TEAS Working Group
Internet-Draft
Intended status: Experimental
Expires: December 28, 2018

A. Wang
China Telecom
X. Huang
C. Kou
BUPT
Z. Li
China Mobile
L. Huang
P. Mi
Huawei Technologies

June 26, 2018

CCDR Scenario, Simulation and Suggestion
draft-ietf-teas-native-ip-scenarios-01

Abstract

This document describes the scenarios, simulation and suggestions for the "Centrally Control Dynamic Routing (CCDR)" architecture, which integrates the merit of traditional distributed protocols (IGP/BGP), and the power of centrally control technologies (PCE/SDN) to provide one feasible traffic engineering solution in various complex scenarios for the service provider.

Traditional MPLS-TE solution is mainly used in static network planning scenario and is difficult to meet the QoS assurance requirements in real-time traffic network. With the emerge of SDN concept and related technologies, it is possible to simplify the complexity of distributed control protocol, utilize the global view of network condition, give more efficient solution for traffic engineering in various complex scenarios.

Status of This Memo

This Internet-Draft will expire on December 28, 2018.

Copyright Notice

Table of Contents

1. Introduction

Internet network is composed mainly tens of thousands of routers that run distributed protocol to exchange the reachability information between them. The path for the destination network is mainly calculated and controlled by the traditional IGP protocols. These distributed protocols are robust enough to support the current evolution of Internet but has some difficulties when the application requires the end-to-end QoS performance, or the service provider wants to maximize the links utilization within their network.

MPLS-TE technology is one perfect solution for the finely planned network but it will put heavy burden on the router when we use it to solve the dynamic QoS assurance requirements within real time traffic network.

SR(Segment Routing) is another prominent solution that integrates some merits of traditional distributed protocol and the advantages of centrally control mode, but it requires the underlying network, especially the provider edge router to do label push and pop action in-depth, and need some complex solutions for co-exist with the Non-SR network. Finally, it can only maneuver the end-to-end path for MPLS and IPv6 traffic via different mechanisms.

The advantage of MPLS is mainly for traffic isolation, such as the L2/L3 VPN service deployments, but most of the current application requirements are only for high performances end-to-end QoS assurance. Without the help of centrally control architecture, the service provider almost can’t make such SLA guarantees upon the real time traffic situation.

This draft gives some scenarios that the centrally control dynamic routing (CCDR) architecture can easily solve, without adding more extra burdening on the router. It also gives the PCE algorithm results under the similar topology, traffic pattern and network size to illustrate the applicability of CCDR architecture. Finally, it gives some suggestions for the implementation and deployment of CCDR.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

3. CCDR Scenarios.

The following sections describe some scenarios that the CCDR architecture is suitable for deployment.

3.1. Qos Assurance for Hybrid Cloud-based Application.

With the emerge of cloud computing technologies, enterprises are putting more and more services on the public oriented service infrastructure, but keep still some core services within their network. The bandwidth requirements between the private cloud and the public cloud are occasionally and the background traffic between these two sites varied from time to time. Enterprise cloud applications just want to invoke the network capabilities to make the end-to-end QoS assurance on demand. Otherwise, the traffic should be controlled by the distributed protocol.

CCDR, which integrates the merits of distributed protocol and the power of centrally control, is suitable for this scenario. The possible solution architecture is illustrated below:

```
                            +------------------------+
                            | Cloud Based Application|
                            +------------------------+
                                        |
                                  +-----------+
                                  |    PCE    |
                                  +-----------+
                                        |
                                        |
                               //--------------\\
                          /////                  \\\\\
     Private Cloud Site ||       Distributed          |Public Cloud Site
                         |       Control Network      |
                          \\\\\                  /////
                               \\--------------//
```

Fig.1 Hybrid Cloud Communication Scenario

By default, the traffic path between the private cloud site and public cloud site will be determined by the distributed control network. When some applications require the end-to-end QoS assurance, it can send these requirements to PCE, let PCE compute one e2e path which is based on the underlying network topology and the real traffic information, to accommodate the application’s QoS requirements. The proposed solution can refer the draft [I-D.ietf-teas-pce-native-ip]. Section 4 describes the detail simulation process and the results.

3.2. Increase link utilization based on tidal phenomena.

Currently, the network topology within MAN is generally in star mode as illustrated in Fig.2, with the different devices connect different customer types. The traffic pattern of these customers demonstrates some tidal phenomena that the links between the CR/BRAS and CR/SR will experience congestion in different periods because the subscribers under BRAS often use the network at night and the dedicated line users under SR often use the network during the daytime. The uplink between BRAS/SR and CR must satisfy the maximum traffic pattern between them and this causes the links underutilization.

```
                              +--------+
                              |   CR   |
                              +----|---+
                                   |
                       --------|--------|-------|
                       |       |        |       |
                    +--|-+   +-|-    +--|-+   +-|+
                    |BRAS|   |SR|    |BRAS|   |SR|
                    +----+   +--+    +----+   +--+
```

Fig.2 STAR-style network topology within MAN

If we can consider link the BRAS/SR with local loop, and control the MAN with the CCDR architecture, we can exploit the tidal phenomena between BRAS/CR and SR/CR links, increase the efficiency of them.

```
                                       +-------+
                                   -----  PCE  |
                                   |   +-------+
                              +----|---+
                              |   CR   |
                              +----|---+
                                   |
                       --------|--------|-------|
                       |       |        |       |
                    +--|-+   +-|-    +--|-+   +-|+
                    |BRAS-----SR|    |BRAS-----SR|
                    +----+   +--+    +----+   +--+
```

Fig.3 Increase the link utilization via CCDR

3.3. Traffic engineering for IDC/MAN asymmetric link

The operator’s networks are often comprised by tens of different domains, interconnected with each other, form very complex topology that illustrated in Fig.4. Due to the traffic pattern to/from MAN and IDC, the links between them are often in asymmetric style. It is almost impossible to balance the utilization of these links via the distributed protocol, but this unbalance phenomenon can be overcome via the CCDR architecture.

```
                    +---+                +---+
                    |MAN|-----------------IDC|
                    +-|-|       |       +-|-+
                      |     ---------|     |
                      ------|BackBone|------
                      |     ----|----|     |
                      |         |          |
                    +-|--       |        ----+
                    |IDC|----------------|MAN|
                    +---|                |---+
```

Fig.4 TE within Complex Multi-Domain topology

3.4. Network temporal congestion elimination.

In more general situation, there are often temporal congestion periods within part of the service provider’s network. Such congestion phenomena will appear repeatedly and if the service provider has some methods to mitigate it, it will certainly increase the satisfaction degree of their customer. CCDR is also suitable for such scenario that the traditional distributed protocol will process most of the traffic forwarding and the controller will schedule some traffic out of the congestion links to lower the utilization of them. Section 4 describes the simulation process and results about such scenario.

4. CCDR Simulation.

The following sections describe the topology, traffic matrix, end-to-end path optimization and congestion elimination in CCDR simulation.

4.1. Topology Simulation

The network topology mainly contains nodes and links information. Nodes used in simulation have two types: core nodes and edge nodes. The core nodes are fully linked to each other. The edge nodes are connected only with some of the core nodes. Fig.5 is a topology example of 4 core nodes and 5 edge nodes. In CCDR simulation, 100 core nodes and 400 edge nodes are generated.

```
                +----+
               /|Edge|\
              | +----+ |
              |        |
              |        |
+----+    +----+     +----+
|Edge|----|Core|-----|Core|---------+
+----+    +----+     +----+         |
        /  |    \   /   |           |
  +----+   |     \ /    |           |
  |Edge|   |      X     |           |
  +----+   |     / \    |           |
        \  |    /   \   |           |
+----+    +----+     +----+         |
|Edge|----|Core|-----|Core|         |
+----+    +----+     +----+         |
            |          |            |
            |          +------\   +----+
            |                  ---|Edge|
            +-----------------/   +----+
```

Fig.5 Topology of simulation

The number of links connecting one edge node to the set of core nodes is randomly between 2 to 30, and the total number of links is more than 20000. Each link has its congestion threshold.

4.2. Traffic Matrix Simulation.

The traffic matrix is generated based on the link capacity of topology. It can result in many kinds of situations, such as congestion, mild congestion and non-congestion.

In CCDR simulation, the traffic matrix is 500*500. About 20% links are overloaded when the Open Shortest Path First (OSPF) protocol is used in the network.

4.3. CCDR End-to-End Path Optimization

The CCDR end-to-end path optimization is to find the best end-to-end path which is the lowest in metric value and each link of the path is far below link’s threshold. Based on the current state of the network, PCE within CCDR architecture combines the shortest path algorithm with penalty theory of classical optimization and graph theory.

Given background traffic matrix which is unscheduled, when a set of new flows comes into the network, the end-to-end path optimization finds the optimal paths for them. The selected paths bring the least congestion degree to the network.

The link utilization increment degree(UID) when the new flows are added into the network is shown in Fig.6. The first graph in Fig.6 is the UID with OSPF and the second graph is the UID with CCDR end-to-end path optimization. The average UID of graph one is more than 30%. After path optimization, the average UID is less than 5%. The results show that the CCDR end-to-end path optimization has an eye-catching decreasing in UID relative to the path chosen based on OSPF.

```
           +-----------------------------------------------------------+
           |                *                               *    *    *|
         60|                *                             * * *  *    *|
           |*      *       **     * *         *   *   *  ** * *  * * **|
           |*   * ** *   * **   *** **  *   * **  * * *  ** * *  *** **|
           |* * * ** *  ** **   *** *** **  **** ** ***  **** ** *** **|
         40|* * * ***** ** ***  *** *** **  **** ** *** ***** ****** **|
     UID(%)|* * ******* ** ***  *** ******* **** ** *** ***** *********|
           |*** ******* ** **** *********** *********** ***************|
           |******************* *********** *********** ***************|
         20|******************* ***************************************|
           |******************* ***************************************|
           |***********************************************************|
           |***********************************************************|
          0+-----------------------------------------------------------+
          0    100   200   300   400   500   600   700   800   900  1000
           +-----------------------------------------------------------+
           |                                                           |
         60|                                                           |
           |                                                           |
           |                                                           |
           |                                                           |
         40|                                                           |
     UID(%)|                                                           |
           |                                                           |
           |                                                           |
         20|                                                           |
           |                                                          *|
           |                                     *                    *|
           |        *         *  *    *       *  **                 * *|
          0+-----------------------------------------------------------+
          0    100   200   300   400   500   600   700   800   900  1000
                                Flow Number
```

Fig.6 Simulation result with congestion elimination

4.4. Network temporal congestion elimination

Different degree of network congestion is simulated. The congestion degree (CD) is defined as the link utilization beyond its threshold.

The CCDR congestion elimination performance is shown in Fig.7. The first graph is the congestion degree before the process of congestion elimination. The average CD of all congested links is more than 10%. The second graph shown in Fig.7 is the congestion degree after congestion elimination process. It shows only 12 links among totally 20000 links exceed the threshold, and all the congestion degree is less than 3%. Thus, after schedule of the traffic in congestion paths, the degree of network congestion is greatly eliminated and the network utilization is in balance.

```
                          Before congestion elimination
           +-----------------------------------------------------------+
           |                *                            ** *   ** ** *|
         20|                *                     *      **** * ** ** *|
           |*      *       **     * **       **  **** * ***** *********|
           |*   *  * *   * **** ****** *  ** *** **********************|
         15|* * * ** *  ** **** ********* *****************************|
           |* * ******  ******* ********* *****************************|
     CD(%) |* ********* ******* ***************************************|
         10|* ********* ***********************************************|
           |*********** ***********************************************|
           |***********************************************************|
          5|***********************************************************|
           |***********************************************************|
           |***********************************************************|
          0+-----------------------------------------------------------+
              0            0.5            1            1.5            2

                        After congestion elimination
          +-----------------------------------------------------------+
          |                                                           |
        20|                                                           |
          |                                                           |
          |                                                           |
        15|                                                           |
          |                                                           |
    CD(%) |                                                           |
        10|                                                           |
          |                                                           |
          |                                                           |
        5 |                                                           |
          |                                                           |
          |        *        **  * *  *  **   *  **                 *  |
        0 +-----------------------------------------------------------+
           0            0.5            1            1.5            2
                            Link Number(*10000)
```

Fig.7 Simulation result with congestion elimination

5. CCDR Deployment Consideration.

With the above CCDR scenarios and simulation results, we can know it is necessary and feasible to find one general solution to cope with various complex situations for the most complex optimal path computation in centrally manner based on the underlay network topology and the real time traffic.

[I-D.ietf-teas-pce-native-ip] gives the principle solution for above scenarios, such thoughts can be extended to cover requirements that are more concretes in future.

6. Security Considerations

This document considers mainly the integration of traditional distributed protocol and the global view of central control. It certainly can ease the management of network in various traffic-engineering scenarios described in this document, but the central control manner may also bring the new point be easily attacked. Solutions for CCDR scenarios should keep these in mind and consider more for the protection of SDN controller and their communication with the underlay devices, which described in document 1 and [RFC8253]

7. IANA Considerations

This document does not require any IANA actions.

8. Normative References

[I-D.ietf-teas-pce-native-ip]
Wang, A., Zhao, Q., Khasanov, B., and K. Mi, "PCE in Native IP Network", draft-ietf-teas-pce-native-ip-00 (work in progress), February 2018.

[I-D.ietf-teas-pcecc-use-cases]
Zhao, Q., Li, Z., Khasanov, B., Ke, Z., Fang, L., Zhou, C., Communications, T., and A. Rachitskiy, "The Use Cases for Using PCE as the Central Controller(PCECC) of LSPs", draft-ietf-teas-pcecc-use-cases-01 (work in progress), May 2017.

[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>.

[RFC8253] Lopez, D., Gonzalez de Dios, O., Wu, Q., and D. Dhody, "PCEPS: Usage of TLS to Provide a Secure Transport for the Path Computation Element Communication Protocol (PCEP)", RFC 8253, DOI 10.17487/RFC8253, October 2017, <https://www.rfc-editor.org/info/rfc8253>.

[RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <https://www.rfc-editor.org/info/rfc8283>.

Authors’ Addresses

Aijun Wang
China Telecom
Beiqijia Town, Changping District
Beijing, Beijing 102209
China

Email: wangaj.bri@chinatelecom.cn

Xiaohong Huang
Beijing University of Posts and Telecommunications
No.10 Xitucheng Road, Haidian District
Beijing
China

Email: huangxh@bupt.edu.cn

Caixia Kou
Beijing University of Posts and Telecommunications
No.10 Xitucheng Road, Haidian District
Beijing
China

Email: koucx@lsec.cc.ac.cn

Zhenqiang Li
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing 100053
China

Email: li_zhenqiang@hotmail.com

Lu Huang
Huawei Technologies
Unit 7 NO 8.XiBinHe Road,YongDingMen
Beijing, Dongcheng District 100077
China

Email: hlisname@yahoo.com

Penghui Mi
Huawei Technologies
Tower C of Bldg.2, Cloud Park, No.2013 of Xuegang Road
Shenzhen, Bantian,Longgang District 518129
China

Email: mipenghui@huawei.com