PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 29, 2010

D. Wang
Z. Wang
X. Fu
ZTE Corporation
October 26, 2009

Path Computation Element (PCE) Extensions for Inter-Layer Path Computation
draft-wdj-pce-for-inter-layer-path-computation-01

Abstract

This document mainly describes the extensions of multiple PCE inter-

layer path computation with inter-PCE communication.

Table of Contents

1. Introduction

The Path Computation Element (PCE) defined in [RFC4655] is an entity that is capable of computing a network path or route based on a network graph and applying computational constraints.

A network may comprise multiple layers. In [PCE-INTER-LAYER-FRWK], there are three inter-layer path computation models defined to perform PCE-based inter-layer path computation, namely Single PCE Computation, Multiple PCE Computation with inter-PCE communication, and Multiple PCE Computation without inter-PCE communication. Mainly by scalability, automation and confidentiality considerations, multiple PCE computation instead of single PCE computation is more concerned. In this mode, a PCE has the topology visibility of one or more layers (but not all layers, in this document, only one layer being considered).

This document mainly describes the extensions of multiple PCE inter-layer path computation with inter-PCE communication.

1.1. 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 [RFC2119].

This document uses terminology from the PCE-based path computation architecture [RFC4655] and also common terminology from Multi Protocol Label Switching (MPLS) [RFC3031], Generalized MPLS (GMPLS) [RFC3945], Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering [PCE-INTER-LAYER-FRWK], OSPF Protocol Extensions for Path Computation Element (PCE) Discovery [RFC5088], and Multi-Layer Networks [RFC5212].

2. Scenario Statement and Requirements Overview

2.1. Scenario Statement

In the model of multiple PCE inter-layer path computation, there is at least one PCE in each layer. Each PCE has the topology visibility restricted to its own layer. When an ingress LSR tries to set up an end-to-end LSP to an egress LSR is set up in the higher-layer network, as a PCC it will select a PCE located in its layer and send a path computation request to the PCE. At this time, the PCE will be confronted with such secenarios:

- As absence of essential TE link in the higher-layer network, the network topology of the whole layer is divided into at least two independent partitions.

- The border LSRs of each partition MAY be produced and a lower-layer network exists between the pairs of border LSRs belonging to two different partitions.

- The border LSRs MAY originate TE links connected to the lower-layer network.

- The ingress LSR and the egress LSR MAY be located in different partitions.

- When receiving a path computation request from the ingress LSR in the higher-layer, this higher-layer PCE will select one border LSR of the partition that the ingress LSR is located in and another border LSR of the partition that the egress LSR is located in respectively. This pair of LSRs will be taken as the head-end and the tail-end nodes which are included in the inter-layer path computation request sent to the lower-layer PCE. As the number of the border LSRs in each partition is usually more than one, higher-layer PCE will produce such pairs of border LSRs.

- As there can be at least one PCE in each layer,,the higher-layer PCE SHOULD select one from the lower-layer PCEs according to the information advertised by them and send a inter-layer path computation request about the pairs of border LSRs to the selected lower-layer PCE.

- In order to allow for effective PCE selection by PCCs, furthermore, a PCC needs to know the layer that the PCE is located in and the neighbor layer toward which the PCE can compute paths. Here, the neighbor layer is usually lower than its own layer according to multi-layer technology.

2.2. Requirements Overview

The multiple PCE inter-layer path computation mechanism MUST satisfy the following requirements.

For the higher-layer PCE, it MUST identify connectivity of its own layer topology by some algorithm. When the topology is no-connected, if the head-end node and the tail-end node lie in two independent partitions respectively, inter-layer path solution will have to be considered.

As there is no direct TE link between partitions in non-connected

topology, the higher-layer PCE SHOULD designate border LSRs to the lower-layer. Generally in one partition, obtainment of border LSR lies in whether the LSR has a TE link connected to the lower-layer network. If so, this LSR will be considered as a border LSR of the partition. An independent partition SHOULD usually have at least one border LSR. Similarly, the lower-layer PCE should also designate border LSRs which should have a TE link connected to the higher-layer network.

In each layer, PCE discovery information should include the layer into which the PCE has visibility.

In each layer, PCE discovery information should include the neighbor layer to which the PCE can send path computation request. In this case, the local-layer is usually higher, and the neighbor layer is lower. Since a NE receives discovery information of PCEs in the local-layer, as a PCC, it should know about which PCEs are in the same layer and which neighbor layer the PCEs can communicate with. When a path computation request is originated from this NE, firstly it should estimate the neighbor layer that the path passes through if possible.

A PCE discovery mechanism MUST allow that PCEs in different layers know about their discovery information each other, especially their local layer, neighbor layer, and the computation capability. Such information MAY be got by flooding mechanism of some routing protocol. So PCE discovery mechanism permits PCEs belonging to different layers get discovery information each other, especially the information about local layer, adjacent layer, and PCE's computing capability etc. Some of PCEs in different layers and with inter-layer computing capability can be chosen and consist of a PCE topology. Then in this PCE topology the PCEs' discovery information can be gotten by starting an instance of some routing protocol with flooding mechanism. According to the PCEs' discovery information, as a PCC, a higher-layer PCE may choose a suitable one from many adjacent lower-layer PCEs with inter-layer path computing capability in the PCE topology to compute an inter-layer path.

3. Terminology

BRPC: Backward-Recursive PCE-Based Computation.

FSC: Fiber Switching Capability.

GMPLS: Generalized Multi-Protocol Label Switching.

LSC: Lambda Switching Capability.

L2SC: Layer2 Switching Capability.

LSP: Label Switched Path.

LSR: Label Switching Router.

MPLS: Multi-Protocol Label Switching.

PCC: Path Computation Client.

PCE: Path Computation Element.

PCReq: Path Computation Request.

PCRsp: Path Computation Respond.

PSC: Packet Switching Capability.

TDM: Time Division Multiplex.

TE: Traffic Engineering.

4. Inter-Layer Optimal Path Computation Procedure

4.1. Confirmation of The Topology Connectivity

The PCE in the local layer will compute the optimal path from the ingress LSR to the egress LSR according to the information included by the PCReq message belonging to PCEP.

Generally, only when the network topology of the whole layer is divided into two or more partitions resulting from lack of TE links, and the ingress LSR and the egress LSR are respectively located in the different partitions, inter-layer path computation will be considered by the PCE.

4.2. Inter-layer TE Links

Whether it is higher-layer or lower-layer PCE needs to get the TE attribute of inter-layer TE links during the process of inter-layer TE LSP computation.

Between the local port belonging to one inter-layer border LSR in the side of higher-layer and the remote port belonging to the other one in the side of lower-layer, one-way inter-layer TE links usually appear in pairs. Among them, one is form the local port to the remote port, and the other is in the opposite direction. In other

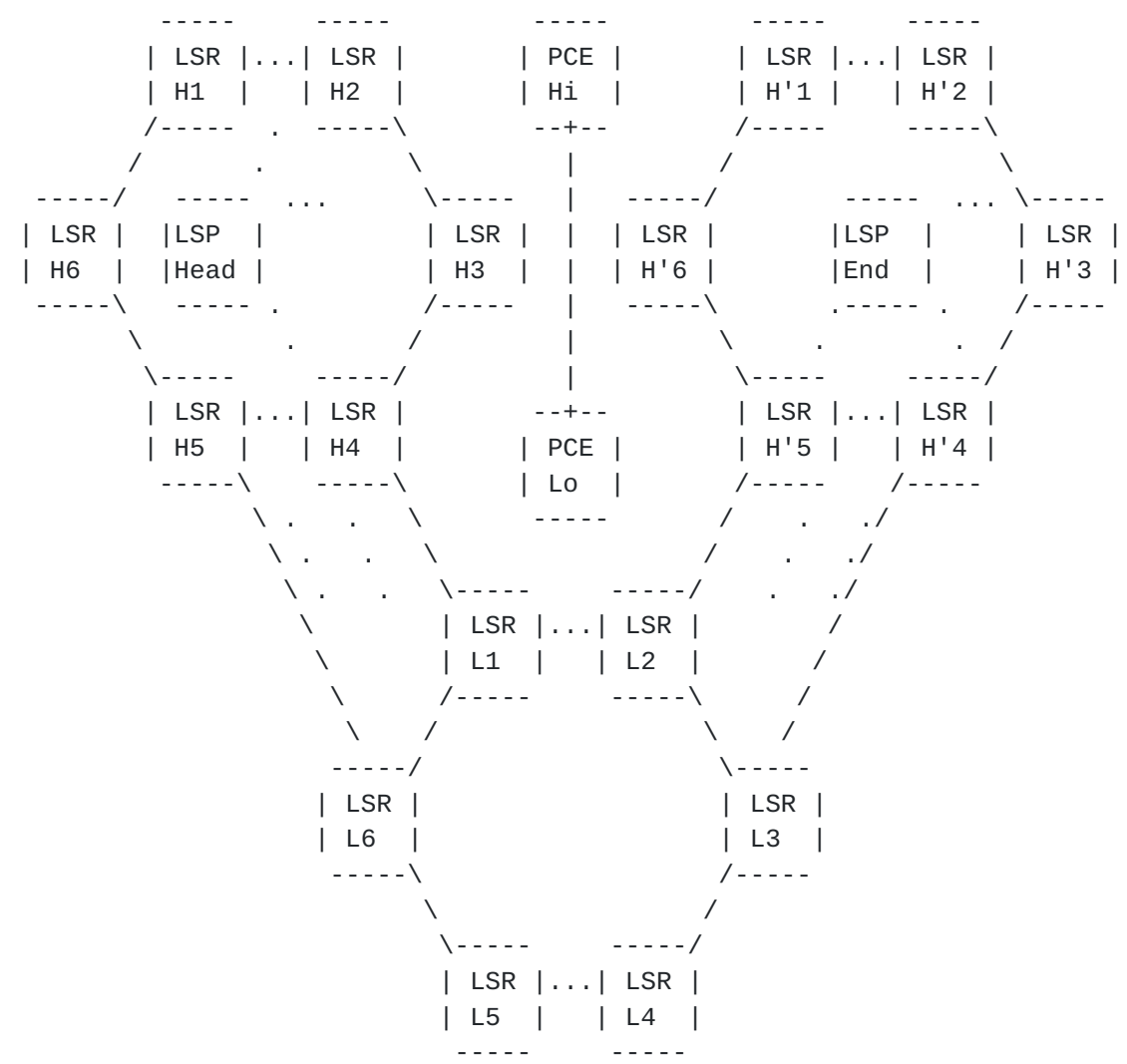
words, both of them consist of a bidirectional TE link. The former is originated from an inter-layer border LSR in the higher-layer, and it belongs to the scope of higher-layer topology into which only the higher-layer PCEs have visibility. Similarly the latter is originated from an inter-layer border LSR in the lower-layer, and it belongs to the scope of lower-layer topology into which only the lower-layer PCEs have visibility.

By an inter-layer TE link, an adaptation is completed from one type of switch capability to another one(e.g., electrical switching to optical switching, etc.).

4.3. Confirmation of Inter-Layer Border LSR

Inter-layer border LSRs are respectively located on the boundary of the high-layer and lower-layer networks. Each of them has at least one unidirectional TE link connected to the other inter-layer border LSR in the adjacent layer.

Before a higher-layer PCE begins to compute an inter-layer path, it SHOULD collect the information about all of inter-layer border LSRs in its local layer. And it MAY combine the inter-layer border LSRs belonging to the partition in which the ingress LSR is located with the inter-layer border LSRs belonging to the partition in which the egress LSR is located into many pairs of inter-layer border LSRs. By this combination, many pairs of inter-layer border LSRs are formed and as end LSRs provided for the lower-layer PCE to compute end-to-end paths across the lower-layer network. That is, a PCReq message which is sent to lower-layer PCE and is used to do inter-layer path computation will include many pairs of ingress and egress LSRs.



4.4. Discovery of Inter-layer PCEs

From PCEP's point of view, inter-layer PCE communication is also a kind of communication between PCC and PCE. The requesting entity will be looked as a PCC, and the requested entity will be looked as a PCE.

So PCE discovery mechanism permits PCEs belonging to different layers get discovery information each other, especially the information about local layer, adjacent layer, and PCE's computing capability etc. Some of PCEs in different layers and with inter-layer computing capability can be chosen and consist of a PCE topology. Then in this PCE topology the PCEs' discovery information can be gotten by starting an instance of some routing protocol with flooding mechanism. Only in this PCE topology can this routing protocol instance run. Its flooding contents are limited to the discovery information of those PCEs that are added to the PCE topology and can

not be beyond the scope of the PCE topology. According to the PCEs' discovery information, as a PCC, a higher-layer PCE may choose a suitable one from many adjacent lower-layer PCEs with inter-layer path computing capability in the PCE topology to compute an inter-layer path.

4.5. Communication of Inter-layer PCEs

If the higher-layer PCE decides to start inter-layer path computation, it will confirm as many pairs of end inter-layer LSRs as possible firstly. Then as a PCC, the higher-layer PCE will choose a suitable lower layer PCE according to the discovery information flooded by the lower layer PCEs and send a request to it for computing paths of the pairs of end inter-layer LSRs.

Since a inter-layer TE link includes the information about inter-layer LSRs on the both ends of it, when the lower-layer PCE receives the PCReq message, it can find the correspondent inter-layer TE links in its local-layer topology according to the message which includes many pairs of end inter-layer LSRs in the side of the higher-layer. By each inter-layer TE link, pairs of end inter-layer LSRs in the side of the lower-layer can be found.

Thus the lower-layer PCE can compute at least one inter-layer path by the above information and respond the computation result to the higher-layer PCE.

4.6. Optimal Inter-layer Path Computation by BRPC

In [PCE-INTER-LAYER-FRWK], BRPC is recommended to solve the problem of multiple PCE inter-layer path computation by considering layers as separate domains. BRPC is always started by the higher-layer PCE according to the response of lower-layer PCE about lower-layer path computation results. In addition, if multiple paths about a common pair of end LSRs are needed, all the complete paths will be ordered in some optimal strategy(e.g., minimum hops number, minimum aggregate te-metrics, etc.).

5. Routing Protocol Extensions for Inter-Layer PCE Discovery

In a multi-layer network model, multiple PCEs MAY have the function of inter-layer path computation. It may be necessary or desirable for a higher-layer PCE to choose a lower-layer PCE by the PCE discovery information originated from the lower-layer PCE.

Since the number of local-layer PCEs may be usually more than one, the PCEs' discovery information to which now is added local-layer and

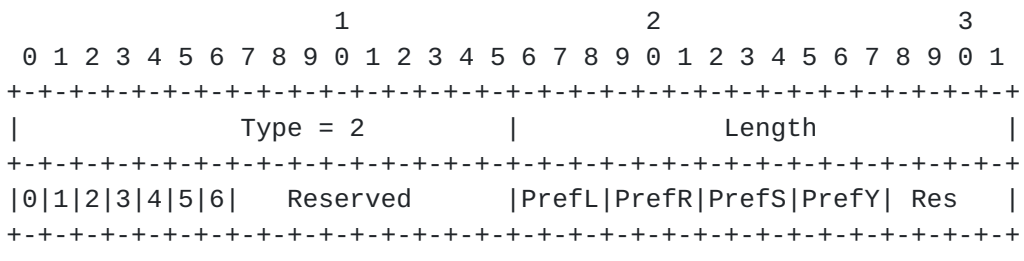
adjacent layer information needs to be flooded through all the layers unless safety is considered. By this way, as a PCC, when a higher-layer PCE needs to send a path computation request to a lower-layer PCE, it can dynamically choose a suitable one but not be allocated a fixed one according to the flooded layers' information from all the PCEs in the lower-layer. That is, all the PCEs belonging to all different layers will form a PCE topology and MAY participate in an independent IGP (OSPF or IS-IS) instance. This PCE topology's visibility MUST be opened only for these PCEs.

In [RFC5088] and [RFC5089], PCE discovery information is referred to. It is usually composed of the PCE location, the PCE path computation scope, the set of PCE-Domain(s), the set of neighbor PCE-Domain(s), and the set of communication capabilities etc. Here according to the above description, for the purpose of layer information automatic discovery and dynamically choosing among PCEs from different layers, the set of PCE-Layer(s) and the set of neighbor PCE Layer(s) SHOULD be added to the PCE discovery information.

Considering the two new types of PCE discovery information, the sub-TLVs that may be carried within the PCED TLV will be extended. The following sections describe the new extended sub-TLVs.

5.1. Extension of PATH-SCOPE Sub-TLV

In view of layer information discovery, here PATH-SCOPE Sub-TLV in [RFC5088] and [RFC5089] is extended. Yd bit is added to PATH-SCOPE Sub-TLV and the new sub-TLV has the following format:



Type: 2
Length: 4
Value: This comprises a 2-octet flags field where each bit represents a supported path scope, as well as four preference fields used to specify PCE preferences.

The following bits are defined:

Bit	Path Scope
0	L bit: Can compute intra-area paths.
1	R bit: Can act as PCE for inter-area TE LSP computation.
2	Rd bit: Can act as a default PCE for inter-area TE LSP computation.
3	S bit: Can act as PCE for inter-AS TE LSP computation.
4	Sd bit: Can act as a default PCE for inter-AS TE LSP computation.
5	Y bit: Can act as PCE for inter-layer TE LSP computation.
6	Yd bit: Can act as a default PCE for inter-layer TE LSP computation.

PrefL field: PCE's preference for intra-area TE LSP computation.

PrefR field: PCE's preference for inter-area TE LSP computation.

PrefS field: PCE's preference for inter-AS TE LSP computation.

PrefY field: PCE's preference for inter-layer TE LSP computation.

Res: Reserved for future use.

The L, R, S, and Y bits are set when the PCE can act as a PCE for intra-area, inter-area, inter-AS, or inter-layer TE LSP computation, respectively. These bits are non-exclusive.

Similarly, when set, the Yd bit indicates that the PCE can act as a default PCE for inter-layer TE LSP computation (that is, the PCE can compute a path toward an adjacent layer). When the Yd bit is set,

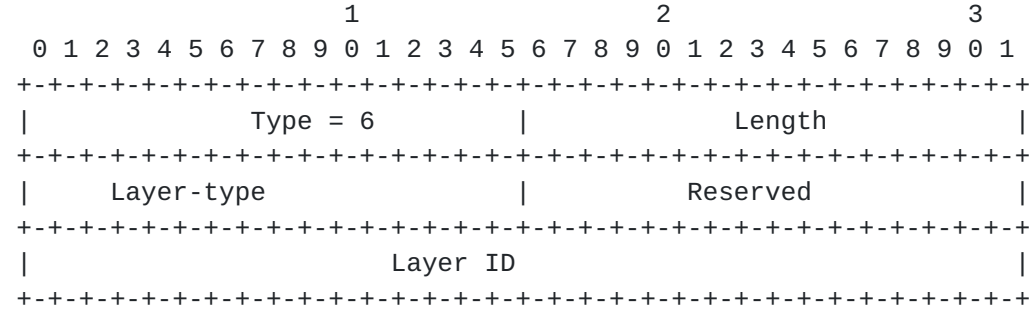
the PCED TLV MUST NOT contain a NEIG-PCE-LAYER sub-TLV.

When the Y bit is clear, the Yd bit SHOULD be clear on transmission and MUST be ignored on receipt.

5.2. PCE-LAYER Sub-TLV

The PCE-LAYER sub-TLV specifies a PCE-LAYER into which the PCE has topology visibility and through which the PCE can compute paths. The PCE-LAYER sub-TLV SHOULD be present when PCE-Layers for which the PCE can operate cannot be inferred by other IGP information: for instance, when the PCE is inter-layer capable (i.e., when the Y bit is set) and the flooding scope is the entire visible layer topology and the entire PCE topology that is composed of all the PCEs from all of layers. A PCED TLV may include multiple PCE-LAYER sub-TLVs when the PCE has visibility into multiple PCE-Layers.

The PCE-LAYER sub-TLV has the following format:



PCE-LAYER sub-TLV format

Type: 6
Length: 8
Current five layer-type values are defined:

1 PSC layer
2 L2SC layer
3 TDM layer
4 LSC layer
5 FSC layer

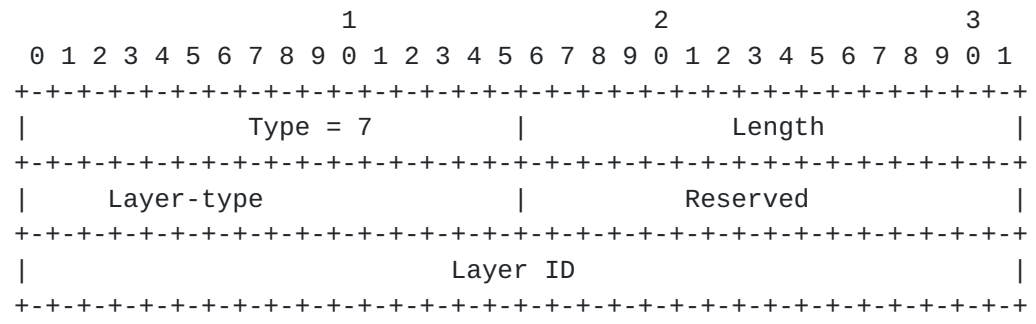
Layer ID: This indicates the 32-bit Layer ID of a layer where the PCE has visibility and can compute paths.

5.3. NEIG-PCE-LAYER Sub-TLV

The NEIG-PCE-LAYER sub-TLV specifies a neighbor PCE-Layer toward which a PCE can compute paths. It means that the PCE can take part in the computation of inter-layer TE LSPs with paths that transit this neighbor PCE-Layer.

For each PCE-LAYER sub-TLV, this neighbor layer usually refers to the lower adjacent layer.

The NEIG-PCE-LAYER sub-TLV has the following format:



NEIG-PCE-LAYER sub-TLV format

Type: 7
Length: 8
Current five layer-type values are defined:

1 PSC layer
2 L2SC layer
3 TDM layer
4 LSC layer
5 FSC layer

Layer ID: This indicates the 32-bit Layer ID of a neighbor layer toward which the PCE can compute paths.

The NEIG-PCE-LAYER sub-TLV MUST be present at least once if the Y bit is set and the Yd bit is cleared.

6. Security Considerations

The inter-layer path computation procedure relies on the use of the PCEP and as such is subjected to the potential attacks listed in Section 10 of [RFC5440]. In addition to the security mechanisms described in [RFC5440] with regards to spoofing, snooping, falsification, and denial of service, an implementation MAY support a

policy module governing the conditions under which a PCE should participate in the inter-layer path computation.

In addition, each layer topology's confidentiality should be especially considered.

For the purpose of network topology safety and confidentiality, the following points should be paid attention to:

- For each PCE in the PCE topology, the flooding scope of its discovery information is limited only to the local-layer topology and the PCE topology.

- For the PCEs outside the PCE topology, the flooding scope of its discovery information is limited only to the local-layer topology.

- For the PCEs of each layer, not all the PCEs can be added to the PCE topology. Only those which have the capability of inter-layer path computation are qualified for being added to the PCE topology.

- For confidentiality consideration, users can add the PCEs to the PCE topology or move them out of it dynamically. But if inter-layer path computation need to be desired, in each layer at least a PCE with the capability of inter-layer path computation should be add to the PCE topology. As a result, some of the PCEs in each layer are invisible to the PCEs in other layers, that is, the PCEs' visibility in one layer is partially opened to other layers.

7. Acknowledgments

The RFC text was produced using Marshall Rose's xml2rfc tool.

The authors would like to extend their warmest thanks to the ZTE's PCE software engineers.

8. Normative References

[ITUT-G709] ITU-T, "Interface for the Optical Transport Network (OTN)", G.709 Recommendation (and Amendment 1) , October 2001.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

[RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.

[RFC4328] Papadimitriou, D., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.

[RFC5150] Ayyangar, A., Kompella, K., Vasseur, JP., and A. Farrel, "Label Switched Path Stitching with Generalized Multiprotocol Label Switching Traffic Engineering (GMPLS TE)", RFC 5150, February 2008.

Authors' Addresses

Dajiang Wang
ZTE Corporation
12F,ZTE Plaza,No.19,Huayuan East Road,Haidian District
Beijing 100191
P.R.China

Phone: +8613811795408
Email: wang.dajiang@zte.com.cn
URI: http://wwwen.zte.com.cn/

Zhenyu Wang
ZTE Corporation
12F,ZTE Plaza,No.19,Huayuan East Road,Haidian District
Beijing 100191
P.R.China

Phone: +8613911266628
Email: wang.zhenyu1@zte.com.cn

Xihua Fu
ZTE Corporation
West District,ZTE Plaza,No.10,Tangyan South Road,Gaoxin District
Xi An 710065
P.R.China

Phone: +8613798412242
Email: fu.xihua@zte.com.cn
URI: http://wwwen.zte.com.cn/