

Chapter 2

Optical Switching in Datacenters: Architectures Based on Optical Circuit Switching

Liam P. Barry, Jingyan Wang, Conor McArdle, and Dan Kilper

2.1 Introduction

The proliferation of modern computer applications such as cloud computing, social networking services, multimedia streaming, and the Internet of Things (IoT) creates thousands of terabytes of heterogeneous data every minute [1], creating huge volumes of complex workload in datacenter servers and vast aggregate traffic loads across datacenter networks. Internet content providers such as Google, Facebook, Microsoft, Amazon, and Apple have installed mega datacenters which house hundreds of thousands, even millions, of servers in very large-scale layouts. As well as this scale-out, server network interfaces and top-of-rack (ToR) switches are being scaled up, from 10 to 25 Gbps and 100 Gbps data rates, to keep pace with the required workload rates and increasing service capacity [2]. Big data and cloud computing in particular have been shifting datacenter workloads from north-south to east-west, meaning that more data is flowing within the datacenter than in and out of the datacenter, driving up the internal datacenter network capacities. These trends are presenting substantial challenges to future-proofing of datacenter-scale networking, in terms of the required latency, bandwidth capacity, server connectivity, energy and cost efficiency, network configuration, and control complexity. Furthermore, as datacenters run increasingly mixed applications with more complex workloads, traffic complexity will evolve rapidly and exhibit more diverse and unpredictable communication patterns [3–7]. Thus, dynamic bandwidth-on-demand provisioning needs to be made possible for future datacenter network fabrics. Solutions toward highly scalable, high data rate, low-latency, agile datacenter

L.P. Barry (✉) • J. Wang • C. McArdle
School of Electronic Engineering, Dublin City University, Dublin 9, Ireland
e-mail: liam.barry@dcu.ie

D. Kilper
College of Optical Sciences, The University of Arizona,
1630 E. University Blvd., Tucson, AZ 85721-0094, USA

network infrastructure, with on-demand service provisioning capacity that responds quickly and dynamically to changing application requirements, are urgently required in order to maintain communication performance for building-scale and distributed multisite server installations of the future.

Today's large-scale datacenter networks are constructed by interconnecting a massive number of commodity hardware components, such as low-radix electronic packet switches, transceivers, and optical-electronic-optical (O/E/O) converters and optical fibers, into multi-tier interconnection topologies. Datacenter networks are typically organized as three-layer fat-tree [8] or Clos [9] topologies, where racks of servers are attached directly to a network access layer composed of top-of-rack (ToR) switches. A network aggregation layer collects traffic flows from the access layer and then forwards them to a core layer. Commonly, a datacenter network employs a Clos-based topology with centralized control, where commodity switches based on merchant silicon are arranged into multiple stages to form a large multi-stage switching fabric [10]. For example, Google has exploited a modular design concept, where a pool of machine servers and network devices are grouped into a cluster networking block to enable fine-grained resource provisioning and boosted operation efficiency [10, 11]. Facebook's current datacenter networks follow a hierarchical treelike networking model composed of three layers of electronic packet switches, but are now migrating toward a three-layer folded-Clos networking model with no oversubscription between racks [12, 13]. This network infrastructure is also designed with a modular structure using clusters or pods. Furthermore, a new two-tier datacenter network platform, the leaf-spine architecture, is emerging as an alternative to the traditional three-tier network topology. The leaf-spine architecture is a two-layer folded-Clos network comprising leaf-and-spine layers with each leaf switch connected to each of the spine switches [14]. Other electronic network architectures such as DCell [15], BCube [16], and FiConn [17] have been proposed to alleviate the bottlenecks of the three-layer fat-tree networks by building a recursive interconnection architecture.

The infrastructure scaling of these hierarchical networks, to accommodate the ever-growing number of servers, is potentially viable by adding more commodity network components and increasing the number of switching stages. For example, a new Facebook datacenter network with 100,000 servers placed in racks of 48 servers requires 2084 rack switches, each configured with 64×10 Gbps port capacity and 176 aggregation switches, each with 384×10 Gbps port capacity and 192 core switches, and each of which has at least 176×10 Gbps port capacity. Networking such a large number of servers, using electronic packet switches and links, imposes significant challenges on performance scalability, energy consumption cost, flexibility, network configuration, and management complexity. These issues are further exacerbated by the unprecedented datacenter traffic growth and the exploding bandwidth demands, which are rapidly outpacing the capacity of the state-of-the-art electronic switches, thereby necessitating the scaling of data rates, connectivity, and bandwidth capacity in the electronic switches. Electronic switch design currently faces switch port and capacity constraints due to high costs and excessive amounts of power consumption [18]. Ultimately, scaling of electronically switched

datacenters calls for a significant upgrading/replacement of the interconnection networks in order to support heterogeneous hardware and speeds. Furthermore, with server capacity scaling and communication speed rising, the higher interconnection level of the multilayer datacenter topology faces substantial resource stranding. That is, the aggregated traffic demands at higher levels of the system hierarchy (aggregation and core layers) potentially require progressively larger bandwidth capacities and consequently put the statically provisioned and oversubscribed upper tiers under strain. These limitations are stressing the existing datacenter networks and motivating the network to evolve toward more advanced networking solutions, augmenting innovative networking technologies, topological structures, and network management techniques.

A constructive approach toward the aforementioned design challenges involves application of optical switching and networking technologies. The most important attribute of optical switching is ultrahigh bandwidth capacity, which would provide unprecedented data carrying capacity for datacenter networking. Additionally, optical switching is potentially a low-latency communication technology, possibly achieving nanosecond-scale switching speed [19]. Further, optical switching can potentially future-proof datacenter network infrastructure, due to transparency to data rates, modulation formats, and protocols, and thus, except at the edge, does not require upgrading/replacement of the underlying physical network components as network link capacities evolve from 10 to 25 Gbps and 100 Gbps and beyond. Low power consumption is another prominent feature of optical switching technology. For example, a commercial 320×320 3D MEMS-based optical switch typically consumes 45 watts [20], whereas an Ethernet packet switch with a maximum 224×10 Gbps ports requires a typical power consumption of 1363 watts [21]. This contrast suggests that, compared to electronic switching, optical switching potentially offers substantial economic and technical benefits. Also at the link level, all-optical switching eliminates the need for expensive, power-consuming optical-electronic-optical (O/E/O) conversions which are necessary to deploy point-to-point optical links in electronically switched networks. The advantages of low loss, low power, and the high capacity of optics, coupled with the rapid development of photonic integration technology, highlight the opportunity to fabricate compact, high-stability, low-cost, and energy-efficient large-radix optical switching fabrics. Deploying high-radix optical switches in massive-scale datacenter networks would significantly flatten the network topology and simplify the network architecture and control complexity. From the network management perspective, software-defined networking (SDN) is becoming progressively common in datacenter networks and in the optical switching domain in general. This emerging trend facilitates the realization of an agile, flexible, and scalable optical datacenter interconnection network, with SDN enabling coordinated control of network resources and support for on-demand capacity provisioning by flexibly allocating network capacities to dynamically changing application demands. The abovementioned benefits indicate that an agile optically switched datacenter network has the potential to address current and future datacenter application demands, making it a suitable substitution for electronic networks. Indeed, research effort in this area has seen considerable growth in recent years.

In this chapter, the main focus is on reviewing the previous approaches to integrating mature optical circuit switching technologies into datacenter networks, to facilitate near- and medium-term expansion of datacenter and cloud computing capacity. In these approaches, high-radix optical circuit switches are interconnected into massive-scale datacenter interconnects using various fundamental networking structures such as Clos [9], Spanke [29], Benes [30], and flattened butterfly [31]. The current proposals are reviewed and discussed in the next section. Following that, a novel, agile datacenter network architecture, which supports dynamic and efficient sharing of network resources and flow-level provisioning, is presented. By exploiting the potential benefits of the passive fast-speed, low-radix flexible arrayed waveguide grating (AWG) switches coupled with high port-count optical circuit switches (OCSs), the proposed dynamic network offers the prospect of building a highly scalable, highly flexible, efficient large-scale datacenter network. We conclude the chapter with a perspective on how the current trends in OCS-based networks are expected to shape the future directions and capabilities of datacenter networking.

2.2 Optical Circuit Switching in Datacenter Networks

Optical switching technologies can be generally classified into two categories, optical circuit switching (OCS) and optical packet (or burst) switching (OPS), depending on the switching granularity. Optical circuit switching (OCS) is a relatively mature commercially available technology. Typically, an OCS is built based on optical 2D or 3D microelectromechanical systems (MEMS) technology. Currently, commercial 3D MEMS OCS supports up to 320 ports [20], and Polatis beam-steering OCS supports 384 ports [22], with high capacity, low loss, and low energy consumption. Port counts of more than 1000 have been offered previously for telecom applications and may become viable again for datacenters. OCS is a coarse-grained switching technique operating at the granularity of a full optical fiber [23]. That is, a dedicated point-to-point optical connection between an input and output fiber pair is set up for the entire data transmission session, providing guaranteed uncontended bandwidth [24] and ensuring quality of service (QoS). Nonetheless, the circuit-oriented configuration of the 3D MEMS-based OCS exhibits slow switch reconfiguration time, in the order of milliseconds. Microsecond switching times have been reported in experimental demonstrations of smaller-port-count optical space switches, such as a 64×64 compact silicon photonic MEMS module [25] and a 24×24 OCS built of 2D MEMS wavelength selective switch (WSS) modules [34]. Slow OCS potentially creates significant configuration overhead for traffic flows, and, as such, high-capacity optical circuit switching has previously been proposed for transferring only highly aggregated bulk datasets where the data transmission period is significantly longer than the switch setup time overhead. Slow switching time optical circuit switches exhibits relatively low flexibility and potentially low average utilization if employed in lower network tiers, due to the inability to efficiently handle dynamic, bursty traffic streams. In comparison, optical packet switching (OPS) is a

fine-grained, flexible switching paradigm which supports fast optical switching at packet (or burst) level. Optical packet switching is expected to be able to fully exploit the advantages and potential of optical switching, and OPS has been extensively studied in a number of large-scale research projects including WASPNET [26], LIGHTNESS [27, 28], and in [19]. However, the required optics and photonic technologies are generally considered not yet mature enough to fabricate a commercial large-scale optical packet switch. The main technical challenges faced by OPS involve high-speed optical packet header processing, lack of efficient optical packet buffering mechanisms, optical wavelength conversion and regeneration, scalability, and reliability. These difficulties would seem to make the wide deployment of OPS in commercial, cost-sensitive networks unachievable in the very near future. The deployment of optical circuit switches as optical cross-connects (OXCs) in datacenter-scale networking has instead been considered the first step to realizing scalable, transparent optical datacenter infrastructure. Extensive research efforts have been devoted to expanding the applicability of OCS from high-speed core telecommunication networks (which is implemented with wavelength switching) to datacenter networking. The major research directions include hybrid electronic/OCS datacenter networks [32–38]; microsecond OCS networks [34–36]; wavelength-, space-, or time-division multiplexing (WDM/SDM/TDM) OCS datacenter networks [39–41]; software-defined networking (SDN)-controlled OCS networks [20, 22, 39–44]; flexible fixed-grid OCS networks [45, 46]; elastic OCS datacenter networks [48, 49]; hybrid OCS/OPS interconnects [27, 28, 51]; hybrid wireless/wired datacenter networks; and various combinations of these emerging trends.

Hybrid electronic/OCS networks [32, 33] integrate advanced OCS technology into existing well-established electronic packet switching (EPS) network architecture, with EPS accommodating short-lived bursty traffic and slow, high-port-count OCS targeting large data transfers. In OCS-based datacenter network design [34–36], the driving goal is to achieve relatively fast microsecond switching speeds. This flexibility allows the OCS to efficiently handle more dynamic traffic patterns and route a larger fraction of the datacenter traffic, in comparison to traditional OCS networks which are expected to switch only in the millisecond range. Advances in WDM/SDM/TDM transmission and switching technologies motivate the exploitation of these technologies in datacenter networking [39–41]. Deploying WDM/SDM/TDM technologies in high-capacity OCS networks has led to significant benefits regarding network capacity, flexibility, and scalability. This flexibility is supported by SDN, a unified network control and management paradigm, which decouples the network control plane from the underlying switching and routing data plane, places the network intelligence into a logically centralized control system, and applies software-programmable functionality in the networking devices to facilitate the deployment of new network applications [42]. SDN promises network configuration optimization, high-level network flexibility, efficient capacity utilization, and guaranteed application performance.

Flexible fixed-grid OCS networks [45, 46] attempt to promote wavelength assignment capacity by augmenting the OCS with flexible fixed-grid optical wavelength switching and scheduling technologies, i.e., wavelength selective switching (WSS),

which allow an arbitrary wavelength to be switched to any output port of a WSS device. To further incorporate an even higher level of network elasticity, elastic OCS datacenter networks [48, 49] supporting arbitrary modulation formats and dynamic optical spectrum allocation with channel capacity ranging from subchannels to super-channels have been designed to optimize flexibility and efficiency. Transparent hybrid OCS/OPS datacenter networks [27, 28, 51] rely on both OPS and OCS to realize the central all-optical switching matrix. This scheme aims to combine the merits of OCS and OPS so as to support finer switching granularity. Another emerging trend is the possible deployment of the above-described wired network and the advanced wireless technologies such as free-space optics (FSO) [52] and 60 GHz wireless technology [53–54]. Recently, a MEMS-based approach was used to implement a high-radix OCS in free space across a datacenter [55]. With radix above 10,000, this approach enables a flat architecture with the benefits of OCS switching. In this section, several OCS-based datacenter architectures in the research literature are reviewed, and the opportunities and challenges of these proposals are discussed.

c-Through [32] and Helios [33] are two examples of hybrid EPS/OCS datacenter networks. The system-level structures of the c-Through and Helios networks are illustrated in Figs. 2.1 and 2.2, respectively. These hybrid network configurations combine a traditional hierarchical electronic packet switching network with a complementary optical circuit switching network, with the EPS network mainly managing small traffic flows and the OCS supporting large-volume, aggregated data transfers with guaranteed bandwidth. As demonstrated in Fig. 2.1, the c-Through electronic network follows a three-tier spanning-tree topology, whereas Helios deploys a two-level multi-rooted networking structure (Fig. 2.2). The optical circuit-switched network is built of slow, high-capacity MEMS-based OCS which provides direct interconnections between access switches. Helios also makes use of wavelength-division multiplexing (WDM), which is highly advantageous as it expands the capacity and flexibility of optical circuit provisioning. An important implication of the hybrid designs is that the collaborative operations of two independent switch fabrics to optimize network resources necessitate the need for traffic estimation and demultiplexing, which is conducted in c-Through end hosts and Helios switches, respectively. The hybrids c-Through and Helios offer the prospect of efficient networking by exploiting the best of EPS and OCS technologies, and, to some extent, they alleviate the oversubscription, reliability, scalability, flexibility, and capacity issues in current datacenters without requiring a complete replacement of hardware equipment (Fig. 2.3).

The Helios research group has further designed Mordia [34, 35] and REACToR [36]. Mordia is a prototype demonstration of a hybrid EPS/OCS network. One of the most significant characteristics of Mordia is that it adopts microsecond OCS networking, which is designed based on multiple 2D MEMS wavelength selective switches (WSSs). The microsecond OCS supports substantially higher switching speed, 2–3 orders of magnitude faster than commercial 3D MEMS OCS, and thus can potentially mitigate the slow switching issue and high buffering/aggregation requirement in traditional hybrid EPS/OCS datacenter networks. The topological structure of the Mordia OCS prototype follows a unidirectional ring supporting all-to-all connectivity and arbitrary input/output mapping. Additionally, a novel control algorithm called traffic matrix scheduling (TMS) is proposed, which is compatible with the

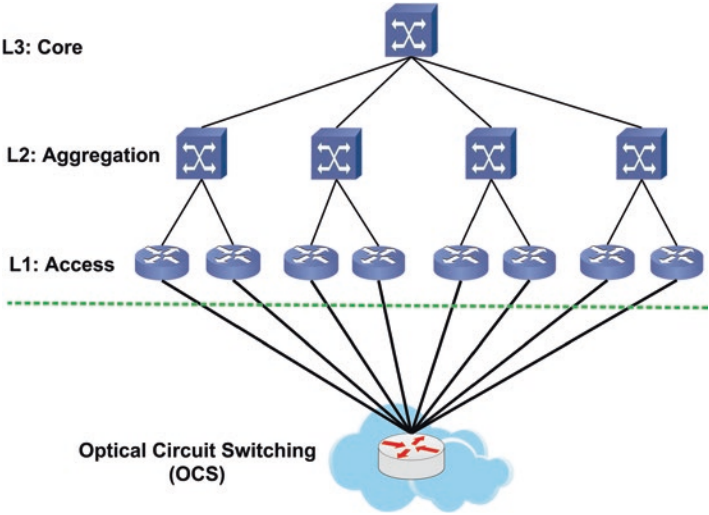


Fig. 2.1 Hybrid c-Through network

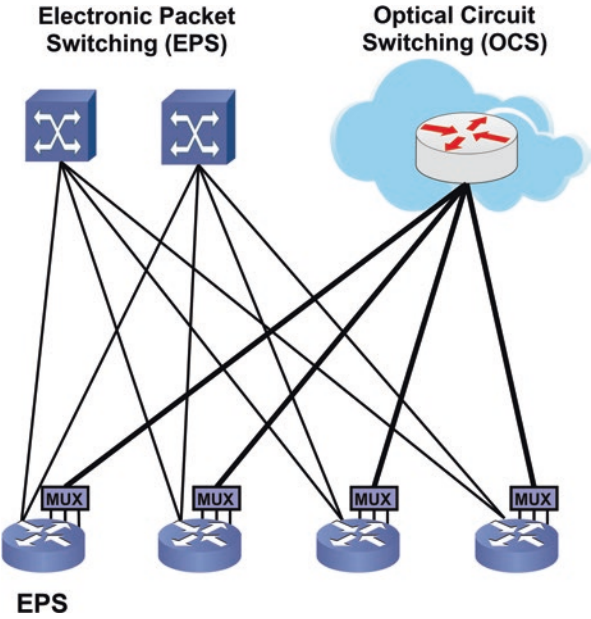


Fig. 2.2 Hybrid Helios network

microsecond-latency OCS. The TMS predicts traffic demands and then schedules short-time circuits to carry these demands. The short-lived traffic scheduling, coupled with microsecond circuit switching, greatly boosts the switching flexibility and capacity, making the OCS react rapidly to changing traffic patterns. The 24-port architectural demonstration is tested on a small-size network comprising 23 nodes.

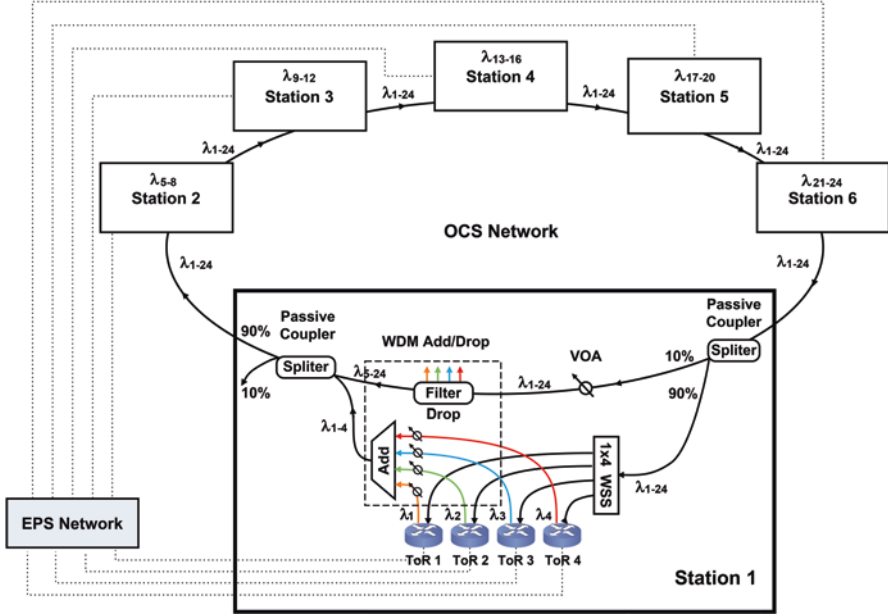


Fig. 2.3 The Mordia network [34, 35]

In large-scale datacenter networks, the implementations of the WSS-based OCS and its corresponding control algorithm may not be sufficiently scalable, due to the fact that the Mordia WSS-based OCS is a DWDM ring assigning one wavelength per port. As a consequence, scaling beyond 88 DWDM channels imposes significant technical and economic challenges.

The ideas behind Mordia, using WSS-based OCS delivering optical circuit switching at microsecond speeds, bring OCS closer to emulating electronic packet switching speeds, capable of supporting more dynamic traffic patterns, compared to the commercial OCS. Motivated by this, the prototype hybrid REACToR architecture [36] has been demonstrated on the basis of the Mordia datacenter network. In REACToR, the buffering and scheduling are mainly performed in the end hosts where the data packets are aggregated, grouped, and stored in transmission queues on a per-destination-ToR basis. Once an optical circuit path has been set up from the REACToR to an end host, the REACToR schedules the appropriate source end host to constantly pump data packets into the network through the established end-to-end optical circuit. In this architecture, direct, high-capacity, flow-level optical connections are provisioned between end hosts to serve high-volume server-to-server traffic, rather than serve rack-to-rack traffic and being shared by many flows as in traditional hybrid networks. The high-speed WSS-based OCS can be reconfigured to adapt to the changing traffic patterns of datacenter applications. During the reconfiguration times of the OCS, the EPS network is used to transmit host-to-host data traffic. This hybrid operation ensures non-degraded performance of the core network.

Thus, by integrating high-speed OCS supporting microsecond switching times with the existing high-speed, flexible EPS network, the REACToR hybrid EPS/OCS network, coupled with an efficient control plane, functions similarly to packet-switched ToR switches and supports fine-grained scheduling and thus can effectively schedule dynamic, rack-level traffic demands and achieve high link bandwidth utilization. This facilitates an upgrade path for existing EPS networks to higher data rates, higher flexibility, and better performance with significant economic benefits. Nevertheless, in addition to the previously described design challenges of a large-scale Mordia OCS network, a major problem that remains is how to efficiently interconnect a large number of the REACToR switches in a massive-scale datacenter, with effective global scheduling, control, and synchronization.

The single-stage shuffle-exchange (SSX) architecture [37, 38] is another hybrid EPS/OCS datacenter architecture. Unlike c-Through and Helios, the SSX scheme augments EPS and OCS in the same interconnection framework where the OCS functions as the central switching fabric that directly interconnects edge EPSs, as shown in Fig. 2.4. The SSX architecture operates on the principle that the *exchange* is performed first and then the *shuffling*. More precisely, an optical signal injected to an EPS is first switched to the appropriate output port of the EPS which directly leads to an input port of the OCS, and then, through the OCS, the optical signal is delivered to the destination EPS.

With the development of high-radix OCS, the SSX architecture now has the port densities required for large-scale datacenters. Nevertheless, the MEM-based OCS, as discussed previously, exhibits coarse-grained switching granularity and long reconfiguration times, which partially offset the benefits of high-capacity optical switching. To avoid constant reconfiguration and to expand the connectivity of the OCS, SSX deploys a hop-by-hop routing strategy which allows communications between two EPSs which are not currently being interconnected by the OCS. However, to do this, the input signal needs to travel through multiple EPSs and undergoes multiple optical-electronic-optical conversions before reaching its destination EPS, which consequently imposes increased communication latency and power consumption. Hence, a trade-off is presented by the SSX architecture.

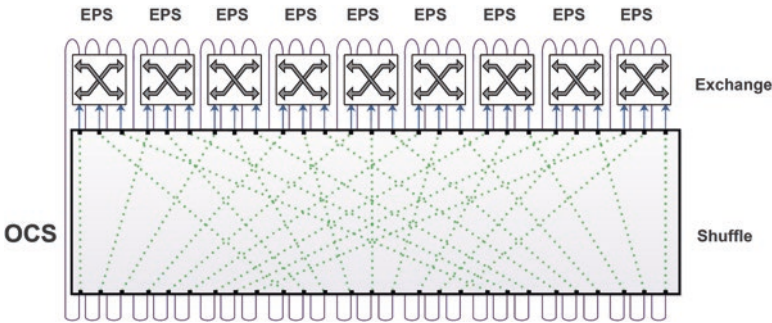


Fig. 2.4 Hybrid SSX architecture [37, 38]

[39, 40] propose a converged EPS/OCS datacenter network with SDN control, where an all-optical circuit-switched core network interconnects ToR electronic switches (EPSs), as illustrated in Fig. 2.5. The OCS network is designed with a flattened butterfly topology [31], where the basic building modules—optical virtual switches (OvS) enabled by optical wavelength switching—are arranged into a 2D array. In the OvS architecture, the optical wavelengths injected from the transceivers in a ToR are combined into a DWDM channel by an optical multiplexer (MUX) and subsequently forwarded to two passive tap coupler tree modules called passive routing fabric (PRF) blocks, which broadcast the DWDM signal in two dimensions. The receiving block of the OvS is mainly based on an optical wavelength switching component, i.e., WSS, which is flexibly configured to allow the desired wavelengths to pass through and block all other wavelengths. The passing wavelengths are then routed through an optical demultiplexer (DEMUX) toward the ToR. Essentially, the OvS is a broadcast-and-select (B&S) building block exploiting the advantages of

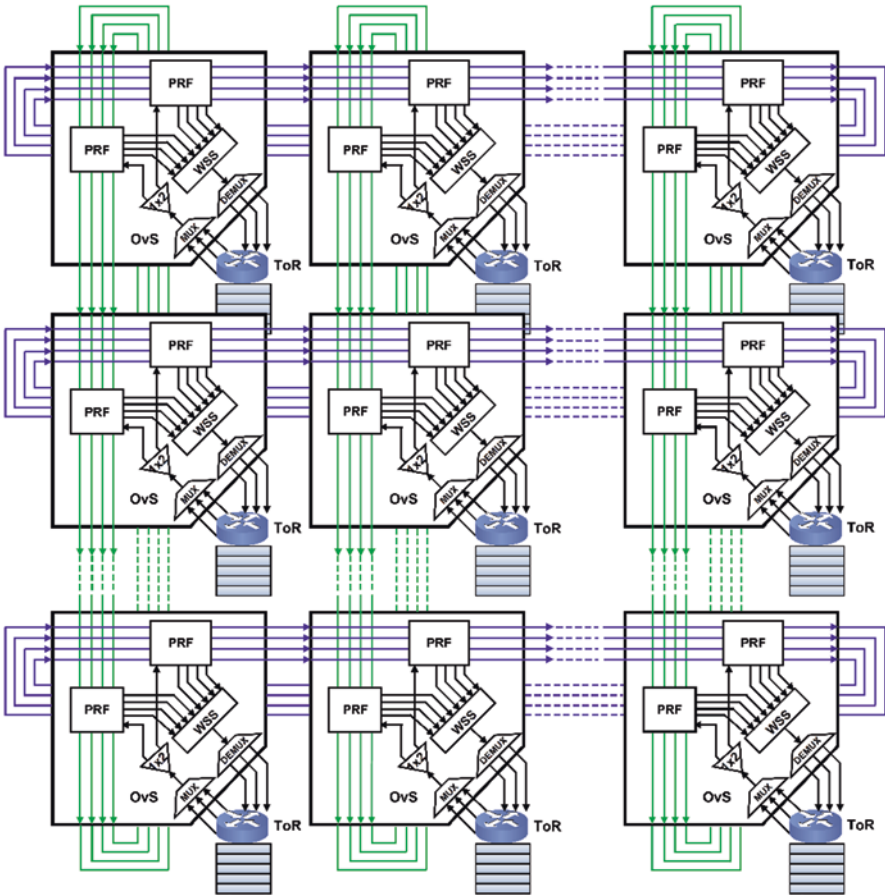


Fig. 2.5 The DWDM/SDM OCS network [39, 40]

DWDM, space-division multiplexing (SDM), and flexible optical wavelength switching technologies. The 2D-positioned OCS network, when utilized in conjunction with the ToR EPS, supports fully meshed interconnectivity among all rack switches, thus facilitating building large-scale networks with good system performance. Particularly, the integration of the B&S architecture and dynamic optical wavelength switching yields high network flexibility, and as such the datacenter network can dynamically support diverse traffic patterns such as multi-cast, in-cast, and all-to-all cast. Nonetheless, the end-to-end reconfiguration speed largely relies on the optical wavelength switching speed, which could be multi-milliseconds with current technologies. In addition, the further scaling of the network is limited by the number of DWDM channels, the port density of the optical wavelength switching component (WSS), and the costs of the DWDM transceivers.

Archon [41] is a transparent optical circuit-switched intra- and inter-datacenter network which adopts solely optical switching for core cross-connection and eliminates electronic switching, as shown in Fig. 2.6. In comparison with the DWDM/SDM datacenter architecture in [39, 40], the Archon network takes advantages of both space- and time-division multiplexing (SDM/TDM) technologies to provide high-capacity interconnection. In Archon, racks of servers are organized into clusters, with each cluster comprising a group of racks. Inside a cluster, all ToR switches are directly connected to a large-scale, high-capacity OCS via high-capacity SDM (multielement fiber (MEF)) optical links. The intra-cluster OCS is mainly used to handle long-lived bulky intra-cluster data flows. To efficiently switch short-lived, bursty intra-cluster traffic streams, an additional high-speed TDM switch with relatively low bandwidth capacity is adopted to support intra-cluster communications with variable link capacity. The TDM switch, as well as amplifiers, optical splitters, and optical couplers, is flexibly connected to a group of ToR switches through the reconfigurable intra-cluster OCS. Alternatively, the intercluster network is constructed by directly interconnecting the intra-cluster OCSs through a large-port-count intercluster OCS via SDM optical links. As for the inter-datacenter connectivity, the SDM signal from the central OCS is routed to the metro/core network through a SDM-to-WDM converter. In the Archon architecture, the unstable, shifting intra-cluster traffic is supported by two complementary switching paradigms, a high-capacity, slow OCS and a low-capacity, flexible TDM switch, while the relatively stable aggregated intercluster traffic is transferred over a centralized flexible OCS. The highly integrated SDM technology promises high capacity, high scalability, and simplified fiber connection complexity. However, as each cluster is equipped with a high-port-count OCS, a large-scale datacenter network composed of many clusters requires a large number of OCSs, which potentially results in high network costs. Another challenge is the practical implementation of the TDM switches and the slow-speed, limited flexibility intra-/intercluster OCS to support excessive traffic demands and heterogeneity in datacenter-scale networking, limiting the network's ability to scale to very large sizes.

Innovative SDN-enabled OCS datacenter architectures have been demonstrated, such as Calient [20], Polatis [22], proposal in [39, 40], Archon [41], SDN implementations in hybrid OCS/EPS datacenters [43], and C-Share [44]. Calient's SDN-based

support dynamic topology and link capacities. A prominent example of this category is OSA [45, 46], where the backplane of the network features a high-port-count OCS and reconfigurable wavelength selective switches (WSSs). In OSA, a rack switch can communicate with one or multiple rack switches simultaneously using a number of wavelength channels through the OCS. By exploiting dynamically reconfigurable WSS technology, these wavelength channels are adaptively assigned to the communication requests based on their traffic demands. This flexible bandwidth assignment scheme introduces a high degree of network flexibility and boosts communication performance. Similar to the SSX architecture, OSA also employs a multi-hop routing policy so as to alleviate the inherent connectivity limitations of the OCS, but potentially at the expense of increased routing complexity, communication latency, and optical-electronic-optical conversions.

Elastic OCS-based datacenter networking is a more flexible networking concept where flexible-grid OCS modules are deployed to set up adaptable wavelength channels [47–49]. The network elasticity is mainly reflected in two aspects: mixed line rates and multiple modulation formats. Elastic OCS networks support adaptive optical spectrum assignments with almost an arbitrary spectral width, which allows optical signals to be transmitted with a wide range of line rates depending on traffic requirements. Alternatively, the adoption of adaptive modulation formats gives the network the ability to change the signal’s modulation formats according to transmission distances and traffic profiles [47]. The very fine-grid spectrum allocation and variable modulation formats together optimize the spectrum resource utilization and facilitate on-demand datacenter service provisioning. The STRAUSS project [48], the Archon architecture [41], and the elastic ring-based OCS datacenter network proposed in [49] focus on introducing elasticity in optical circuit provisioning to support finer and multiple granularities in large-scale networks. The key enabling technologies include bandwidth-variable transceivers; bandwidth-variable OCS components, i.e., bandwidth-variable WSS; and SDN control technology [50]. Compared to the conventional fixed-grid OCS networks, the flexible-grid OCS designs future-proof the network with respect to channel baud rate and modulation formats, and enable greatly enhanced capacity, spectrum efficiency, and flexibility, which could better serve diverse datacenter applications. However, the flexible-grid software and hardware development, the design of an efficient spectrum allocation algorithm, and the control complexity involved need to be resolved [47]. Advances in integrated photonics are needed to bring the cost and energy of these components down to levels required for datacenters.

The hybrid OCS/OPS datacenter interconnect is another important area of research—replacing the EPS with an optical solution. An example of a hybrid OCS/OPS interconnect is the SDN LIGHTNESS architecture [27, 28] with highly distributed control, which has been proposed by the EC FP7 LIGHTNESS project. In this network configuration, the core layer includes transparent OCS and OPS networks, both providing fully meshed interconnectivity among rack switches. In [51] an energy-efficient torus datacenter structure combining OPS and agile OCS, where a novel flow management scheme is deployed to support optical path on-demand, is introduced. Nevertheless, a cost-effective hybrid OPS/OCS network is still elusive

and faces substantial challenges related to optical packet switching technologies, traffic prediction and classification, scalability, complexity, and costs.

An alternative to the hybrid packet and circuit switching is the new “seamless” switching approach used in the ProjectToR architecture [55]. In this case electronic switching and aggregation are performed entirely at the ToR, and all connections above the ToR are implemented using an OCS space switch. With high port counts in the OCS, including multiple ports per ToR, the OCS switch can be thought of as a set of reconfigurable fibers between the ToRs—carrying all traffic elephant and mice. The majority of OCS ports are reconfigured on a slow time scale of hours or even daily, while a small subset is reserved for cases in which the fly reconfiguration is needed—such as long-lived deterministic elephant flows or congestion relief. Multi-hop routing is used in the ToRs when needed as well. This seamless switching approach was adopted in part because analysis of Microsoft production datacenter traffic showed that separating elephant flows through traffic analysis is not feasible based on the measured statistics [55].

It is worth mentioning that development in free-space optics (FSO) [52, 55] and 60 GHz [53–54] wireless technologies enables the possible implementation of wireless links combined with OCS technology to realize a high-capacity, dynamically reconfigurable hybrid wireless/wired datacenter network. The wireless links hold significant advantages with respect to interconnectivity, power consumption, and low cabling complexity. Using MEMS switching through optical diffraction in free space, as opposed to just reflection, has been further shown to enable port counts exceeding 10,000 with switching transition times in the microsecond range. These free-space technologies can be envisioned to provide additional dynamic bandwidth capacity to relieve the congested hotspots in datacenter networks.

2.3 Agile Optical Datacenter Network Architecture

Having considered recent advances in the application of optical circuit switching to future large-scale datacenter networks, in this section we propose a novel architecture for a highly agile optical datacenter network which supports adaptive topologies and efficient sharing of network resources at flow level to provide a dynamic configuration response to changing datacenter application and traffic requirements. By exploiting the potential benefits of the passive, low-radix flexible arrayed waveguide grating (AWG) switch together with high-port-count optical circuit switches (OCSs), the proposed dynamic network offers the prospect of building a highly scalable, highly flexible, efficient large-scale datacenter network. It also supports pulling the metro network into the datacenter network using the OCS fabric for efficient metro implementations.

The proposed network infrastructure follows a flat topology comprising two layers of switch elements (Fig. 2.7). Server racks are directly attached to the top-of-rack (ToR) switches, which provide server-to-server connectivity within racks. ToR switches are connected to the large-scale reconfigurable optical space switch (OSS) which is constructed by interconnecting optical circuit switches (OCSs) with hundreds

of ports in a multistage topology such as Clos [9], Spanke [29], Benes [30], or flattened butterfly [31]. On top of the large OSS lies an additional layer of high-speed switching modules including optical wavelength-routed arrayed waveguide grating (AWG) switches, direct fibers (DFs), and small electronic packet switches (EPSs). Metro ROADMs can also be incorporated directly above the OSS. To make the most of the high-capacity optical circuits and the high-speed interconnect connections such as AWGs, DFs, and EPSs, so as to efficiently serve diverse network-wide demands, a large OSS is appropriately configured to connect the high-speed switching modules, mainly the AWG switches, with a collection of frequently communicating ToR switches, based on traffic demand estimations over relatively long time scales. For that, the designed optical network is given the ability to support circuit-switched provisioning carrying flows of packets, thereby achieving high-capacity, high-speed all-optical switching. Importantly, by flexibly reconfiguring the OSS, various combinations of network connections can be established, thus adapting the embedded network topology to the changing traffic patterns. Strictly speaking, the proposed novel dynamic datacenter network is akin to a flattened two-layer leaf-and-spine architecture consisting of a ToR switch layer and a fast optical switching layer with the layers interconnected via the OSS. This could be implemented using the ProjecToR free-space switching for the OSS function [55]. Because of the high port counts in ProjecToR, the OSS can be flattened and simplified.

Figure 2.7 demonstrates the optical schematic design of the modular datacenter network where the large reconfigurable space switch (OSS) logically groups the ToR switches into N clusters, each comprising M ToR switches and a $P \times P$ fast optical AWG switch. In total, this distributed network hosts $M \cdot N$ ToR switches. In a multi-rack cluster, each of the M ToR switches is equipped with one or multiple tunable optical transmitters and then directly connects to an input/output port pair of the $P \times P$ fast AWG switch through the reconfigurable OSS. The tunable optical transmitters and the $P \times P$ AWG switch together feature a reconfigurable, high-speed wavelength router, where the tunable optical transponder tunes the traffic flows originated from the connected ToR switch on the appropriate wavelengths based on the cyclic wavelength routing characteristic of the AWG switch, and subsequently the AWG switch directly routes these traffic flows to their destined AWG output ports. Of particular note is that the wavelength-routed AWG switch allows an optical signal entering from any AWG input port to be switched to any AWG output port by accordingly reconfiguring the output wavelength of the associated tunable transmitter [56]. It also allows multiple optical signals carried on different wavelengths injected from different ingress ToRs to be transmitted concurrently to an egress ToR. In doing so, a fully interconnected cluster network is realized, which moves data across the M ToR switches within the cluster quickly. A local cluster AWG controller is adopted to schedule the traffic flows, configure the tunable transmitters, and manage the network resources. Connections through the AWG are shown as duplex channels with a forward and reverse path through the AWGs. Although many applications will generate highly asymmetric traffic and may not require a high-capacity duplex connection as shown, the reverse channel is included here so that any handshaking or acknowledgments (ACKs) have a low-latency connection available to them so that a traffic flow does not stall due to delays in the

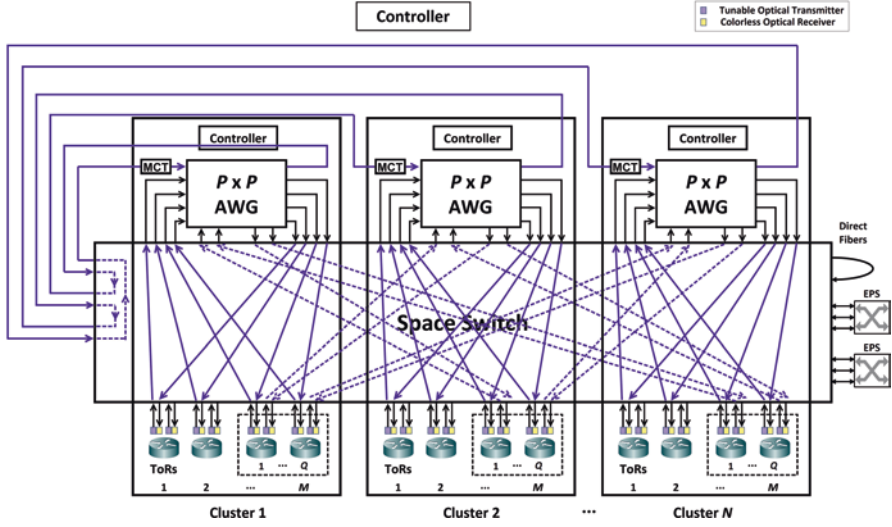


Fig. 2.7 The agile datacenter network architecture. The datacenter network is composed of N clusters. Each cluster contains M ToR switches and a $P \times P$ fast optical AWG switch. The number of cluster heads in each cluster is configurable and denoted Q

reverse channel. The direct fiber (DF) connections shown in the architecture are shown as simplex channels for simplicity and may be used by applications that require long-lived, high-capacity, low-latency links in one direction. The reverse channel may be carried by the electrical network to allow control of this direct fiber connection. Applications that may use such links are typically storage area network applications that may have traditionally been based on Fiber Channel (FC) links. One advantage of this approach is that the AWG switching fabrics can be replaced by metro ROADMs switches to connect to a metro network. The wavelength selective switches used in ROADMs can be used to provide the same functionality as the AWG and thereby flexibly set up connections between nearby datacenters.

The N cluster AWG switches, in conjunction with the multichannel tuning (MCT) blocks, can be flexibly interconnected into various optically transparent network topologies, such as a linear topology, a ring topology, a mesh topology, and a partially connected mesh topology, depending on the switch configuration of the OSS. Figure 2.7 illustrates an AWG ring topology where the interconnection network of the AWGs is a unidirectional ring. Each AWG is dedicated one or multiple multichannel tuning (MCT) modules. The primary functionality of the MCT is to facilitate high-speed optical wavelength switching in the network of AWGs. That is, the MCT allows a group of desired optical signals to pass through and then tunes these signals on the appropriate carrying wavelengths so that through the cluster AWG switch, they can either be directly dropped to the destination ToRs within the cluster or be transmitted to the next cluster hop. The multichannel tuning (MCT) architecture is demonstrated in Fig. 2.8. The MCT is composed of a wavelength selective switch (WSS), K , tunable wavelength converters (TWCs), and optical multiplexers (MUXs).

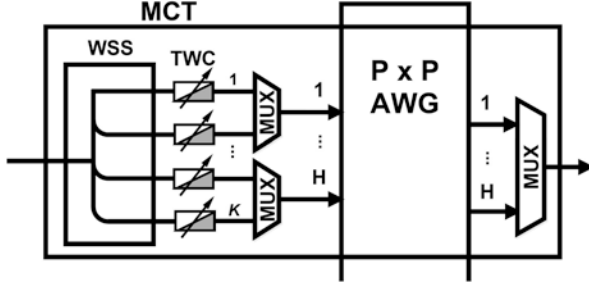


Fig. 2.8 The multichannel tuning (*MCT*) consists of a wavelength selective switch (*WSS*), K , tunable wavelength converters (*TWC*s), and optical multiplexers (*MUX*s)

The *WSS* supports highly flexible wavelength resource allocation, due to its dynamic wavelength route/selection property, which allows each input channel to be individually switched to an arbitrary *WSS* output port and allows each output signal to include an arbitrary number of wavelengths [57]. In the *MCT*, the *WSS* is configured to select the desired wavelengths at desired output ports and then forwards these optical wavelengths to the *TWC*s for wavelength conversion. The converted optical signals traverse through the core *AWG* switch and then reach the required *AWG* ports. The *MCT* architecture in Fig. 2.8 occupies 1 to K *AWG* ports. Note that there is a trade-off between the network performance and the number of *AWG* ports used. The *AWG*-to-*AWG* network enables high-speed cluster-to-cluster on-demand connections; thus, the network connectivity is substantially expanded.

One of the most significant aspects of the proposed distributed datacenter network is the design and implementation of the multi-hop routing in managing cluster-to-cluster communications. As illustrated in Fig. 2.7, two distinct fast *AWG* switches in two separate clusters may be interconnected by one or multiple hops of the *ToR* switches utilizing bidirectional optical connections through the *OSS*, so an *AWG*-to-*AWG* multi-hop routing path can be established through these intermediate *ToR* switches. This allows for the exchange of data between two remote clusters that are not directly interconnected. The intermediate *ToR* switch, referred to as a *cluster head*, is used to aggregate and redirect the received intercluster traffic streams. The main idea behind the multi-hop routing is the deployment of traffic relays, which exploits the routing and data forwarding capabilities of the *AWG* switches and the *ToR* switches. That is, a traffic flow containing a sequence of optical packets is propagated along an established path by hopping from one cluster header to another. In doing so, the multi-hop connectivity manages the communications between clusters; thus, it facilitates the realization of network-wide traffic communication and also avoids the constant reconfigurations of the large space switch *OSS*.

There are two extreme topologies possible with the multi-hop network connection for the proposed datacenter network. In the first network configuration, each cluster contains one cluster head that is directly attached to the logically adjacent cluster. This network design attempts to use only one cluster head per cluster to support intercluster connections. The scaling of the architecture is realized by adding

more clusters into the network, without affecting the rest of the network architecture. Nonetheless, the multi-hop routing path is severely unbalanced, as the shortest path includes only two hops, whereas the longest path length grows linearly with the number of clusters in the network. At the other extreme, each cluster has a dedicated cluster head directly connected to each of the other clusters, resulting in balanced multi-hop routing with at most two hops. This network model is highly symmetric and has a low network diameter, but the scalability is limited, as it can accommodate at most $P/2$ clusters, each containing $P/2$ ToR switches, yielding a network totaling at most $P^2/4$ ToR switches. Assuming that the AWG switch has a port count of 64×64 ($P = 64$), this topology can interconnect 32 clusters, each of which comprises 32 ToR switches, so a large-scale interconnection network including 1024 ToR switches is then obtained. Although the former network architecture can be scaled out easily, multi-hop routing may include a large number of traffic relays, which leads to performance degradation in terms of loss and communication latency. In contrast, the latter network architecture has limited scalability, but it is symmetric, requires low network diameter, and achieves loss uniformity, due to the fact that there exists a two-hop routing path between any source-destination cluster pair. The comparison indicates that there are obvious trade-offs between the two extreme cases, in terms of scalability, network diameter, loss uniformity, delay, and intra-/intercluster connection capacity. Given the flexibility provided by the physical architecture, arbitrary logical network topologies between these two extremes can be provisioned to balance among these trade-offs to ensure optimal throughput. In practice, the optimized connection structure is selected based on datacenter application requirements over a medium-term time scale.

In the proposed network architecture, the datacenter traffic is classified into two types of communications: intra-cluster communication and intercluster communication. Since the intra-cluster communication is exchanged between two ToRs residing in the same cluster, it can be switched from its source ToR directly to its destination ToR through a high-speed cluster AWG switch which directly interconnects all ToR switches in a cluster. Some heavily communicating ToR switches may be interconnected by the high-speed, high-capacity, low-latency direct fiber connections. A flexible EPS network is also utilized to provide additional intra-cluster connections. Potentially, the intra-cluster connectivity can be supported by three types of connections: fast AWG switches, direct fibers, and EPS modules. Differently, the intercluster connectivity can be provisioned by four types of connections: direct fibers, EPS modules, AWG-to-AWG network, and multi-hop routing. Similarly to the intra-cluster communication, the intercluster communication may be sustained by the direct fibers or the electronic switches, depending on the traffic requirements. In addition to that, an interconnection network of the AWGs is designed to provide high-speed all-optical intercluster connectivity. Furthermore, the multi-hop switching allows allocation of traffic flows on the existing multi-hop paths composed of multiple cluster heads and fast AWG switches. Clearly, the cluster fast AWG switches function as the most essential building blocks in the presented datacenter network and facilitates the routing operations of both intra- and intercluster traffic flows.

To maximize the network performance and resource utilization, an efficient control and scheduling plane is required, which is able to coordinate with the local AWG controllers in managing network resources and performing end-to-end scheduling for flow-level data traffic. Importantly, it needs to compute the most efficient network configuration that guarantees a fully connected graph of the cluster-based network and also fully exploits the capacity of the network connections, in accordance with the predicted traffic requirements. More specifically, the most heavily communicating source/destination ToR switches are provided with direct cross-connections using the fast AWG switches, the direct fibers, and the electronic switches, whereas the remaining communication pairs are interconnected by the AWG-to-AWG connections or the multi-hop paths.

2.4 Conclusions

Optical circuit switching is a promising technology for enabling future expansion of datacenter interconnection networks, in terms of scaling the number of network nodes and scaling link and switch port bandwidths well beyond the levels of today's electronically switched networks. The current research into this area, as reviewed in this chapter, shows a wide range of approaches that generally leverage the potential bandwidths supported by optical circuit switching to off-load aggregated traffic, or longer-lived bulk data flows between servers, from the electronically switched interconnection network. This approach is driving research into faster, submicrosecond optical switching speeds, through integrated photonic optical space switching. As optical switching times reduce, more and more shorter-lived traffic can be switched purely optically, further alleviating traffic overloading and energy consumption and cooling constraints for buffered electronic packet switches. Ultimately, achieving high-port-count nanosecond optical switching fabrics could pave a future path toward elimination of electronic switching altogether above the ToR or server connection level. A further driving factor for high-bandwidth optical switches is to enable future scaling-out of datacenter server numbers. Already, there is a trend in electronically switched datacenter network design toward flattening hierarchical network topologies to two layers, to reduce the requirement for very large core switches and to reduce routing and network configuration and management complexity. As datacenters scale, the aggregation center-stage switching will inevitably come under bandwidth requirement pressures, particularly to maintain high bisection bandwidths needed in cloud and big data applications. With transparency to format and bit-rate, replacing these core switches with fast optically switched alternatives would offer significant future-proofing advantages. Another increasing trend is the requirement for flexible provisioning of inter-datacenter network capacity. As datacenter requirements grow beyond the capacity of a single installation site, seamless distribution of datacenters across multiple sites is becoming more and more desirable. As the required interconnecting metro area and core networks have already been deployed based on optical circuit switching and WDM technologies, a

more seamless integration of intra- and inter-datacenter networks is possible if intra-datacenter networks were to be migrated to the same fundamental optically switched technologies. Indeed optical switch vendors are already adapting and targeting these technologies toward datacenter interconnection, which would seem to align well with datacenter operators' future needs.

Acknowledgment The datacenter network project described is a work in progress. The authors would like to acknowledge the support of the Science Foundation of Ireland (SFI) under the IPIC Programme (12/RC/2276), the CONNECT Programme (13/RC/2077), and the US-Ireland R&D Partnership Programme (SFI project code 15/US-C2C/I3132) and the NSF grants #EEC-0812072 and CNS1650669 and DOE grant #DE-SC0015867.

References

1. Cisco, *Cisco Global Cloud Index: Forecast and Methodology, 2014–2019*, White paper, October 2015
2. Dell'Oro Group, *Dell'Oro Group: Server Report*. White paper, March 2016
3. S. Kandula et al., Nature of datacenter traffic: measurements and analysis. in *Proceedings of ACM SIGCOMM conference on Internet Measurement*, November 2009
4. T. Benson, A. Akella, D.A. Maltz, Network traffic characteristics of data centers in the wild. in *Proceedings of ACM SIGCOMM conference on Internet measurement*, November 2010
5. P. Bodík, I. Menache, M. Chowdhury, Surviving failures in bandwidth-constrained datacenters. in *Proceedings of ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, August 2012
6. F.P. Tso, D. Pezaros, Improving data center network utilization using near-optimal traffic engineering. *IEEE Trans. Parallel Distrib. Syst.* **24**(6), 1139–1148 (June 2013)
7. A. Roy, H. Zengy, J. Baggay, G. Porter, A.C. Snoeren, Inside the social network's (Datacenter) network. in *Proceedings of ACM SIGCOMM Conference on Special Interest Group on Data Communication*, August 2015, pp. 123–137
8. M. Al-Fares, A. Loukissas, A. Vahdat, A scalable, commodity data center network architecture. in *Proceedings of ACM SIGCOMM Conference on Data Communication*, August 2008
9. C. Clos, A study of non-blocking switching networks. *Bell Syst. Tech. J.* **32**(2), 406–424 (1953)
10. A. Singh et al., Jupiter rising: a decade of clos topologies and centralized control in Google's datacenter network. in *Proceedings of ACM SIGCOMM*, August 2015
11. L.A. Barroso, J. Dean, U. Holzle, Web search for a planet: The Google cluster architecture. *IEEE Micro* **23**(2), 22–28 (2003)
12. N. Farrington, A. Andreyev, "Facebook's data center network architecture. in *Proceedings of IEEE Optical Interconnects Conference (OI)*, 2013
13. A. Andreyev, *Introducing data center fabric, the next-generation Facebook data center network*. White paper, November 2014
14. M. Alizadeh, T. Edsall, *On the data path performance of Leaf-Spine datacenter fabrics*. High-Performance Interconnects (HOTI), August 2013, pp. 71–74
15. C. Guo et al., DCell: a scalable and fault-tolerant network structure for data centers. in *Proceedings of ACM SIGCOMM Conference on Data Communication*, August 2008
16. C. Guo et al., BCube: a high performance, server-centric network architecture for modular data centers. in *Proceedings of ACM SIGCOMM Conference on Data Communication*, August 2009

17. D. Li et al., FiConn: using backup port for server interconnection in data centers. in *Proceedings of IEEE International Conference on Computer Communications*, April 2009, pp. 2276–2285
18. C. Kachris, I. Tomkos, A survey on optical interconnects for data centers. *IEEE Commun. Surv. Tut.* **14**(4) (2012)
19. D. Nikolova et al., Scaling silicon photonic switch fabrics for data center interconnection networks. *Opt. Exp.* **23**(2), 1159–1175 (Jan. 2015)
20. CALIENT, *S320 Photonic Switching*. White paper, March 2013
21. ARISTA, *7280R Series Data Center Switch Router*. White paper, July 2016
22. Polatis, *Series 7000 – 384x384 port Software-Defined Optical Circuit Switch*. White paper, March 2016
23. M. Chen, H. Jin, Y. Wen, V.C.M. Leung, Enabling technologies for future data center networking: a primer. *IEEE Netw.* **27**(4), 8–15 (Aug. 2013)
24. K.J. Barker et al., On the feasibility of optical circuit switching for high performance computing systems. in *Proceedings of ACM/IEEE SC 2005 Conference*, November 2005, pp. 1–22
25. T. Joon Seok et al., Large-scale broadband digital silicon photonic switches with vertical adiabatic couplers. *Optica* **3**(1) (Jan. 2016)
26. D.K. Hunter et al., WASPNET: a wavelength switched packet network. *Commun. Magaz.* **37**(3), 120–129 (1999)
27. S.D. Lucente, N. Calabretta, J.A.C. Resing, H.J.S. Dorren, Scaling low-latency optical packet switches to a thousand Ports. *J. Opt. Commun. Netw.* **4**(9), 17–28 (Sep. 2012)
28. N. Calabretta et al., On the performance of a large-scale optical packet switch under realistic data center traffic. *J. Opt. Commun. Netw.* **5**(6), 565–573 (June 2013)
29. R.A. Spanke, Architectures for guided-wave optical space switching systems. *IEEE Commun. Mag.* **25**(5) (1987)
30. V.E. Benes, *Mathematical Theory of Connecting Networks and Telephone Traffic* (Academic Press, New York, 1965)
31. J. Kim, W.J. Dally, D. Abts, Flattened butterfly: a cost-efficient topology for high-radix networks. *Int. Symp. Comput. Archit. (ISCA)* (2007)
32. G. Wang et al., c-Through: part-time optics in data centers. in *Proceedings of ACM SIGCOMM*, October 2010
33. N. Farrington et al., Helios: a hybrid electrical/optical switch architecture for modular data centers. in *Proceedings of ACM SIGCOMM*, October 2010
34. G. Porter et al., Integrating microsecond circuit switching into the data center. in *Proceedings of ACM SIGCOMM*, 2013, pp. 1–12
35. N. Farrington et al., Multiport microsecond optical circuit switch for data center networking. *IEEE Photon. Technol. Lett.* **25**(16) (2013)
36. H. Liu et al., Circuit switching under the radar with REACToR. in *Proceedings of USENIX Conference on Networked Systems Design and Implementation*, April 2014
37. D. Lugones, K. Katrinis, M. Collier, A reconfigurable optical/electrical interconnect architecture for large-scale clusters and datacenters. in *Proceedings of ACM International Conference on Computing Frontiers*, May 2012, pp. 13–22
38. D. Lugones et al., A reconfigurable, regular-topology cluster/datacenter network using commodity optical switches. *Futur. Gener. Comput. Syst.* **30**, 78–89 (2014)
39. Z. Zhu, S. Zhong, Scalable and topology adaptive intra-data center networking enabled by wavelength selective switching. in *Proceedings of OFC*, 2014
40. Z. Zhu, S. Zhong, L. Chen, K. Chen, Fully programmable and scalable optical switching fabric for petabyte data center. *Opt. Exp.* **23**(3), 3563–3580 (2015)
41. S. Yan et al., Archon: A function programmable optical interconnect architecture for transparent intra and inter data center SDM/TDM/WDM networking. *J. Lightw. Technol.* **33**(8), 1586–1595 (2015)
42. B.A.A. Nunes et al., A survey of software-defined networking: past, present, and future of programmable networks. *IEEE Commun. Surv. Tut.* **16**(3), 1617–1634 (2014)

43. K. Katrinis, G. Wang, L. Schares, SDN control for hybrid OCS/electrical datacenter networks: an enabler or just a convenience? IEEE Photon. Soc. Summer Top. Meet. Ser., 242–243 (2013)
44. Y. Ben-Itzhak, C. Caba, L. Schour, S. Vargaftik, *C-Share: Optical Circuits Sharing for Software-Defined Data-Centers*. eprint arXiv, September 2016
45. A. Singla et al., Proteus: a topology malleable data center network. in *Proceedings of ACM SIGCOMM Workshop on Hot Topics in Networks*, October 2010
46. K. Chen et al., OSA: an optical switching architecture for data center networks with unprecedented flexibility. Northwestern University, Technical Report, 2012
47. S.L. Woodward, M.D. Feuer, Benefits and requirements of flexible-grid ROADMs and networks. *J. Opt. Commun. Netw.* **5**(10), A19–A27 (2013)
48. M. Schlosser et al., Requirements analysis of technology enablers for the flexi-grid optical path-packet infrastructure for Ethernet transport. STRAUSS Project, May 2014
49. Z. Zhang et al., Elastic optical ring with flexible spectrum ROADMs: An optical switching architecture for future data center networks. *Opt. Switch. Netw.* **19**(part 1), 1–9 (2016)
50. J. Zhang et al., Experimental demonstration of OpenFlow-based control plane for elastic light-path provisioning in flexi-grid optical networks. *Opt. Exp.* **21**(2), 1364–1373 (Jan. 2013)
51. K.-I. Kitayama et al., Torus-topology data center network based on optical packet/agile circuit switching with intelligent flow management. *J. Lightw. Technol.* **33**(5) (Mar. 2015)
52. N. Hamedazimi, et al., “Firefly: a reconfigurable wireless data center fabric using free-space optics,” *SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 4, pp. 319–330, Aug. 2014.
53. X. Zhou et al., Mirror mirror on the ceiling: Flexible wireless links for data centers. in *Proceedings of ACM SIGCOMM*, 2012
54. H. Vardhan et al., 60 GHz wireless links in data center networks. *Comput. Netw.* **58**, 192–205 (2014)
55. M. Ghobadi, R. Mahajan, A. Phanishayee, N. Devanur, J. Kukarni, G. Ranade, P.-A. Blanche, H. Rastegarfar, M. Glick, D. Kilper, ProjecToR: Agile reconfigurable data center interconnect. in *Proceedings of ACM SIGCOMM*, 2016, pp. 216–229
56. S. Cheung, T. Su, K. Okamoto, S.J.B. Yoo, Ultra-compact silicon photonic 512 x 512 25-GHz arrayed waveguide grating router. *IEEE J. Select. Top. Quant. Electron.* **20**(4), 310–316 (2014)
57. R. Ryf et al., Wavelength-selective switch for few-mode fiber transmission. in *Proceedings of 39th European Conference on Optical Communications (ECOC)*, September 2013

Optical Switching in Next Generation Data Centers

Testa, F.; Pavesi, L. (Eds.)

2018, XIV, 336 p. 179 illus., 139 illus. in color.,

Hardcover

ISBN: 978-3-319-61051-1