Characteristics of Field Operation Data for Optical Transceivers in Hyperscale Data Centers

Chongjin Xie[1], Chunxiao Wang[2], Qin Chen[3], Zhicheng Wang[3], Peng Wang[3], Rui Lu[3], Lei Wang[2]
1. Alibaba Cloud, Sunnyvale, CA, USA, 2. Alibaba Cloud, Beijing, China, 3. Alibaba Cloud, Hangzhou, China
chongjin.xie@alibaba-inc.com

Abstract: We collect and analyze field operation data of optical transceivers in hyperscale data centers, including temperatures, infant mortality and causes of failures, to shed some light on reliability of optical transceivers in data centers.

1. Introduction

Most internet services and cloud computing are running in data centers, and data centers have become a crucial information infrastructure in our society. Data centers are buildings that house massive numbers of servers, and these servers are connected through data center networks (DCNs) to form a large distributed computing system to provide required computing power for various internet and cloud computing services [1]. Today virtually all the connections in DCNs use optical interconnect technology, which connects almost all the routers, switches, and even servers, and a typical hyperscale data center operator has millions of optical transceivers in their networks. Even a small optical transceiver failure rate may become a big threat to the health of DCNs and could cause service interruptions. Today the qualification and reliability requirements of optical transceivers are still based on Telcordia GR-468-CORE, which was developed for optoelectronic devices used in telecommunications equipment and hasn't been revised since 2004 [2]. The document does not reflect the recent technology advances and operation requirements of optical transceivers in data center environment. For example, most optical transceivers in data centers use chip-on-board (COB) non-hermetic packaging technology and have less than 5-year lifetimes. Considering most of the optical transceivers in the industry are consumed by data centers, there have been some industry efforts to define reliability requirements of optical transceivers for data center [3]. Therefore, to understand the reliability behavior of optical transceivers in data centers is not only important for data center operators, but may have a profound impact on our industry as well.

In this paper, we present the statistical operation data in hyperscale data centers collected and analyzed by an optical transceiver data analytics system, including temperatures of multi-mode and single-mode transceivers with different speeds and in different locations, infant mortality failure data of some transceivers, and root causes of optical transceiver failures in our data centers. Some techniques to prevent optical transceiver infant mortality and to detect optical transceiver failures in advance are discussed.

2. Optical Transceiver Data Analytics System

Figure 1 shows the architecture of the optical transceiver data analytics system we developed to monitor and analyze the operation of optical transceivers in our production DCNs [4][5]. The system collects, stores and analyzes online operation data of optical transceivers from servers, switches, routers, and optical transport networks (OTNs) equipment across entire networks covering all our data centers. In the data collection layer, multiple running collectors pull or subscribe different information data of optical transceivers at frequencies of up to minute levels. Optical transceiver manufacturing data can also be directly collected from vendors by the system. The most recent raw data is temporally stored in an online collection database, and periodically transferred to the MaxCompute platform with the help of integrated extract, transform, load tools for data consistency checking and subsequent processing. Note that MaxCompute is a general purpose, multi-tenancy data processing cloud platform for large-scale data warehousing. To have a better understanding of failures and reliability of optical transceivers, return merchandise authorization (RMA) and failure analysis (FA) data are also input to the system and stored in an RMA/FA database. In addition, several other online databases are integrated into the data warehouse, including inventory and configuration databases. Then data processing is implemented and a machine learning (ML) platform is also used in the data storage and processing layer to accelerate big data analyzing and processing. Data processing and analyzing results are provided to the data application layer for various services, such as resource management, reliability evaluation, risk and fault assessment, ticketing, visualization et al.

3. Statistics of Operation Data

There are three generations of DCNs in our production networks, 40G, 100G and 200G. Vertical-cavity surface-emitting laser (VCSEL) based multi-mode 40G SR4 are mainly used in the 40G DCNs, and in 100G and 200G DCNs, we use mixed multi-mode technology (100G SR4 and 200G SR4) and single-mode technology (100G CWDM4 and 200G FR4). Note that we have two versions of 200G SR4 transceivers, one is based on analogy technology without using any digital signal processing (DSP), and the other uses DSP chips. We observed that multi-mode optical transceivers have a much lower failure rate than single-mode ones in our DCNs.

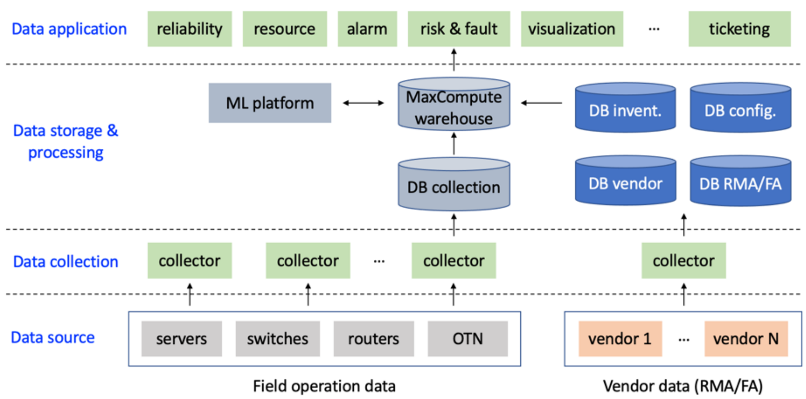

Fig. 1. Architecture of the optical transceiver data analytics system. OTN: optical transport network. DB: database, ML: machine learning. RMA: return merchandise authorization, FA: failure analysis.

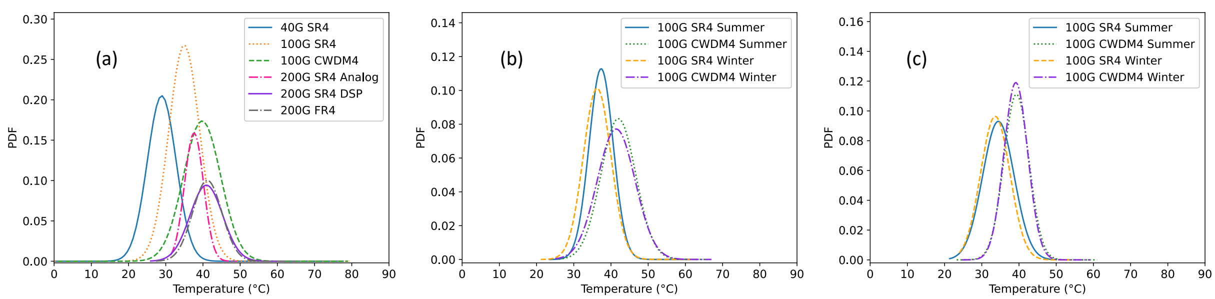

Fig. 2. Optical transceiver temperature distributions. (a) temperatures of six different types of optical transceivers in all data centers in China, (b) summer and winter temperatures of 100G SR4 and CWDM4 in a data center in northern China, (c) summer and winter temperature of 100G SR4 and CWDM4 in a data center in southern China. PDF: probability density function.

One of the most important operation environment parameters is temperature. Fig. 2 depicts the statistical distributions of optical transceiver case temperatures collected in the period of 2021-01-01 ~ 2021-10-09. In Fig. 2 (a), we plot the probability density functions (PDFs) of six different types of optical transceivers in all our data centers in China. It shows that the temperatures increase with increasing transceiver speeds, and the temperatures of single-mode transceivers are higher than those of multi-mode transceivers. The average temperatures of 40G SR4, 100G SR4 and 100G CWDM4 are 28.9 ºC, 35.0 ºC, 39.8 ºC, respectively. This is because power consumption of higher speed optical transceivers is larger and single-mode optical transceivers' power consumption is higher than multi-mode ones. As analog based 200G SR4 consumes less power than DSP based 200G SR4, its temperature is lower. DSP based 200G SR4 has similar temperature as 200G FR4 since they consume similar power (200G FR4 uses DSP too and DSP chips consume most of the power in an optical transceiver). The figure clearly shows that even in data center environment, optical transceiver operation temperatures have a wide range, and high temperatures can be larger than 70.0 ºC. Figs. 2 (b) & (c) show that in an air-conditioning environment, there is little difference on temperatures between seasons and locations.

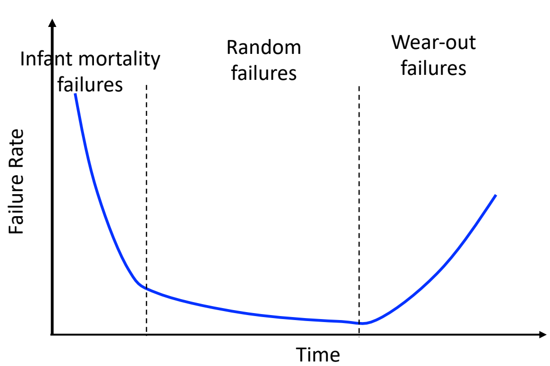

Fig. 3. "Bathtub" failure rare function of optoelectronic devices

The failure rate characteristics of optoelectronic devices change over time and a failure rate function typically has a "bathtub" shape, as shown in Fig. 3. There are three types of failures over a device life: 1) infant mortality failures, which occur early in life and at a declining rate; 2) random failures, which occur at a relatively constant rate when the device is in stable working conditions; 3) wear-out failures, the rate of which generally increases with time due to aging [2]. As lifetimes of optical transceivers in data centers are typically short, to increase optical transceiver reliability in their DCNs, data center operators in general do not need to worry too much about wear-out failures, but need to find ways to prevent infant mortality failures and reduce random failures.

We observed some infant mortality failures in our networks, and Fig. 4 gives two examples. The figure depicts the failure numbers of optical transceivers over field operation time for two batches of 100G CWDM4 transceivers. In the case of Fig. 4 (a), most of the failures occurred within the first year after deployment, and in Fig. 4 (b), most failures occurred within the first two years after deployment, and after the infant mortality failure period, the optical transceivers became stable. The infant mortality could be reduced by enhanced screening such as extending burn-in time in device manufacture processes.

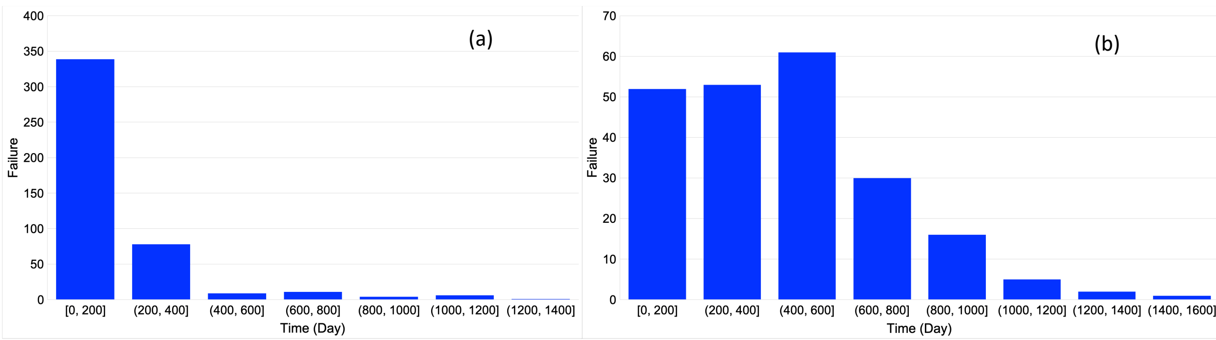

Fig. 4. Two infant mortality failure cases for 100G CWDM4 transceivers.

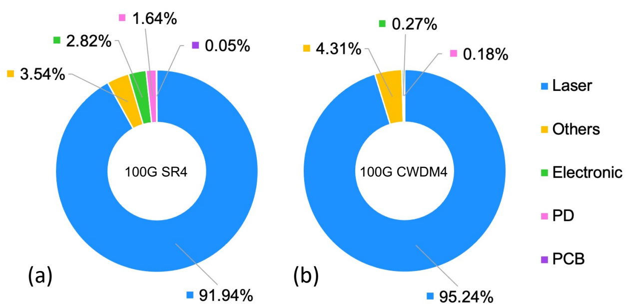

Fig. 5. Distributions of failure reasons for 100G SR4 (a) and 100G CWDM4 (b). PD: photo-detector, PCB: printed circuit board.

We found that most failures of optical transceivers in our DCNs are caused by lasers. Fig. 5 gives the distributions of reasons that cause the failures of 100G SR4 and CWDM4 optical transceivers in our data centers. It shows that lasers caused about 92% of the failures for 100G SR4 and 95% for 100G CWDM4, respectively. Electronics in Fig. 5 mean high-speed electronic chips such as drivers and trans-impedance amplifiers (TIAs), and the figures show that electronics have a much lower failure rate than optics in the transceivers.

We observed that most of the laser failures can be detected in advance by monitoring laser output optical powers. For example, a continuous decrease of output optical power typically indicates that a laser will fail soon. There are many power variation patterns. By combining large volumes of real time collected data with machine learning technologies, problematic optical transceivers can be proactively detected before they fail.

4. Summary

We presented statistical field operation data of optical transceivers in hyperscale data centers, including operation temperatures, failure behaviors, and failure reason distributions, which can be of some help to understand optical transceiver reliability in data centers.

References

[1]. L. A. Barroso, U. Hölzle, and P. Ranganathan, The Datacenter as a Computer, Third edition, Morgan & Claypool Publishers, 2018.

[2]. “Generic Reliability Assurance Requirements for Optoelectronic Devices Used in Telecommunications Equipment,” GR-468-CORE, Issue 2, Sep. 2004.

[3] “Datacenter Optics Reliability: Can We Standardize Requirements, and Can They be Relaxed Given Redundancies and <~5-Year Lifetimes?”, workshop S1D, OFC’2019.

[4]. J. Li, Z. Wang, C. Wang, Q. Chen, P. Wang, R. Lu, C. Xie, “Data Analytics Practice for Reliability Management of Optical Transceivers in Hyperscale Data Centers,” OFC’2020, paper T3K.6, 2020.

[5]. C. Wang, L. Wang, Z. Wang, Q. Chen, P. Wang, R. Lu, C. Xie, “Data Analytics and Unsupervised Learning Enabled Proactive Maintenance for Optical Transceivers in Hyperscale Data Centers,” OECC’2021, paper M4A.3, 2021.