Network work group
Internet Draft
Expires: October 2007
Category: Standards Track

Mach Chen
Renhai Zhang
Huawei Technologies Co.,Ltd
April 2, 2007

OSPF Traffic Engineering (OSPF-TE) Extensions in Support of Inter-AS Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineering

draft-chen-ccamp-ospf-interas-te-extensions-02.txt

Status of this Memo

Abstract

This document describes extensions to the OSPF Traffic Engineering (OSPF-TE) mechanisms to support Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineering (TE) for multiple Autonomous Systems (ASes). It defines OSPF-TE extensions for the flooding of TE information about inter-AS links which can be used to perform inter-AS TE path computation.

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].

Table of Contents

1. Introduction

[OSPF-TE] defines extensions to the OSPF protocol [OSPF] to support intra-area Traffic Engineering (TE). The extensions provide a way of encoding the TE information for TE-enabled links within the network (TE links) and flooding this information within an area. Type 10 opaque LSAs [RFC2370] are used to carry such TE information. Two top-level TLVs are defined in [OSPF-TE]: Router Address TLV and Link TLV. The Link TLV has several nested sub-TLVs which describe the TE attributes for a TE link.

[OSPF-TE-V3] defines similar extensions to OSPFv3 [OSPFV3].

Requirements for establishing Multiprotocol Label Switching (MPLS) TE Label Switched Paths (LSPs) that cross multiple Autonomous Systems (ASes) are described in [INTER-AS-TE-REQ]. As described in [INTER-AS-TE-REQ], a method SHOULD provide the ability to compute a path spanning multiple ASes. So a path computation entity that may be the head-end Label Switching Router (LSR), an AS Border Router (ASBR), or a Path Computation Element (PCE [PCE]) needs to know the TE information not only of the links within an AS, but also of the links that connect to other ASes.

In this document, some extensions to OSPF-TE are defined in support of carrying inter-AS TE link information for inter-AS Traffic Engineering. A new sub-TLV is added to the Link TLV and a new link type is introduced. The extensions are equally applicable to OSPFv2 and OSPFv3 as identical extensions to [OSPF-TE] and [OSPF-TE-V3]. The detailed definitions and procedures are discussed in the following sections.

2. Problem statement

As described in [INTER-AS-TE-REQ], in the case of establishing an inter-AS TE LSP traversing multiple ASes, the Path message [RFC3209] may include the following elements in the Explicit Route Object (ERO) in order to describe the path of the LSP:

- a set of AS numbers as loose hops; and/or

- a set of LSRs including ASBRs as loose hops.

Two methods for determining inter-AS paths are currently discussed. The per-domain method [PD-PATH] determines the path one domain at a time. The backward recursive method [BRPC] uses cooperation between PCEs to determine an optimum inter-domain path. The sections that follow examine how inter-AS TE link information could be useful in both cases.

2.1. A Note on Non-Objectives

It is important to note that this document does not make any change to the confidentiality and scaling assumptions surrounding the use of ASes in the Internet. In particular, this document is conformant to the requirements set out in [INTER-AS-TE-REQ].

The following lists of features are explicit exclusions.

o There is no attempt to distribute TE information from within one AS to another AS.

o There is no mechanism proposed to distribute any form of TE reachability information for destinations outside the AS.

o There is no proposed change to the PCE architecture or usage.

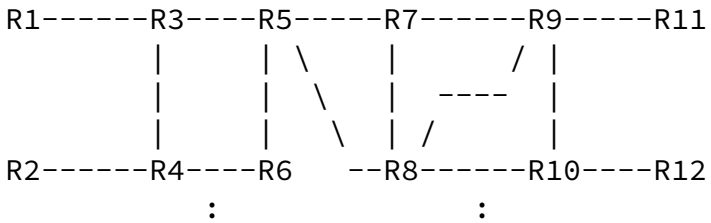
o TE aggregation is not supported or recommended.

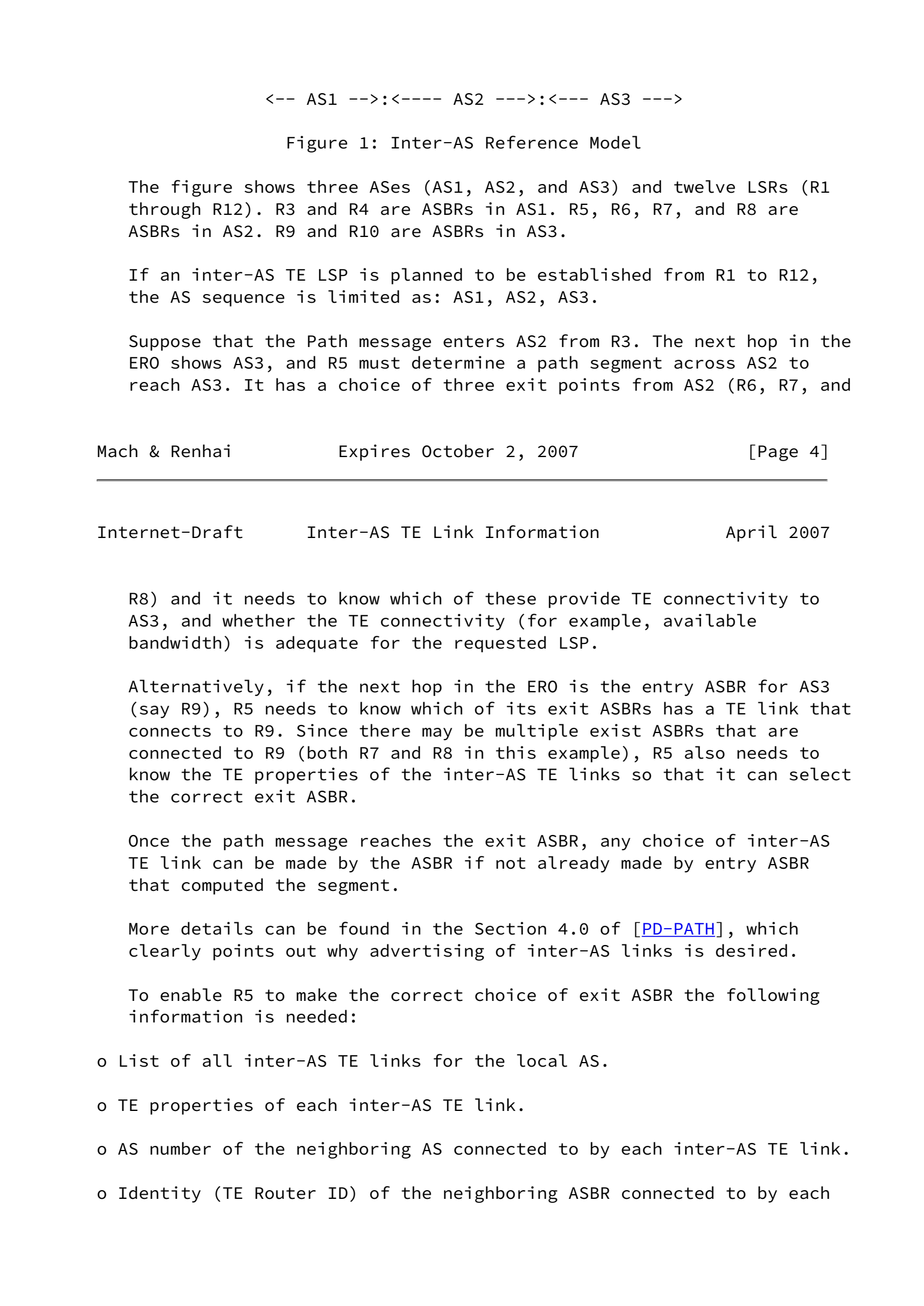
o There is no exchange of private information between ASes.

Note further that the extensions proposed in this document are limited to use for information about inter-AS TE links. L1VPN Auto-Discovery [L1VPN-OSPF-AD] defines how TE information about links between Customer Edge (CE) equipment and Provider Edge (PE) equipment can be advertised in OSPF alongside the auto-discovery information for the CE-PE links. That is separate functionality and does not overlap with the function defined in this document.

2.2. Per-Domain Path Determination

In the per-domain method of determining an inter-AS path for an MPLS-TE LSP, when an LSR that is an entry-point to an AS receives a PATH message from an upstream AS with an ERO containing a next hop that is an AS number, it needs to find which LSRs within the local AS are connected to the downstream AS so that it can compute a TE LSP segment across the AS to that LSR and forward the PATH message to the LSR and hence into the next AS. See the figure below for an example:





The figure shows three ASes (AS1, AS2, and AS3) and twelve LSRs (R1 through R12). R3 and R4 are ASBRs in AS1. R5, R6, R7, and R8 are ASBRs in AS2. R9 and R10 are ASBRs in AS3.

If an inter-AS TE LSP is planned to be established from R1 to R12, the AS sequence is limited as: AS1, AS2, AS3.

Suppose that the Path message enters AS2 from R3. The next hop in the ERO shows AS3, and R5 must determine a path segment across AS2 to reach AS3. It has a choice of three exit points from AS2 (R6, R7, and

R8) and it needs to know which of these provide TE connectivity to AS3, and whether the TE connectivity (for example, available bandwidth) is adequate for the requested LSP.

Alternatively, if the next hop in the ERO is the entry ASBR for AS3 (say R9), R5 needs to know which of its exit ASBRs has a TE link that connects to R9. Since there may be multiple exist ASBRs that are connected to R9 (both R7 and R8 in this example), R5 also needs to know the TE properties of the inter-AS TE links so that it can select the correct exit ASBR.

Once the path message reaches the exit ASBR, any choice of inter-AS TE link can be made by the ASBR if not already made by entry ASBR that computed the segment.

More details can be found in the Section 4.0 of [PD-PATH], which clearly points out why advertising of inter-AS links is desired.

To enable R5 to make the correct choice of exit ASBR the following information is needed:

o List of all inter-AS TE links for the local AS.

o TE properties of each inter-AS TE link.

o AS number of the neighboring AS connected to by each inter-AS TE link.

o Identity (TE Router ID) of the neighboring ASBR connected to by each

inter-AS TE link.

In GMPLS networks further information may also be required to select the correct TE links as defined in [GMPLS-TE].

The example above shows how this information is needed at the entry point ASBRs for each AS (or the PCEs that provide computation services for the ASBRs), but this information is also needed throughout the local AS if path computation function is fully distributed among LSRs in the local AS, for example to support LSPs that have start points (ingress nodes) within the AS.

2.3. Backward Recursive Path Computation

Another scenario using PCE techniques has the same problem. [BRPC] defines a PCE-based TE LSP computation method (called Backward Recursive Path Computation) to compute optimal inter-domain constrained MPLS-TE or GMPLS LSPs. In this path computation method, a specific set of traversed domains are assumed to be selected before

computation starts. Each downstream PCE in domain(i) returns a multipoint-to-point tree of potential paths to its upstream neighbor PCE in domain(i-1). Each tree consists of the set of paths from all Boundary Nodes located in domain(i) to the destination where each path satisfies the set of required constraints for the TE LSP (bandwidth, affinities, etc.).

So a PCE needs to select Boundary Nodes (that is, ASBRs) that provide connectivity from the upstream AS. In order that the tree of paths provided by one PCE to its neighbor can be correlated, the identities of the ASBRs for each path need to be referenced, so the PCE must know the identities of the ASBRs in the remote AS reached by any inter-AS TE link, and, in order that it provides only suitable paths in the tree, the PCE must know the TE properties of the inter-AS TE links.

Thus, to support Backward Recursive Path Computation the same information as listed in Section 2.2 is required.

3. Extensions to OSPF-TE

Note that this document does not define mechanisms for distribution of TE information from one AS to another, does not distribute any

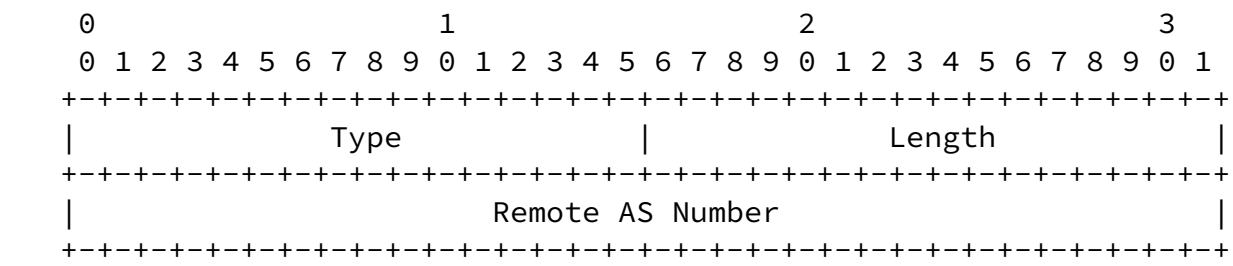
form of TE reachability information for destinations outside the AS, does not change the PCE architecture or usage, does not suggest or recommend any form of TE aggregation, and does not feed private information between ASes. See section 2.1.

The extensions defined in this document allow an inter-AS TE link advertisement to be easily identified as such by the use of a new link type. A new sub-TLV to the Link TLV is defined to carry the information about the neighboring AS. The extensions are equally applicable to TE distribution using OSPFv2 and OSPFv3.

3.1. Remote AS Number Sub-TLV

As described in [OSPF-TE], the Link TLV describes a single link and consists of a set of sub-TLVs. A new sub-TLV, the Remote AS Number sub-TLV is added to the Link TLV when advertising inter-AS links. The Remote AS Number sub-TLV specifies the AS number of the neighboring AS to which the advertised link connects.

The Remote AS number sub-TLV is TLV type 21 (which needs to be confirmed by IANA), and is four octets in length. The format is as follows:



The Remote AS number field has 4 octets. When two octets are used for the AS number, as in current deployments, the left (high-order) two octets MUST be set to zero.

3.2. Inter-AS Link Type

To identify a link as an inter-AS link and allow easy identification of these new advertisements, a new Link Type value is defined for use in the Link Type sub-TLV. The value of the Link Type for an inter-AS point-to-point link is 3 (which needs to be confirmed by IANA). The

use of multi-access inter-AS TE links is for future study.

3.3. Link ID

For an inter-AS link, the Link ID carried in the Link ID sub-TLV is the TE Router ID of the remote ASBR reached through this inter-AS link.

4. Procedure for Inter-AS TE Links

When TE is enabled on an inter-AS link and the link is up, the ASBR SHOULD advertise this link using the normal procedures for OSPF-TE [OSPF-TE]. When either the link is down or TE is disabled on the link, the ASBR SHOULD withdraw the advertisement. When there are changes to the TE parameters for the link (for example, when the available bandwidth changes) the ASBR SHOULD re-advertise the link, but the ASBR MUST take precautions against excessive re-advertisements as described in [OSPF-TE].

The information advertised comes from the ASBR's knowledge of the TE capabilities of the link, the ASBR's knowledge of the current status and usage of the link, and configuration at the ASBR of the remote AS number and remote ASBR TE Router ID.

The TE link advertisement SHOULD be carried in a Type 10 Opaque LSA if the flooding scope is to be limited to within the single IGP area to which the ASBR belongs, or MAY be carried in a Type 11 Opaque LSA if the information should reach all routers (including area border routers, ASBRs, and PCEs) in the AS.

Legacy routers receiving an advertisement for an inter-AS TE link are able to ignore it because the Link Type carries an unknown value. They will continue to flood the LSA, but will not attempt to use the information received as if the link were an intra-AS TE link.

Routers or PCEs that are capable of processing advertisements of inter-AS TE links SHOULD NOT use such links to compute paths that exit an AS to a remote ASBR and then immediately re-enter the AS through another TE link. Such paths would constitute extremely rare occurrences and SHOULD NOT be allowed except as the result of specific policy configurations at the router or PCE computing the path.

5. Security Considerations

The protocol extensions defined in this document are relatively minor and can be secured within the AS in which they are used by the existing OSPF security mechanisms.

It should be noted, however, that some of the information included in these new advertisements(the remote AS number and the remote ASBR ID) are obtained from a neighboring administration and cannot be verified in anyway. Since the means of delivery of this information is likely to be part of a commercial relationship, the source of the information should be carefully checked before it is entered as configuration information at the ASBR responsible for advertising the inter-AS TE links.

6. IANA Considerations

IANA is requested to make the following allocations from registries under its control.

6.1. OSPF LSA Sub-TLVs type

IANA maintains the "Open Shortest Path First (OSPF) Traffic Engineering TLVs" registry with sub-registry "Types for sub-TLVs in a TE Link TLV". IANA is requested to assign a new sub-TLV as follows. The number 21 is suggested.

Value	Meaning
21	Remote AS Number sub-TLV.

6.2. OSPF TE Link Type

IANA is requested to create a new sub-registry "TE Link Types" of the registry "Open Shortest Path First (OSPF) Traffic Engineering TLVs" to track TE Link Types.

The sub-registry should read as follows:

[OSPF-TE] defines the Link Type sub-TLV of the Link TLV. The following values are defined.

Value	Meaning	Reference
1	Point-to-point link	[OSPF-TE]
2	Multi-access link	[OSPF-TE]
3	Inter-AS link	[this document]

New allocations from this registry are by IETF Standards Action.

7. Acknowledgments

The authors would like to thank Adrian Farrel, Acee Lindem, JP Vasseur, and Dean Cheng for their review and comments to this document.

8. References

8.1. Normative References

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

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC2370] R. Coltun, "The OSPF Opaque LSA Option", RFC2370, July 1998.

[OSPF] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

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

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

8.2. Informative References

[INTER-AS-TE-REQ] Zhang and Vasseur, "MPLS Inter-AS Traffic Engineering Requirements", RFC4216, November 2005.

[PD-PATH] Ayyangar, A., Vasseur, JP., and Zhang, R., "A Per-domain path computation method for establishing Inter-domain", draft-ietf-ccamp-inter-domain-pd-path-comp, (work in progress).

[BRPC] JP. Vasseur, Ed., R. Zhang, N. Bitar, JL. Le Roux, "A Backward Recursive PCE-based Computation (BRPC) procedure to compute shortest inter-domain Traffic Engineering Label Switched Paths ", draft-ietf-pce-brpc, (work in progress)

[PCE] Farrel, A., Vasseur, JP., and Ash, J., "A Path Computation Element (PCE)-Based Architecture", RFC4655, August 2006.

[OSPF-TE-V3] Ishiguro K., Manral V., Davey A., and Lindem A. "Traffic Engineering Extensions to OSPF version 3", draft-ietf-ospf-ospfv3-traffic, {work in progress}.

[OSPFV3] Coltun, R., Ferguson, D., and J. Moy, "OSPF for IPv6", RFC 2740, April 1998.

[L1VPN-OSPF-AD] Bryskin, I., and Berger, L., "OSPF Based L1VPN Auto-Discovery", draft-ietf-l1vpn-ospf-auto-discovery-02.txt, (work in progress).

Author's Addresses

Mach Chen
Huawei Technologies Co.,Ltd
KuiKe Building, No.9 Xinxi Rd.,
Hai-Dian District
Beijing, 100085
P.R. China

Email: mach@huawei.com

Renhai Zhang
Huawei Technologies Co.,Ltd
KuiKe Building, No.9 Xinxi Rd.,
Hai-Dian District
Beijing, 100085
P.R. China

Email: zhangrenhai@huawei.com

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment