Open access • Proceedings Article • DOI:10.1109/GLOCOM.2014.7037133

Free space optical data center architecture design with fully connected racks

— Source link

Abdelbaset S. Hamza, Jitender S. Deogun, Dennis R. Alexander

Institutions: University of Nebraska–Lincoln

Published on: 09 Feb 2014 - Global Communications Conference

2014

Free Space Optical Data Center Architecture Design with Fully Connected Racks

Abdelbaset S. Hamza
University of Nebraska-Lincoln

Jitender S. Deogun
University of Nebraska-Lincoln, jdeogun1@unl.edu

Dennis R. Alexander
University of Nebraska-Lincoln, dalexander1@unl.edu

Globecom 2014 - Optical Networks and Systems Symposium
Pages: 2192 - 2197, DOI: 10.1109/GLOCOM.2014.7037133

Free Space Optical Data Center Architecture Design with Fully Connected Racks

Abdelbaset S. Hamza, Jitender S. Deogun
Computer Science and Engineering Department
University of Nebraska-Lincoln

Dennis R. Alexander
Electrical Engineering Department
University of Nebraska-Lincoln

Abstract—**In this paper, we propose a novel design of a data center (DC) using free space optical (FSO) technology. The proposed FSO-DC design is based on fixed, non-mechanical, FSO links facilitating the realization of fully connected FSO racks and rows/columns of racks. Each rack becomes a point of intersection of three fully connected sub-networks. We investigate requirements, advantages and challenges of the proposed design. We develop and analyze a switch-free, fully connected FSO rack, present its link budget analysis and validate it by simulation. Results establish the feasibility of a switch-free FSO rack based on the proposed FSO-DC design. A cost estimate for the proposed FSO-DC design is also presented and compared to three well-known conventional DC designs.**

I. Introduction

Future data centers (DCs) must support huge network traffic that is continuously increasing. Moreover, DCs must satisfy minimum performance requirements of latency, reliability, flexibility and scalability. However, cabling technologies (i.e., copper-cables and fiber optics) currently deployed in conventional wired data centers (CDCs), require larger number of cables to support higher data rates. Besides limiting the possible topologies and connectivity, this results into design and development problems for the infrastructure of wire ducting and maintenance, heat dissipation, and power consumption [1].

As a solution for the capacity problem of the CDCs, the possibility of integrating 60 GHz radio frequency (RF) technology into CDCs is being investigated [1]–[3]. 60 GHz links are capable of adding multi-Gbps rates [3]. On the other hand, to address the cabling problems, a completely wireless DC design using the 60 GHz technology is proposed in [1].

Although promising, 60 GHz deployment in DCs has its limitations as it has lower practical bandwidth, and suffers from high attenuation and propagation loss [2]. Radiation patterns of 60 GHz impose additional restrictions on the activity of wireless modules in close proximity to avoid interference. This increases the complexity of routing and network management, and reduces the throughput [1].

The absence of the atmospheric impairments of free space optical (FSO) or optical wireless (OW) links in indoor systems, has motivated us to consider deploying FSO in DCs. Moreover, the speed of light in FSO is approximately 1.5 times faster than that of in fiber optics, which mean less latency. Finally, unlike copper cables and optical fibers, FSO is zero sunk (i.e., links can be re-deployed after deployment). Thus, we believe that FSO leads to a high performance and cost effective infrastructure for DCs. Accordingly, we investigate the use of FSO is DCs and propose a complete FSO-DC design (i.e., intra/inter-rack communications) and cost analysis.

II. Related Work

There are only few papers and patents that discuss application of FSO to DCs [4]–[8]. In [4], authors suggest the realization of FSO links inside DCs using pedestal platform mounted to the top of the rack. The arm holding a transceiver and connected to the pedestal allows vertical and rotational movement such that line-of-sight (LOS) links are established between different racks (see Figure 1-(a)). Incorporating a mechanical system to establish FSO links significantly adds to the complexity and latency of the system and increases risk of failure. In order to avoid the mechanical reconfiguration, the authors in [5] consider a DC network in which FSO links are used for *inter-rack* communications by connecting TOR switches using FSO. As shown in Figure 1-(b), FSO transceivers and switchable mirrors (SMs) are placed on top of each rack and pre-aligned to connect to different racks. According to the states of the SMs (i.e, glass or mirror), a link is directed and reflected off of a ceiling mirror to other racks. The communication and network reconfigurability is controlled using a centralized topology and routing managers. The limited number of active links at any time and the time delay associated with the change in the mirror's state add to the complexity of the centralized managers and routing.

A bi-directional point-to-point FSO link design is proposed in a patent [7]. The inventors suggest using it in DCs for intra-rack communication using a top-of-rack (TOR) optical switch which is linked to servers in the rack as shown in Figure 1-(c). The optical switch then directs the information back to the servers using data shower beams. In this design, the optical switch must be equipped with number of transceivers equal to the number of servers co-located on the bottom of the switch to have point-to-point FSO links, otherwise, problem of interference or limited connectivity might arise. For large number of servers, this design is intractable. In another patent [8], the inventors present an extensive theoretical discussion of a DC using FSO but do not discuss any mean of connecting multiple components.

It might be noted that a common impediment of all designs is the difficulty of connecting multiple adjacent components using FSO links. This is because LOS links can not be

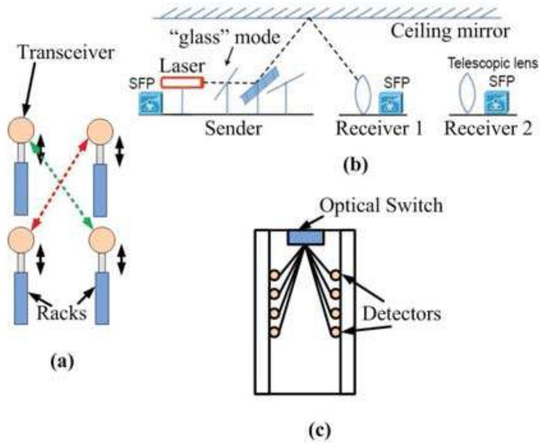


Fig. 1: Designs Presented in: (a) [4] (b) [5] (c) [7]

easily maintained as other components get in between the source and destination need to be connected leading to risk of link blocking. In this paper, we address these limitations and propose an FSO-DC (OWDC) that uses fixed links to achieve a fully connected FSO rack (i.e., no mechanically or adaptively reconfigured links).

III. Proposed FSO-DC Design

Figure 2-(a) shows a DC in which racks are deployed in a row-based arrangement with J rows. Each row contains k racks. A rack can be identified using its row and column numbers (j, k), where, $(1 \leq j \leq J \; and \; 1 \leq k \leq K)$. For the ease of demonstration, each rack is represented by a rectangle labeled with the rack's coordinates $Rack(j, k)$.

We start with the design of an FSO rack, in order to understand the design of the proposed FSO-DC.

A. Switch-Free FSO Rack

Figure 2-(b) shows a *switch-free* FSO rack comprising S servers. The servers are numbered from 1 to S from top to bottom of the rack. Therefore, a server in the $Rack(j, k)$ can be identified by its coordinates $Server(j, k, s)$, where, $(1 \leq s \leq S)$, is the number of the server. In order to achieve high data rate communication between servers within the same rack, servers must be connected using point-to-point FSO links.

In our design, each server is equipped with an optical transmitter on one side of the server, and an optical receiver comprising a photodetector (or an array of photodetectors) on the opposite side. Servers are mounted on the FSO rack such that all transmitters (receivers) of the servers are on the same side of the rack. The main idea is to direct the transmitted beams either for intra-rack, inter-rack, or both communications, using the intra/inter-rack selector. For intra-rack communication, the beams are directed to the other side of the rack where receivers are placed. Using a beam distributer, beams are distributed to all servers allowing *switch-free* intra-rack communication. On the other hand, for inter-rack communication, the combined beam is directed to the *Rack Optical Controller (ROC)*.

Directing the beams around the rack can be done using a set of mirrors mounted to the structure of the rack. Any server can receive a copy of the S beams using beam splitters placed in front of the server to be able to intercept the beams.

Figure 2-(a) shows three rows (i.e., rows 1, 2, and J), and the first and last columns (i.e., columns 1 and K). ROCs within the same row (and similarly, ROCs within the same column) can be connected together using a method similar to the method used to connect servers within the same rack.

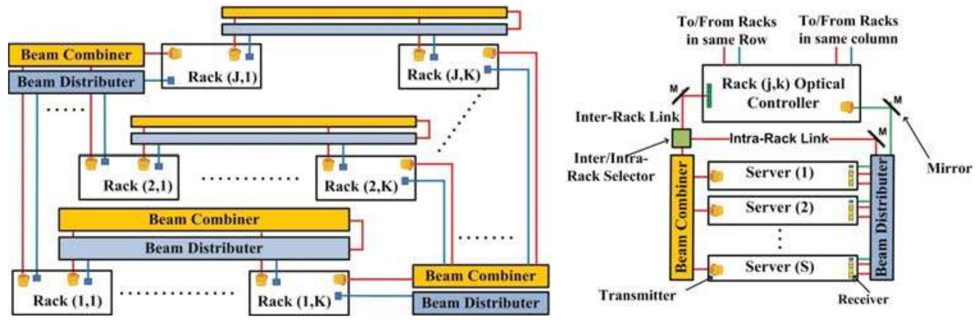
In case of intra-rack communication, S light beams from the S servers can be transmitted and received by all servers, simultaneously. Each transmitter has a separate optical path connecting it to all other servers. Therefore, there are no collision domains, instead, each server has its broadcast domain which must be managed efficiently so that, data are delivered to the intended destination(s) only. Many networking and addressing schemes can be used. A network topology of the rack can be changed according to the scheme selected. In the following, we briefly discuss three of such schemes:

- Time Division Multiple Access (TDMA), where, the frame of any source server s is divided into $S - 1$ time slots (TSs). The server s transmits data to the server i using the TS i, where, $1 \leq i \leq S \; and \; i \neq s$. The intra-rack network can be considered as S subnetworks, each subnetwork is a *bus network* with a single transmitter.
- Using a technique similar to optical burst switching (OBS), source server s sends a short optical packet prior to the data transmission. The short packet contains addresses of the destination(s) and any other necessary information. The $S - 1$ servers receive the short packet, intended destinations receive incoming data, and other servers ignore it. The topology of the network is similar to that of in TDMA case.
- Wavelength Division Multiplexing (WDM). Can help boosting the capacity of intra-rack links. Each receiver is assigned a wavelength, or multiple wavelengths. Using tunable transmitters and receivers, signals transmitted to other servers are delivered using the same beam at different wavelengths. In this case, the rack topology is a fully connected (complete mesh) network.

B. Rack Optical Controller (ROC)

For inter-rack communication, an ROC receives data from other racks to deliver to the servers in its rack, communicate with other racks in the same row/column, and relay the data received from any of the ROCs in the same row/column to any of the ROCs in the same column/row.

Racks are arranged in rows and columns, and it is possible to connect ROCs using the same method as for servers within the same rack. Moreover, communications between ROCs can follow same schemes discussed in the intra-rack communication. The functions performed by ROCs are very similar to a regular switch, however, it might be noted that unlike TOR switches, intra-rack communication is not dependent on the ROC. Moreover, each ROC is the intersection of three fully connected networks. This can be efficiently utilized in



(a) (b)

Fig. 2: (a) Proposed Design of an FSO-DC. (b) Proposed Fully Connected, Switch-Free FSO Rack of Servers.

routing and DC network management. An ROC is expected to handle large amount of traffic compared to servers, therefore, we envision the use of WDM/DWDM to increase inter-rack link capacities. Two cases in inter-rack communication:

1) The source ROC is located on the same row/column of the destination ROC, and two-hops link is needed to perform the communication. The source server sends data to the source ROC, which forwards the data to the destination ROC, and finally, to the destination server.
2) The source and destination ROCs are neither located on the same row nor same column. In this case, a link with a minimum of three hops is needed. Source server sends data to source ROC, which in turn forwards the data to the ROCs on the same row or column. The ROC located at the row/column intersection of the source and destination ROCs will forward the data to the destination ROC then to the destination server. However, due to the full connectivity, other paths can be used for routing. The decision of transmitting the data to the row or the column ROCs depends on the used routing algorithm.

In order to realize rack topology equivalent to S bus sub-networks using current technology, S^2 wires are needed (i.e., 1600 wires/rack for $S = 40$). Similarly, a fully connected rack using the current technology requires a total of $(S^2 - S)/2$ full-duplex wire segments (i.e., 780 wires/rack for $S = 40$) where each server is equipped with, at least, S ports. This is almost impossible to manage and basically one of the main reasons why the star topology was adopted in the first place.

The small size of FSO components, and the ability to split a beam among S servers using a set of passive optical elements, help realize a fully connected rack using only S beams.

IV. Indoor Point-to-Point LOS FSO Link

In this section, we provide a brief description of an indoor point-to-point, LOS FSO link.

The proposed design uses Laser Diodes (LDs) because they have high optical power outputs and can support transmission at high bit rates. We adopt Avalanche photodetectors (APDs) since they are preferred in systems that require high data rates and where the noise induced by ambient light is negligible because APDs have high cost and require high bias [9].

A. Optical Noise Sources

FSO communication links are deployed in a wide range of environments (e.g., indoor and outer space). Different noise sources affect the performance of the FSO link with varying degrees of severity depending on the environment. We first discuss the noise from different sources affecting indoor FSO links that can be mitigated in our design.

The absence of the noise due to background radiation (e.g., the sun) makes the ambient artificial light the dominant source of noise in DCs [9].

Point-to-point LOS links utilize transmitter and receivers with narrow field of view (FOV), therefore, these links are capable of rejecting majority of the ambient artificial light [9]. Moreover, using high pass filters (HPF), fluorescent lights driven by a conventional ballast can be mitigated, whereas, fluorescent lights driven by electronic ballast are harder to mitigate [9]. Since LEDs have narrower PSDs as compared to that of other light sources, a more efficient solution for the artificial ambient light would be to illuminate the DC using LED sources that are out of band of the LDs used in the DC. This way, the ambient artificial light can be easily mitigated.

On the other hand, there are three inevitable noise sources, namely; quantum (shot) noise, dark noise and thermal (Johnson) noise. The shot noise is due to the random arrival rate of photons from the transmitter and has a variance σ_q^2. On the other hand, dark noise is due to a very small current from the PD which is a combination of two currents: bulk (I_D) and surface leakage (I_L) currents with variances (σ_{db}^2) and (σ_{dl}^2), respectively. Finally, the thermal noise exists in any circuit of equivalent resistance R_L and temperature T_e and modeled as a white Gaussian noise with zero mean and variance σ_{th}^2. The total noise variance σ_N^2 is given by:

$$N_T = \sigma_q^2 + \sigma_{db}^2 + \sigma_{dl}^2 + \sigma_{th}^2 \quad (1)$$

$$N_T = 2qRP_RBFM^2 + 2qI_DBFM^2 + 2qI_LB + \frac{4\kappa T_e B}{R_L} \quad (2)$$

where q is the electric charge, R is the PD's responsivity, P_R is the power received, B is the electronic bandwidth, M is the PD's gain factor, and F is the excess noise factor.

B. Link Budget

Using LDs, we can get very narrow beam with concentrated power, however, any beam propagating in the free space experiences a slight divergence. Short links might not be affected by this problem since it is possible to use PDs with light collecting areas that matches the spot size of the light beam. For long FSO links, collimators can be placed at certain points along the path to re-collimate the beams.

The diameter of the spot size of a beam that has a very small beam width angle θ and travels a distance D is approximately equal to $(\theta \cdot D)$ [10],

Assume a point-to-point, LOS link with a transmitted power P_t, transmitter and receiver optics efficiencies, η_T and η_R, respectively. The received unfaded power P_R is given by,

$$P_R = \eta_T \eta_R \mathcal{L}_{GL} P_T \tag{3}$$

where, $\mathcal{L}_{GL}$, denotes the geometrical loss which is the ratio between aperture area of the receiver (A_R) and the spot size area of the beam at the receiver (A_{im}), and is given by,

$$\mathcal{L}_{GL} = \frac{A_R}{A_{im}} = \left(\frac{D_R}{\theta D}\right)^2 \tag{4}$$

where, $D_R \leq \theta D$, and hence, Equation 3 becomes,

$$P_R = \eta_T \eta_R \left(\frac{D_R}{\theta D}\right)^2 P_T \tag{5}$$

V. Design and Analysis of an FSO Rack

In our design, servers are connected using point-to-point, non-LOS (NLOS) links formed using specular reflections (i.e., a set of mirrors and BSs). The difference between the link budget of a point-to-point, LOS link and that of NLOS with specular reflection is that, mirrors and BSs absorb light, and hence, might have efficiencies less than 100%. Moreover, a BS is used to split the light beam into two perpendicular beams: transmitted beam (along the path of the original incident beam), and reflected beam. Based on the design, transmitted and reflected beams may or may not have the same power. Therefore, in case of point-to-point, NLOS link, Eq. 5 must be extended to include the efficiencies and power reductions caused by mirrors and BSs. The losses and factors depend on the number and arrangement of mirrors and BSs in the design.

Figure 3 depicts the proposed design of a fully connected FSO rack. A typical FSO rack consists of S servers. Each server, s (for $1 \leq s \leq S$), is equipped with an optical transmitter T_s operating at wavelength λ_s. The power transmitted by a transmitting server s is denoted as P_{T_s} and the power efficiency of the transmitter's lens is η_s.

An optical receiver is placed on the other side of the server with a PD array R_s to receive signals transmitted by the S servers. Each array contains S PDs, numbered from 1 to S. A PD s within the PD array R_s has a diameter D_{R_s}, power efficiency of the optical lens η_{R_s} and operates at the corresponding wavelength λ_s. It is assumed that the receiver is capable of handling the S input signals, using multiple receivers, a control plane, or a scheduler.

A mirror M_s is associated with each server s on the transmitter's side. On the other side of the server, a beam

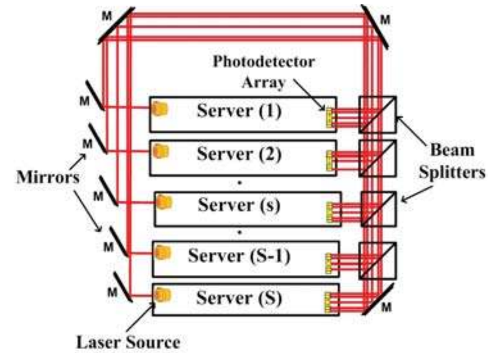


Fig. 3: A Fully Connected FSO Rack of Servers.

splitter BS_s is placed except for the server number S where the BS is replaced by the mirror M_S. Each mirror M_s has an efficiency of η_{M_s}, whereas, each BS has an efficiency of η_{BS_s}. A beam splitter BS_s splits the incident beam into two beams: reflected and transmitted with powers α_s and $(1-\alpha_s)$, respectively. As shown in Figure 3, two mirrors are used to direct the beams from transmission side to receiver side.

We consider the case where a single wavelength is used i.e., $\lambda_s = \lambda$, $1 \leq s \leq S$. In order to distinguish between transmitter and receiving servers, we use notation s for the transmitting server and s' for the receiving server, $1 \leq s, s' \leq S$. The received power by a server s' from transmitter server s is,

$$P_R(s,s') = P_{T_s} \cdot \eta_{T_s} \cdot \eta_{R_{s'}} \cdot \left(\frac{D_{R_{s'}}}{\theta_{s,s'} \cdot D_{s,s'}}\right)^2 \cdot \eta_M(s') \cdot \eta_{BS}(s') \cdot \Delta_{BS}(s') \tag{6}$$

where, $\theta_{s,s'}$, and $D_{s,s'}$ are the angle width, and the distance of the link between transmitter s and receiver s'. $\eta_M(s')$ and $\eta_{BS}(s')$ are the mirrors and BSs aggregated power efficiency functions, respectively. $\Delta_{BS}(s')$ is the aggregate power splitting function of BSs.

In a rack of 40 servers the maximum distance between a transmitter and a receiver does not exceed 5 meters. Therefore, it is possible to use PDs with light collecting area equal to the area of the beam at the PD. It is assumed that $D_R = \theta D$ for all transmitter-receiver combinations. This is a reasonable assumption since the beam diameter at the receiver is 2.5 mm, assuming a beam width angle of 0.5 $mrad$.

It is assumed that all transmitters and receivers are identical with the same power transmitted and optical efficiencies. It is also assumed that all mirrors have the same efficiency η_M and all BSs has the same efficiency η_{BS}. Then $P_R(s,s')$ becomes,

$$P_R(s,s') = P_T \cdot \eta_T \cdot \eta_R \cdot \eta_M(s') \cdot \eta_{BS}(s') \cdot \Delta_{BS}(s') \tag{7}$$

Figure 3 depicts that a link between two servers is reflected from three mirrors, except for server S, the number of mirrors is four. So, mirrors aggregated power efficiency function is,

$$\eta_M(s') = \begin{cases} \eta_M^3, & \text{for } 1 \leq s' \leq S-1 \\ \eta_M^4, & \text{for } s' = S \end{cases} \tag{8}$$

A link to a destination server s' traverses s' BSs except for server S, where the beam traverses $(S-1)$ BSs. Accordingly, BSs aggregated power efficiency function is given by,

$$\eta_{BS}(s') = \begin{cases} \eta_{BS}^{s'}, & \text{for } 1 \le s' \le S-1 \\ \eta_{BS}^{S-1}, & \text{for } s' = S \end{cases} \tag{9}$$

The number of BSs and the power ratio of the transmitted/reflected beams at each of the BSs affect the received power at each server. We assume that a BS s' reflects $\alpha_{s'}\%$ of the incident beam's power, and transmits $(1-\alpha_{s'})\%$. Hence, the power splitting function of BSs is,

$$\Delta_{BS}(s') = \begin{cases} \alpha_{s'} \prod_{j=1}^{s'-1} (1-\alpha_j), & \text{for } 1 \le s' \le S-1 \\ \prod_{j=1}^{S-1} (1-\alpha_j), & \text{for } s' = S \end{cases} \tag{10}$$

The signal-to-noise ratio (SNR) is given by [9],

$$SNR_{IM\text{-}DD} = \frac{I_p^2}{N_T} = \frac{(RMP_T)^2}{\sigma_q^2 + \sigma_{dl}^2 + \sigma_{db}^2 + \sigma_{th}^2} \tag{11}$$

VI. Numerical Results and Analysis

The electronic charge q is equal to 1.602×10^{-19}. PD responsivity (R) and gain factor (M) are assumed to be 0.9 and 3, respectively. Both, dark current and leakage currents are assumed to be $15nA$. The temperature and equivalent resistance of the receiver are assumed to be $290K$ and 1 $K\Omega$. We assume that the number of servers $S = 40$. Optical efficiency of all transmitters/receivers optics, mirrors and BSs are assumed to be 99%. For a BS s', the power of the reflected light beam $\alpha_{s'}$ is 10%, and hence, the power of the transmitted light beam is 90%. The wavelength is assumed to be 1500 nm.

The power received as a function of the position of the server in the rack is shown in Fig. 4. The power reception falls as we move towards the bottom of the rack. Power levels are of the order of 10^{-3} W for transmitted power in the range of 0.5 mW to 10 mW. There is a sudden improvement in the received power by the server S compared to the server $S-1$. This is due to the ratios of the BSs used where the last server receives 0.99×0.9 of the power incident to the BS number $S-1$ while the server $S-1$ receives 10% of that power.

In order to evaluate the performance of an FSO link within the rack, OptiSystem software was used. An FSO link was implemented with an FSO channel of five meters. We created a link with the same characteristics, however, it deploys a fiber optic instead of FSO. Both transmitters use OOK NRZ modulation scheme for simplicity, however, for higher data rates, other modulation schemes such as pulse position modulation (PPM) are preferred [9]

Figure 5 depicts the eye diagrams of the FSO and fiber optical links at 2.5 Gbps. Three servers are selected (i.e., server 1, 25 and 39). It is clear that as we move towards the bottom of the rack, the power received decreases, degrading the performance of the FSO link. On the other hand, it is difficult to notice any variation in the fiber optical link since the link is too short, and the received power is not affected by BSs or mirrors as in the FSO link.

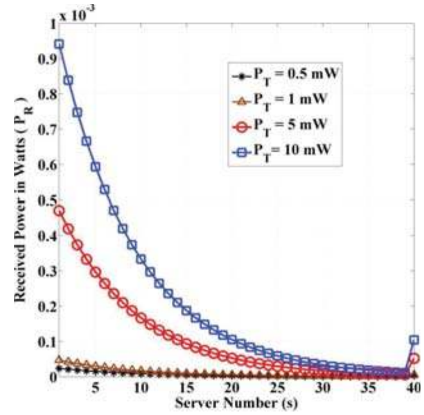


Fig. 4: Received power by Servers.

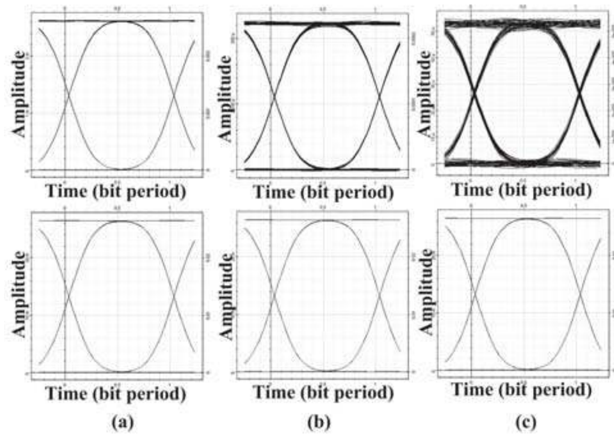


Fig. 5: Eye Diagrams of FSO (top) and Fiber Optics (bottom) at 2.5 Gbps and $P_T = 10\ mW$ (a) s=1. (b) s=25. (c) s=39.

Table I summarizes the performance of FSO and fiber optical links at different servers and different transmitting powers. As we move towards servers at the bottom of the rack, when transmitted power is low, Q-factor, eye height, and threshold all degrade and minimum BER increases. On the other hand, increasing the transmitted power improves the performance of the link allowing error-free communication.

Results in Table I suggest that for higher bit rates, using low power (near 1mW) is sufficient to realize low BER for servers near the top of the rack. On the other hand, it might be difficult to establish a link with servers near the bottom of the rack at this low transmitted power. Therefore, system optimization can be realized by setting the power transmitted by each server based on the intended destination. This way, the power consumption can be minimized.

VII. Cost Estimate

Comparing FSO-DC to CDC is challenging. CDCs have been the main interest of the academic and industrial communities for long time. This implies that the cost-performance tradeoff of the wired technology has been improving over the last decades. On the other hand, FSO components for DCs may

TABLE I: Summary of the FSO Link Performance Compared to the Optical Fiber Link.

		$P_T = 1\ mW$		$P_T = 5\ mW$		$P_T = 10\ mW$	
	s	FSO	Fiber	FSO	Fiber	FSO	Fiber
Max. Q Factor	1	110.4	253.8	227.0	276.9	274.8	284.4
	25	15.8	248.3	61.9	269.5	105.5	293.0
	39	3.84	257.0	18.8	286.7	33.1	276.0
Min BER	1	0	0	0	0	0	0
	25	1.5E-6	0	0	0	0	0
	39	6.1E-5	0	0	0	0	0
Eye Height	1	250E-6	2.6E-3	1.3E-3	13.0E-3	2.6E-3	26.0E-3
	25	1.8E-5	2.6E-3	100E-6	13.0E-3	210E-6	26.0E-3
	39	1.1E-6	2.6E-3	2.2E-5	13.0E-3	4.8E-5	26.0E-3
Thresh-old	1	3.0E-5	190E-6	100E-6	940E-6	180E-6	1.9E-3
	25	9.7E-6	180E-6	2.5E-5	960E-6	2.8E-5	1.9E-3
	39	2.4E-6	190E-6	1.1E-5	930E-6	2.3E-5	1.9E-3

not exist yet. Therefore, we only aim to have an approximate sense of the FSO-DC cost. We consider the price of the TOR, aggregate and core switches following [1], however, we also include the cost of the network interface cards (NICs). Tables II and III depict the prices used in our calculations and the costs of three reference CDC configurations used to connect 10K servers for comparison, respectively [1].

TABLE II: Cost of Different Components used in CDC

Component	Price ($)	Minimum Unit
NIC	80	1
TOR	8,000	1
Aggregate Switch (AS)	9,000	1
Core Switch (CS) Subunit	60,000	1
Core Switch (CS) Chassis	12,000	1
Core Switch (CS) Power Supply	3,500	3

TABLE III: Total Cost of Different CDC Configurations

Configuration	# TOR	# AS	# CS Subunit	# CS Chassis	Total Cost ($)
CDC1	250	52	16	2	4,162,500
CDC2	250	48	12	2	3,886,500
CDC3	250	26	8	1	3,436,500

In case of the FSO-DC, we consider cost of FSO transceivers and ROCs. Since there is no reference for these prices, we estimate cost by refereing to the prices of CDC devices, e.g., we assume that the price of FSO transceiver is γ times the price of an NIC (C_{NIC}), and the price of an ROC is β times the price of an aggregate switch (C_{agg}), where $0.1 \leq \gamma,\ \beta \leq 2$. Therefore, the total cost of an FSO-DC is:

$$C_{FSO-DC} = J \cdot K \cdot [\gamma \cdot S \cdot C_{NIC} + \beta \cdot C_{agg}] \qquad (12)$$

Figure 6-(b) depicts the cost function of the FSO-DC. It might be noted that, at $\gamma = \beta = 2$, the cost of FSO-DC is approximately 1.7 times the CDC3, and around 1.4 times the price of CDC1. However, there is still a range where, γ and β are greater than one, and yet, the price of the FSO-DC is cheaper or comparable to that of the CDC. We expect that, the cost of FSO technology will decrease as it is commercialized, leading to further reduction in the cost FSO-DC.

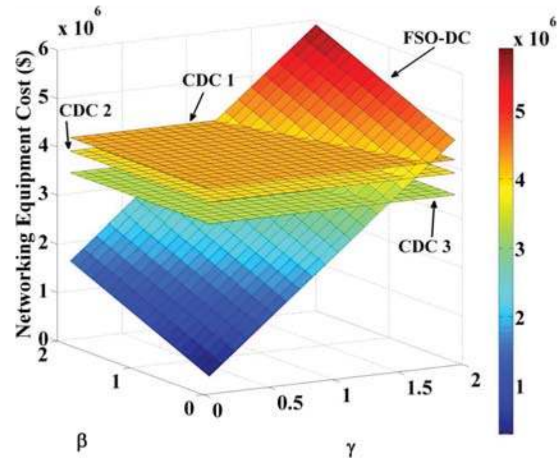


Fig. 6: FSO-DC Cost Function Compared to CDCs.

FSO-DC has another advantage over CDC is that an upgrade in a DC (e.g. from 10 Gbps to 40/100/400 Gbps or higher) will require huge investment and changes in the CDC as cables and switches must be replaced. FSO-DC presents a more modular architecture that is highly scalable with little upgrade required.

VIII. Conclusions and Future Work

We present a road map for the realization of an FSO-DC and investigate challenges presented by the proposed design. An FSO-DC design and associated link budget analysis for a fully-connected rack of servers is presented. Simulation shows that the proposed design realizes high data rates within a rack. Our cost analysis shows that the cost of the proposed FSO-DC design is comparable to that of conventional wired DCs. It is expected that the cost of the proposed design will decrease as FSO technology is commercialized. The proposed design is highly suitable for scaling and upgrading DCs.

The proposed design addresses many problems and limitations of the current art, but several issues remain to be investigated. Currently, we are extending our models in different ways, e.g., to take into consideration the effect of beam misalignment, heat, air flow and vibration on the FSO links.

References

[1] J. Shin *et al.*, "On the feasibility of completely wireless datacenters," *Proc. Archit. for Netw. and Commun. Sys.*, pp. 3–14, 2012.

[2] K. Ramachandran *et al.*, "60 ghz data-center netw.: wireless => worry less?" NEC Laboratories America,, Tech. Rep., 2008.

[3] S. Kandula *et al.*, "Flyways to de-congest data center networks," Proc. ACM Hotnets-VIII, 2009.

[4] N. Riza and P. Marraccini, "Power smart indoor optical wireless link applications," *Wireless Comm. & Mobile Comp.*, pp. 327–332, 2012.

[5] N. Hamedazimi *et al.*, "Patch panels in the sky: A case for free-space optics in data centers," in *Proc. ACM Hotnets*, 2013, pp. 23:1–23:7.

[6] H. Kuo and R. Walmsley, "Free space optical interconnect," US Patent App. 12/864,231, Nov. 25 2010.

[7] J. Joseph *et al.*, "High speed free-space optical communications," WO Patent App. PCT/US2012/052,397, Mar. 7 2013.

[8] H. Davidson *et al.*, "Data center with free-space optical communications," US Patent 8,483,569, Jul. 9 2013.

[9] Z. Ghassemlooy *et al.*, *Optical Wireless Communications: System and Channel Modelling with MATLAB*, 1st ed. Taylor & Francis, 2012.

[10] S. Bloom *et al.*, "Understanding the performance of free-space optics [invited]," *J. Opt. Netw.*, vol. 2, no. 6, pp. 178–200, 2003.