Workgroup: CCAMP Working Group
Internet-Draft:
draft-liu-ccamp-optical2cloud-problem-
statement-03
Published: 21 October 2022
Intended Status: Standards Track
Expires: 24 April 2023
Authors: S. Liu, China Mobile
H. Zheng, Huawei Technologies
A. Guo, Futurewei Technologies
Y. Zhao, China Mobile
D. King, Old Dog Consulting

Problem Statement and Gap Analysis for Connecting to Cloud DCs via Optical Networks

Abstract

Many applications, including optical leased line, cloud VR and computing cloud, benefit from the network scenario where the data traffic to cloud data centers (DCs) is carried end-to-end over an optical network. This document describes the problem statement and requirements for connecting to cloud DCs over optical networks, and presents a gap analysis for existing control plane protocols for supporting this network scenario.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Cloud applications are becoming more popular and widely deployed in enterprises and vertical industries. Organizations with multiple campuses are interconnected together with the remote cloud for storage and computing. Such cloud services demand that the underlying network provides high quality of experience, such as high availability, low latency, on-demand bandwidth adjustments, and so on.

Cloud services have been carried over IP/Ethernet-based aggregated networks for years. MPLS-based VPNs with traffic engineering (TE) are usually used to achieve desired service quality. Provisioning and management of MPLS VPNs is known to be complicated and typically involves manual TE configuration across the network.

To improve the performance and flexibility of aggregated networks, Optical Transport Network (OTN) technology is introduced to complement the IP/Ethernet-based aggregation networks to enable full-fiber connections. This scenario is described in the Fifth

Generation Fixed Network Architecture by the ETSI F5G ISG [ETSI.GR.F5G.001]. OTN can be used to provide high quality carrier services in addition to the traditional MPLS VPN services. OTN provides Time Division Multiplexing (TDM) based connections with no queueing or scheduling needed, with an access bandwidth granularity of 1.25Gbps, i.e., ODU0 (Optical Data Unit 0) and above. This bandwidth granularity is typically more than what a single application would demand, therefore, user traffic usually needs to be aggregated before being carried forward through the network. However, advanced OTN technologies developed in ITU-T work items have aimed to enhance OTN to support services of much finer granularity. These enhancements, when implemented, will make OTN an even more suitable solution for bearing cloud traffic with high quality and bandwidth granularity close to what an IP/Ethernet-based network could offer.

Many cloud-based services that require high bandwdith, deterministic service quality, and flexible access could potentially benefit from the network scenario of using OTN-based aggregation networks to interconnect cloud data centers (DCs). For example, intra-city Data Center Interconnects (DCIs), which communicate with each other to supports public and/or private cloud services, can use OTN for via intra-city DCI networks to ensure ultra-low latency and on-demand provisioning of large bandwidth connections for their Virtual Machine (VM) migration services. Another example is the high quality private line, which can be provided over OTN dedicated connections with high security and reliability for large enterprises such as financial, medical centers, and education customers. Yet another example is the Cloud Virtual Reality (VR) services, which typcially require high bandwidth (e.g., over 1Gbps for 4K or 8k VR) links with low latency (e.g., 10ms or less) and low jitter (e.g., 5ms or less) for rendering with satisfactory user experience. These network properties required for cloud VR services can typically be offered by OTNs with higher quality comparing to IP/Ethernet based networks.

[I-D.ietf-rtgwg-net2cloud-problem-statement] and [I-D.ietf-rtgwg-net2cloud-gap-analysis] present a detailed analysis of the coordination requirements between IP-based networks and cloud DCs. This document complements that analysis by further examining the requirements and gaps from the control plane perspective when accessing cloud DCs through OTNs. Data plane requirements are out of the scope of this document.

2. Requirements and Gap Analysis

2.1. Multi-cloud Access

Cloud services are deployed in geographically distributed locations for scalability and resilliency, and they are usually hosted by multiple interconnected DCs. DCs have usually been interconnected through Layer 2/3 switches or routers with full mesh connectivity. To improve interaction efficiency as well as service experience, OTN is also considered as an option for DC interconnection. This network scenario is illustrated in Figure 1.

```
   __________                                             ________
  /          \                                           /        \
 | Enterprise |          ___________                    | Vertical |
 |    CPE     |\        /           \          +-----+ /|  Cloud   |
  \__________/  \ +---+/             \+---+    |Cloud|/  \________/
                 \|O-A*               *O-E|----+ GW  |
                  +---+               +---+    +-----+
    ________          |       OTN     |                    _______
   /        \     +---+               +---+    +-----+    /       \
  | Vertical |----+O-A*               *O-E|----+Cloud|---| Private |
  |   CPE    |    +---+\             /+---+    | GW  |   | Cloud   |
   \________/           \___________/          +-----+    \_______/
```

Figure 1: Multi-cloud access through an OTN

A customer application is connected to the cloud via one of the Customer Premises Equipment (CPE), and access cloud services are hosted in multiple clouds that are attached to different cloud gateways. Layer 2 or Layer 3 Virtual Private Networks (L2VPN or L3VPN) are used as overlay services on top of the OTN to support multi-cloud access. Serving as an overlay, the OTNs should provide the capability to create different types of connections, including point-to-point (P2P), point-to-multipoint (P2MP) and multipoint-to-multipoint (MP2MP) connections to support diverse L2VPN or L3VPN services.

In the data plane, OTN connections are P2P by nature. To support P2MP and MP2MP services, multiple P2P OTN connections can be established between each source and destination pair. The routing and signaling protocols for OTN need to coordinate these OTN connections to ensure they are routed with proper diverse paths to meet resilliency and path quality constraints.

[RFC4461] defines the requirements for establishing P2MP MPLS traffic engineering label switched paths (LSPs). [RFC6388] describes extensions to the Label Distribution Protocol (LDP) for the setup of P2MP and MP2MP LSPs in MPLS Networks. The generic rules introduced by those documents work also apply to OTNs, however, the protocol extensions are missing and are required for establishing P2MP and MP2MP connetctions with OTN resources, i.e., time slots.

2.2. Service Awareness

Cloud-oriented services are dynamic in nature with frequent changes in bandwidth and quality of service (QoS) requirements. However, in typical OTN scenarios, OTN connections are preconfigured between provider edge (PE) nodes, and client traffic like IP or Ethernet is fixed-mapped onto the payload of OTN frames at the ingress PE node. This makes the OTN connections rather static and they cannot accommodate the dynamicity of the traffic unless they are permanently over-provisioned, resulting in slow and inefficient use of the OTN bandwidth resources. To address this issue and to make the OTN more suitable for carrying cloud-oriented services, it needs to be able to understand the type of traffic and its QoS requirements, so that OTN connections can be dynamically built and selected with the best feasible paths. The mapping of client services to OTN connections should also be dynamically configured or modified to better adapt to the traffic changes.

New service-aware capabilties are needed for both the control plane and data plane to address this challenge for OTNs. In the data plane, new hardware that can examine cloud traffic packet header fields (such as the IP header source and destination IP address and/or the type of service (TOS) field, virtual routing and forwarding (VRF) identifiers, layer 2 Media Access Control (MAC) address or virtual local area network (VLAN) identifiers) are introduced to make the PE node able to sense the type of traffic. This work for the data plane is out of the scope of this document.

Being service aware allows the OTN network to accurately identify the characteristics of carried client service flows and the real-time traffic of each flow, making it possible to achieve automated and real-time operations such as dynamic connection establishment and dynamic bandwidth adjustment according to preset policies. Those capabilities help to optimize the resource utilization and significantly reduce the operational cost of the network.

Upon examining the client traffic header fields and obtaining client information such as the cloud destination and QoS requirements, the OTN PE node needs to forward such information to the control entity of the OTN to make decisions on connection configurations, and map the client packets of different destination/QoS to different ODU connections The client information could include, but is not limited to, the destination IP addresses, type of cloud service, and QoS information such as bandwidth, latency bounds, and resiliency factors. The control entity may be an SDN controller or a control plane instace: in the former case communications are established between each of the PE nodes and the controller, and the controller serves as a central authority for OTN connection configurations; whereas in the latter case, all of the PE nodes need to disseimate client information to each other using control plane protocols or possibly through some intermediate reflectors. It is desirable that the protocols used for both cases are consistent, and ideally, the same. A candidate protocol is the PCE communication Protocol (PCEP) [RFC5440], but there are currently no extensions defined for describing such client traffic information. Extensions to PCEP could be defined outside this document to support the use case. It is also possible to use the BGP Link State (BGP-LS) protocol [RFC7752] to perform the distribution of client information. However, an OTN PE node does not typically run BGP protocols due to that BGP lacks protocol extensions to support optical networks. Therefore, PCEP seems to be a better protocol choice in this case.

3. Framework

3.1. Service Identification and Mapping

The OTN PE node should support the learning and identification of the packet header carried by client services. The identification content may include but not limited to the following content:

*Source and destination MAC addresses

*Source and destination IP addresses

*VRF identifier

*VLAN (S-VLAN and/or C-VLAN) identifier

*MPLS label

The OTN PE node should support reporting the above identified client services to the management and control system, which can obtain the client-side addresses reported by each node in the entire network to build up a global topology. Some of the learnt content, such as the VLAN identifier, are not required to be reported since VLAN is of only local significance.

The management and control system should be able to calculate the corresponding ODU connection route based on the source and destination addresses of the service, and create the mapping between service address and the ODU cnnection according to preset policies. The mapping table can be generated through management plane configuration or control plane protocol.

3.2. Reporting Service Identification

The control plane protocol extension should report to the controller service identification contents, which should include at least the following content:

* A private network or network slice identifier, which is a globally unique identifier to identify different tenants or applications supported by the private network

* OTN node identifier, which identify the OTN PE node that reported this packet

* The IP/MAC address of the client side learned by the OTN PE node

When the PCEP protocol is used, this extension may be defined as a PCEP Report message.

3.3. Configuring Service Mapping

The control plane protocol extension may be defined to push the mapping table between service address to ODU connections from the controller to the OTN PE nodes. The message should include at least the following content:

* A private network or network slice identifier, which is a globally unique identifier to identify different tenants or applications supported by the private network

* A mapping table of {service address, ODU connection identifier}, with each entry of the table contains at least the information of {remote OTN node, remote service address}, where the concept of "remote" is based on the perspective of the OTN device that receives this packet

When the PCEP protocol is used, this extension may be defined as a PCEP Update message.

4. Manageability Considerations

TBD

5. Security Considerations

This document analyzes the requirements and gaps in connecting to cloud DCs over optical networks without defining new protocols or interfaces. Therefore, this document introduces no new security considerations to the control or management plane of OTN. Risks presented by existing OTN control plane are described in [RFC4203] and [RFC4328], and risks presented by existing northbound and southbound control interfaces in general are described in [RFC8453]. Moreover, the data communication network (DCN) for OTN control plane protocols are encapsulated in fibers, which providers a much better security environment for running the protocols.

6. IANA Considerations

This document requires no IANA actions.

7. References

7.1. Normative References

[ETSI.GR.F5G.001] European Telecommunications Standards Institute (ETSI), "Fifth Generation Fixed Network (F5G);F5G Generation Definition Release 1", ETSI GR F5G 001 , December 2020, <https://www.etsi.org/deliver/etsi_gr/F5G/001_099/001/01.01.01_60/gr_F5G001v010101p.pdf>.

[RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <https://www.rfc-editor.org/info/rfc4203>.

[RFC4328] Papadimitriou, D., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, DOI 10.17487/RFC4328, January 2006, <https://www.rfc-editor.org/info/rfc4328>.

[RFC4461] Yasukawa, S., Ed., "Signaling Requirements for Point-to-Multipoint Traffic-Engineered MPLS Label Switched Paths (LSPs)", RFC 4461, DOI 10.17487/RFC4461, April 2006, <https://www.rfc-editor.org/info/rfc4461>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <https://www.rfc-editor.org/info/rfc5440>.

[RFC6388] Wijnands, IJ., Ed., Minei, I., Ed., Kompella, K., and B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6388, DOI 10.17487/RFC6388, November 2011, <https://www.rfc-editor.org/info/rfc6388>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC

7752, DOI 10.17487/RFC7752, March 2016, <https://www.rfc-editor.org/info/rfc7752>.

[RFC8453] Ceccarelli, D., Ed. and Y. Lee, Ed., "Framework for Abstraction and Control of TE Networks (ACTN)", RFC 8453, DOI 10.17487/RFC8453, August 2018, <https://www.rfc-editor.org/info/rfc8453>.

7.2. Informative References

[I-D.ietf-rtgwg-net2cloud-gap-analysis] Dunbar, L., Malis, A. G., and C. Jacquenet, "Networks Connecting to Hybrid Cloud DCs: Gap Analysis", Work in Progress, Internet-Draft, draft-ietf-rtgwg-net2cloud-gap-analysis-09, 15 June 2022, <https://www.ietf.org/archive/id/draft-ietf-rtgwg-net2cloud-gap-analysis-09.txt>.

[I-D.ietf-rtgwg-net2cloud-problem-statement]
Dunbar, L., Malis, A. G., Jacquenet, C., Toy, M., and K. Majumdar, "Dynamic Networks to Hybrid Cloud DCs Problem Statement", Work in Progress, Internet-Draft, draft-ietf-rtgwg-net2cloud-problem-statement-15, 25 July 2022, <https://www.ietf.org/archive/id/draft-ietf-rtgwg-net2cloud-problem-statement-15.txt>.

Acknowledgments

TBD

Authors' Addresses

Sheng Liu
China Mobile

Email: liushengwl@chinamobile.com

Haomian Zheng
Huawei Technologies

Email: zhenghaomian@huawei.com

Aihua Guo
Futurewei Technologies

Email: aihuaguo.ietf@gmail.com

Yang Zhao
China Mobile

Email: zhaoyangyjy@chinamobile.com

Daniel King
Old Dog Consulting

Email: daniel@olddog.co.uk