An Open Line System with Ultra-fast Protection Switching for Data Center Interconnect

Juan Wang*[,1], Yu Jin[1], Chen Zhu[1], Feng Gao[1], Yongxin Cui[1], Gang Cheng[1], Xu Zhou[1]
[1]Baidu, Network Infrastructure Department, Kuike Technology Building, No.9 Shangdi Information Road, Haidian District, Beijing, China
wangjuan28@baidu.com

Abstract: We present a DCI OLS with 5ms ultra-fast protection switching. By optimizing the DSP traffic rebuild time with a magneto-optic switch, we are able to improve the current ITU standard by an order of magnitude. © 2023 The Author(s)

1. Introduction

In the cloud computing era, businesses are deployed over a large-scale distributed computing platform composed of millions of servers in disperse data centers (DCs) that are interconnected via optical network. The metro DC interconnect (DCI) optical networks are rapidly evolving to support exponentially growing regional traffic demand [1]. Today's DCI metro optical network mainly features an open line system (OLS) with the flexibility to adopt latest transceiver innovation, unified management interface and smart software-defined-control functionalities [2]. Low-complexity point-to-point system with modular equipment is widely deployed as the underneath DCI metro OLS architecture, which enables simple replication for flexible and scalable network expansion. Furthermore, OLS is typically equipped with automatic optical line protection (OLP) switching to provide guaranteed service level agreement (SLA) as optical fiber cuts and degradations are the major cause for service disruption.

Due to its open nature, DCI metro OLS consistently enjoys hardware and software innovations [3], however there is few works been announced for the OLP aspect. As the advanced opto-electro semiconductor developments keep pushing a higher total bit rate carried by each fiber, the OLP switching time plays more and more important role for customer experience since the packet loss during traffic re-establishment linearly scales with the per-fiber capacity, which is particularly critical for time-sensitive applications such as high-performance computing. Industry has been holding to the ITU standard [4] of completing the switching within 50ms for decades, in this work we have proposed an ultra-fast-protection-switching (UFPS) subsystem, based on optimized coherent digital signal processing (DSP) traffic recovery algorithms and a magneto-optic switch, service can be reestablished within 5ms. We have experimentally demonstrated the ultra-fast switching time with extraordinary consistency in hundreds of thousands tests over multi-vendors' coherent transceivers, and UFPS has already been integrated in our OLS to widely serve Baidu's DCI metro optical network. To the best of our knowledge, this is the first time that an order of magnitude of switching time improvement is realized for successful field deployment compared with ITU standard.

2. Proposed UFPS Transmission System

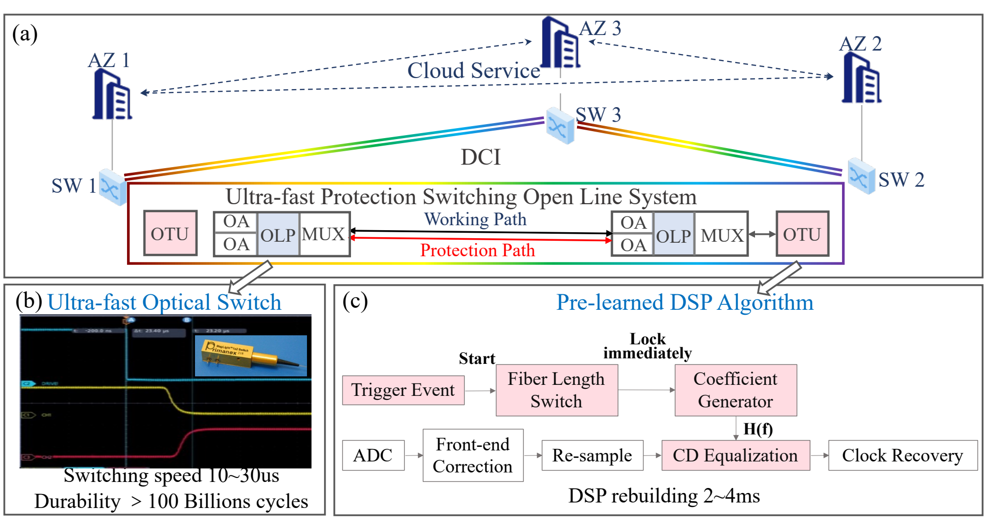

Fig. 1 (a) Architect, (b) ultra-fast optical switch and (c) optimized DSP rebuild algorithm of an UFPS transmission system

Fig.1(a) shows a Baidu's typical metro DCI network, in order to meet service latency requirement of less than 2ms, our metro distance is restricted to below 200km between two available zones (AZs). The data communication between AZs are carried by separate point-to-point DWDM OLS optical transmission systems, that consists of multi-vendor optical transponders units (OTUs), wavelength multiplexers, OLP and optical amplifiers. Cost-effective Optical Multiplex Section Protection (OMSP) is used in most of our metro DCI links.

UFPS follows similar traffic rebuild flow as the traditional optical protection switching subsystem. The first step is to trigger the protection switching by continuous optical power monitoring of the working and protection path; upon switching completion the coherent DSP would reinitialize the entire state machine for data recovery. We have been working on both the optical and digital step to achieve ultra-fast switching. Conventionally, mechanical optical switch is used to perform the optical switching that generally takes 5~10ms, we replaced it with a magneto-optic switch that switches with a solid-state all-crystal core based on Faraday Effect. The new optical switch can complete the switching in 10-30us as depicted in Fig.1(b), it also has great reliability and durability of more than 100 billions switch cycles that is even suitable for the field of aviation and national defense, furthermore, it also has low-power consumption and low drive voltage advantages compared to the mechanical switch. On the digital side, Fig.1(c) depicts the optimized DSP traffic recovery flow for UFPS. During traditional DSP traffic recovery, chromatic dispersion (CD) estimation is triggered to re-estimate the fiber length and set the corresponding CD equalizer coefficients, which dominates the total rebuilding time (e.g. 10-30ms), as the CD estimation algorithm is normally based on a lengthy CD scan process [5]. In UFPS we pre-learn the working and protection fiber lengths and stored them into the DSP ASIC in Day 1 (i.e. as part of the installation process), and while the CD estimation is triggered by service disruption events (e.g. LOS), the fiber length are immediately 'switched' and CD equalizer coefficients are set instantly, that reduce the total coherent DSP rebuilding time to 2-4ms. By combining the proposed optical and digital approaches, UFPS-assisted OLS is able to achieve a 5ms service recovery time.

3. Experimental Demonstration

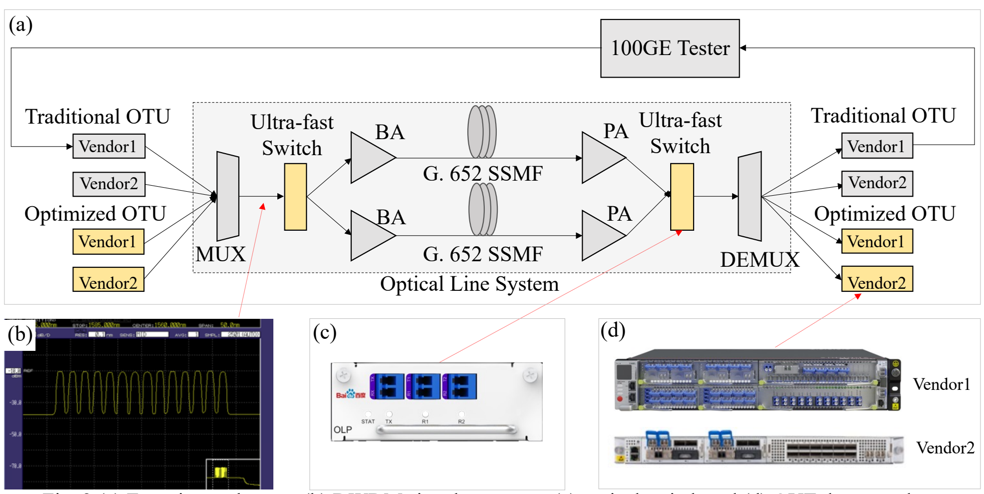

Fig. 2 (a) Experimental setup, (b) DWDM signal spectrum, (c) optical switch and (d) OUT demo cards.

Fig.2(a) shows the experimental setup of the UFPS-assisted DCI metro OLS network. Fourteen OTUs are multiplexed into 100GHz ITU grids via a fixed-grid Array Waveguide Multiplexer and its spectrum is shown in Fig.2(b), there are two OTU vendors both use CFP2-DCO module to convert 4x100GE client payloads onto a 400Gb/s 62-Gbaud dual-polarization 16-QAM line-side signal for each wavelength (Fig.2(d)). The DWDM signal was split equally into two paths with the 50:50 splitter inside the OLP card, amplified by a pair of booster amplifiers (BAs), and propagate into two G.652 SSMFs. At the receiver side, after pre-amplifiers (PAs) compensates for the span loss, the magneto-optic switch selects the proper fiber path to be forwarded towards the Rx. A VIAVI MTS-6000A Ethernet test platform was connected to one of the 100GE client ports to measure the exact traffic recovery time during protection switching.

In order to maintain DCI OLS' key benefit of multi-vendor OTU compatibility and interoperability, we have performed same rebuild optimization procedure on two vendors' DSP ASIC, the successful implementations prove that our proposed optimization method can be adopted by wide range DSP vendors therefore is suitable for OLS. Both the OTUs and the OLS are controlled via unified Baidu Yang software models, which also contains a configuration node to set the pre-learned working and protection fiber lengths. In order to emulate the real field deployment scenarios and test the robustness of UFPS, we have selected different fiber lengths for working/protection fiber paths: 40km/80km, 40km/120km and 100km/200km. The protection switching was triggered by manually switching the optical switch throughout all our tests.

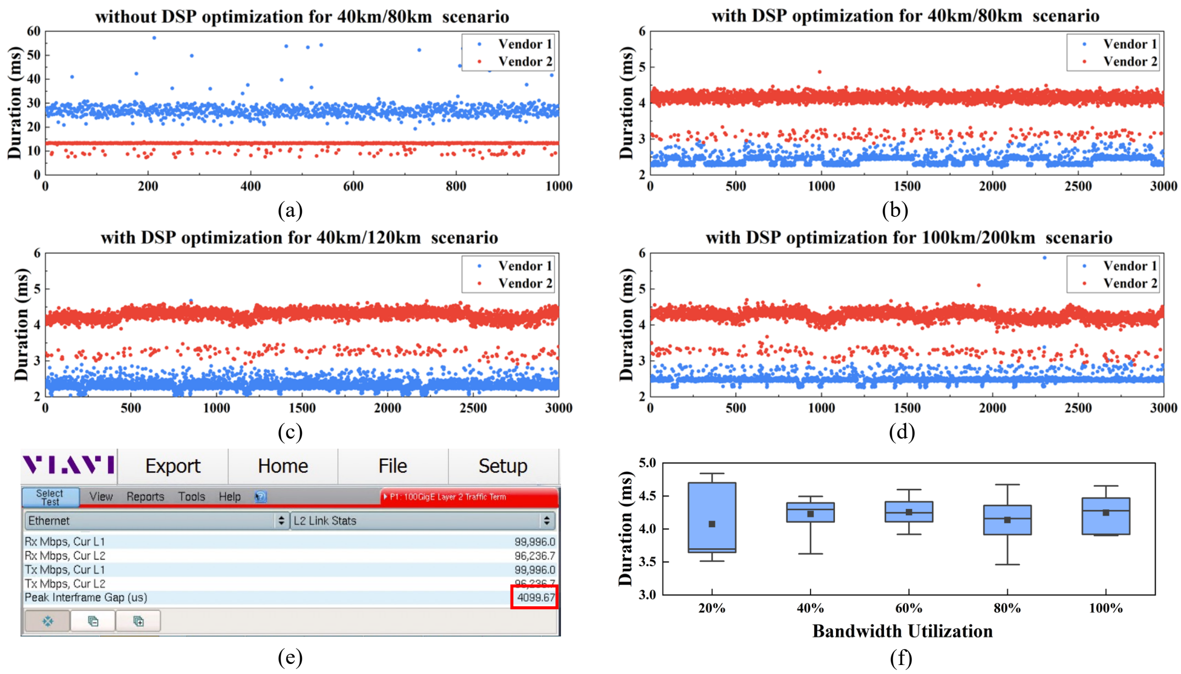

Fig. 3 (a) Traditional DSP switching time in the 40km/80km scenario; optimized DSP switching time in (b) 40km/80km, (c) 40km/120km and (d) 100km/200km of the working/protection path fiber length scenarios; (e) traffic recovery time measurement on ethernet tester; (f)average switching time measured by network tester with different 100GE bandwidth utilizations.

Fig.3(a) depicts the baseline results of two OTU vendors without DSP optimization under 1000 switching protection tests with 40km/80km setup, an average of 30ms and 13ms are obtained for vendor 1 and 2, respectively. The worst case performance of vendor 1 even exceeded the 50ms standard. Same vendors' OTU performance with optimized DSP firmware are shown in Fig. 3(b)-(d) for 40km/80km, 40km/120km and 100km/200km test setups, respectively. Clearly, the average protection switching time for both vendors are below 5ms within total 9000 test cycles. The snapshot of reading one switching time measurement by Ethernet tester is shown in Fig.3(e). Although not shown in this paper, we have performed more than 100,000 experimental tests on each vendor's OTU and the statistics show that the UFPS can achieve a sub-5ms probability of >99.99%.

Fig.3(f) shows a real-field monitoring result of one of the 100GE port of a deployed UFPS-assisted OLS, the protection switching time is represented by measured average dropped packet time. At all different bandwidth utilization levels, the service can be rebuilt within ~4.2ms, fully meets our expectation.

4. Conclusions

We have proposed an ultra-fast protection switching method that employs a magneto-optic switch and optimizes coherent DSP traffic rebuild algorithm to bring the optical switching time down to 5ms. The proposed UFPS subsystem is developed and successfully deployed in our metro DCI OLS to greatly enhance our cloud customer experience. Our experimental and field-deployment results prove that the current ITU standard of 50ms protection switching time could potentially be reduced by an order of magnitude for metro DCI in the near future.

5. References

[1] Cisco Global Cloud Index: Forecast and Methodology, 2014–2019", Cisco, April 2016
[2] M. Filer, et. al, "Toward transport ecosystem interoperability enabled by vendor-diverse coherent optical sources over an open line system," J. Opt. Commun. Netw. 10, A216–A224 (2018)
[3] N. Sambo, et. al, "Experimental demonstration of fully disaggregated white box including different types of transponders and monitors, controlled by NETCONF and YANG," in Optical Fiber Communication Conference (2018), paper M4A.3
[4] ITU-T G.873.1 : Optical transport network: Linear protection
[5] F. N. Hauske, et. al, "Precise, Robust and Least Complexity CD estimation," in Optical Fiber Communication Conference (2021), paper JWA032.