

Optical Wireless Data Center Networks: Potentials, Limitations, and Prospects

Abdulkadir Celik, Basem Shihada, and Mohamed-Slim Alouini

Computer, Electrical, and Mathematical Sciences and Engineering (CEMSE) Division
King Abdullah University of Science and Technology (KAUST), Thuwal, KSA.

ABSTRACT

Data centers (DCs) are intrinsic to emerging technologies which require to store and process massive amounts of versatile data through large-scale networks of computing and storage units. Conventional wired DC networks (DCNs) with static links and finite network interfaces suffer from cabling cost and complexity as scalability increases, lack the flexibility to handle the dynamic and large volume of traffic outbursts (i.e., hotspots), and have limited bisection bandwidths mainly due to the switching speeds. Therefore, optical wireless DCNs have recently attracted attention by their capability to augment the inherent restrictions of wired DCNs. In addition to the reduced cost of wiring, wireless communication can provide a flexible topology to overcome oversubscriptions and hotspots. Moreover, it is possible to adapt the link capacities in accordance with the quality of service demands of different services and flow classes. Furthermore, wireless communication can also make it possible to eliminate switches by establishing direct links among the servers. In this paper, we first present potential state-of-art optical wireless technologies. Second, practical challenges of design and provisioning of optical wireless communications in real-life DCNs are outlined along with a survey of recent advances and implementations. Finally, we motivate researchers with the exciting prospects of optical wireless DCNs including physical and virtual topology design, interference management, multiple access techniques, traffic management and grooming, and flow classification.

1. INTRODUCTION

Data centers (DCs) are indispensable infrastructures to store and process highly diversified data deluge generated by the emerging technologies such as fifth-generation networks, big data, cloud services, content delivery, Internet of things, artificial intelligence, neural networks, etc. Ever increasing quality of service (QoS) and quality of experience (QoE) demands of these applications push the limits for available resources in terms of bandwidth, latency, processing power, and storage capacity. These needs can be met either by centralized (or remote) mega DCs scaling up to hundreds of thousands of servers (i.e., processing and computing units) or by a network of distributed modular micro DCs. To provide high-speed and low-latency interconnections within and across DCs, which are also referred to as intra-DC and inter-DC networking, respectively, there is an urge to develop novel architectures, powerful networking solutions, and highly competent communication technologies. Noting that the plethora of servers and switches which magnify with the scale of the DCN, it is of the paramount importance to guarantee scalable, sustainable, and resilient DCNs which shrinks the capital and operational expenditures, utilize the available network resources efficiently, and adapts itself to dynamically changing traffic conditions.¹

Among many possible DCN architectures, conventional tree-based DCNs are generally constructed in a hierarchical topology where racks interconnect its servers via top-of-rack (ToR) switches. For the inter-rack communications, racks are linked to each other via aggregate switches (ASs) in the second layer, which is then interconnected by core switches (CSs) at the third layer. In wired DCNs, all these connections are realized by uniform and fixed capacity cables, e.g., coaxial, twisted-pair, fiber, etc. Albeit its scalability, this hierarchical tree-based topology results in a multi-root tree where more powerful links and switches are needed in the branches nearer to the third layer. However, a limited number of CSs and related links in the top layer become a bottleneck

Further author information: A.C.: E-mail: abdulkadir.celik@kaust.edu.sa (Corresponding Author), B.S.: E-mail: bshihada@kaust.edu.sa, M-S.A.: E-mail: slim.alouini@kaust.edu.sa

of the DCN. As a result, inter-rack communications experience a throughput lower than the actual available bandwidth, which is also known as the *oversubscription*.²

On the contrary of the widely-held belief that DCNs must ensure a full bisection bandwidth among a considerable amount of servers,³ empirical studies have reveal the fact that 80% of the flows are mice while 95% of the traffic volume is carried out by the top 10% of the elephant flows.^{4,5} As a remedy for insufficiency of digital switches to handle elephant flows, hybrid digital-optic DCNs are proposed to provide on-demand high-speed inter-rack communications. Hybrid digital-optic DCNs (e.g., Helios³ and c-Through⁶) generally use electrical packet switching (EPS) and optical circuit switching (OCS) technologies to handle bursty and long duration elephant flows, respectively. Although EPS limits the actual benefit of the OCS, it is still necessary because of the high switching delay of OCSs technology.³ Furthermore, the high energy consumption of fault-prone switches poses extra cost and complexity.

On the other hand, measurements of real DCN traces show that some applications generate unpredictable and asymmetrical traffic patterns, which mainly contain common data required by many servers across the DCNs and results in *hotspots*.⁴ Unfortunately, it is a challenge for DCNs to tackle such erratic and unbalanced traffic due to the inflexible hierarchical topology with uniform and fixed capacity links. Proposed solutions either focus on extending the aforementioned switch-centric solution (e.g, Fat-Tree,⁷ VL2,⁸ Portland,⁹ etc.) or designing server-centric architectures (e.g., BCube,¹⁰ DCell,¹¹ Jellyfish,¹² FiConn,¹³ etc.) to reduce the negative impacts of the switches. Nevertheless, these approaches are still not sufficient to mitigate chronic oversubscription and hotspot problems. Even if deploying extra on-demand wired links is a possible solution, it is hard to predict deployment locations because of the unpredictable and unbalanced traffic conditions. On the other hand, installing extra cables for every rack pair is not practical since scaling up the DCNs increases the cabling cost and complexity along with indirect consequences such as inefficiency in cooling and space utilization.

Fortunately, wireless DCNs (WDCNs) have recently received considerable attention with its ability to support reconfigurability via flexible and adaptive links which can be adjusted as per the dynamically changing traffic patterns and loads; to deliver higher throughputs than their wired counterparts thanks to the more efficient bandwidth utilization; to reduce the capital investment and operational costs by eliminating cables and switches; to offer convenient deployment, management, and maintenance; and to provide efficient utilization of space and cooling systems. Nonetheless, implementing WDCNs in real-life poses formidable challenges in designing physical and logical topology optimization. In this regard, this paper first surveys virtues and drawbacks of optical wireless communications which are followed by research challenges and advances in the literature. Finally, we provide open research problems and relish exciting opportunities in WDCNs and conclude the paper with a few remarks.

2. POTENTIALS OF OPTICAL WIRELESS COMMUNICATIONS

The light-beam transmission by using signal carriers in the unguided medium of visible, infrared (IR), and ultraviolet is referred to as optical wireless communications (OWCs).¹⁴ In particular, the OWC operates in visible light and near-IR bands are known as visible light communication (VLC) and free-space optical (FSO) communications, respectively. Laser diodes (LDs) and Light-emitting diodes (LEDs) are well known FSO and VLC transmitters widely used in practice, respectively. Since the LEDs are cheaper and reliable light sources, VLC is typically considered for indoor communication due to its short-range and wide-beam coverage area with data rates in the order of Gbps.¹⁵ Therefore, VLC is a natural candidate for provisioning the intra-rack or short inter-rack communications. Thanks to its ability to provide razor-sharp light beams, the LDs are powerful devices to reach very long destinations with desirable high-speed connections in the order of Tbps.¹⁶ Therefore, recent research efforts on FSO typically concentrates on tackling the outdoor FSO channel impairments such as misalignment and pointing errors, scintillation, atmospheric turbulence, etc. However, reasonable ranges and acclimatized indoor environment of DCNs provide a much more friendly playground for the use of FSO communications. Obviously, this also eliminates the need for costly and bulky FSO transceivers equipped with sophisticated subsystems to mitigate the outdoor effects. Moreover, when FSO is combined with wavelength division multiplexing (WDM), it can provide very high data rates and high fanout required by DCNs. Ciaramella *et. al.* have shown that an outdoor WDM-FSO link can reach 1.28 Tbps capacity on 32 wavelengths (32x40 Gbps) over 212 m distance.^{16,17}

The main distinction of the OWC from its RF counterparts is its highly directed propagation nature which necessitates the establishment of the line of sight (LoS) links between the transceivers. Since LEDs emit collimated light beams, the directivity is more significant in FSO communications which helps in compensating the transmission losses and mitigating the interference. On the other hand, the relatively large divergence angle of LEDs requires multiple access schemes for a VLC receiver to distinguish signals coming from multiple LED sources. Directivity and high penetration loss of OWC especially pave the way for an enhanced physical layer security. The dominant interference source in the indoor OWC receivers is caused by the ambient light sources. The virtues of optical wireless DCNs can be summarized as follows: ,

- Deploying wires to interconnect a massive number of servers is a quite demanding engineering task that worsens as the size of the DCNs increases. Since wiring is generally error-prone, necessary modifications, as well as maintenance, also requires a considerable amount of time and efforts. However, replacing cables with plug-and-play wireless modules is convenient and can ease such hard operations to a great extent.
- In addition to cabling complexity, wireless links can also decrease the cabling related capital investments and operational costs. Moreover, thick cable bundles traveling across the DCNs reduce the space utilization and cause heat dissipation, which must be compensated with extra power consumed for cooling.
- DCN switches are known to be power hungry and fault-prone devices. Thus, wireless technologies can enable a more energy efficient DCN by partially or eliminating the need for the switches. Since error-prone switches typically constitute the bottleneck of DCNs, their elimination can help to improve the overall bisection bandwidth and fault-tolerance.
- Thanks to the flexibility of wireless links, logical topology of DCNs can be reconfigurable to handle oversubscription and hotspot problems. Adjusting the logical topology can provide higher throughputs by efficient utilization of the available capacity, which is important to meet the ever-increasing QoS and QoE demands of diverse DCN traffic.

Albeit these appealing advantages, there are also daunting challenges in design and provisioning of WDCNs. In what follows, we present advances and challenges of optical WDCNs.

3. OPTICAL WIRELESS DCNS: CHALLENGES AND ADVANCES

In wired DCNs, traditional row-based arrangements of racks are preferable because of its efficient space utilization. However, LoS requirement of OWC transceivers requires a novel rack placement solution to free the communication space from obstacles. Therefore, the physical topology design must consider the following design considerations:

- The degree of LoS connectivity: The degree of LoS connectivity is the most important metric to judge the suitability of a physical topology for optical WDCNs. It also delimits the achievable performance by reconfiguring the logical topology.
- Scalability and reconfigurability: No matter how high the degree of LoS connectivity is, scalability is a critical issue to expand a physical topology design to a mega DCN setting. Moreover, the reconfigurability is another parameter since employing hardware and techniques that require a long reconfiguration delay does not comply with the highly dynamic traffic characteristics of DCNs.
- Fault tolerance: A fault-tolerant optical WDCN design should provide a high fanout to provide the required redundancy in case of node and link failures. Noting that the switches are among the most fault-prone network entities, a rack placement that reduce the role of switches inherently enhance the fault tolerance of the DCN.
- Space utilization: It is also important to use available space efficiently while ensuring the LoS connectivity. Otherwise, it will cost the waste of the cooling. Moreover, the size of the wireless hardware and modules must comply with the small form factor of the DCN components (e.g., servers, switches, racks, etc.).

- Cost-performance index: An in-depth cost-performance tradeoff analysis is necessary. That is, employing numerous bulky, costly, and power-hungry OWC modules may naturally yield high-speed wireless links. However, their hardware cost and power consumption must also be factored in via a practical cost-performance index.

In the remainder, we present the recent advances in optical WDCNs: Riza *et. al.* propose an indoor power smart link which employs an electrically controlled variable focus lens.¹⁸ They consider installation of transceivers at the top of racks using a pedestal platform which allows vertical and horizontal alignment towards other racks. Even though authors keep their focus on containerized DCNs, the extension of this approach to a mega DCN is not possible due to the limit top of rack space. Moreover, fault tolerance and adaptivity of the proposed scheme is questionable because of the complexity and reconfiguration latency of the fault-prone mechanical components.

Similar to *Flyways*² which consider 60 GHz wireless transceivers, *Firefly* establishes inter-rack connections via both LoS links and non-LoS links via a mirror on the ceiling.¹⁹ For the beam steering purpose, authors consider switchable mirrors which are prealigned to the transceiver in another rack and this link is switched by changing the state of the mirror, i.e., glass or mirror. They also employ Galvo mirrors with limited rotation ability for beam steering. Using three-state switches (i.e., reflective (R), transmissive (T), and splitting (S)), *FlyCast* proposes a similar setup with multi-cast link capability. A common limitation of these approaches is the lack of scalability due to the limited top of rack space. Moreover, the imperfection of the ceiling mirror and reflection losses can cause performance degradation in the non-LoS links. Hamza *et. al.* proposes the concept of FSO-Bus²⁰ where intra-rack communication is realized by an optical controller which handles the signals coming from beam combiner and gong to the distributor in the transmitter and receiver sides, respectively. On the other hand, the optical controller receives the inter-rack traffic and forwards it to its servers via beam distributor. Unlike the intra-rack communication, the distance of the inter-rack buses can cause performance loss and scalability of the FSO-Bus concept may not be suitable for mega DCNs.

Arnon *et. al.* proposes a cellular arrangement of racks which can be interconnected via LoS links between transceivers integrated to the ToR switches.²¹ ToR switches are also equipped with extra transceivers to connect CSs located above the cell. In this way, racks within a cell can communicate with other cells. Likewise, OWCell²² arranges racks in regular polygonal topologies, develops its theoretical underpinnings, and proposes a routing geographical protocol to communication among different OWCells. These cellular concepts are much more suitable for mega DCNs due to their scalability. They also do not suffer from reconfiguration delay as no electro-mechanical tools are employed to realize LoS links. However, the space utilization decreases with the increasing number of racks within a cell. In particular, switch-less nature of OWCell requires multiple hops to reach another cell at the other end of a mega DCN, which might yield undesirable effects of congestion and delay.

It is obvious that early research efforts^{18, 19, 20, 21, 22} focus on physical topology design to prove the optical WDCN concept by ensuring the connectivity of OWC transceivers. However, logical topology design is still unexplored to show how reconfigurability and flexibility of WDCN can be used to meet the demands of the high volume and dynamically changing traffic characteristics of DCNs. Logical topology optimization involves many interesting open research problems such as resource allocation, interference management, routing, load balancing, congestion control, etc. Notice that the complexity of these problems increases with the number of network components (i.e., servers, racks, switches, links, etc.) which exacerbates with the number of generated flows and their diverse QoS requirements. Although the logical topology optimization has no control over the size of DCN, it can classify flows and manage the traffic as per the needs of each class.

In this regard, we have recently studied a traffic grooming approach which combine the mice flows to create rack-to-rack (R2R) jumbo flows.²³ Based on a spine-leaf DCN architecture where ToR switches at the leaf layer are connected to the CSs at the spine layer via WDM-FSO links, R2R jumbo flows are routed over dedicated lightpaths between each rack pair. On the contrary of R2R approach, elephant flows are routed over server-to-server (S2S) express lightpaths whose capacity is determined based on the available capacity after ensuring the low-latency requirement of mice flows. Our main motivation in separating these two classes was to avoid the performance degradation of mice flows caused by the bandwidth hungry elephant flows. Since DCNs are also required to serve mission-critical flows which require ultra-reliability and low-latency, proposed traffic grooming approach handles these critical flows with the highest priority over the S2S express paths and provides the

delay analysis using queuing theory. Obtained results have shown that throughput and timely flow completion performance of the proposed scheme is much higher than its counterparts.

Even though proposed traffic grooming assumes that flow classification is readily available, a flow classification mechanism is required to detect the class of flows, which is especially important to prevent elephant flows to congest the mice flow paths. However, existing flow detection mechanisms, e.g., packet sampling²⁴ and port mirroring,²⁵ are not sufficient to classify flows within a reasonable time frame. Therefore, we propose LightFD²⁶: A lightweight, fast, and accurate flow detection mechanism. LightFD benefits from the TCP behavior of the DCN flows and execute monitoring to detect elephant flows if a certain packet size threshold is reached. Numerical results show that it delivers superior performance in terms of the accuracy and detection speed in comparison with the traditional methods. It has also shown that flow detection speed has a distinguishable impact on the DCN throughput especially with the conventional detection mechanisms.

4. CONCLUSIONS AND FUTURE PROSPECTS

In this paper, we presented potentials and limitations of the optical WDCNs. Unlike their wired counterparts, flexibility and reconfigurability of the wireless links offer novel solutions to resolve chronic DCN problems such as oversubscription and hotspots. As a candidate wireless technology, OWC has great potentials to provide the required high-speed low-latency links for meeting the ever-increasing demand of DCN applications. Thanks to efficient utilization of the bandwidths, optical WDCNs offer much higher throughput which can be tailored to dynamically changing traffic conditions in DCNs. Although early research efforts focus on realizing optical WDCNs by developing hardware topologies which ensure the LoS connectivity among the OWC transceivers, there exists ample opportunity to reap the full benefits via optimization of the logical topology. In addition to aforementioned open research problems, we believe following future prospects of WDCNs are promising and interesting research directions:

Hybrid Wireless DCNs: Instead of a pure optical WDCNs, hybridizing the architecture with wires and other wireless technologies (e.g., millimeter wave, terahertz communication) might be less complicated and more cost-efficient.^{18,19,27} For instance, instead of sophisticated designs to ensure LoS OWC links, wired intra-rack communications can be cost-efficient yet technically sufficient thanks to low cabling complexity and adequate bandwidth over the short cable lengths.

Containerized DCNs: Installing DCNs in containers has recently received attention with their attractive attributes of modularity and mobility. It is even possible to employ containers within the umbrella of a mega DCN framework because companies prefer to rent a portion of the mega DCNs to avoid costs related to updates and upgrades. However, they store and process sensitive data which requires extra care for the security and privacy of the transmitted data. Due to the penetration in the ability of light beams, it is possible to isolate this rented portion from the rest of DCN. At this point, proposed solutions should be extended to provide inter-DCN networking to keep the holistic functionality of the entire DCN.

Energy and Cost Efficiency: Energy efficiency is one of the most critical design parameters of DCN architectures. While a significant portion of the power is consumed for cooling, the remaining consumption is mostly related to power-hungry devices such as switches. Therefore, there is a dire need for an in-depth analysis of the energy efficiency that can be offered by WDCNs. Moreover, potential architecture must also be evaluated carefully based on the metrics listed in Section 3.

Real-life Implementation: Except for some prototyping efforts in laboratory, WDCNs are not implemented in real-life to reveal its exact potentials. We believe a medium scale realization of WDCNs can promote the advancement of these potential and practical field of research and development to a great extent.

REFERENCES

- [1] Celik, A., Shihada, B., and Alouini, M.-S., “Wireless data center networks: advances, challenges, and opportunities,” *preprint available at arXiv* (2018).
- [2] Kandula, S., Padhye, J., and Bahl, V., “Flyways to de-congest data center networks,” in [*Proc. ACM Hotnets-VIII*], 202–208 (2009).
- [3] Farrington, N., Porter, G., Radhakrishnan, S., Bazzaz, H. H., Subramanya, V., Fainman, Y., Papen, G., and Vahdat, A., “Helios: a hybrid electrical/optical switch architecture for modular data centers,” *SIGCOMM Comput. Commun. Rev.* **41** (Aug. 2010).
- [4] Kandula, S. et al., “The nature of data center traffic: measurements & analysis,” in [*proc. the 9th ACM SIGCOMM*], 202–208 (2009).
- [5] Benson, T., Anand, A., Akella, A., and Zhang, M., “Understanding data center traffic characteristics,” *SIGCOMM Comput. Commun. Rev.* **40**, 92–99 (Jan. 2010).
- [6] Wang, G., Andersen, D. G., Kaminsky, M., Papagiannaki, K., Ng, T. E., Kozuch, M., and Ryan, M., “c-through: Part-time optics in data centers,” *SIGCOMM Comput. Commun. Rev.* **40**, 327–338 (Aug. 2010).
- [7] Al-Fares, M., Loukissas, A., and Vahdat, A., “A scalable, commodity data center network architecture,” in [*ACM SIGCOMM Computer Communication Review*], **38**(4), 63–74, ACM (2008).
- [8] Greenberg, A., Hamilton, J. R., Jain, N., Kandula, S., Kim, C., Lahiri, P., Maltz, D. A., Patel, P., and Sengupta, S., “V12: A scalable and flexible data center network,” in [*ACM SIGCOMM computer communication review*], **39**(4), 51–62, ACM (2009).
- [9] Niranjana Mysore, R., Pamboris, A., Farrington, N., Huang, N., Miri, P., Radhakrishnan, S., Subramanya, V., and Vahdat, A., “Portland: A scalable fault-tolerant layer 2 data center network fabric,” in [*ACM SIGCOMM Computer Communication Review*], **39**(4), 39–50, ACM (2009).
- [10] Guo, C., Lu, G., Li, D., Wu, H., Zhang, X., Shi, Y., Tian, C., Zhang, Y., and Lu, S., “Bcube: A high performance, server-centric network architecture for modular data centers,” *ACM SIGCOMM Computer Communication Review* **39**(4), 63–74 (2009).
- [11] Guo, C., Wu, H., Tan, K., Shi, L., Zhang, Y., and Lu, S., “Dcell: A scalable and fault-tolerant network structure for data centers,” in [*ACM SIGCOMM Computer Communication Review*], **38**(4), 75–86, ACM (2008).
- [12] Singla, A., Hong, C.-Y., Popa, L., and Godfrey, P. B., “Jellyfish: Networking data centers randomly,” in [*NSDI*], **12**, 1–6 (2012).
- [13] Li, D., Guo, C., Wu, H., Tan, K., Zhang, Y., and Lu, S., “Ficonn: Using backup port for server interconnection in data centers,” in [*IEEE INFOCOM 2009*], 2276–2285, IEEE (2009).
- [14] Khalighi, M. A. and Uysal, M., “Survey on free space optical communication: A communication theory perspective,” *IEEE Commun. Surveys Tuts.* **16**, 2231–2258 (Fourthquarter 2014).
- [15] Azhar, A. H., Tran, T., and O’Brien, D., “A Gigabit/s indoor wireless transmission using MIMO-OFDM visible-light communications,” *IEEE photonics technology letters* **25**(2), 171–174 (2013).
- [16] Esmail, M. A., Ragheb, A., Fathallah, H., and Alouini, M.-S., “Experimental demonstration of outdoor 2.2 Tbps super-channel fso transmission system,” in [*2016 IEEE International Conference on Communications Workshops (ICC)*], 169–174 (May. 2016).
- [17] Ciaramella, E. et al., “1.28 Terabit/s (32x40 Gbit/s) WDM transmission system for free space optical communications,” **27**(9) (2009).
- [18] Riza, N. A. and Maccarrini, P. J., “Power smart in-door optical wireless link applications,” in [*2012 8th International Wireless Communications and Mobile Computing Conference (IWCMC)*], 327–332 (Aug. 2012).
- [19] Hamedazimi, N. et al., “Firefly: A reconfigurable wireless data center fabric using free-space optics,” *SIGCOMM Comput. Commun. Rev.* **44**, 319–330 (Aug 2014).
- [20] Hamza, A. S., Deogun, J. S., and Alexander, D. R., “Free space optical data center architecture design with fully connected racks,” in [*2014 IEEE Global Communications Conference*], 2192–2197 (Dec. 2014).
- [21] Arnon, S., “Next-generation optical wireless communications for data centers,” in [*Broadband Access Communication Technologies IX*], **9387**, 938703, SPIE (2015).
- [22] Hamza, A. S. et al., “Owcell: Optical wireless cellular data center network architecture,” in [*Proc. IEEE Intl. Conf. Commun.(ICC)*], 1–6 (May. 2017).

- [23] Celik, A., Al-Ghadhban, A., Shihada, B., and Alouini, M.-S., “Design and provisioning of optical wireless data center networks: A traffic grooming approach,” in [*Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*], 1–6 (Apr. 2018).
- [24] McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., and Turner, J., “Openflow: Enabling innovation in campus networks,” *ACM SIGCOMM Computer Commun. Review* **38**(2), 69–74 (2008).
- [25] Rasley, J., Stephens, B., Dixon, C., Rozner, E., Felter, W., Agarwal, K., Carter, J., and Fonseca, R., “Planck: Millisecond-scale monitoring and control for commodity networks,” in [*ACM SIGCOMM Computer Communication Review*], **44**(4), 407–418, ACM (2014).
- [26] Al-Ghadhban, A., Celik, A., Shihada, B., and Alouini, M.-S., “LightFD: A lightweight flow detection mechanism for traffic grooming in optical wireless dcns,” in [*Proc. IEEE Global Commun. Conf. (GLOBECOM)*], (2018).
- [27] Halperin, D. et al., “Augmenting data center networks with multi-gigabit wireless links,” in [*Proc. of the ACM SIGCOMM*], 38–49 (2011).