Characterizing Large-Scale Production Reliability for 100G Optical Interconnect in Facebook Data Centers

Abhijit Chakravarty, Srinivasan Giridharan, Matt Kelly, Ashwin Poojary and Vincent Zeng

Facebook Inc., 1601 Willow Road, Menlo Park, CA 94025

ABSTRACT:

Facebook is deploying cost effective 100G CWDM4 transceivers in data centers. This paper describes the post production performance monitoring system which is being implemented to identify optical interconnect early failure modes.

INTRODUCTION:

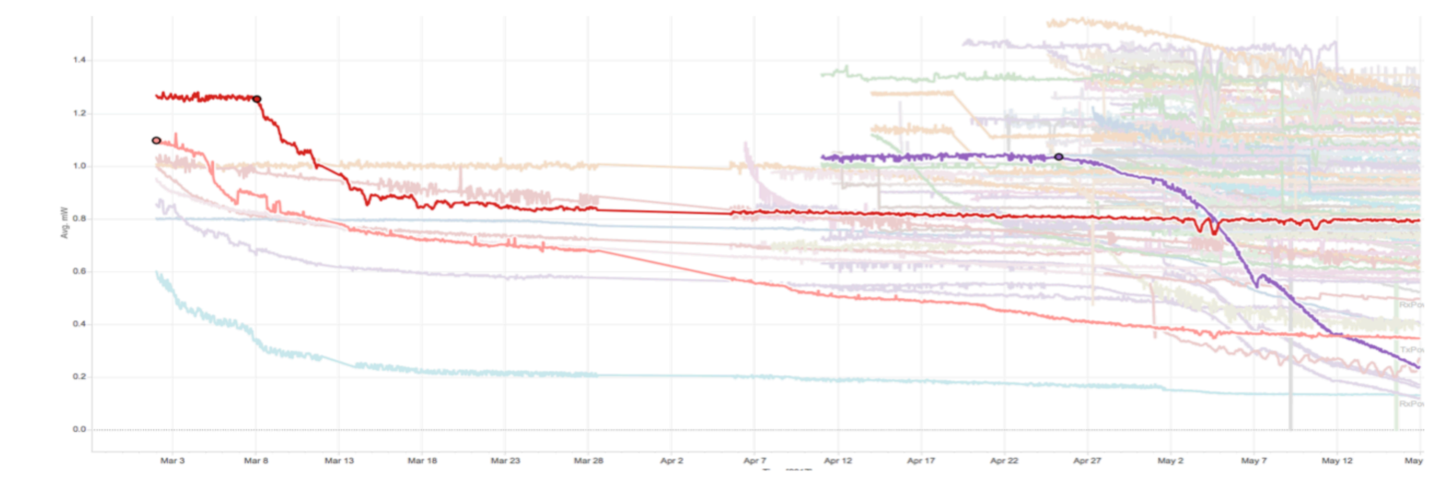
To meet current requirements and future bandwidth demands, Facebook (FB) data centers are continuously pursuing ways to handle data more efficiently and at ever-faster speeds. As the data-rate requirements increase due to large number of users (currently over 1.8 billion people use FB every month) and increased data usage per user, higher bandwidth optical interconnects are needed to support the demand. FB is migrating the TOR and Fabric switch interconnect bandwidth of data centers to 100Gbps. To realize this migration, FB is using 100G CWDM4 technology [1] for intra data center applications with single mode fiber installations. Since 100G CWDM4 is a new technology, our objective is to proactively manage avalanche failures induced by unexpected optical power degradation of the 100G CWDM4-QSFP28 transceivers during active operation in the FB data center. To serve this objective we have developed a production performance monitoring framework to measure DOM parameters for the 100G CWDM4 transceivers in active deployment and assess their operational health. Given the operational complexity of responding to avalanche failure events in remote regions, our monitoring includes the necessary slice and dice options to quickly characterize the field events (DOM parameters) by vendor, data center region, host, component (Tx/Rx) and channel.

We are experiencing several new early failure modes such as side mode suppression ratio (SMSR), channel broadening, optical power degradation, PCB delamination at high temperature, DOA laser failures, etc. which we are trying to identify at an early stage to avoid the infant mortality failure at Facebook's data centers. We are closely monitoring the trend in time of all BOL and EOL. This paper discusses Facebook's monitoring methodology to capture these early failure modes and ensure a smooth deployment of 100G CWDM4 technology. DOM parameters and are trying to understand the CWDM4 DML laser degradation, case temp and optical power behavior. Based on this data, we are trying to understand the 100G CWDM4 reliability including swap rate, MTBF and FIT rate across vendors. As one of the first few companies who have deployed 100G CWDM4 optical interconnect technology at scale for intra data center applications, our monitoring methodology shows the maturity of the process, and our trends are a good indicator of the maturity of this technology. We are developing a mechanism to predict failures based on the degradation slope defined from BOL and EOL. This paper discusses Facebook's monitoring methodology to capture these early failure modes and ensure a smooth deployment of 100G CWDM4 technology.

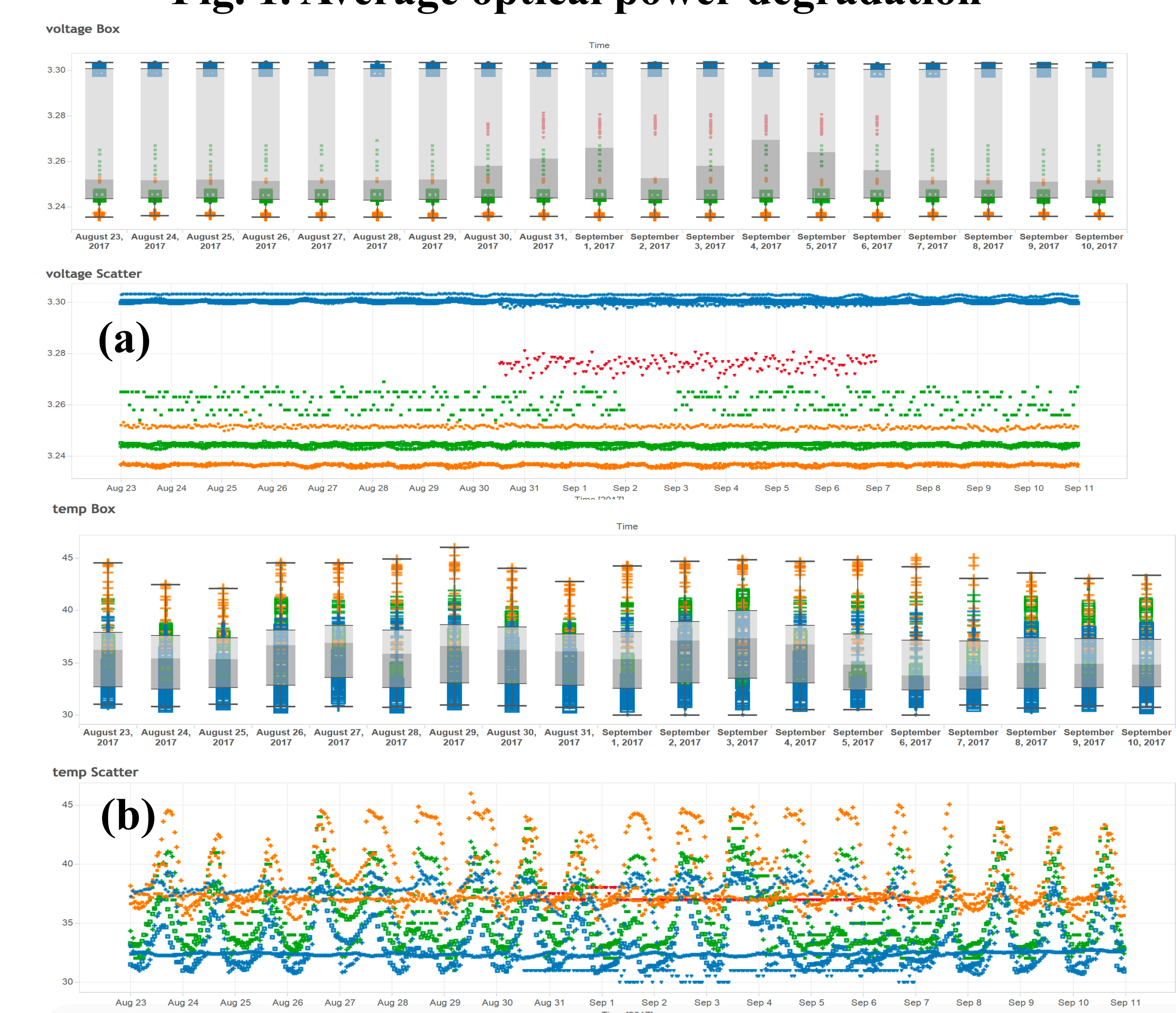
DATA COLLECTION AND MONITORING SYSTEM:

Custom scripts [2] were added to periodically collect module case temperature, supply voltage, optical power and Tx bias from the DOM. A process parses and stores these into a time series tracking database. Alarms are set to capture individual transceivers exceeding Facebook 100G CWDM4-OCP specifications [3]. Supply voltage alarms are set at ±10% of 3.3V. Temperature alarms are set at 55°C and 15°C (FB relaxed spec). Power alarms are set at .251mw. Trend data is aggregated hourly by data center region, network switch platform, vendor, component, port and channel. Transceiver count provides a quick verification that the entire process is working, matches the actual install base across the various aggregations, and measures the impact of any trends found. A separate process is used to monitor swap rate, MTBF, FIT rate and cause of failure to detect trends and help our vendors improve their product design.

Most transceivers are operating within spec and do not show notable power degradation with a goal of no more than ±1dB change over the life of the transceiver. However, there is an important subset of modules showing significant power degradation, and we have observed two general problematic scenarios: (1) modules that degrade quickly but then stabilize after a few days of operation and stay within spec (the red, orange, blue lines in Fig. 1 - we hypothesize this could be relate to epoxy curing) and (2) modules that degrade continually and fall out of spec (the purple line in Fig. 1). With this data, we are then able to discuss potential causes/remediation with our suppliers.



"Fig. 1. Average optical power degradation"



"Fig. 2.(a) Supply voltage (b) QSFP28 case temperature"

CONCLUSION:

We have successfully implemented this monitoring system across all our data centers and are closely observing the performance. Through this monitoring system, we may be able to identify failures modes such as optical power degradation, temperature instability and anticipate the time to failure for infant mortality parts and notify operations of forthcoming part swaps and RMA. As next steps, we are developing a mechanism to predict failures based on the degradation slope defined from BOL and EOL.

References:

[1] Abhijit Chakravarty, Katharine Schmidtke, Srinivasan Giridharan, John Huang, and Vincent Zeng, "100G CWDM4 SMF Optical Interconnects for Facebook Data Centers", Conference on Lasers and Electro-Optics, OSA Technical Digest, paper STu4G.1, 2016.

[2] Venkataraman and Thayer, Tang, "Facebook's Large Scale Monitoring System Built on HBase", Strata + Hadoop World Conference, 2012.

[3] Katharine Schmidtke, Vincent Zeng, Abhijit Chakravarty "Facebook: CWDM4-OCP, 100G Optical Transceiver Specification", Open Compute Project © Facebook, 2017.