



Towards standardization of compact data center switches

Gonzalo Guelbenzu*, Nicola Calabretta, Oded Raz

Electrical Engineering, Institute for Photonic Integration, Eindhoven University of Technology, Eindhoven, The Netherlands

Received 28 February 2017; received in revised form 10 May 2017; accepted 23 May 2017

Available online xxxx

Abstract

Data Centers, the essential infrastructure for the increasing amount of cloud applications, are required to expand in number of devices even further. However, the scaling is limited by both space and power constraints. It is therefore essential to shrink the size and power consumption of the devices. This work explores how On-Board Optics provides a solution for data center switches. We experimentally demonstrate a proof-of-concept data center switch with integrated On-Board Optics modules, power consumption below 150W and measuring only 0.25U. We further discuss the standards required to fully exploit this technology, and the future of On-Board Optics.

© 2017 The Korean Institute of Communications Information Sciences. Publishing Services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Society is increasingly relying on cloud applications such as Amazon, Google, or Facebook. These cloud applications are commonly executed in large data centers, which are a collection of switches and servers optically interconnected with a chosen topology [1,2]. For instance, a data center implementing a 3-level fat-tree topology requires 5120 128-ports switches to interconnect 131072 servers with full bi-section bandwidth [3]. Assuming one rack unit (1U) devices, 50 U/rack, and 300 W/device [4], this example requires above 2700 racks and 40 MW only for the devices. It is clear that adding more devices with similar characteristics, in order to continue expanding the computational capacity of these data centers, is reaching the point where it is impractical in terms of space and power.

The solution to this problem is to reduce the size and power consumption of the devices. For instance, by cutting by one half the size and power of the switches and servers, it is possible to duplicate the amount of devices in the data center without requiring extra space or power. For the servers, there is a

clear trend towards more modular and compact solutions [5]. However, the dominant design consideration for switches is that optical interfaces are front panel pluggable transceivers. The form factor of these standard optical modules leads to switches which are 1U in size having a front panel fully populated with transceivers, and consuming around 300 W [4]. On-Board Optics (OBO) transceivers are a promising alternative to achieve smaller data center switches with reduced power consumption. They are compact and not limited by the front panel bottleneck to increase the bandwidth density of the rack unit. Besides, OBO modules are more power efficient because they are located closer to the network switch chip. However, OBO modules lack an updated standard [6].

This work explores, experimentally, how the future release of updated standards for OBO transceivers could influence the design of data center switches. As a proof-of-concept, we have designed and implemented a 128×10 Gbps switch exploiting OBO. With a size of less than one fourth of the standard 19" rack case, it achieves power consumption of less than one half of the traditional switches. In addition, in this work we examine the need for new standards to maximize the benefits of this technology and future developments of OBO.

* Corresponding author.

E-mail address: G.Guelbenzu@tue.nl (G. Guelbenzu).

Peer review under responsibility of The Korean Institute of Communications Information Sciences.

<http://dx.doi.org/10.1016/j.ictexpress.2017.05.003>

2405-9595/© 2017 The Korean Institute of Communications Information Sciences. Publishing Services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

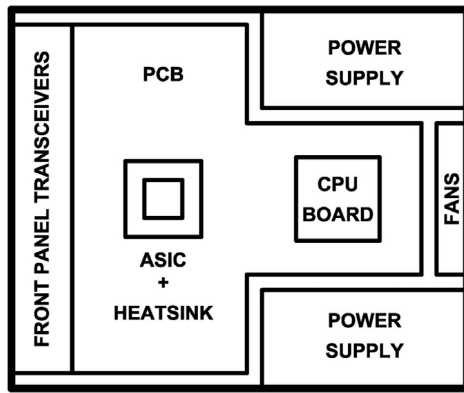


Fig. 1. Typical switch rack case layout.

2. Data center switches with front panel transceivers

Fig. 1 shows a typical 1U rack case layout of a data center switch [4]. The main components of the Printed Circuit Board (PCB) are one single switching ASIC (Application-Specific Integrated Circuit), the required receptacles for the front panel transceivers, and the socket for the CPU control board. In addition, the rack unit includes a redundant power supply, and hot air extracting fans located in the back.

Regarding transceivers choice, industry seems to favor 4-ports front panel transceivers that lead to 1U devices. For instance, the 10G version of the 128-ports switching ASIC [7] is packaged with 32 QSFP (4 × 10G) transceivers, and the 25G version [8] with 32 QSFP28 (4 × 25G) transceivers. And the recently announced 50G (2 × 25G) version of the 128-ports switching ASIC will be probably assembled with 32 DD-QSFP transceivers, whose standard has not been released yet. While it is clear that pluggable modules offer ease of use, they cannot compete on density and power consumption with OBO as it is shown in the following sections.

3. On-Board Optics transceivers

On-Board Optics is a promising alternative to front panel transceivers. Both types of transceivers provide the same functionality. They are responsible for the electro-optical conversion (E/O and O/E) required to electronically switch the data being transmitted through optical fibers. For instance, a basic VCSEL-based (Vertical-Cavity Surface-Emitting Laser) transceiver includes in the transmitter side the VCSEL driver and the VCSEL array, and in the receiver side, the PD (Photodiode) array, and the TIA/LA (TransImpedance Amplifier/Limiting Amplifier) [9].

However, OBO modules can be designed to be more compact and power efficient than the corresponding front panel transceivers. The main reason for this is that, being placed closer to the switching ASIC, OBO modules deal with shorter PCB channels (and reduced losses), which simplifies its design. For instance, a QSFP (4 × 10G) front panel transceiver requires approximately the same power but double area than the commercial 12 × 10G OBO module [10] integrated in the prototype introduced in the next section. Even more compact

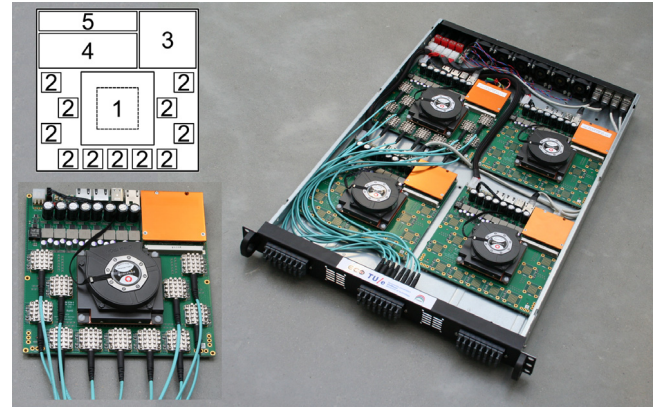


Fig. 2. Experimental prototype diagram, picture, and packaging. 1—ASIC and fan, 2—OBO transceivers, 3—CPU control board, 4—12V input power supply, 5—connectors.

devices have been reported, such as a 24×20 G transceiver in just $14 \text{ mm} \times 14 \text{ mm}$ area [11].

4. Data center switches with On-Board Optics

4.1. Experimental prototype and packaging

Fig. 2 presents the proof-of-concept design of a data center switch with OBO transceivers in the lower left corner. It integrates the 10G version of a 128-ports switching ASIC [7], covered by a fan in the figure. Eleven 12-ports OBO transceivers [10] surround the switching ASIC, providing the optical interfaces through MPO (Multi-Fiber Push On) connectors. The top left corner of the board includes the interface connectors, and the voltage converter blocks. A total of nine phases generate the internal voltages required by the board from a single 12 V power input. The 12 V input is suggested by the Open Compute Project (OCP) to enable power distribution engineers to optimally design the amount and placement of power supplies in the data center, and also, to fully exploit the rack unit space. The top right corner of the board includes the CPU control board, with an orange heatsink, that runs the operating system, and converts the openflow messages received from the data center controller into adequate instructions for the data plane switching ASIC. Below the CPU board are all the rest of the components. The final size is $200 \text{ mm} \times 200 \text{ mm}$, and makes it one of the most compact, fully integrated, single 128-ports switches ever implemented.

Fig. 2 also shows the packaging proof-of-concept of four of those switches in a single rack unit. We did not attempt to build a larger switch in agreement with [12].

Each one of the compact boards requires only half of the rack width. Besides, thanks to the single 12 V power supply input, it is possible to remove the power supplies from the rack unit, and fully populate the space available with four boards. The front panel, free from transceivers, accommodates the MPO connectors array. While this design provides four times the bandwidth density, compared with a switch implementation using QSFP modules, there is still enough room in the front

Download English Version:

<https://daneshyari.com/en/article/4966330>

Download Persian Version:

<https://daneshyari.com/article/4966330>

[Daneshyari.com](https://daneshyari.com)