Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 16, 2014

G. Chen
C. Li
China Mobile

July 15, 2013

The Applicability Analysis of Traffic Engineering over Trill
draft-chen-trill-te-applicability-00

Abstract

This memo intended to describe the needs of Traffic Engineering (TE) in a Data Center Network (DCN) in which Transparent Interconnection of Lots of Links (Trill) has been adopted. Several use cases have been proposed to demonstrate the usage while the feasibility of Trill extension has been evaluated. It serves to motivate the normative work subsequently.

Table of Contents

1. Introduction

Transparent Interconnection of Lots of Links (Trill) protocol is targeted to improve link utilizations of switch networks. With the growth of Data Center Network (DCN), this technology is widely used and approved to be an essential component building layer-two networks. The services in DCN could benefit from multi-path forwarding for good flexibility and resilience. Shortest Path First (SPF) is used to forward data frames. The selection of forwarding path is inherited from a routing protocol, which may be optimized based on a simple additive metric. Those algorithms are generally topology-driven, so link utilizations and traffic features can't be factors considered in data path decisions.

Trafﬁc Engineering (TE) provided a useful feature to improve the performance of the telecommunication networks. The ultimate objective is to avoid congestion in the network by keeping its links from being overloaded. Current uses are exclusively focused on the applications of Multiprotocol Label Switching (MPLS). It's normally deployed in a large Internet backbone. The adoption in a DCN environment is worth to be investigated given the scale of large Ethernet networks has been dramatically expanded.

This document describes service characteristics in a DCN. The deployment of traffic engineering would likely optimize the Quality of Experiences. It motivates the work to populate TE functions into Trill protocol. Several use cases help to elaborate the usages of those technologies. The considerations of technical principle can also be deduced from actual uses. This memo also provided the possibilities of protocol design.

2. Motivations

Data centers are being heavily employed in enterprise and consumer settings to run a variety of applications and cloud-based services. These range from Internet-facing "sensitive" applications, such as, Web services, instant messaging, stock updates, financial applications and gaming, to storage intensive applications, such as, data analysis, file sharing and scientific computing. The performance of these applications crucially depends on the functioning of the data center's network infrastructure. For example, a congested data center network, where internal traffic is routinely subjected to losses and poor throughput, could lead search queries to take longer to complete, instant message to get delayed, gaming experience to deteriorate, and POP mail services and Web transactions to hang. The dissatisfied end-users and subscribers could choose alternate providers, resulting in a significant loss in revenues for the data center. TE mechanism likely could help the situation by monitoring the link's loads and traffic growth trends. It's desirable to introduce the technology into DCN even there is constituted with a number of switch device.

Trill has been adopted in DCN in recent years. SPF-based algorithm still can't provide satisfactory performance if congestions are occurred on the links. Our preliminary study reveals that SPF-based algorithm achieve only 80-85% of the performance achieved by an optimal TE algorithm. It indicates that there is significant room for improvement if TE functions could be integrated with Trill. Those benefits could motivate new works on Trill-TE.

3. Use Cases

3.1. Bandwidth Ensuring

The Fig. 1 shown the case, in which different services, including gaming, stroge backup and web browse are relied on a switch network. There are two paths between ingress RB1 and egress RB3, i.e. RB1---RB2---RB3 and RB1---RB4---RB5--RB6. Trill-TE is used to guarantee the performance of gaming with low delay. And other services would bypass the route gaming traversed so as to avoid congestions with gaming. Therefore, RB1---RB2---RB3 are served as to transmit the

gaming data. RB1---RB4---RB5--RB6 has been selected for other data services.

```
         +-----+       +-----+        +-----+
         | RB4 |-------| RB5 |--------| RB6 |
         +-----+       +-----+        +-----+
            |                             |
         +-----+       +-----+        +-----+
         | RB1 |-------| RB2 |--------| RB3 |
         +-----+       +-----+        +-----+
    Ingress |                             | Egress
```

Figure 1: Bandwidth Ensuring with Trill-TE

3.2. Link Protections

The Fig.2 shows the topology. Ingress RB1 would go through the way of RB1-- RB2--RB3--RB4 to RB8. If there is failure detected on the path. Trill-TE would switch the traffic to the backup link, i.e. RB1-- RB5--RB6--RB7--RB8. Once the above link is recovered, the traffic would be switched back to below link.

```
           +-----+       +-----+        +-----+
   +------ | RB2 |-------| RB3 |--------| RB4 |-------+
   |       +-----+       +-----+        +-----+       |
+-----+                                            +-----+
| RB1 |Ingress                                     | RB8 |Egress
+-----+                                            +-----+
   |       +-----+       +-----+        +-----+       |
   +------ | RB5 |-------| RB6 |--------| RB7 |-------+
           +-----+       +-----+        +-----+
```

Figure 2: Link Protections with Trill-TE

3.3. Traffic Matrix

Various traffic flows can be well gauged when the Trill is capable of Traffic Engineering. Operators could get accurate traffic statistics because the traffic is deterministically transmitted over the planned data paths. The deployment of traffic engineering in a DCN helps to form a traffic matrix, which provides the volume of traffic between every pair of ingress and egress points over a given time interval. Such information is critical inputs to network design, capacity planning and business planning.

3.4. Unequal Load Balancing

Unequal load balancing could help to utilize bandwidth resource properly. Each link is vary in the exchange capability, for example 10GbE, 40GbE or 100GbE have been deployed on the route paths. The traffic engineering in Trill could distribute the flow on the matched link. It could optimize the network resource uses as a whole.

4. Applicabilty Analysis

The following is the features Trill-TE could leverage. Those can verify the feasibility of this work.

- o Multipath: Trill-TE must take advantage of the path diversity available in the data center's network. Based on that, the forwarding path can be decided via an global view of traffic.

- o Metric transmission: IS-IS Extensions for Traffic Engineering is already defined in [RFC5305]. Trill-TE may need to insert Nicknames to the forwarding table for the TE routing path. The existing Trill process would immune to those changes.

5. IANA Considerations

This document makes no request of IANA.

6. Security Considerations

TBD

7. Normative References

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

Authors' Addresses

Gang Chen
China Mobile
53A,Xibianmennei Ave.,
Xuanwu District,
Beijing 100053
China

Email: phdgang@gmail.com

Chen Li
China Mobile
53A,Xibianmennei Ave.,
Xuanwu District,
Beijing 100053
China

Email: lichen@chinamobile.com