Optical Fiber Technology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Optical Fiber Technology

journal homepage: www.elsevier.com/locate/yofte

Invited Papers

Data center links beyond 100 Gbit/s per wavelength

Jose Krause Perin*, Anujit Shastri, Joseph M. Kahn

E. L. Ginzton Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA

ARTICLE INFO

Keywords:
Data center interconnect
Direct detection
Stokes vector receiver
Coherent detection
Differentially coherent detection

ABSTRACT

Increased traffic demands within and between data centers now necessitate low-cost and low-power systems with per-wavelength bit rates beyond what can be easily achieved using conventional on/off keying. We review spectrally efficient links based on direct detection, Stokes vector detection, coherent detection and differentially coherent detection for data center applications. We show that limited spectral efficiency and power margin will inhibit scaling of direct detection-compatible formats beyond 100 Gbit/s. Stokes vector receivers can provide higher spectral efficiency without requiring a local oscillator laser, but require power-hungry analog-to-digital converters (ADCs) and digital signal processing (DSP). Similarly, existing DSP-based coherent systems designed for long-haul transmission may be excessively complex and power-hungry for short-reach data center links. We present low-power DSP-free coherent and differentially coherent alternatives that avoid high-speed ADCs and DSP and achieve similar performance to their DSP-based counterparts in intra-data center links and dispersion-compensated inter-data center links.

1. Introduction

Global data center internet protocol (IP) traffic is estimated to grow at a compound annual rate of 26.8% from 2015 to 2020, corresponding to a threefold increase in five years [1]. Data center-to-data center IP traffic is expected to grow at an even faster rate of 31.9% [1]. This poses a significant challenge to continuously scaling the capacity of data center links while meeting strict constraints in cost and power consumption, particularly for intra-data center applications, where 77% of the traffic is expected to reside in 2020. User-destined traffic will account for 14% of global data center IP traffic, and the remaining 9% will be between data centers [1].

Scaling the capacity of data center links has long relied on using multiple wavelengths or multiple fibers to carry conventional on-off keying (OOK) signals. This strategy cannot scale much further, however, as 400 Gbit/s links, for instance, would require 16 lanes of 25 Gbit/s, resulting in prohibitively high cost, complexity and power consumption. Recent research has focused on spectrally efficient modulation formats compatible with intensity modulation and direct detection (IM-DD) [2–5] to minimize power consumption. These efforts led to the adoption of four-level pulse amplitude modulation (4-PAM) by the IEEE 802.3bs task force to enable 50 and 100 Gbit/s per wavelength. Nevertheless, 4-PAM systems already face tight optical signal-to-noise ratio (OSNR) and power margin constraints in amplified and unamplified systems, respectively. Moreover, next-generation interconnects will likely need to accommodate increased optical losses due to fiber plant, wavelength demultiplexing of more channels, and possibly optical switches. To alleviate some of these constraints, both mature and emerging technologies can help on a number of fronts. High-bandwidth, low-power modulators [6] will reduce intersymbol interference (ISI) and improve signal integrity. Segmented modulators [7] may simplify the transmitter-side electronics. Avalanche photodiodes (APD) and semiconductor optical amplifiers (SOA) can improve receiver sensitivity of 100 Gbit/s 4-PAM systems by 4.5 and 6 dB [8], respectively. Improved laser frequency stability, either using athermal lasers [9] or frequency combs [10], will enable dense wavelength-division multiplexing (DWDM) within the data center, possibly yielding a multi-fold increase in capacity.

These technologies will extend the lifetime of 4-PAM, but they do not address the fundamental problem of such IM-DD systems, which is that they only exploit one degree of freedom of optical signals, namely, their intensity. Stokes vector detection has been proposed to enable up to three independent dimensions [11], while avoiding a local oscillator (LO) laser and coherent detection. Nonetheless, Stokes vector receivers rely on power-hungry analog-to-digital converters (ADCs) and digital signal processing (DSP) and do not address the problem of high required OSNR in amplified links or poor receiver sensitivity in unamplified links. Coherent detection allows four degrees of freedom, namely two quadratures in two polarizations, and significantly improve receiver sensitivity due to the LO laser gain. Coherent receivers based on analog signal processing [12] are particularly promising architectures because of their low power consumption, as they avoid high-

* Corresponding author.
E-mail address: jkperin@stanford.edu (J. Krause Perin).

https://doi.org/10.1016/j.yofte.2017.12.006
Received 28 April 2017; Received in revised form 4 November 2017; Accepted 10 December 2017
1068-5200/

Please cite this article as: Krause Perin, J., Optical Fiber Technology (2017), https://doi.org/10.1016/j.yofte.2017.12.006

speed ADCs and DSP. DSP-based coherent receivers may also become attractive in the future, as demand for even higher spectral efficiency increases, and as those systems are optimized for low-power, short-reach applications by leveraging more advanced complementary metal-oxide semiconductor (CMOS) integrated circuit processes. The high spectral efficiency enabled by coherent detection, combined with its improved receiver sensitivity, will potentially blur distinctions between intra- and inter-data center links.

In this paper, we review and compare these different detection techniques and their enabling technologies. In Section 2, we start by reviewing data center networks and important characteristics of intra- and inter-data center links. In Section 3, we review recent research on modulators, in particular electro-optic Mach-Zehnder modulators (MZMs). In Section 4, we discuss optical fiber requirements. In Section 5, we discuss direct detection (DD)-compatible techniques including M-PAM and orthogonal frequency-division multiplexing (OFDM), also commonly referred to as discrete multitone (DMT). We present comparative results in terms of receiver sensitivity and required OSNR. In Section 6, we review Stokes vector receivers that allow utilization of more than one degree of freedom of the optical channel. In Section 7, we review digital and analog coherent receivers, as well as differentially coherent receivers. In Section 8, we compare the different modulation formats and detection techniques according to their overall complexity and DSP power consumption. In Section 9, we conclude the paper.

2. Data center networks

2.1. Network architectures

In recent years, large internet content providers (ICPs) have begun to host and process large amounts of information in massive, hyperscale data centers. Evolving traffic patterns due to virtualization and cloud computing have led to shifts from north-south traffic, i.e., traffic from outside data centers to servers, to east-west traffic, i.e., traffic from servers to other servers within the same data center or another one nearby.

A traditional data center architecture, as shown in Fig. 1, consists of three tiers. In this scheme, servers connect to access switches that then connect to two aggregation routers for redundancy. These aggregation routers are then connected to core routers with redundancy. While this is an efficient structure to manage north-south traffic, it is inefficient for east-west traffic. Traffic from one server to another in the same data center may travel up to the core layer and then back down, traversing two access switches, two aggregation routers and a core router.

Hyperscale data centers have shifted to a flatter architecture consisting of two tiers [13], as shown in Fig. 2. In this configuration, servers are connected to leaf switches or to top-of-rack (TOR) switches that are connected to leaf switches, which in turn are connected to every spine switch, resulting in a multitude of paths. East-west traffic must now only travel to a spine switch before traveling back down to the desired leaf switch, resulting in low and predictable latency. Expanding the network is readily done by adding more leaf switches or spine switches, as needed. Fault tolerance is also improved, as a single spine switch failing will only result in a marginal decrease in performance. Achieving the full connectivity of the leaf-spine architecture does require more transceivers, as every leaf switch is connected to every spine switch.

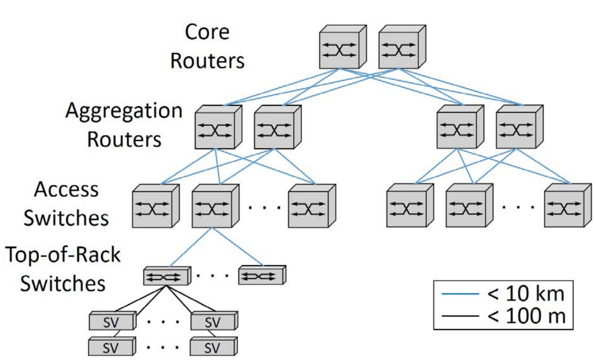

Fig. 1. A traditional three-tier data center architecture. Traffic from one server to another within the data center may need to travel up and through a core router.

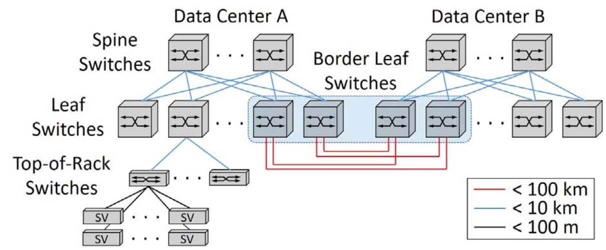

Fig. 2. A newer two-tier data center architecture. Intra-data center links are shown in blue and black, while inter-data center links are shown in red. Every leaf switch is connected to every spine switch. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this article.)

Interconnection between nearby (< 100 km) data centers is achieved by interconnecting their border leaf switches, as illustrated in Fig. 2. These inter-data center links have different constraints and impairments than the intra-data center links used within data centers.

2.2. Intra- and inter-data center links

Table 1 summarizes the different constraints and impairments of intra- and inter-data center, in contrast with long-haul systems. In long-haul systems, the high cost and power consumption of complex designs are amortized, as a 3-dB improvement in receiver sensitivity may double the reach and nearly halve the number of required repeaters. Intra- and inter-data center links, however, have other design priorities such as cost, power consumption, and port density, and they face fewer propagation impairments, as polarization mode dispersion (PMD) and nonlinearities are typically negligible over these short propagation distances.

Fig. 3a shows an example system model for an intra-data center link. The transceivers in these links can use multiple wavelengths to achieve high bit rates, but they are typically multiplexed and demultiplexed within the module. Intra-data center links reach up to 10 km and typically operate near 1310 nm to minimize total chromatic dispersion (CD). In this small-CD regime, receiver-side electronic equalization is effective, as shown in the performance curves of Section 5. Moreover, intra-data center links are typically unamplified, resulting in low power margin. APDs and SOAs may improve the receiver sensitivity, as shown in [8] and discussed in Section 5. Current intra-data center links employ either coarse wavelength-division multiplexing (CWDM) with wavelength spacing of 20 nm, or LAN-WDM with wavelength spacing of 4.5 nm to avoid power-hungry laser temperature control. Dense WDM (DWDM) may become commercially viable by leveraging advances in athermal lasers and frequency combs.

Fig. 3b shows an example system model for an inter-data center link. Inter-data center links reach up to 100 km and operate near 1550 nm to leverage erbium-doped fiber amplifiers (EDFAs). CD is significant and must be compensated. As CD is a nonlinear operation in IM-DD systems, simple receiver-side electronic equalization is not effective. Nevertheless, there are other effective electronic CD compensation techniques, which are discussed and compared in Section 5. Alternatively, CD may be compensated optically by dispersion-shifted fibers (DCFs) or tunable fiber Bragg gratings (FBGs) [14], depicted in Fig. 3b by the block CD^{-1}. Though they are less flexible than electronic equalization, they are more power-efficient.