TEAS Working Group
Internet-Draft
Intended status: Informational
Expires: April 1, 2020

A. Wang
China Telecom
X. Huang
C. Kou
BUPT
Z. Li
China Mobile
P. Mi
Huawei Technologies
September 29, 2019

Scenarios and Simulation Results of PCE in Native IP Network
draft-ietf-teas-native-ip-scenarios-09

Abstract

Requirements for providing the End to End(E2E) performance assurance are emerging within the service provider network. While there are various technology solutions, there is no one solution which can fulfill these requirements for a native IP network. One universal (E2E) solution which can cover both intra-domain and inter-domain scenarios is needed.

One feasible E2E traffic engineering solution is the use of a Path Computation Elements (PCE) in a native IP network. This document describes various complex scenarios and simulation results when applying a PCE in a native IP network. This solution, referred to as Centralized Control Dynamic Routing (CCDR), integrates the advantage of using distributed protocols and the power of a centralized control technology.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 1, 2020.

Table of Contents

1. Introduction

A service provider network is composed of thousands of routers that run distributed protocols to exchange the reachability information. The path for the destination network is mainly calculated, and controlled, by the distributed protocols. These distributed protocols are robust enough to support most applications, but have some difficulties supporting the complexities needed for traffic

engineering applications, e.g. E2E performance assurance, or maximizing the link utilization within an IP network.

Multiprotocol Label Switching (MPLS) using Traffic Engineering (TE) technology (MPLS-TE)[RFC3209]is one solution for traffic engineering network but it introduces an MPLS network and related technology which would be an overlay of the IP network. MPLS-TE technology is often used for Label Switched Path (LSP) protection and complex path set-up within a domain.

It has not been widely deployed for meeting E2E (especially in inter-domain) dynamic performance assurance requirements for an IP network.

Segment Routing [RFC8402] is another solution that integrates some advantages of using a distributed protocol and a centrally control technology, but it requires the underlying network, especially the provider edge router, to do a label push and pop action in-depth, and adds complexity, when coexisting with the Non-Segment Routing network. Additionally, it can only maneuver the E2E paths for MPLS and IPv6 traffic via different mechanisms.

Deterministic Networking (DetNet)[RFC8578] is another possible solution. It is primarily focused on providing bounded latency for a flow and introduces additional requirements on the domain edge router. The current DetNet scope is within one domain. The use cases defined in this document do not require the additional complexity of deterministic properties and so differ from the DetNet use cases.

This draft describes scenarios for a native IP network that a Centralized Control Dynamic Routing (CCDR) framework can easily solve, without requiring a change of the data plane behaviour on the router. It also provides path optimization simulation results to illustrate the applicability of the CCDR framework.

2. Terminology

This document uses the following terms defined in [RFC5440]: PCE.

The following terms are defined in this document:

- BRAS: Broadband Remote Access Server
- CD: Congestion Degree
- CR: Core Router
- CCDR: Centralized Control Dynamic Routing

- E2E: End to End
- IDC: Internet Data Center
- MAN: Metro Area Network
- QoS: Quality of Service
- SR: Service Router
- UID: Utilization Increment Degree
- WAN: Wide Area Network

3. CCDR Scenarios

The following sections describe various deployment scenarios for applying the CCDR framework.

3.1. QoS Assurance for Hybrid Cloud-based Application

With the emergence of cloud computing technologies, enterprises are putting more and more services on a public oriented cloud environment, but keeping core business within their private cloud. The communication between the private and public cloud sites will span the Wide Area Network (WAN) network. The bandwidth requirements between them are variable and the background traffic between these two sites varies over time. Enterprise applications require assurance of the E2E Quality of Service(QoS) performance on demand for variable bandwidth services.

CCDR, which integrates the merits of distributed protocols and the power of centralized control, is suitable for this scenario. The possible solution framework is illustrated below:

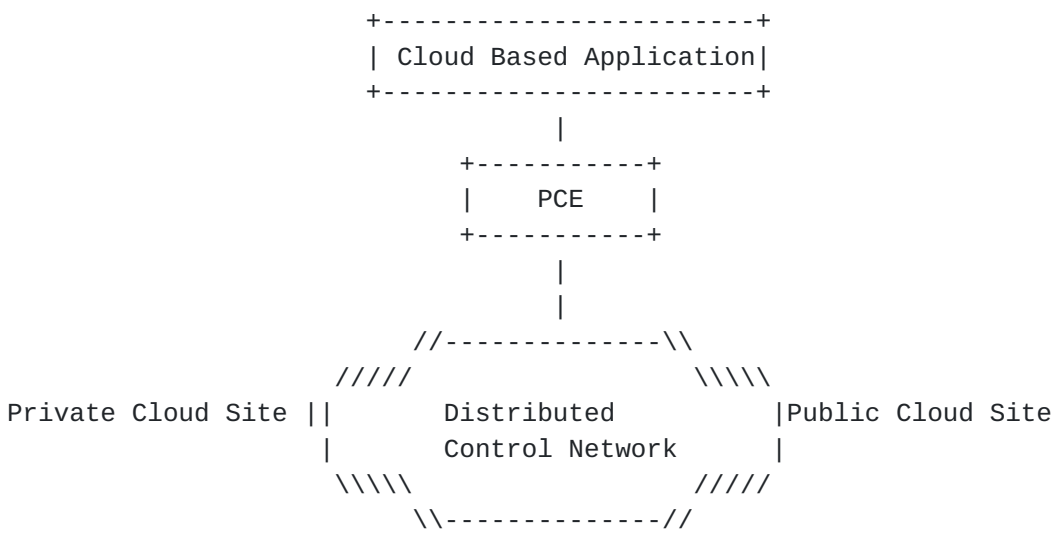


Figure 1: Hybrid Cloud Communication Scenario

By default, the traffic path between the private and public cloud site will be determined by the distributed control network. When applications require the E2E QoS assurance, it can send these requirements to the PCE, and let the 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. Section 4 of this document describes the simulation results for this use case.

3.2. Link Utilization Maximization

Network topology within a Metro Area Network (MAN) is generally in a star mode as illustrated in Figure 2, with different devices connected to different customer types. The traffic from these customers is often in a tidal pattern, with the links between the Core Router(CR)/Broadband Remote Access Server(BRAS) and CR/Service Router(SR), experiencing 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 volume between them respectively and this causes these links often to be under-utilized.

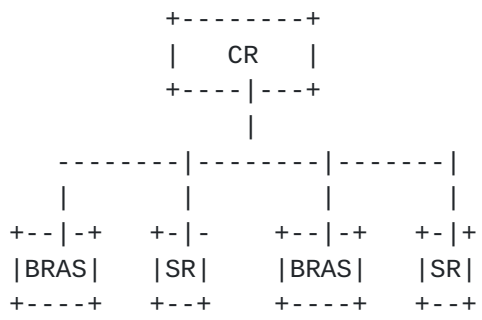


Figure 2: Star-mode Network Topology within MAN

If we consider connecting the BRAS/SR with a local link loop (which is usually lower cost), and control the overall MAN topology with the CCDR framework, we can exploit the tidal phenomena between the BRAS/ CR and SR/CR links, maximizing the utilization of these links (which are usually higher cost).

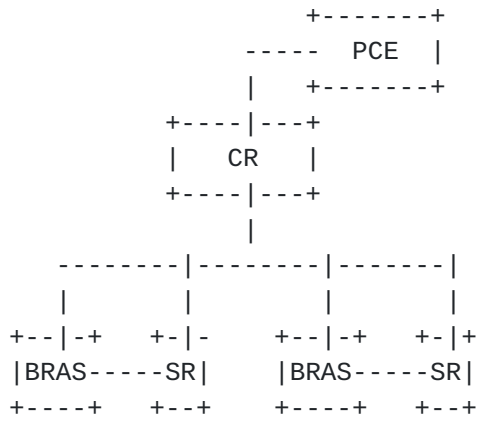


Figure 3: Link Utilization Maximization via CCDR

3.3. Traffic Engineering for Multi-Domain

Service provider networks are often comprised of different domains, interconnected with each other,forming a very complex topology as illustrated in Figure 4. Due to the traffic pattern to/from the MAN and IDC, the utilization of the links between them are often asymmetric. It is almost impossible to balance the utilization of these links via a distributed protocol, but this unbalance can be overcome utilizing the CCDR framework.

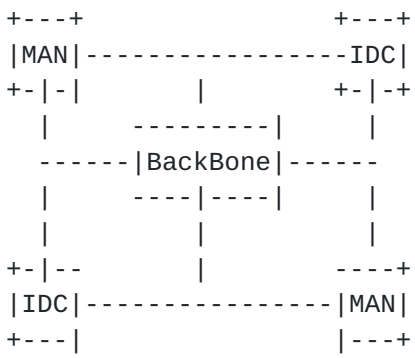


Figure 4: Traffic Engineering for Complex Multi-Domain Topology

A solution for this scenario requires the gathering of NetFlow information, analysis of the source/destination AS, and determining what is the main cause of the congested link. After this, the operator can use the external Border Gateway Protocol(eBGP) sessions to schedule the traffic among the different domains.

3.4. Network Temporal Congestion Elimination

In more general situations, there are often temporal congestions within the service provider's network. Such congestion phenomena often appear repeatedly, and if the service provider has methods to mitigate it, it will certainly improve their network operations capabilities and increase satisfaction for their customers. CCDR is also suitable for such scenarios, as the controller can schedule traffic out of the congested links, lowering the utilization of them during these times. Section 4 describes the simulation results of this scenario.

4. CCDR Simulation

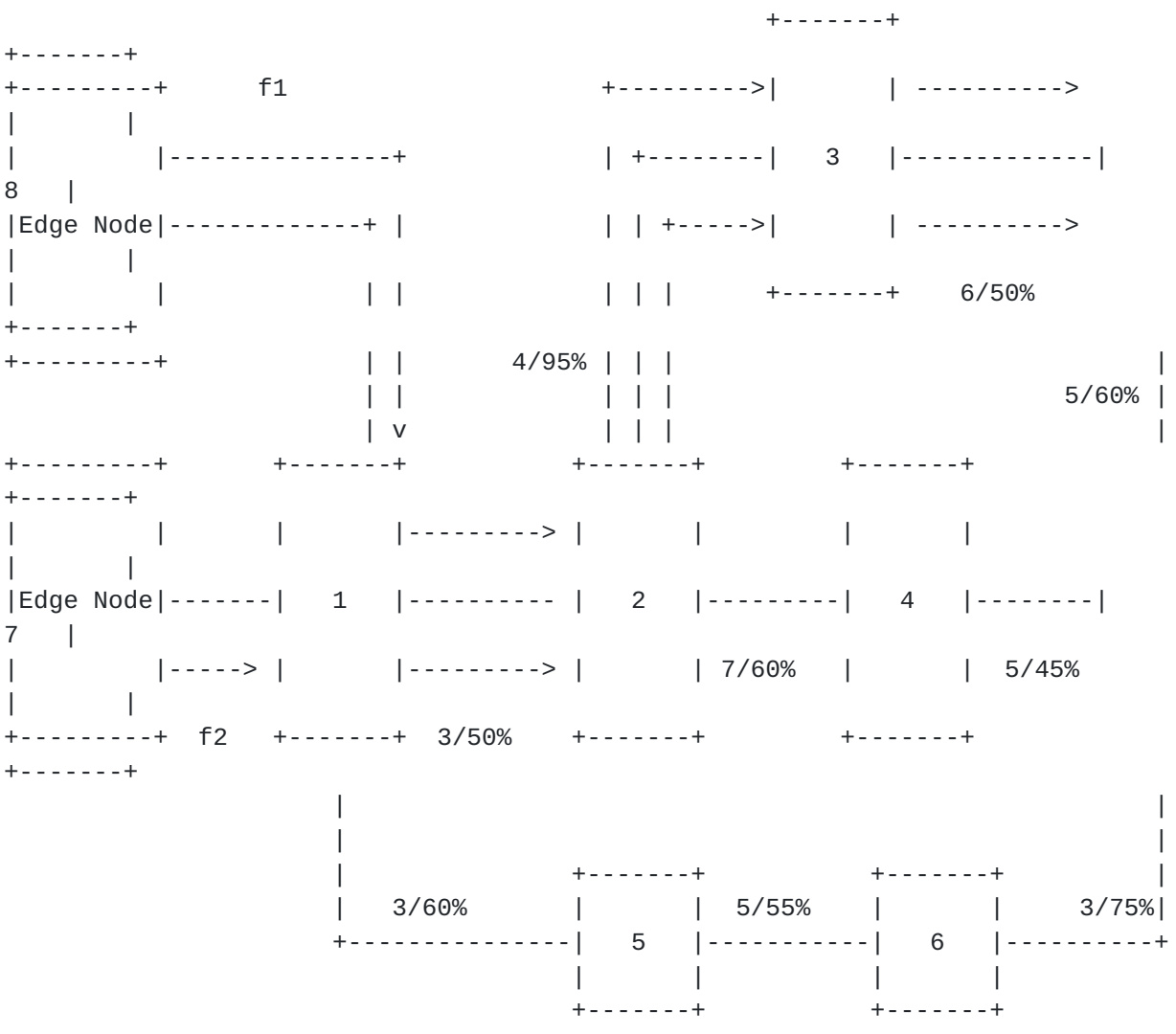
The following sections describe one case study to illustrate CCDR algorithm, the topology and traffic matrix generation process and the optimization results for E2E QoS assured path and congestion elimination in applied scenarios.

4.1. Case Study

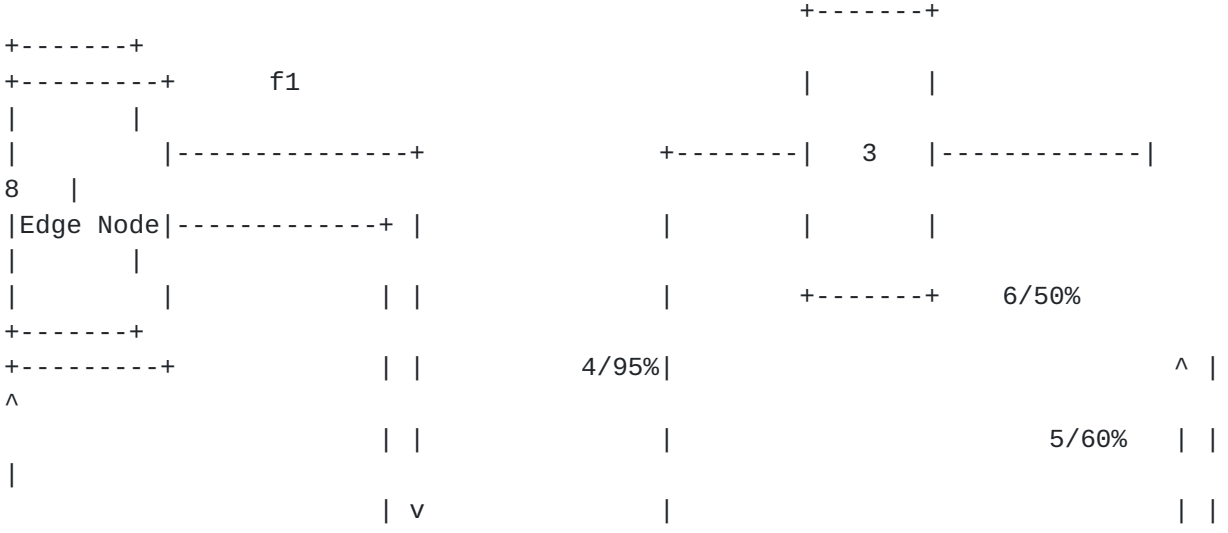
Figure 5 depicts the topology of the network for the case study. There are 8 forwarding devices in the network. The original cost and utilization are marked on it, as shown in the figure. For example, the original cost and utilization for the link (1,2) are 3 and 50% respectively. There are two flows: f1 and f2. Both of these two flows are from node 1 to node 8.For simplicity, it is assumed that the bandwidth of the link in the network is 10Mb/s.The flow rate of

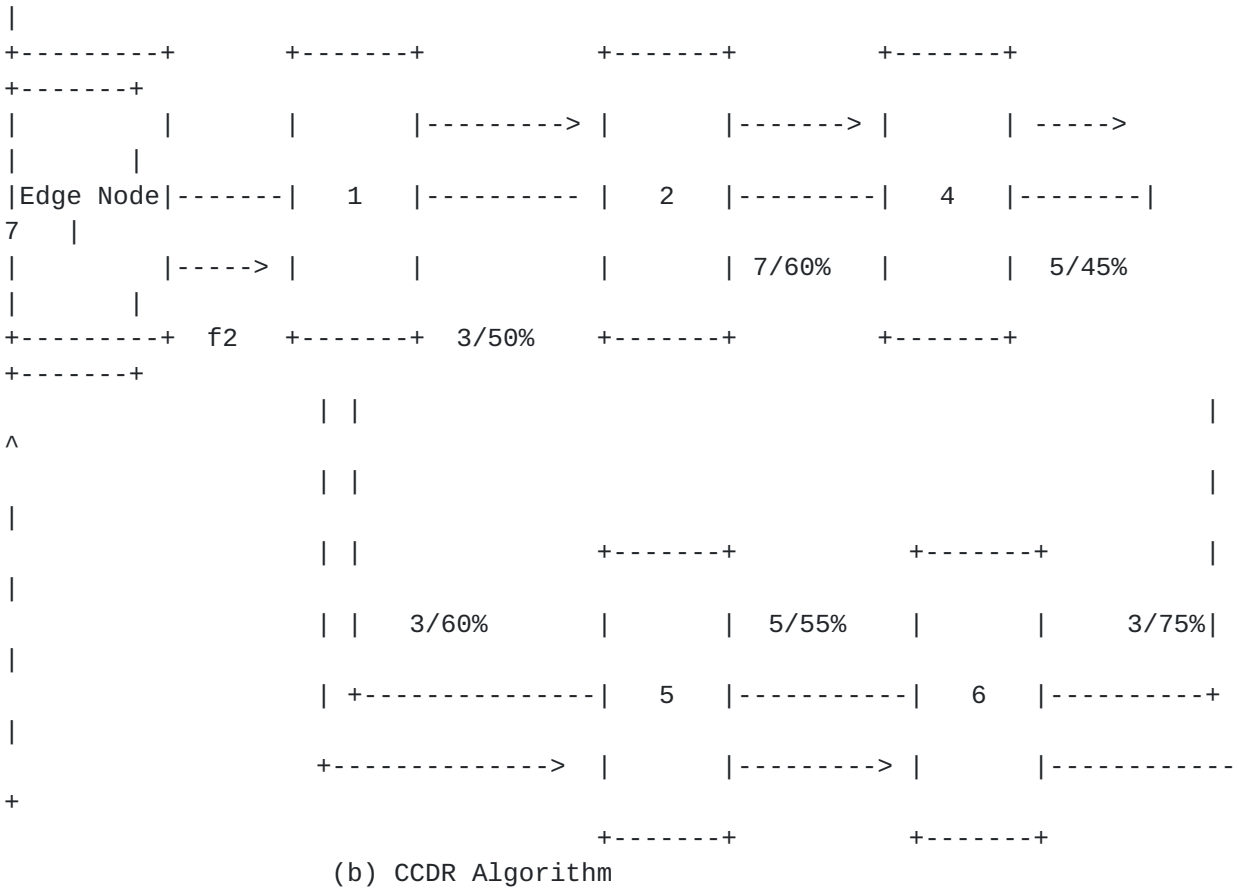
f1 is 1Mb/s, and the flow rate of f2 is 2Mb/s.The threshold of the link in congestion is 90%.

If OSPF protocol is applied in the network, which adopts Dijkstra's algorithm, the two flows from node 1 to node 8 can only use the OSPF path (p1: 1->2->3->8). It is because Dijkstra's algorithm mainly considers original cost of the link.Since CCDR considers cost and utilization simultaneously, the same path with OSPF will not be selected due to the severe congestion of the link (2,3). In this case, f1 will select the path (p2: 1->5->6->7->8) since the new cost of this path is better than that of OSPF path.Moreover, the path p2 is also better than the path (p3: 1->2->4->7->8) for for flow f1. However,f2 will not select the same path since it will cause the new congestion in the link (6,7). As a result, f2 will select the path (p3: 1->2->4->7->8).



(a) Dijkstra's Algorithm





(b) CCDR Algorithm

Figure 5: Case Study

4.2. Topology Simulation

The network topology mainly contains nodes and links information. Nodes used in the simulation have two types: core node and edge node. The core nodes are fully linked to each other. The edge nodes are connected only with some of the core nodes. Figure 6 is a topology example of 4 core nodes and 5 edge nodes. In this CCDR simulation, 100 core nodes and 400 edge nodes are generated.

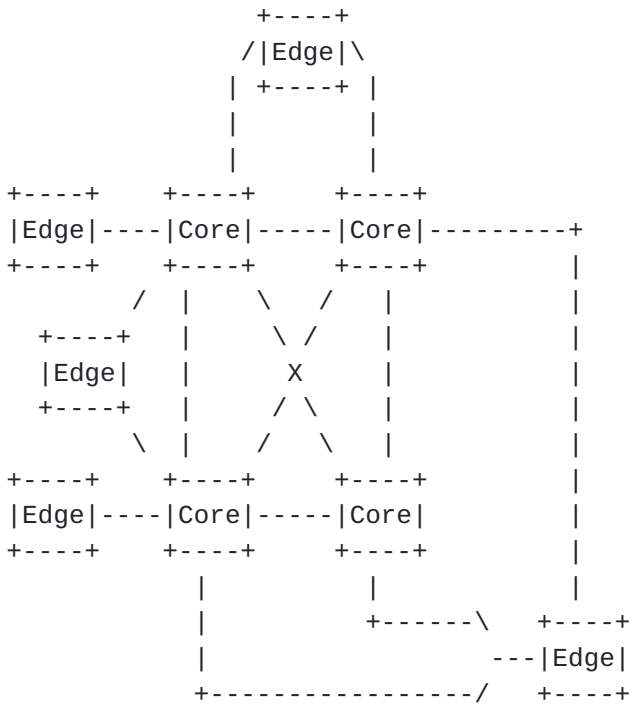


Figure 6: Topology of Simulation

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 a congestion threshold.

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

4.4. CCDR End-to-End Path Optimization

The CCDR E2E path optimization is to find the best 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, the PCE within CCDR framework combines the shortest path algorithm with a penalty theory of classical optimization and graph theory.

Given a background traffic matrix, which is unscheduled, when a set of new flows comes into the network, the E2E 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 Figure 7. The first graph in Figure 7 is the UID with OSPF and the second graph is the UID with CCDR E2E path optimization. The average UID of the first graph is more than 30%. After path optimization, the average UID is less than 5%. The results show that the CCDR E2E path optimization has an eye-catching decrease in UID relative to the path chosen based on OSPF.

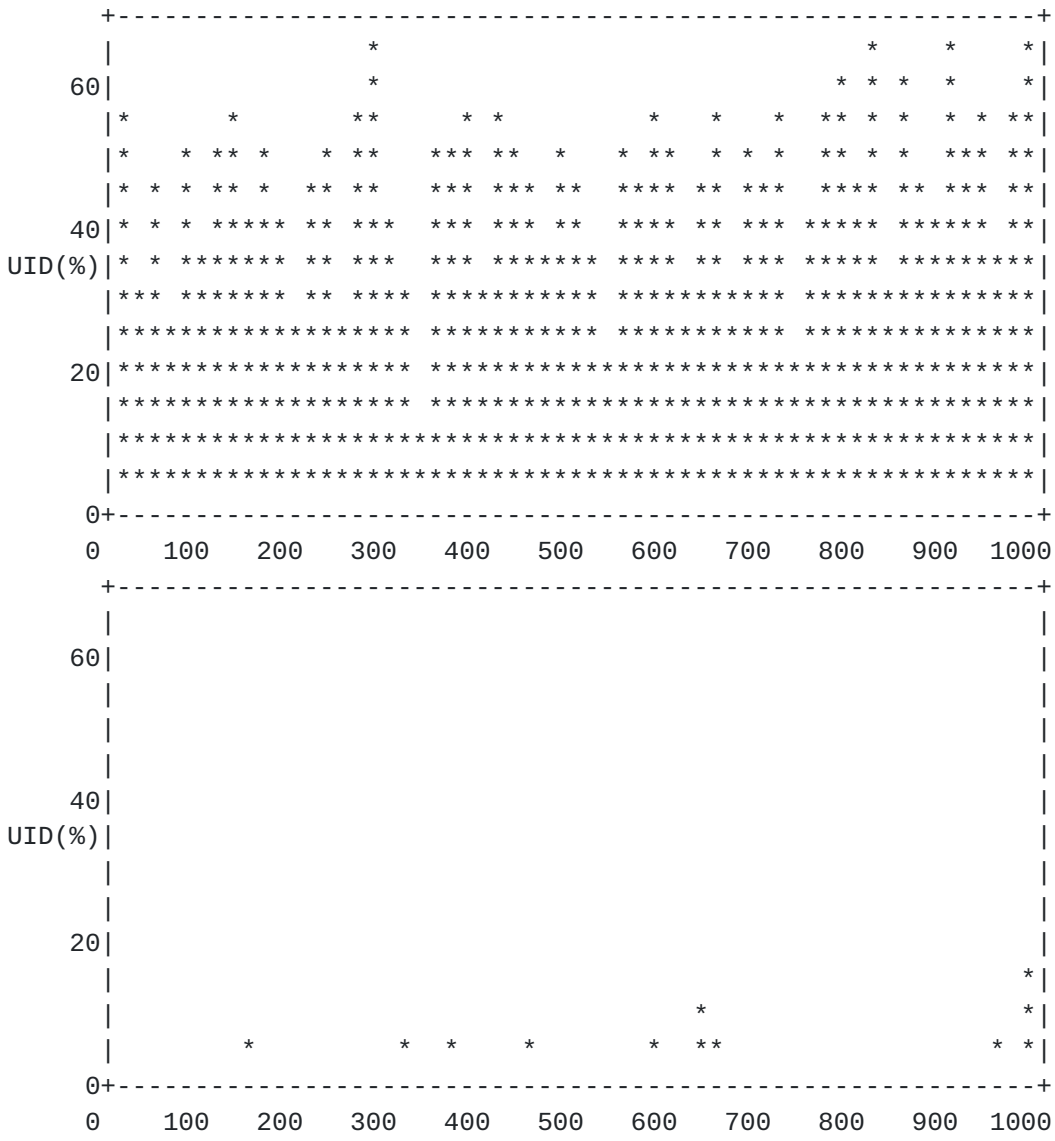


Figure 7: Simulation Result with Congestion Elimination

4.5. Network Temporal Congestion Elimination

Different degrees of network congestions were simulated. The Congestion Degree (CD) is defined as the link utilization beyond its threshold.

The CCDR congestion elimination performance is shown in Figure 8. The first graph is the CD distribution before the process of congestion elimination. The average CD of all congested links is more than 10%. The second graph shown in Figure 8 is the CD distribution after using the congestion elimination process. It shows only 12 links among totally 20000 links exceed the threshold, and all the CD values are less than 3%. Thus, after scheduling of the

traffic away from the congested paths, the degree of network congestion is greatly eliminated and the network utilization is in balance.

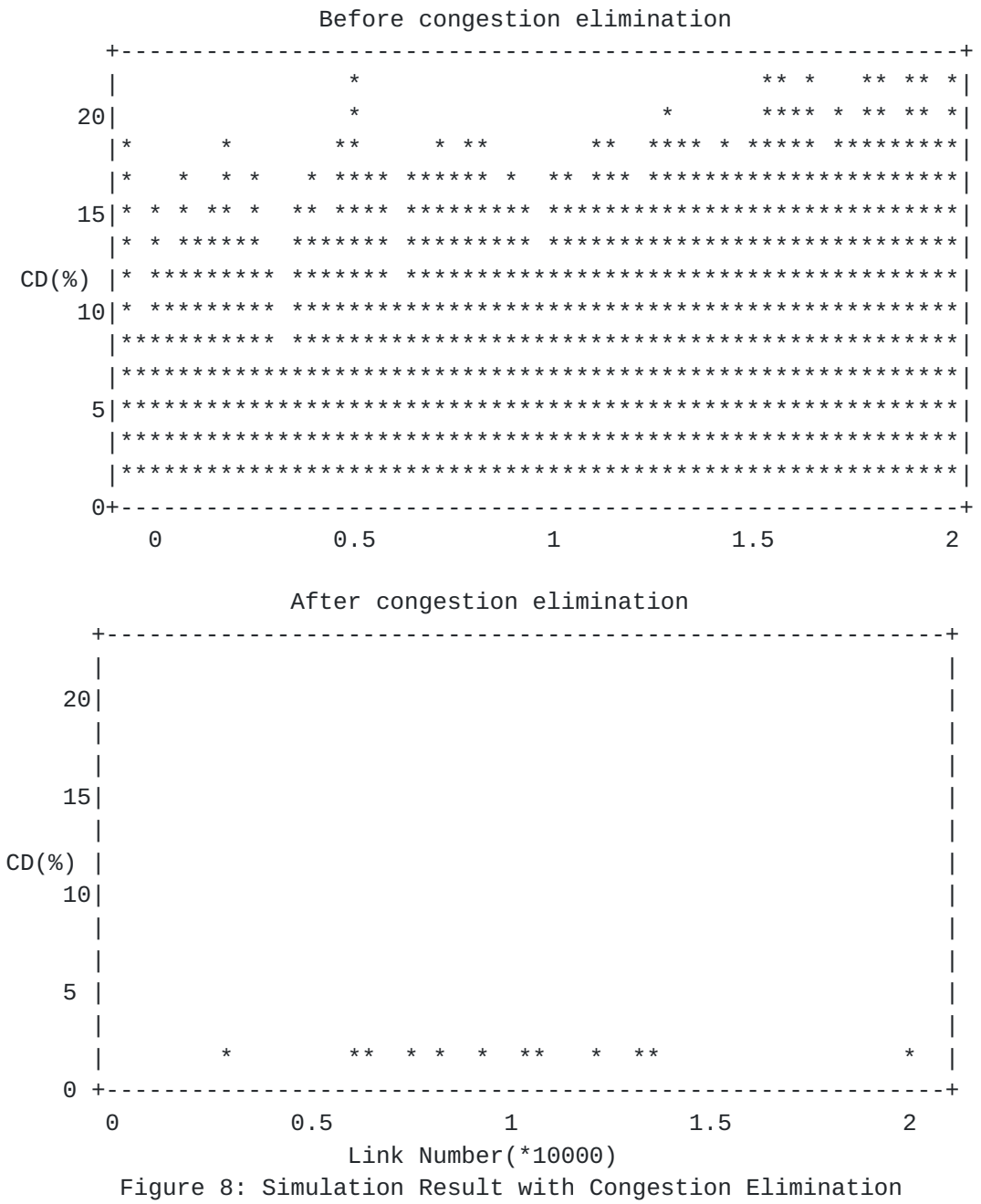


Figure 8: Simulation Result with Congestion Elimination

5. CCDR Deployment Consideration

With the above CCDR scenarios and simulation results, we demonstrate it is feasible to find one general solution to cope with various complex situations. Integrated use of a centralized controller for the more complex optimal path computations in a native IP network

results in significant improvements without impacting the underlay network infrastructure. A proposed solution is described in draft[I-D.ietf-teas-pce-native-ip] .

More detailed information about the algorithm can refer to the IEEE document " A Practical Traffic Control Scheme With Load Balancing Based on PCE Architecture"

6. Security Considerations

This document considers mainly the integration of distributed protocols and the central control capability of a PCE. While It certainly can ease the management of network in various traffic engineering scenarios as described in this document, the centralized control also bring a new point that may be easily attacked. Solutions for CCDR scenarios need to consider protection of the PCE and communication with the underlay devices. [RFC5440] and [RFC8253] provide additional information.

7. IANA Considerations

This document does not require any IANA actions.

8. Contributors

Lu Huang contributed to the content of this draft.

9. Acknowledgement

The author would like to thank Deborah Brungard, Adrian Farrel, Huaimo Chen, Vishnu Beeram and Lou Berger for their support and comments on this draft.

10. References

10.1. Normative References

[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>.

10.2. Informative References

[I-D.ietf-teas-pce-native-ip]
Wang, A., Zhao, Q., Khasanov, B., Chen, H., and R. Mallya, "PCE in Native IP Network", draft-ietf-teas-pce-native-ip-04 (work in progress), August 2019.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8578] Grossman, E., Ed., "Deterministic Networking Use Cases", RFC 8578, DOI 10.17487/RFC8578, May 2019, <https://www.rfc-editor.org/info/rfc8578>.

Authors' Addresses

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

Email: wangaj3@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

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