Unavailability Analyses of Hyperscale Data Center Interconnect Optical Networks with Optical Layer Protection

Lingling Wang[1,*], Lei Wang[1], Chunxiao Wang[1], Chongjin Xie[2]
1. Alibaba Cloud, Alibaba Group, Beijing, China, 2. Alibaba Cloud, Alibaba Group, New York, USA
*wanglingling.wll@alibaba-inc.com

Abstract: With massive field operation data collected from our production optical networks, we analyze the network unavailability of metro data center interconnect networks where optical layer protection is used, and the main factors affecting network unavailability are quantified. © 2024 The Author(s)

1. Introduction

Data center interconnect (DCI) optical networks, which offer ultra-high-speed, low-latency, and high-bandwidth network connections, have been widely used in cloud data centers to support fast growing data-intensive workloads of cloud operators. Driven by growing demands, the scale and topology of optical networks become increasingly complex, and services can be disrupted by any failures in the networks. As the most basic infrastructure for data exchange and transmission among data centers, it is of great significance to accurately identify the main factors leading to the unavailability of optical networks to ensure the quality of cloud services [1].

According to previous work, we quantified the factors affecting the unavailability of backbone optical network (BN) without protection for the 1st time and showed more than 95% was caused by transmission fibers [2]. Thus, in this work, we focus on a production DCI metropolitan area network (MAN) with optical multiplex section protection (OMSP) for all links, which has a hierarchical structure composed of optical transmission sections (OTSes), optical multiplex sections (OMSes), and optical channels (OCHes), as shown in Fig. 1(a). Each OCH in the network has an average length of 63 km and usually traverses 1 or 2 OMSes. For every OMS, there are two parallel paths between each pair of network nodes, working path OMS_W and protection path OMS_P, and every path generally contains 1 or 2 OTSes. The same traffic is transmitted over both paths simultaneously and at the receiver side, the traffic from one path is selected according to a certain rule. When a failure occurs on the path, the traffic is automatically switched to the other path by an automatic protection switch (APS), ensuring that the network remains operational and that there is no disruption to the services being provided [3].

In this paper, based on daily network performance data and case tickets collected by Alibaba Optical Network Analytics (AONA) platform, we analyze nearly eight months of operation data for more than 5000 OCHes in the optical network with optical layer protection, quantify the effect of OMSP on network availability, and study the main factors affecting network unavailability in this network.

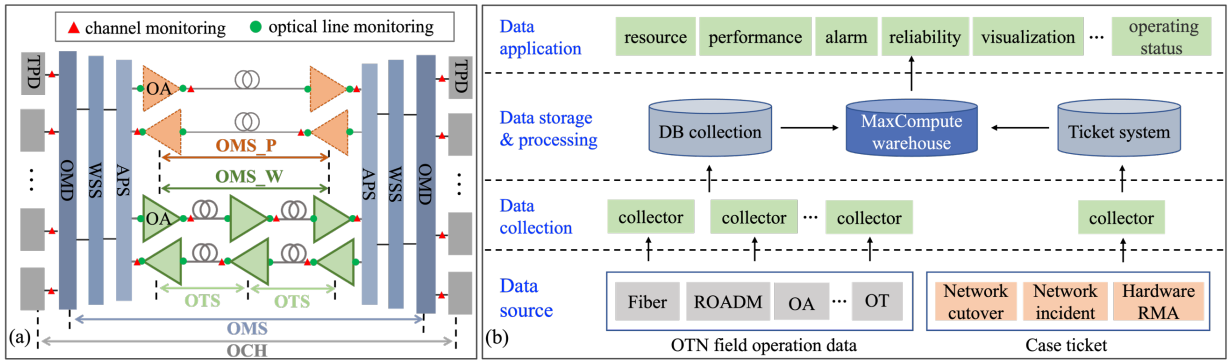

Fig. 1. Architecture of optical link with OMSP and Alibaba Optical Network Analytics (AONA). (a) architecture of optical link with OMSP. (b) architecture of AONA. TPD: transponder; OMD: optical multiplexer and demultiplexer; WSS: wavelength selective switch; APS: automatic protection switch; OA: optical amplifier; OTS: optical transmission section; OMS: optical multiplex section; OCH: optical channel; OMS_P/OMS_W: protection/working path of OMS; ROADM: reconfigurable optical add–drop multiplexers; OT: optical terminal; RMA: return merchandise authorization; DB: database.

2. Alibaba data analytics platform of DCI optical network

Fig. 1(b) shows the data analytics platform architecture of our DCI optical networks for data monitoring and analyzing. In the data collection and storage layer, we collect network field operation data and store them in the online collection database (DB), and the change or failure events are saved in the case ticket system. In the processing layer, data in DB and case ticket system are transferred to the MaxCompute platform, a home-grown big data analysis platform, for long-term data storage and analysis. Finally, MaxCompute provides data support for various services at the application layer, including visual monitoring, network alarm, network status evaluation, and so on [2].

3. End-to-end network unavailability analysis result

As shown in Fig. 1(a), there are two levels of performance monitoring in our network. One is channel monitoring of OCH, including optical power, pre-forward error correction (FEC) bit error rate (BER), error seconds (ES), unavailable seconds (UAS) [4], and so on. The other is optical line monitoring, including optical power, gain, and attenuation of optical amplifier (OA), which can be used to evaluate the real-time status of fiber. We defined end-to-end network unavailability as abnormal seconds (AS) generated within a 15-minute period on an OCH. Note that, AS=ES+UAS. Combined with performance data, alarm, and case tickets stored in MaxCompute for nearly eight months of more than 5000 OCHes, we classify the causes of network unavailability, as shown in Table 1.

Table 1. Unavailable cause information.

Cause	Data source	Description
OMSP switching	Alarm	Path automatically switching of OMSP.
Device hardware RMA	Ticket	Equipment hardware failure.
Fiber incidents	Performance	Fiber failure that OMSP can't work out, such as both OMS paths break or degrade simultaneously.
Config/Operation incidents	Performance/ Ticket	Misconfiguration or incorrect maintenance operation.

Table 2. Comparison of the two networks.

	BN	MAN
Number of OCHes	400+	5000+
Average OTSes of OCHes	13	1.2
Average length of OCHes (km)	905	63
Annualized fiber failure rate (%)	0.068	0.077
Normalized network unavailability (%)	100	3.5

We compared the detailed topology of BN and MAN, as shown in Table 2. Compared to BN, MAN has 10 times the number of OCHes, while the average length and span number are less than 1/10. Moreover, the annualized fiber failure rate of these two networks is similar, which is 0.068% and 0.077%, respectively. However, due to shorter link lengths and the use of OMSP, the normalized network unavailability of MAN is only 3.5% of that in BN. The distribution of AS caused by various causes in these two networks is shown in Fig. 2. As previous work analyzed, due to the absence of optical layer protection in the BN [2], fiber incidents are the primary factor affecting optical network unavailability, accounting for more than 95.1%, while the proportion in the MAN is only 35.8%. That is, in the BN, the normalized network unavailability caused by fiber incidents is 95.1%, while 1.2% in the MAN.

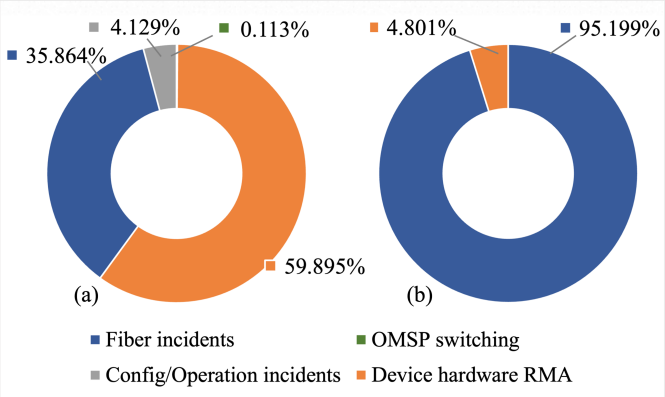

Fig. 2. Comparison of the distribution of abnormal seconds caused by various causes in two networks. (a) MAN, (b) BN.

The analysis focuses on network unavailability in MAN. Since the AS lengths of different causes are large, we counted the frequency of network unavailability due to various causes in Fig. 3. Further, we also studied the abnormal time length in continuous statistical periods about various causes and plotted the probability density functions (PDFs) separately, as shown in Fig. 4. The analysis results indicate that there are four types of causes for the network unavailability:

1) OMSP switching. OMSP switching can be completed within 50 milliseconds, but the average AS caused by OMSP switching is 1.3 seconds due to second-level acquisition accuracy limit of AS. The proportion of AS caused by OMSP switching is only 0.11% but the frequency can reach 88%, which greatly reduces long periods of network unavailability caused by fiber incidents.

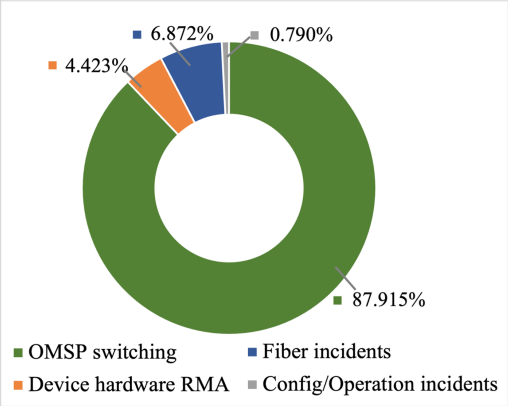

Fig. 3. Frequency distribution of various causes.

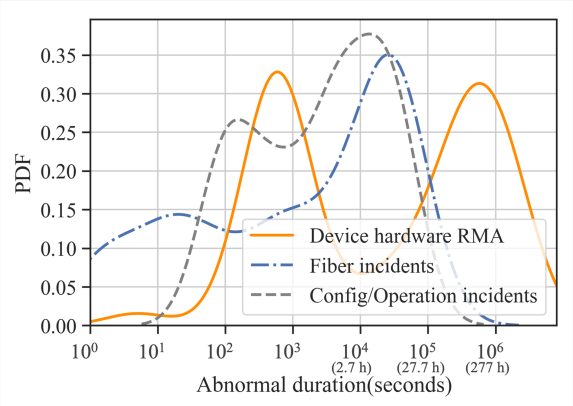

Fig. 4. Abnormal duration distribution of various causes.

2) Fiber incidents. OMSP can only cover the majority of OMS single-path incidents. As can be seen from Fig. 3 and Fig. 2(a), there are still 6.8% of fiber incidents frequency but lead to 35.8% of AS. According to the duration distribution in Fig. 4, the shorter part is mainly caused by two scenarios. One is due to the degradation of both OMS paths at the same time. The other is only single-path degradation, but the APS is generally set to the relative threshold switching, that is, when the difference between the received optical power of the two paths is greater than the threshold, an automatic switch is triggered to ensure that services always work on the better path. In the above case, although degradation occurs, the power difference between these two paths does not reach the switching threshold. The longer part that is greater than one hour to several hours mainly corresponds to the break of both OMS paths, and some are caused by fiber routes overlap of both the OMS paths. These cases are difficult to be solved by OMSP.

3) Device hardware return merchandise authorization (RMA). During the statistical period, there was no hardware RMA for APS and wavelength selective switch (WSS), and OA failures can be protected by OMSP, so no network unavailability was caused by optical layer equipment. The hardware RMA is mainly caused by electrical layer devices, i.e., transponders. As can be seen from Fig. 3 and Fig. 2(a), the frequency of hardware RMA is 4.4%, but causes 59.8% of network AS, indicating that the average duration of RMA is longer. Since electrical card generally only affects one or a few OCH channels and the impact on services can be mitigated by IP protection, the discovery and RMA time is around 4 days on average. The AS duration of RMA is shown in Fig. 4, it is mainly concentrated in two ranges, which reflects two characteristics of AS, occasional performance fluctuation caused by hardware degradation or continuous severe errors caused by hardware failure.

4) Config/Operation incidents. These mainly refer to the optical switch being incorrectly forced to a certain path, meanwhile a fiber incident occurs on this path. As can be seen in Fig. 3 and Fig. 4, the frequency of this cause is extremely low, but the duration is mostly concentrated around 2 to 3 hours, which mainly corresponds to some cases of fiber breaks. In addition, there are a few cases of misconnections after fiber maintenance.

The above analysis shows that the network unavailability of MAN has been greatly reduced, but there is still space for improvement. 1. Detect and repair potential risks in fiber routes such as overlapping route in advance to prevent simultaneous break of the working and protection route. 2. Establish standardized fiber maintenance procedures and implement automated acceptance to prevent misconnections and misconfigurations. 3. Adopt sub-network connection protection (SNCP) to avoid the unavailability caused by electrical layer devices.

4. Summary

In conclusion, the AONA data analytics platform enables us to collect various data in production data center optical networks and study various factors that lead to network unavailability. The result showed that in the case of a similar fiber failure rate, optical layer protection can effectively reduce the unavailability caused by fiber incidents, thereby greatly reducing the network unavailability.

5. References

[1] C. Xie, L. Wang, L. Dou, M. Xia, S. Chen, H. Zhang, Z. Sun, and J. Cheng, "Open and disaggregated optical transport networks for data center interconnects [Invited]," JOCN, vol. 12, no. 6, pp. C12-C22(2020).
[2] L. Wang, L. Wang, C. Wang, and C. Xie, "Unavailability Analyses of Hyperscale Data Center Interconnect Optical Networks," ECOC'2023, paper We.B.1.2, 2023.
[3] "Optical transport network: Linear protection," ITU-T Recommendation G.873.1.
[4] "Error performance parameters and objectives for multi-operator international paths within optical transport networks," ITU-T Recommendation G.8201.