigate parameter, can vary three

orders of magnitude from milliseconds to nanoseconds, determining the granularity of a switch and therefore, its application. The comparison has shown in Table 2.

Slow (milliseconds and microseconds) Optical Switches: Micro-electrical mechanical switches (MEMS) switches are micrometer-scale devices that rely on mechanical moving micro-mirrors to switch the optical signal from input ports to output ports [69]. An array of N2 micro-mirrors is needed to build a N×N direct switching MEMS switch. The area to lay the mirrors out and the mirror size limits the port-radix to N≤32. By steering a pair of micro-mirrors, indirect switching MEMS is implemented in a three-dimensional (3D) plane [70], which requires only 2N mirrors configurable to N discrete angles. The switch reconfiguration time (milliseconds magnitude) is strongly determined by the mirror response speed to the precise movement between these N discrete angles. Large scale (N=1100) 3D-MEMS switches [71] have been investigated with a 4 dB maximum insertion loss and commercial product with 25 ms reconfiguration time supports up to N=320 with a 3 dB maximum insertion loss [72].

Liquid crystal (LC) technology can be used to switch light based on the LC's birefringence to control the polarization of incident light [73]. For the liquid crystal on silicon (LCoS) technique, a large size of 1×N optical switch can be built by depositing a reflective array of LC cells on a silicon backplane [74]. At the LCoS optical switch, all channels focus on a pixel area depositing multiple phase-shifting LCs that collectively form a controllable linear phase grating (in the range 0 to 2π) to steer a reflected channel to an output port. Since each channel is switched separately, multiple channels can be selected and multiplexed into the same output port. Multiple 1×N switches can be stacked in a Spanke topology for N×N switching [75]. The reconfiguration time for this switch type is of the order of tens of milliseconds [76]. [77] demonstrates a 1×40 LCoS switch with 8 dB insertion loss and -40 dB crosstalk.

Micro-ring resonator (MRR) is waveguide topology that can exploit the thermo-optic (to) effect for optical switching. A circular resonant cavity of micrometer-scale radius (r) can be formed by bending back a waveguide onto itself. When the optical path of the ring is an integer multiple (L) of the guided wavelength (λ), the ring starts resonating according to the principle of $(2\pi r)\cdot Ri = L\cdot\lambda$, where Ri is the effective refractive index of the ring [78]. For switching, multi-rings are placed between two bus waveguides for an optical signal, at the resonant wavelength, to be coupled from one waveguide to the other. Heating a ring changes its effective refractive index and shifts its resonant wavelength, thereby enabling wavelength-selective switching [79]. A silicon photonic 8×8 switch, reconfigurable in the order of a few microseconds, is recently reported [80] with -16.75 dB average crosstalk and 8.4 dB average insertion loss.

Fast (nanoseconds) Optical Switches: A Mach-Zehnder interferometer (MZI) waveguide exploiting electro-optic (EO) effect can perform fast switching by changing the waveguide refractive index [81]. In a 2×2 MZI switch unit, a coupler divides the input optical signal to the two MZI arms. In the "on" state, the two arms are in anti-phase leading to constructive interference at the through waveguide. In the "off" state, the two MZI arms are in phase and a signal entering an input waveguide leads to destructive interference at the crossover waveguide, switching the input signal [82]. To date, the largest demonstrated silicon integrated MZI switch is 32×32 built in Benes topology with reconfiguration time of 1.2 ns and the high insertion loss (20.8 dB) and crosstalk (-14.1 dB) [83].

A passive router can be built based on a N×N arrayed waveguide grating (AWG), assuming wavelength tuning (N2 fixed wavelength transceivers or N wavelength tunable transceivers) is used at the input/output ports [84]. In an AWG router (AWGR), each input port at the same time exploits the same N wavelengths to establish a strictly non-blocking all-to-all connectivity [85]. Following the cyclic mechanism, the wavelengths from two adjacent input ports appear at the output ports cyclically rotated by one position. Hence, each output port could receive N wavelength channels, one from each input port. The input waveguides are spaced such that, on any phased array waveguide, signals of the same wavelength from N input ports, have an additional phase difference. Signals are separated again at the output coupler and directed to different output ports. Switch reconfiguration time, which can be less than 10 ns, is determined by the time required for wavelength tuning [86]. A silicon AWGR with N=512 has been fabricated with a high inter-channel crosstalk of -4 dB [87].

Semiconductor optical amplifier (SOA) can be used as an optical switch gate, the "on" state providing broadband amplification and the "off" state blocking the incident signal [88]. Based on a kind of broadcast-and-select (B&S) topology as shown in Figure 4, N×N SOA-based optical switches can be implemented, where each switching path exploits an SOA gate. The input signal is split into N paths, in the B&S architecture, for broadcasting the input signal to all output ports. N parallel SOA gates are used to switch ("on" or "off") the N paths signal, establishing an input-output port connection. The broadcast operation of B&S architecture allows wavelength, space and time switching, and another benefit is that the SOA gain in "on" state inherently compensates the splitting and combining losses [89]. Additionally, the high "on"/"off" extinction ratio brings its excellent crosstalk suppression. Nevertheless, the scalability in the B&S structure is limited by the splitter

caused power loss and the high number of waveguide crossing [90]. Multi-stage topology can be used to arrange smaller B&S modules to scale out the switch to a large size, but the scalability is still limited by the amplified spontaneous emission (ASE) noise. Lossless, integrated 16×16 SOA switches with nanoseconds reconfiguration time have been fabricated [91,92], based on smaller B&S switching modules interconnected in a three-stage Clos architecture. These features make SOA-based switch a suitable candidate for data center applications which requires fast and high-bandwidth transmission.

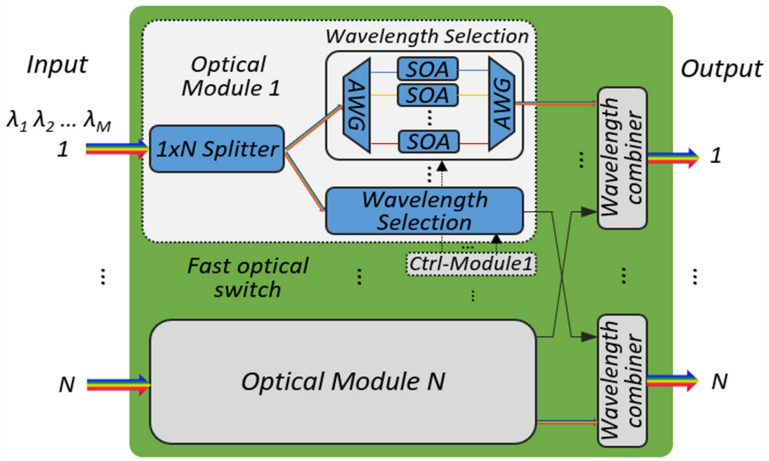

Figure 4. Broadcast-and-select (B&S) topology.

Table 2. Comparison of optical switches.

Switching technology	Switching time	Switch scale	Insertion loss	Crosstalk	Ref.
2D MEMS	O(100) μs	32×32	low	low	[69]
3D MEMS	O(10) ms	1100×1100	low	low	[71]
LC	O(100) μs	2×2	low	low	[73]
LCoS	O(10) ms	1×40	medium	low	[77]
TO-MRR	O(1) μs	8×8	low	medium	[80]
EO-MZI	O(1) ns	32×32	high	high	[83]
AWGR	O(1) ns	512×512	medium	high	[87]
SOA	O(100) ps	16×16	low	low	[91]

3.3 *Optical Data Center Network: State-of-Art*

Various optically switched architecture prototypes, based on the above optical switches, have been proposed to demonstrate the potential of optical data center networks. Optical data center networks are mainly classified into two categories based on the switching techniques used, the electrical/optical hybrid scheme, where electrical along with the optical switches constitute together for the network interconnection, and the full optical scheme, where only optical switches (slow or fast) are employed.

Hybrid Electrical/Optical Data Center Networks: A number of hybrid electrical/optical interconnecting architectures have been reported for data centers [93], such as Helios [94], HydRA [95], c-Through [96] and RotorNet [97]. The electrical switches typically connect all the servers in a multi-level hierarchy and short reconfiguration time with large connectivity to handle small-size and bursty traffic patterns. The ToR are interconnected using both this down-link electrical packet-switched network and an up-link optical circuit-switched network. The optical circuit-switched network, implemented by a single or an array of slow optical switches like MEMS, provides large capacity links for high-volume and slow-changing traffic.

Most of the aforementioned prototypes need a centralized network scheduler to reconfigure the entire architecture, in response to the traffic dynamics except the RotorNet. The RotorNet has predefined schedule which is stored locally. However, it is not sensitive to traffic dynamics. Different from the pre-defined scheduler, the central schedule requires network-wide demand estimation and resource schedule [97,98]. Apart from the long optical switch (MEMS) reconfiguration time, ranging from microseconds to tens of milliseconds, the schedule and control introduce a significant extra processing latency. For instance, in order to achieve high resource utilization in Helios and c-Through, Edmond algorithm needs hundreds of milliseconds to converge to

a network-wide matching [96]. Hence, the proposed architectures are better suited for non-bursty applications such as data migration and storage backup [95] where traffic is aggregated last more than a couple of seconds [94], to compensate the reconfiguration overhead.

All Optical Data Center Networks: This category includes the proposals exploiting the slow (sub-millisecond/milliseconds) or fast (nanoseconds) optical switches. OSA and Proteus [99,100], as the all-optical switching architectures are proposed, where the ToRs are connected to a central MEMS switch, through optical Mux/Demux. A key challenge here is the slow reconfiguration time including both hardware switching time (microseconds to tens of milliseconds) and controlling overhead (hundreds of microseconds to seconds). Mordia [98], Wavecube [101], RODA [102] and OPMDC [103] are all built based on wavelength selective switches utilizing either a ring topology or a multi-dimensional cube structure. The limited port-count of wavelength selective switch requires stacking and cascading multiple switches, deteriorating network performance in terms of packet loss, flexibility and latency [104].

Fast network interconnecting like IRIS [105], DOS [106], Petabit [107], LIONS [108] and Hi-LION [109,110] are reported based on AWGRs, of which the Petabit proposed a three-stage Clos network and Hi-LION demonstrates a mesh-like network exploiting both local and global AWGRs. The network performance and interconnecting scalability of the aforementioned proposals are strongly depended on the port radix [85] and the capability of wavelength tuning components such as tunable lasers or tunable wavelength converters. In addition, the wavelength-related operation as a block on the road to deploy WDM technology further limits the network capacity. An all-optical network, Baldur, is proposed in [111] based on transistor laser (TL) to enable high-speed and power-efficient communications in computing systems. However, the complex current control for the large radix TL array limits the practical deployment in the computing network. SOAs working as switching gates, the OPSquare [112], HiFOST [113], Vortex [114], ROTOS [115] and OSMOSIS [116] are proposed with SOA-based B&S, featuring of nanoseconds switching time. OPSquare uses a parallel-module switch architecture with distributed control and scales by adding more modules and wavelengths. There is only one switching stage, irrespective of the port-count, implemented in the modules. At most two- hop is enough for the traffic forwarding between any two different edge nodes [117]. Utilizing WDM and B&S stages, the OSMOSIS equips high-capacity and low-latency forwarding of the synchronously arrived fixed-length optical packets. Scaling is limited by the data plane and control plane complexity [116]. The scheduler implementation spans multiple interconnected chips, increasing not only network complexity but also the latency.

Given the bursty traffic features and high fan-in/out hotspots patterns in data center networks, slow optical switches providing static-like and pairwise interconnections would only be beneficial as supplementary switching elements. In contrast, fast optical switches with nanosecond-reconfigurable time can handle arbitrary traffic and can be deployed at any layer of the data center network. Considering this, fast optical switches-based network topologies supporting nanoseconds optical packet switching offers a potentially future-proof solution for the fast and high-capacity data center networks.

3.4 *Technical Challenges*

The network and optics communities have been extensively investigating the fast (nanoseconds) optical switching techniques for many years. However, each community tries to address problems and challenges from their own perspective. For instance, the optics communities focus on developing technologies for single components and devices that achieve nanosecond switching configuration time while devoting little attention to solving the network interconnecting challenges, e.g., flexible quality of service (QoS) provisioning or scalable scheduling. In contrast, the networking community proposed a number of solutions that can scale the network interconnecting and address the system challenges such as demand estimation or network bandwidth reconfiguration [118]. Thus, the proposed solutions need to be integrated and, in some cases, need to be redesigned for the cases proposed independently by each community. Following, the main technical challenges that limit the practical deployment of optical data center networks are summarized.

Fast and Scalable Switch Control: Despite the promises held by fast optical switches, the corresponding nanoseconds-scale control mechanisms are required to control the switches and to fast forward the data traffic [119]. Typically, an optical label or header, carrying the destination information of the data packet, is associated with the optical data packet to be processed at the specific switch controller. Based on the received label matrix, the switch controller computes the optical-switch configuration and then accordingly forward the data packets [120]. Short switch configuration time, consisting of both controlling overhead and hardware switching time, is essential as it determines the network latency and throughput performance [121]. Thus, to fully utilize the nanoseconds-magnitude hardware switching time and to reduce the flow completion time of short data packets, the switch controller needs to process the label signals and configure the switch within nanoseconds [122].

Moreover, the switch controlling overhead should be independent of the network scale. Considering the scale of practical DCNs, which typically comprise hundreds of thousands servers, switch control mechanism needs to be performed in parallel and independently for every optical switch, not a network-wide scale schedule [123]. In addition, for the label control mechanism, the edge nodes should be time-synchronized connected to the optical switches at a very fine granularity (ideally few nanoseconds) to align the label signals and corresponding data packets [4]. Even any time inaccuracy in the synchronization phase can be accordingly compensated with a customized interpacket gap, however, this would reduce the overall network throughput. Therefore, the implementation of fast and scalable switch control requires a nontrivial amount of ingenuity and custom hardware support.

Lack of Optical Memory: The lack of optical buffer is one of the main fundamental differences between optical switch and electrical switch. Electrical switches typically employ random access memories (RAM) to buffer the packets that lost contention. Due to the lack of effective RAM in the optical domain, the conflicted packets at the optical switch would be dropped, thereby resulting in packet loss [124]. Thus, packet contention resolution is another unsolved challenge that needs to be solved to guarantee fast switch control. Several solutions have been proposed to address such issue, based either on wavelength conversion [125], optical fiber delay lines (FDLs) [126] or deflection routing [127]. However, none of them is practical for large-scale DCNs, due to the extra hardware deployment of wavelength conversion, fixed buffering time of FDLs and the management complexity of deflection routing. Label control mechanism enabling dropped packet retransmission provides a promising solution to address the packet contention caused packet loss, combining the deployment of RAM at the edge nodes (such as ToRs) [128]. To minimize the introduced retransmission delay, the optical switch should be employed relatively close to the edge nodes, which requires to flat the network topology. Moreover, an efficient scheduler is essential to intrinsically reduce the time-consuming packet retransmission.

Fast Clock Data Recovery (CDR): Unlike the synchronized point-to-point connections between any paired ports in an electrical switch, the optical switch creates momentary physical links between source and destination ports [129]. Therefore, in an optical packet switching network, where the clock frequency and phase of data packets vary packet by packet, new physical connections are created every time reconfiguring the optical switch [130]. The receivers thus have to continuously adjust the local clock (consisting of both frequency and phase) to properly sample the incoming optical packets and thereby recovering the payload, as illustrated in Figure 5. As no payload data can be valuably received before the CDR completing, the long CDR processing time (hundreds of nanoseconds for off-the-shelf transceivers) will deteriorate the network throughput, especially in the intra data center scenarios where applications and services produce short traffic packets [131]. In [9], the clock phase caching technique is used to accelerate the CDR processing at the receiver side. However, to adjust the clock phase interpolator for every packet, extra time at the transmitter side is required which significantly increases the interpacket gap time and whereby offsets the efficacy of short CDR time. Burst-mode receivers enabling nanoseconds CDR processing time based on over-sampling or gated oscillators have been extensively investigated in passive optical networks [132]. These burst-mode techniques, however, introduce the design complexity and increase the deployment cost. These burst-mode receivers also need to be re-evaluated aiming for higher (>25 Gb/s) data rates, not suitable for DCNs with 100Gb/s links to be deployed.

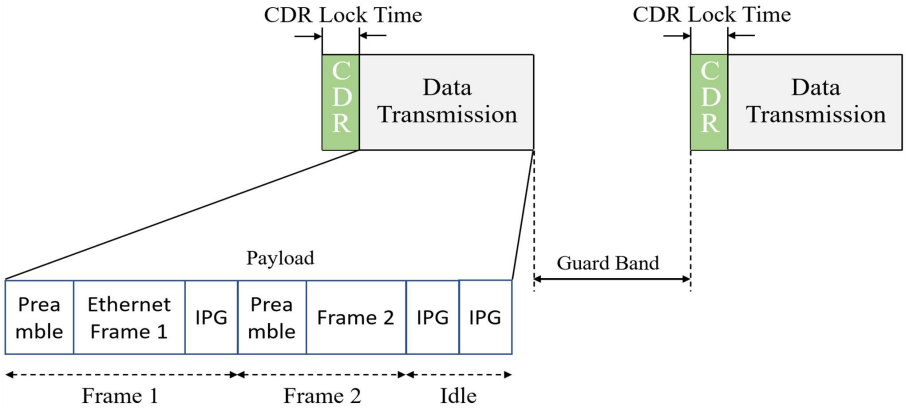

Figure 5. Clock and data recovery step at the receiver.

Reconfigurable Connectivity: The rapid emerging of multi-tenant applications and services with mix data flows impose their own set of various requirements, such as bandwidth and forwarding priority, to the network infrastructure. These specific requirements dynamically change as the application switch-over. Thus, the

reconfigurable and highly flexible connectivity is required in order to optimize the hardware resources, overcoming the limitation of currently deployed static and semi-automated control frameworks [118,133]. One promising strategy is to virtualize the infrastructures in a fully operable and flexible way by the software-defined networking (SDN) to enable such reconfigurable environments [117,134]. The SDN control plane exploits an open, standard, vendor-independent and technology-transparent southbound interface to monitor and configure the underlying data plane. To facilitate the integration of SDN control plane, the components of data plane such as ToRs and optical switches should be compatible with the open interface exhibited by the SDN control plane. Given the specificity and characteristic of the optical switching network, proper extensions and customization on the open protocols (e.g., OpenFlow [135]) have to be performed and implemented. Moreover, efforts need to be made to develop various functional engines running at the SDN control plane to flexibly virtualize the optical infrastructure, supplying reconfigurable connectivity.

4. Conclusion

Variety of applications and services depend on data centers to provide the high reliability and availability of computing and storage resource at minimal costs. Along with the emerging of traffic boosting applications, the bandwidth bottleneck of electrical switches forces the migration of switching from the electrical domain to the optical domain. Data centers are thus moving towards full optical switching with technical evolutions of both optical switches and network topologies to satisfy the demands of massively increasing data center traffic. In this paper, we have presented a summary of data centers traffic characteristics and topologies trends in data center networks which are based on electrical switches. The shortcomings and challenges for electrically switched data centers are also reported to reveal the trends of full optical switching. To that end, we present a brief summary of optical switching technologies that will enable ultra-high bandwidth links, in addition to an overview of optical network topologies that will enable the high utilization of bandwidth and thereby lower cost and power consumption. The full optical switching is expected to deploy in data centers in the next decade, enabling the developments of new applications like artificial intelligence and machine learning, as well as providing the fast, reliable and cost-effective services to users.

Funding

This work was partially supported by the National Natural Science Foundation of China (61821001, 62101065, 62220106002, 62125103, 62171059) and State Key Laboratory of Advanced Optical Communication Systems and Networks, China.

Conflict of Interest

There is no conflict of interest for this study.

References

[1] Cisco. What is a Data Center. Available online: https://www.cisco.com/c/en/us/solutions/data-center-virtualization/what-is-a-data-center.html (accessed on 3 August 2008).

[2] Cui, Y.; Yang, Z.; Xiao, S.; Wang, X.; Yan, S. Traffic-Aware Virtual Machine Migration in Topology-Ad aptive DCN. *IEEE/ACM Trans. Netw.* **2017**, *25*, 3427–3440, https://doi.org/10.1109/tnet.2017.2744643.

[3] Li, Y.; Zhou, X.; Li, K.; Qi, H.; Guo, D. TrafficShaper: Shaping inter-datacenter traffic to reduce the transmission cost. *IEEE/ACM Trans. Netw.* **2018**, *26*, 1193–1206, https://doi.org/10.1109/TNET.2018. 2817206.

[4] Ballani, H.; Costa, P.; Haller, I.; Jozwik, K.; Shi, K.; Thomsen, B.C.; Williams, H. Bridging the Last Mile for Optical Switching in Data Centers. **2018**, W1C.3, https://doi.org/10.1364/ofc.2018.w1c.3.

[5] Dorren, H.J.S.; Wittebol, E.H.M.; de Kluijver, R.; de Villota, G.G.; Duan, P.; Raz, O. Challenges for Optically Enabled High-Radix Switches for Data Center Networks. *J. Light. Technol.* **2015**, *33*, 1117–1125, https://doi.org/10.1109/jlt.2015.2391301.

[6] Xue, X.; Nakamura, F.; Prifti, K. Pan, B.; Wang, F.; Guo, X.; Tsuda, H.; Calabretta, N. Experimental Assessments of a Flexible Optical Data Center Network Based on Integrated Wavelength Selective

Switch. In Proceedings of 2020 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 8–12 March 2020, https://doi.org/10.1364/OFC.2020. W1F.5.
[7] Yan, F.; Xue, X.; Pan, B.; Guo, X.; Calabretta, N. FOScube: a Scalable Data Center Network Architecture Based on Multiple Parallel Networks and Fast Optical Switches. **2018**, 1–3, https://doi.org/10.1109/ecoc. 2018.8535223.
[8] Wang, F.; Liu, B.; Xue, X.; Zhang, L.; Yan, F.; Magalhaes, E.; Zhang, Q.; Xin, X.; Calabretta, N. Demonstration of SDN-Enabled Hybrid Polling Algorithm for Packet Contention Resolution in Optical Data Center Network. *J. Light. Technol.* **2020**, *38*, 3296–3304, https://doi.org/10.1109/jlt.2020.2976549.
[9] Clark, K.A.; Cletheroe, D.; Gerard, T.; Haller, I.; Jozwik, K.; Shi, K.; Thomsen, B.; Williams, H.; Zervas, G.; Ballani, H.; et al. Synchronous subnanosecond clock and data recovery for optically switched data centres using clock phase caching. *Nat. Electron.* **2020**, *3*, 426–433, https://doi.org/10.1038/s41928-020-0423-y.
[10] Xue, X.; Pan, B.; Agraz, F.; Pagès, A.; Guo, X.; Yan, F.; Spadaro, S.; Calabretta, N. SDN-Enabled Reconfigurable Optical Data Center Network with Automatic Network Slicing to Provision Dynamic QoS. In proceedings of 2020 European Conference on Optical Communications (ECOC), Brussels, Belgium, 6–10 December 2020, https://doi.org/10.1109/ECOC48923.2020.9 333213.
[11] IBM. Data Centers. Available online: https://www.ibm.com/cloud/learn/data-centers (accessed on 2 June 2016).
[12] The age of analytics: competing in a data-driven world. Available online: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-age-of-analytics-competing-in-a-data-driven-world (accessed on 7 December 2016).
[13] Rimal, B.P.; Van, D.P.; Maier, M. Mobile edge computing empowered fiber-wireless access networks in the 5G era. *IEEE Commun. Mag.* **2017**, *55*, 192–200, https://doi.org/10.1109/MCOM.2017. 1600156CM.
[14] Marotta, A.; Avallone, S.; Kassler, A. A Joint Power Efficient Server and Network Consolidation approac h for virtualized data centers. *Comput. Netw.* **2018**, *130*, 65–80, https://doi.org/10.1016/j.comnet.2017. 11. 003.
[15] Cisco Global Cloud Index: Forecast and Methodology, 2016–2021. Available online: https://virtualization. network/Resources/Whitepapers/0b75cf2e-0c53-4891-918e-b542a5d364c5_white-paper-c11-738085.pdf (accessed on 18 January 2018).
[16] Wang, T.; Su, Z.; Xia, Y.; Hamdi, M. Rethinking the Data Center Networking: Architecture, Network Pro tocols, and Resource Sharing. *IEEE Access* **2014**, *2*, 1481–1496, https://doi.org/10.1109/access.2014.2383 439.
[17] Li, D.; Wu, J.; Liu, Z.; Zhang, F. Dual-centric Data Center Network Architectures. **2015**, 679–688, https://doi.org/10.1109/icpp.2015.77.
[18] Qu, G.; Fang, Z.; Zhang, J.; Zheng, S.-Q. Switch-Centric Data Center Network Structures Based on Hypergraphs and Combinatorial Block Designs. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *26*, 1154–1164, https://doi.org/10.1109/tpds.2014.2318697.
[19] Zafar, S.; Bashir, A.; Chaudhry, S.A. On implementation of DCTCP on three-tier and fat-tree data center network topologies. *SpringerPlus* **2016**, *5*, 1–18, https://doi.org/10.1186/s40064-016-2454-4.
[20] Conterato, M.d.S.; Ferreto, T.C.; Rossi, F.; Marques, W.d.S.; de Souza, P.S.S. Reducing energy consumption in SDN-based data center networks through flow consolidation strategies. **2019**, 1384–1391, https://doi.org/10.1145/3297280.3297420.
[21] Hammadi, A.; Mhamdi, L. A survey on architectures and energy efficiency in Data Center Networks. *Comput. Commun.* **2014**, *40*, 1–21, https://doi.org/10.1016/j.comcom.2013.11.005.
[22] Malla, S.; Christensen, K. A Survey on Power Management Techniques for Oversubscription of Multi-Tenant Data Centers. *ACM Comput. Surv.* **2019**, *52*, 1–31, https://doi.org/10.1145/3291049.
[23] Guo, Z.; Yang, Y. Multicast fat-tree data center networks with bounded link oversubscription. In Proceedi ng of 2013 Proceedings IEEE INFOCOM, Turin, Italy, 14–19 April 2013, https://doi.org/10.1109/INFCO M.2013.6566793.
[24] Li, X.; Lung, C.-H.; Majumdar, S. Green spine switch management for datacenter networks. *J. Cloud Comput.* **2016**, *5*, 1, https://doi.org/10.1186/s13677-016-0058-8.
[25] Surianarayanan, C.; Chelliah, P.R. Cloud Networking. **2019**, 97–132, https://doi.org/10.1007/978-3-030-13134-0_4.
[26] Valadarsky, A.; Shahaf, G.; Dinitz, M.; Schapira, M. Xpander. **2016**, 205–219, https://doi.org/10.1145/29 99572.2999580.
[27] Joy, S.; Nayak, A. Improving flow completion time for short flows in datacenter networks. **2015**, 700–705, https://doi.org/10.1109/inm.2015.7140358.

[28] Chair-Barcellos, M.G.; Chair-Crowcroft, J.G.; Chair-Vahdat, A.P.; Chair-Katti, S.P. In Proceedings of the 2016 ACM SIGCOMM Conference, Florianópolis, Brazil, 22–26 August 2016, https://doi.org/10.1145/2934872.

[29] Roy, A.; Zeng, H.; Bagga, J.; Porter, G.; Snoeren, A.C. Inside the Social Network's (Datacenter) Network. **2015**, https://doi.org/10.1145/2785956.2787472.

[30] Huang, J.; Huang, Y.; Wang, J.; He, T. Adjusting packet size to mitigate TCP Incast in data center networks with COTS switches. *IEEE Trans. Cloud Comput.* **2018**, *8*, 749–763, https://doi.org/10.1109/TCC.2018.2810870.

[31] Kandula, S.; Menache, I.; Schwartz, R.; Babbula, S.R. Calendaring for wide area networks. *ACM SIGCOMM Comput. Commun. Rev.* **2014**, *44*, 515–526, https://doi.org/10.1145/2740070.2626336.

[32] Zhang, H.; Chen, K.; Bai, W.; Han, D.; Tian, C.; Wang, H.; Guan, H.; Zhang, M. Guaranteeing Deadlines for Inter-Data Center Transfers. *IEEE/ACM Trans. Netw.* **2016**, *25*, 579–595, https://doi.org/10.1109/tnet.2016.2594235.

[33] Dong, X.-D.; Chen, S.; Zhao, L.-P.; Zhou, X.-B.; Qi, H.; Li, K.-Q. More Requests, Less Cost: Uncertain Inter-Datacenter Traffic Transmission with Multi-Tier Pricing. *J. Comput. Sci. Technol.* **2018**, *33*, 1152–1163, https://doi.org/10.1007/s11390-018-1878-4.

[34] Jalaparti, V.; Bodik, P.; Kandula, S.; Menache, I.; Rybalkin, M.; Yan, C. Speeding up distributed request-response workflows. *ACM SIGCOMM Comput. Commun. Rev.* **2013**, *43*, 219–230, https://doi.org/10.1145/2534169.2486028.

[35] Microsoft. Interrupt Moderation. Available online: https://docs.microsoft.com/en-us/windows-hardware/drivers/network/interrupt-moderation (accessed on 15 December 2021).

[36] Kapoor, R.; Snoeren, A.C.; Voelker, G.M. Bullet trains: a study of NIC burst behavior at microsecond timescales. In Proceedings of the ninth ACM conference on Emerging networking experiments and technologies, New York, NY, USA, 27 August 2013, https://doi.org/10.1145/2535372.2535407.

[37] Shan, D.; Ren, F.; Cheng, P.; Shu, R.; Guo, C. Observing and Mitigating Micro-Burst Traffic in Data Center Networks. *IEEE/ACM Trans. Netw.* **2019**, *28*, 98–111, https://doi.org/10.1109/tnet.2019.2953793.

[38] Wu, H.; Feng, Z.; Guo, C.; Zhang, Y. ICTCP: Incast Congestion Control for TCP in Data-Center Networks. *IEEE/ACM Trans. Netw.* **2012**, *21*, 345–358, https://doi.org/10.1109/tnet.2012.2197411.

[39] Han, F.; Wang, M.; Cui, Y.; Li, Q.; Liang, R.; Liu, Y.; Jiang, Y. Future Data Center Networking: From Low Latency to Deterministic Latency. *IEEE Netw.* **2022**, *36*, 52–58, https://doi.org/10.1109/mnet.102.2000622.

[40] Alizadeh, M.; Greenberg, A.; Maltz, D.A.; Padhye, J.; Patel, P.; Prabhakar, B.; Sengupta, S.; Sridharan, M. Data center TCP (DCTCP). *ACM SIGCOMM Comput. Commun. Rev.* **2010**, *40*, 63–74, https://doi.org/10.1145/1851275.1851192.

[41] Greenberg, A.; Hamilton, J.R.; Jain, N.; Kandula, S.; Kim, C.; Lahiri, P.; Maltz, D.A.; Patel, P.; Sengupta, S. VL2. *ACM SIGCOMM Comput. Commun. Rev.* **2009**, *39*, 51–62, https://doi.org/10.1145/1594977.1592576.

[42] Kandula, S.; Sengupta, S.; Greenberg, A.; Patel, P.; Chaiken, R. The nature of data center traffic. **2009**, 202–208, https://doi.org/10.1145/1644893.1644918.

[43] Benson, T.; Akella, A.; Maltz, D.A. Network traffic characteristics of data centers in the wild. In Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, Melbourne, Australia, 1–30 November 2010; pp. 267–280, https://doi.org/10.1145/1879141.1879175.

[44] Fiorani, M.; Tornatore, M.; Chen, J.; Wosinska, L.; Mukherjee, B. Spatial Division Multiplexing for High Capacity Optical Interconnects in Modular Data Centers. *J. Opt. Commun. Netw.* **2017**, *9*, A143–A153, https://doi.org/10.1364/jocn.9.00a143.

[45] Yuang, M.; Tien, P.-L.; Ruan, W.-Z.; Lin, T.-C.; Wen, S.-C.; Tseng, P.-J.; Lin, C.-C.; Chen, C.-N.; Chen, C.-T.; Luo, Y.-A.; et al. OPTUNS: Optical intra-data center network architecture and prototype testbed for a 5G edge cloud. *J. Opt. Commun. Netw.* **2019**, *12*, A28–A37, https://doi.org/10.1364/jocn.12.000a28.

[46] Son, J.; Buyya, R. A Taxonomy of Software-Defined Networking (SDN)-Enabled Cloud Computing. *ACM Comput. Surv.* **2018**, *51*, 1–36, https://doi.org/10.1145/3190617.

[47] Son, J.; Dastjerdi, A.V.; Calheiros, R.N.; Buyya, R. SLA-Aware and Energy-Efficient Dynamic Overbooking in SDN-Based Cloud Data Centers. *IEEE Trans. Sustain. Comput.* **2017**, *2*, 76–89, https://doi.org/10.1109/tsusc.2017.2702164.

[48] Gu, D. Evolving Cloud Data Centers to the 5G Era. Available online: http://www.telcotransformation.com/author.asp?section_id=390&doc_id=741079 (accessed on 23 September 2019).

[49] Chen, H.; Abbas, R. Cheng, P.; Shirvanimoghaddam, M.; Hardjawana, W.; Bao, W.; Li, Y.; Vucetic, B. Ultra-reliable low latency cellular networks: Use cases, challenges and approaches. *IEEE Commun. Mag.* **2018**, *56*, 119–125, https://doi.org/10.1109/MCOM.2018. 1701178.

[50] Chen, Y.; Gmach, D.; Hyser, C.; Wang, Z.; Bash, C.; Hoover, C.; Singhal, S. Integrated management of a pplication performance, power and cooling in data centers. In Proceeding of 2010 IEEE Network Operatio ns and Management Symposium-NOMS 2010, Osaka, Japan, 19–23 April 2010, https://doi.org/10.1109/ NOMS.2010. 5488433.

[51] Wu, K.; Xiao, J.; Ni, L.M. Rethinking the architecture design of data center networks. *Front. Comput. Sci.* **2012**, *6*, 596–603, https://doi.org/10.1007/ s11704-012-1155-6.

[52] Singh, A.; Ong, J.; Agarwal, A.; Anderson, G.; Armistead, A.; Bannon, R.; Boving, S.; Desai, G.; Felderman, B.; Germano, P.; et, al. Jupiter rising: a decade of clos topologies and centralized control in Google's datacenter network. *Commun. ACM* **2016**, *59*, 88–97, https://doi.org/10.1145/ 2829988.2787508.

[53] Zhang, J.; Yu, F.R.; Wang, S.; Huang, T.; Liu, Z.; Liu, Y. Load Balancing in Data Center Networks: A Su rvey. *IEEE Commun. Surv. Tutorials* **2018**, *20*, 2324–2352, https://doi.org/10.1109/comst.2018.2816042.

[54] Zhang, R.; He, Y.; Zhang, Y.; An, S.; Zhu, Q.; Li, X.; Su, Y. Ultracompact and low-power-consumption s ilicon thermo-optic switch for high-speed data. *Nanophotonics* **2020**, *10*, 937–945, https://doi.org/10.1515 /nanoph-2020-0496.

[55] Rumley, S.; Bahadori, M.; Polster, R.; Hammond, S.D.; Calhoun, D.M.; Wen, K.; Rodrigues, A.; Bergman, K. Optical interconnects for extreme scale computing systems. *Parallel Comput.* **2017**, *64*, 65–80, https://doi.org/10.1016/j.parco.2017.02.001.

[56] Fathololoumi, S.; Hui, D.; Jadhav, S.; Chen, J.; Nguyen, K.; Sakib, M.; Li, Z.; Mahalingam, H.; Asl, S.A.; Tang, N.N.; et al. 1.6 Tbps Silicon Photonics Integrated Circuit and 800 Gbps Photonic Engine for Switc h Co-Packaging Demonstration. *J. Light. Technol.* **2020**, *39*, 1155–1161, https://doi.org/10.1109/jlt.2020. 3039218.

[57] Manipatruni, S.; Nikonov, D.E.; Young, I.A. Beyond CMOS computing with spin and polarization. *Nat. Phys.* **2018**, *14*, 338–343, https://doi.org/10.1038/s41567-018-0101-4.

[58] Vahdat, A.; Liu, H.; Zhao, X.; Johnson, C. The Emerging Optical Data Center. **2011**, OTuH2, https://doi.org/10.1364/ofc.2011.otuh2.

[59] Wei, J.L.; Ingham, J.D.; Cunningham, D.G.; Penty, R.V.; White, I.H. Performance and Power Dissipation Comparisons Between 28 Gb/s NRZ, PAM, CAP and Optical OFDM Systems for Data Communication Applications. *J. Light. Technol.* **2012**, *30*, 3273–3280, https://doi.org/10.1109/jlt.2012.2213797.

[60] Benjamin, J.L.; Gerard, T.; Lavery, D.; Bayvel, P.; Zervas, G. PULSE: Optical Circuit Switched Data Center Architecture Operating at Nanosecond Timescales. *J. Light. Technol.* **2020**, *38*, 4906–4921, https://doi.org/10.1109/jlt.2020.2997664.

[61] Tang, Y.; Yuan, T. Effective*-flow schedule for optical circuit switching based data center networks: A comprehensive survey. *Comput. Netw.* **2021**, *197*, 108321, https://doi.org/10.1016/j.comnet.2021. 108321.

[62] Hasan, M.Z.; Hasan, K.Z.; Sattar, A. Burst header packet flood detection in optical burst switching networ k using deep learning model. *Procedia Comput. Sci.* **2018**, *143*, 970–977, https://doi.org/10.1016/j.procs. 2018.10.337.

[63] Poorzare, R.; Abedidarabad, S. A brief review on the methods that improve optical burst switching network performance. *J. Opt. Commun. Netw.* **2019**, https://doi.org/10.1515/ joc-2019-0092.

[64] Pointurier, Y.; Benzaoui, N.; Lautenschlaeger, W.; Dembeck, L. End-to-End Time-Sensitive Optical Net working: Challenges and Solutions. *J. Light. Technol.* **2019**, *37*, 1732–1741, https://doi.org/10.1109/jlt.20 19.2893543.

[65] Waheed, S. Comparing optical packet switching and optical burst switching. *DIU J. Sci. Technol.* **2011**, *6*, 22–32, https://doi.org/10.3329/diujst.v6i2. 9342.

[66] Singh, A.; Tiwari, A.K. Analysis of Hybrid Buffer Based Optical Data Center Switch. *J. Opt. Commun.* **2018**, *42*, 415–424, https://doi.org/10.1515/joc-2018-0121.

[67] Yoo, S.J.B.; Yin, Y.; Wen, K. Intra and inter datacenter networking: The role of optical packet switching and flexible bandwidth optical networking. **2012**, 1–6, https://doi.org/10.1109/ondm.2012.6210261.

[68] Testa, F.; Pavesi, L. Optical Switching in Next Generation Data Centers. Springer, Cham: Switzerland, 2017.

[69] Liu, M.; Wu, X.; Niu, Y.; Yang, H.; Zhu, Y.; Wang, W. Research Progress of MEMS Inertial Switches. *Micromachines* **2022**, *13*, 359, https://doi.org/10.3390/mi13030359.

[70] Keum, H.; Park, J.K.; Kim, S. Micro-Lego of 3D SU-8 structures and its application to a re-entrant surface. *J. Micro-Bio Robot.* **2018**, *14*, 17–23, https://doi.org/10.1007/s12213-018-0105-2.

[71] Furukawa, H. Petabit-class Optical Networks and Switching Technologies. In *OSA Advanced Photonics Congress 2021*, https://doi.org/10.1364/NETWORKS.2021.NeW1C. 1.
[72] Huang, Q. Commercial Optical Switches. **2017**, 203–219, https://doi.org/10.1007/978-3-319-61052-8_11.
[73] Harris, J.M.; Lindquist, R.; Rhee, J.; Webb, J.E. Liquid-Crystal Based Optical Switching. **2008**, 141–167, https://doi.org/10.1007/0-387-29159-8_5.
[74] Frisken, S.; Baxter, G.; Abakoumov, D.; Zhou, H.; Clarke, I.; Poole, S. Flexible and grid-less wavelength selective switch using LCOS technology. In Proceeding of 2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference, Los Angeles, CA, USA, 6–10 March 2011, https://doi.org/10.1364/OFC.2011. OTuM3.
[75] Khan, I.; Masood, M.U.; Tunesi, L.; Bardella, P.; Ghillino, E.; Carena, A.; Curri, V. A Neural Network-B ased Automatized Management of N ×N Integrated Optical Switches. **2021**, NeF2B.2, https://doi.org/10.1 364/networks.2021.nef2b.2.
[76] Wang, M.; Zong, L.; Mao, L.; Marquez, A.; Ye, Y.; Zhao, H.; Caballero, F.J.V. LCoS SLM Study and Its Application in Wavelength Selective Switch. *Photonics* **2017**, *4*, 22, https://doi.org/10.3390/photonics402 0022.
[77] Iwama, M.; Takahashi, M.; Kimura, M.; Uchida, Y.; Hasegawa, J.; Kawahara, R.; Kagi, N. LCOS-based Flexible Grid 1×40 Wavelength Selective Switch Using Planar Lightwave Circuit as Spot Size Converter. **2015**, Tu3A.8, https://doi.org/10.1364/ofc.2015.tu3a.8.
[78] Rakshit, J.; Chattopadhyay, T.; Roy, J. Design of ring resonator based all optical switch for logic and arith metic operations – A theoretical study. *Optik* **2013**, *124*, 6048–6057, https://doi.org/10.1016/j.ijleo.2013.0 4.075.
[79] Sugiyama, H.; Johmoto, K.; Sekine, A.; Uekusa, H. Reversible on/off switching of photochromic properties in *N*-salicylideneaniline co-crystals by heating and humidification. *CrystEngComm* **2019**, *21*, 3170–3175, https://doi.org/10.1039/c9ce00442d.
[80] Huang, Y.; Cheng, Q.; Hung, Y.-H.; Guan, H.; Novack, A.; Streshinsky, M.; Hochberg, M.; Bergman, K. Dual-Microring Resonator Based 8×8 Silicon Photonic Switch. **2019**, W1E.6, https://doi.org/10.1364/ofc.2019.w1e.6.
[81] Mendez-Astudillo, M.; Okamoto, M.; Ito, Y.; Kita, T. Compact thermo-optic MZI switch in silicon-on-insulator using direct carrier injection. *Opt. Express* **2019**, *27*, 899–906, https://doi.org/10.1364/OE.27. 000899.
[82] Lu, L.; Zhou, L.; Li, X.; Chen, J. Low-power 2×2 silicon electro-optic switches based on double-ring assi sted Mach–Zehnder interferometers. *Opt. Lett.* **2014**, *39*, 1633–1636, https://doi.org/10.1364/ol.39.00163 3.
[83] Qiao, L.; Tang, W.; Chu, T. 32 × 32 silicon electro-optic switch with built-in monitors and balanced-status units. *Sci. Rep.* **2017**, *7*, srep42306, https://doi.org/10.1038/srep42306.
[84] Ballani, H.; Costa, P.; Behrendt, R.; Cletheroe, D.; Haller, I.; Jozwik, K.; Karinou, F.; Lange, S.; Shi, K.; Thomsen, B.; et al. Sirius. **2020**, https://doi.org/10.1145/3387514.3406221.
[85] Fu, M.; Liu, G.; Proietti, R.; Zhang, Y.; Yoo, S.J.B. First Demonstration of Monolithic Silicon Photonic Integrated Circuit 32× 32 Thin-CLOS AWGR for All-to-All Interconnections. In Proceeding of 2021 European Conference on Optical Communication (ECOC), Bordeaux, France, 13–16 September 2021, https://doi.org/10.1109/ECOC52684. 2021.9605973.
[86] Matsuo, S.; Segawa, T. Microring-Resonator-Based Widely Tunable Lasers. *IEEE J. Sel. Top. Quantum Electron.* **2009**, *15*, 545–554, https://doi.org/10.1109/jstqe.2009.2014248.
[87] Cheung, S.; Su, T.; Okamoto, K.; Yoo, S.J.B. Ultra-Compact Silicon Photonic 512 × 512 25 GHz Arrayed Waveguide Grating Router. *IEEE J. Sel. Top. Quantum Electron.* **2013**, *20*, 310–316, https://doi.org/10.1 109/jstqe.2013.2295879.
[88] Ju, H.; Zhang, S.; Lenstra, D.; De Waardt, H.; Tangdiongga, E.; Khoe, G.D.; Dorren, H.J.S. SOA-based al l-optical switch with subpicosecond full recovery. *Opt. Express* **2005**, *13*, 942–947, https://doi.org/10.136 4/OPEX.13. 000942.
[89] Schares, L.; Huynh, T.N.; Wood, M.G.; Budd, R.; Doany, F.; Kuchta, D.; Dupuis, N.; Lee, B.G.; Schow, C.L.; Moehrle, M.; et al. A Gain-Integrated Silicon Photonic Carrier with SOA-Array for Scalable Optical Switch Fabrics. **2016**, Th3F.5, https://doi.org/10.1364/ofc.2016.th3f.5.
[90] Yamaguchi, M.; Yukimatsu, K.I.; Hiramatsu, A.; Matsunaga, T. Hyper-media photonic information networks as future network service platforms. *IEICE Trans. Electron.* **1999**, *82*, 170–178.
[91] Wonfor, A.; Wang, H.; Penty, R.V.; White, I.H. Large port count high-speed optical switch fabric for use within datacenters. *J. Opt. Commun. Netw.* **2011**, *3*, A32-A39, https://doi.org/10.1364/ JOCN.3.000A32.

[92] Stabile, R.; Albores-Mejia, A.; Williams, K.A. Monolithic active-passive 16× 16 optoelectronic switch. *Opt. Lett.* **2012**, *37*, 4666–4668, https://doi.org/10.1364/ OL.37.004666.
[93] Balanici, M.; Pachnicke, S. Hybrid Electro-Optical Intra-Data Center Networks Tailored for Different Traffic Classes. *J. Opt. Commun. Netw.* **2018**, *10*, 889–901, https://doi.org/10.1364/jocn.10.000889.
[94] Farrington, N.; Porter, G.; Radhakrishnan, S.; Bazzaz, H.H.; Subramanya, V.; Fainman, Y.; Papen, G.; Va hdat, A. Helios. *ACM SIGCOMM Comput. Commun. Rev.* **2010**, *40*, 339–350, https://doi.org/10.1145/185 1275.1851223.
[95] Kamchevska, V.; Medhin, A.K.; Da Ros, F.; Ye, F.; Asif, R.; Fagertun, A.M.; Ruepp, S.; Berger, M.; Ditt mann, L.; Morioka, T.; et al. Experimental Demonstration of Multidimensional Switching Nodes for All-Optical Data Center Networks. *J. Light. Technol.* **2016**, *34*, 1837–1843, https://doi.org/10.1109/jlt.2016.2 518863.
[96] Wang, G.; Andersen, D.G.; Kaminsky, M.; Papagiannaki, K.; Ng, T.S.E.; Kozuch, M.; Ryan, M. c-Through: Part-time optics in data centers. In Proceedings of the ACM SIGCOMM 2010 conference, New York, NY, USA, 30 August 2010, https://doi.org/10.1145/1851275. 1851222.
[97] Mellette, W.M.; McGuinness, R.; Roy, A.; Forencich, A.; Papen, G.; Snoeren, A.C.; Porter, G. Rotornet: A scalable, low-complexity, optical datacenter network. In Proceedings of the Conference of the ACM Special Interest Group on Data Communication, New York, NY, USA, 07 August. 2017, https://doi.org/10.1145/3098822. 3098838.
[98] Porter, G.; Strong, R.; Farrington, N.; Forencich, A.; Chen-Sun, P.; Rosing, T.; Fainman, Y.; Papen, G.; Vahdat, A. Integrating microsecond circuit switching into the data center. In Proceeding of ACM SIGCOMM Computer Communication Review, New York, NY, USA, 27 August 2013, https://doi.org/10. 1145/2486001.2486007.
[99] Chen, K.; Singla, A.; Singh, A.; Ramachandran, K.; Xu, L.; Zhang, Y.; Wen, X.; Chen, Y. OSA: An Optical Switching Architecture for Data Center Networks With Unprecedented Flexibility. *IEEE/ACM Trans. Netw.* **2013**, *22*, 498–511, https://doi.org/10.1109/tnet.2013.2253120.
[100] Singla, A.; Singh, A.; Ramachandran, K.; Xu, L.; Zhang, Y. Proteus: a topology malleable data center network. In Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks, New York, NY, USA, 20 October 2010, https://doi: .org/10.1145/1868447.1868455.
[101] Chen, K.; Wen, X.; Ma, X.; Chen, Y.; Xia, Y.; Hu, C.; Dong, Q. WaveCube: A scalable, fault-tolerant, high-performance optical data center architecture. In Proceeding of 2015 IEEE Conference on Computer Communications (INFOCOM), Hong Kong, China, 26 April – 1 May 2015, https://doi.org/10.1109/ INFOCOM.2015.7218573.
[102] Pal, A.; Kant, K. RODA: A reconfigurable optical data center network architecture. **2015**, 561–569, https://doi.org/10.1109/lcn.2015.7366371.
[103] Yuang, M.C.; Tien, P.-L.; Chen, H.-Y.; Ruan, W.-Z.; Hsu, T.-K.; Zhong, S.; Zhu, J.; Chen, Y.; Chen, J. OPMDC: Architecture Design and Implementation of a New Optical Pyramid Data Center Network. *J. Light. Technol.* **2015**, *33*, 2019–2031, https://doi.org/10.1109/jlt.2015.2390495.
[104] Khope, A.S.P.; Helkey, R.; Liu, S.; Saleh, A.A.M.; Alferness, R.C.; Bowers, J.E. A Scalable Multicast Hy brid Broadband Crossbar Wavelength Selective Switch for Datacenters. **2021**, 1585–1587, https://doi.org/ 10.1109/ccwc51732.2021.9376020.
[105] Gripp, J.; Simsarian, J.E.; LeGrange, J.D.; Bernasconi, P.; Neilson, D.T. Photonic Terabit Routers: The IRIS Project. **2010**, OThP3, https://doi.org/10.1364/ofc.2010.othp3.
[106] Ye, X.; Yin, Y.; Yoo, S.J.B.; Mejia, P.; Proietti, R.; Akella, V. DOS: A scalable optical switch for datacenters. In Proceedings of the 6th ACM/IEEE Symposium on Architectures for Networking and Communications Systems, La Jolla, CA, USA, 25–26 October 2010, https://doi.org/10.1145/ 1872007.1872037.
[107] Xi, K.; Kao, Y.-H.; Chao, H.J. A Petabit Bufferless Optical Switch for Data Center Networks. **2012**, 135–154, https://doi.org/10.1007/978-1-4614-4630-9_8.
[108] Yin, Y.; Proietti, R.; Ye, Y.; Nitta, C.J.; Akella, V.; Yoo, S.J.B. LIONS: An AWGR-based low-latency optical switch for high-performance computing and data centers. *IEEE J. Sel. Top. Quantum Electron.* **2012**, *19*, 3600409–3600409, https://doi.org/10.1109/ JSTQE.2012.2209174.
[109] Cao, Z.; Proietti, R.; Yoo, S.J.B. Hi-LION: Hierarchical large-scale interconnection optical network with AWGRs. *J. Opt. Commun. Netw.* **2015**, *7*, A97–A105, https://doi.org/10.1364/ JOCN.7.000A97.
[110] Cao, Z.; Proietti, R.; Clements, M.; Ben Yoo, S.J. Experimental Demonstration of Flexible Bandwidth Optical Data Center Core Network With All-to-All Interconnectivity. *J. Light. Technol.* **2015**, *33*, 1578–1585, https://doi.org/10.1109/jlt.2014.2387205.

[111] Jokar, M.R.; Qiu, J.; Chong, F.T.; Goddard, L.L; Dallesasse, J.M.; Feng, M.; Li, Y. Baldur: A power-efficient and scalable network using all-optical switches. In Proceedings of 2020 IEEE International Symposium on High Performance Computer Architecture (HPCA), San Diego, CA, USA, 22–26 February 2020, https://doi: .org/10.1109/hpca47549.2020. 00022.
[112] Xue, X.; Wang, F.; Agraz, F.; Pages, A.; Pan, B.; Yan, F.; Guo, X.; Spadaro, S.; Calabretta, N. SDN-Controlled and Orchestrated OPSquare DCN Enabling Automatic Network Slicing With Differentiated QoS Provisioning. *J. Light. Technol.* **2020**, *38*, 1103–1112, https://doi.org/10.1109/jlt.2020.2965640.
[113] Yan, F.; Xue, X.; Calabretta, N. HiFOST: A scalable and low-latency hybrid data center network architecture based on flow-controlled fast optical switches. IEEE/OSA Journal of Optical Communications Networking, 2018, vol. 10, no. 7, pp. 1-14, DOI: 10.1364/JOCN. 10.0000B1.
[114] Liboiron-Ladouceur, O.; Shacham, A.; Small, B.A.; Lee, B.G.; Wang, H.; Lai, C.P.; Biberman, A.; Bergman, K. The Data Vortex Optical Packet Switched Interconnection Network. *J. Light. Technol.* **2008**, *26*, 1777–1789, https://doi.org/10.1109/jlt.2007.913739.
[115] Xue, X.; Yan, F.; Prifti, K.; Wang, F.; Pan, B.; Guo, X.; Zhang, S.; Calabretta, N. ROTOS: A Reconfigurable and Cost-Effective Architecture for High-Performance Optical Data Center Networks. *J. Light. Technol.* **2020**, *38*, 3485–3494, https://doi.org/10.1109/jlt.2020.3002735.
[116] Luijten, R.P.; Grzybowski, R. The OSMOSIS Optical Packet Switch for Supercomputers. **2009**, OTuF3, https://doi.org/10.1364/ofc.2009.otuf3.
[117] Xue, X.; Wang, F.; Agraz, F.; Pages, A.; Pan, B.; Yan, F.; Spadaro, S.; Calabretta, N. Experimental Assessment of SDN-enabled Reconfigurable OPSquare Data Center Networks with QoS Guarantees. **2019**, M3F.4, https://doi.org/10.1364/ofc.2019.m3f.4.
[118] Xue, X.; Nakamura, F.; Prifti, K.; Pan, B.; Yan, F.; Wang, F.; Guo, X.; Tsuda, H.; Calabretta, N. SDN ena bled flexible optical data center network with dynamic bandwidth allocation based on photonic integrated wavelength selective switch. *Opt. Express* **2020**, *28*, 8949–8958, https://doi.org/10.1364/oe.388759.
[119] Xue, X.; Prifti, K.; Pan, B.; Yan, F.; Guo, X.; Calabretta, N. Fast Dynamic Control of Optical Data Center Networks Based on Nanoseconds WDM Photonics Integrated Switches. **2019**, https://doi.org/10.23919/p s.2019.8817870.
[120] Miao, W.; Luo, J.; Di Lucente, S.; Dorren, H.; Calabretta, N. Novel flat datacenter network architecture based on scalable and flow-controlled optical switch system. *Opt. Express* **2014**, *22*, 2465–2472, https://doi.org/10.1364/oe.22.002465.
[121] Soysouvanh, S.; Jalil, M.A.; Amiri, I.S.; Ali, J.; Singh, G.; Mitatha, S.; Yupapin, P.; Grattan, K.T.V.; Yoshida, M. Ultra-fast electro-optic switching control using a soliton pulse within a modified add-drop multiplexer. *Microsyst. Technol.* **2018**, *24*, 3777–3782, https://doi.org/10.1007/s00542-018-3837-y.
[122] Xue, X.; Calabretta, N. Synergistic Switch Control Enabled Optical Data Center Networks. *IEEE Commun. Mag.* **2022**, *60*, 62–67, https://doi.org/10.1109/mcom.001.2100683.
[123] Xue, X.; Pan, B.; Guo, X.; Calabretta, N. Flow-controlled and Clock-distributed Optical Switch and Control System. *IEEE Trans. Commun.* **2022**, *70*, 3310–3319, https://doi.org/10.1109/TCOMM.2022. 3156613.
[124] Xue, X.; Wang, F.; Chen, S.; Yan, F.; Pan, B.; Prifti, K.; Guo, X.; Zhang, S.; Xie, C.; Calabretta, N. Experimental Assessments of SDN-Enabled Optical Polling Flow Control for Contention Resolution in Optical DCNs. *J. Light. Technol.* **2020**, *39*, 2652–2660, https://doi.org/10.1109/jlt.2020.3042820.
[125] Rangarajan, S.; Hu, Z.; Rau, L.; Blumenthal, D.J. All-optical contention resolution with wavelength conversion for asynchronous variable-length 40 Gb/s optical packets. *IEEE Photon. Technol. Lett.* **2004**, *16*, 689–691, https://doi.org/10.1109/LPT.2003. 819361.
[126] Porzi, C.; Chin, S.; Trita, A.; Fresi, F.; Berrettini, G.; Mezosi, G.; Ghelfi, P.; Giuliani, G.; Poti, L.; Sorel, M.; et al. Application of Brillouin-Based Continuously Tunable Optical Delay Line to Contention Resolut ion Between Asynchronous Optical Packets. *J. Light. Technol.* **2013**, *31*, 2888–2896, https://doi.org/10.11 09/jlt.2013.2275253.
[127] Yao, S.; Mukherjee, B.; Yoo, S.; Dixit, S. A unified study of contention-resolution schemes in optical packet-switched networks. *J. Light. Technol.* **2003**, *21*, 672–683, https://doi.org/10.1109/jlt.2003.809573.
[128] Miao, W.; Di Lucente, S.; Luo, J.; Dorren, H.; Calabretta, N. Low latency and efficient optical flow contr ol for intra data center networks. *Opt. Express* **2014**, *22*, 427–434, https://doi.org/10.1364/oe.22.000427.
[129] Xu, L.; Perros, H.G.; Rouskas, G. Techniques for optical packet switching and optical burst switching. *IEEE Commun. Mag.* **2001**, *39*, 136–142, https://doi.org/10.1109/35.894388.
[130] Xue, X.; Calabretta, N. Nanosecond optical switching and control system for data center networks. *Nat. Commun.* **2022**, *13*, 1–8, https://doi.org/10.1038/s41467-022-29913-1.

[131] Clark, K.; Ballani, H.; Bayvel, P.; Cletheroe, D.; Gerard, T.; Haller, I.; Jozwik, K.; Shi, K.; Thomsen, B.; Watts, P. Sub-nanosecond clock and data recovery in an optically-switched data centre network. In Proceedings of 2018 European Conference on Optical Communication (ECOC), Roma, Italy, 23–27 September 2018, https://doi.org/10.1109/ECOC.2018. 85353 33.
[132] Jafarbeiki, S.; Hajsadeghi, K.; Modir, N. A 20 Gb/s Injection-Locked Clock and Data Recovery Circuit. *Int. J. VLSI Des. Commun. Syst.* **2014**, *5*, 1–12, https://doi.org/10.5121/vlsic.2014.5401.
[133] Xue, X.; Yan, F.; Pan, B.; Calabretta, N. Flexibility assessment of the reconfigurable OPSquare for virtualized data center networks under realistic traffics. In Proceedings of 2018 European Conference on Optical Communication (ECOC), Rome, Italy, 23–27 September 2018, https://doi.org/10.1109/ ECOC.2018.8535451.
[134] Xue, X.; Prifti, K.; Wang, F.; Yan, F.; Pan, B.; Guo, X.; Calabretta, N. SDN-Enabled Reconfigurable Optical Data Center Networks Based on Nanoseconds WDM Photonics Integrated Switches. In Proceedings of 2019 21st International Conference on Transparent Optical Networks (ICTON), Angers, France, 9–13 July 2019, https://doi.org/10.1109/ICTON.2019. 8840293.
[135] Tag: OpenFlow. Available online: https://www.opennetworking.org/tag/openflow/ (accessed on 3 March 2021).