Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: November 5, 2013

H. Chen
R. Li
Huawei Technologies
G. Cauchie

N. So
Tata Communications
L. Liu
UC Davis
A. Retana
Cisco Systems, Inc.

May 4, 2013

OSPF Topology-Transparent Zone
draft-chen-ospf-ttz-05.txt

Abstract

This document presents a topology-transparent zone in a domain. A topology-transparent zone comprises a group of routers and a number of links connecting these routers. Any router outside of the zone is not aware of the zone. The information about the links and routers inside the zone is not distributed to any router outside of the zone. Any link state change such as a link down inside the zone is not seen by any router outside of the zone.

Status of this Memo

This Internet-Draft is submitted to IETF 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 http://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 November 5, 2013.

Table of Contents

1. Introduction

The number of routers in an Autonomous System (AS) becomes larger and larger as the Internet traffic keeps growing. Thus the Open Shortest Path First (OSPF) Link State Database (LSDB) and OSPF routing table are bigger and bigger. Any link state change in an AS leads to a number of link state distributions to every router in the AS. This triggers every router in the AS to re-calculate its OSPF routes, update its Routing Information Base (RIB) and Forwarding Information Base (FIB). All these will consume network resource including network bandwidth and Central Process Unit (CPU) time. This blocks further expansions of a network.

RFC 2328 "OSPF Version 2" describes OSPF areas in an AS. Each area has a number of area border routers connected to the backbone area. Each area border router summarizes the topology of its attached non backbone areas for transmission on the backbone, and hence to all other area border routers.

Through splitting a network into multiple areas, we can extend the network further. However, there are a number of issues when a network is split further into more areas.

At first, dividing an AS or an area into multiple areas is a very challenging task since it is involved in significant network architecture changes.

Secondly, it is complex for a Multi-Protocol Label Switching (MPLS) Traffic Engineering (TE) Label Switching Path (LSP) crossing multiple areas to be setup. In general, a TE path crossing multiple areas is computed by using collaborating Path Computation Elements (PCEs) [RFC5441] through the PCE Communication Protocol (PCEP)[RFC5440], which is not easy to configure by operators since the manual configuration of the sequence of domains is required. Although this issue can be addressed by using the Hierarchical PCE, this solution may further increase the complexity of network design. Especially, the current PCE standard method may not guarantee that the path found is optimal.

Thirdly, some policies need to be configured on area border routers for reducing the number of link states such as summary Link-State Advertisements (LSAs) to be distributed to other routers in other areas.

Furthermore, route convergence may be slower. A router in an OSPF area can see all other routers in the same area. A link-state change anywhere in an OSPF area will be populated everywhere in the same area, and may even be distributed to other areas in the same AS

indirectly. For example, all the routers and links in a Point-Of-Presence (POP) in an OSPF area will be seen by all the other routers in the same area. Any link state change in the POP will be distributed to all the other routers in the same area and may be distributed to routers in other areas indirectly.

A link state change in an area will lead to every router in the same area to re-calculate its OSPF routes, update its RIB and FIB. It may also lead to a number of link state distributions to other areas. This will trigger routers in other areas to re-calculate their OSPF routes, update their RIBs and FIBs. Thus the route convergence is slower.

This document presents a topology-transparent zone in a domain or an area and describes extensions to OSPF for supporting the topology-transparent zone, which may resolve the issues above.

A topology-transparent zone comprises a group of routers and a number of links connecting these routers. Any router outside of the zone is not aware of the zone. The information about the links and routers inside the zone is not distributed to any router outside of the zone. Any link state change such as a link down inside the zone is not seen by any router outside of the zone.

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.

3. Requirements

Topology-Transparent Zone (TTZ) may be deployed for resolving some ctricial issues such as scalability in existing networks and future networks. The requirements for TTZ are listed as follows:

- TTZ MUST be backward compitable. When a TTZ is deployed on a set of routers in a network, the routers outside of the TTZ in the network do not need to know or support TTZ.

- TTZ MUST support at least one more levels of network hierarchies, in addition to the hierarchies supported by existing routing protocols.

- Users SHOULD be able to easily set up an end to end service crossing TTZs.

- The configuration for a TTZ in a network SHOULD be minimum.

- The changes on the existing protocols for supporting TTZ SHOULD be minimum.

4. Topology-Transparent Zone

4.1. Overview of Topology-Transparent Zone

A Topology-Transparent Zone (TTZ) is identified by an Identifier (ID), and it includes a group of routers and a number of links connecting the routers. A Topology-Transparent Zone is in an OSPF domain.

The ID of a Topology-Transparent Zone (TTZ) or TTZ ID is a number that is unique for identifying an entity such as a node in an OSPF domain. It is not zero in general.

In addition to having the functions of an OSPF area, an OSPF TTZ makes some improvements on an OSPF area, which include:

- An OSPF TTZ is virtualized as a group of TTZ edge routers connected.

- An OSPF TTZ receives the link state information about the topology outside of the TTZ, stores the information in the TTZ and floods the information through the TTZ to the routers outside of TTZ.

- No Policy configuration is needed on any edge router of a TTZ.

4.2. An Example of TTZ

4.2.1. Creation of a TTZ

The figure below illustrates an example of a routing domain containing a topology-transparent zone: TTZ 600.

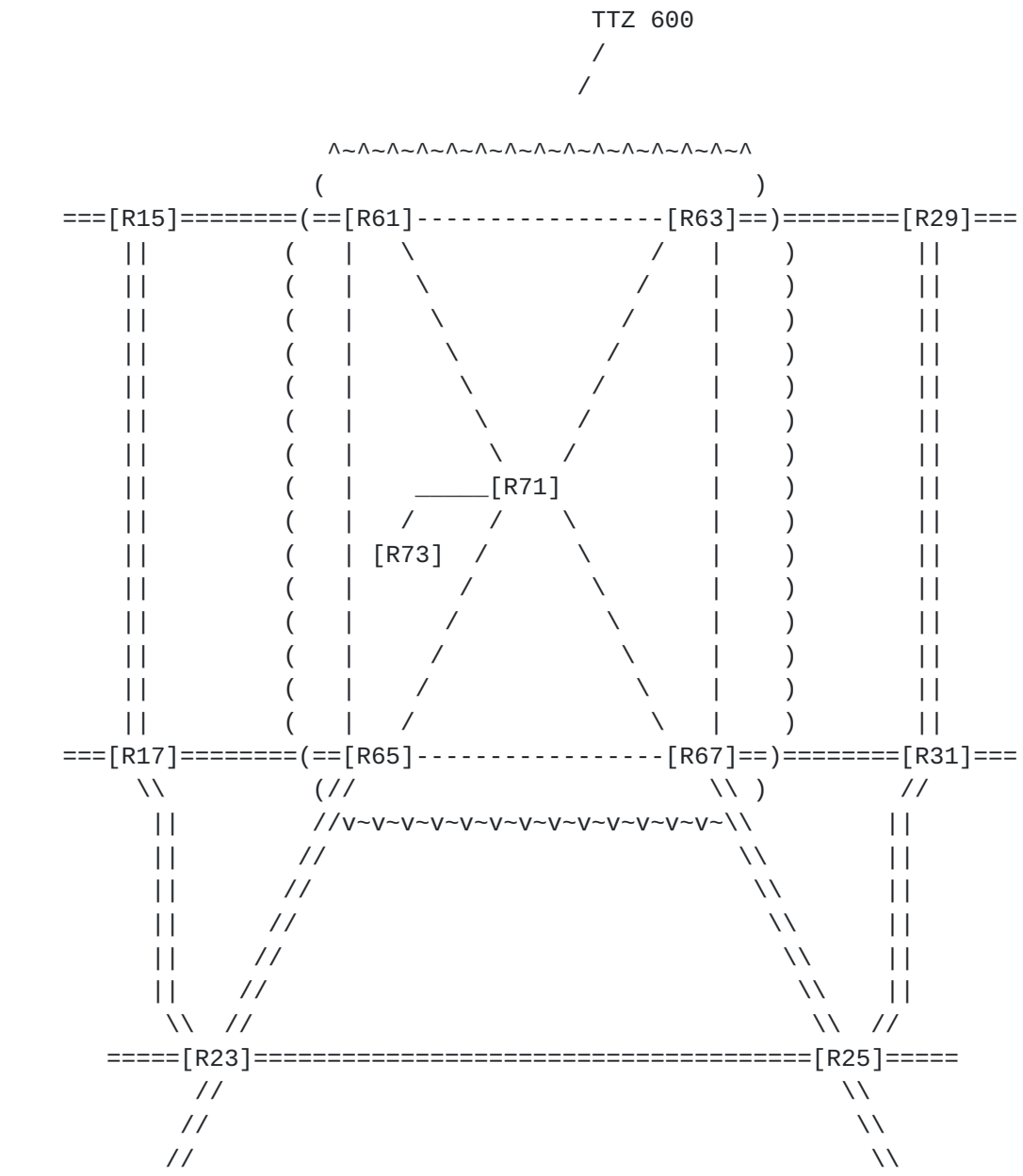


Figure 1: An Example of TTZ

The routing domain comprises routers R15, R17, R23, R25, R29 and R31. It also contains a topology-transparent zone TTZ 600. The TTZ 600 comprises routers R61, R63, R65, R67, R71 and R73, and the links connecting them.

There are two types of routers in a Topology-Transparent Zone (TTZ): TTZ internal routers and TTZ edge routers. A TTZ internal router is a router inside the TTZ and every adjacent router of the TTZ internal router is a router inside the TTZ. A TTZ edge router is a router

inside the TTZ and has at least one adjacent router that is outside of the TTZ.

The TTZ in the figure above comprises four TTZ edge routers R61, R63, R65 and R67. Each TTZ edge router is connected to at least one router outside of the TTZ. For instance, router R61 is a TTZ edge router since it is connected to router R15, which is outside of the TTZ.

In addition, the TTZ comprises two TTZ internal routers R71 and R73. A TTZ internal router is not connected to any router outside of the TTZ. For instance, router R71 is a TTZ internal router since it is not connected to any router outside of the TTZ. It is just connected to routers R61, R63, R65, R67 and R73 inside the TTZ.

A TTZ MUST hide the information inside the TTZ from the outside. It MUST NOT directly distribute any internal information about the TTZ to a router outside of the TTZ.

For instance, the TTZ in the figure above MUST NOT send the information about TTZ internal router R71 to any router outside of the TTZ in the routing domain; it MUST NOT send the information about the link between TTZ router R61 and R65 to any router outside of the TTZ.

In order to create a Topology-Transparent Zone (TTZ), we MUST configure the same TTZ ID on the edge routers and identify the TTZ internal links on them. In addition, we SHOULD confiure the TTZ ID on every TTZ internal router which indicates that every link of the router is a TTZ internal link.

From a router outside of the TTZ, a TTZ is seen as a group of routers fully connected. For instance, router R15 in the figure above, which is outside of TTZ 600, sees TTZ 600 as a group of TTZ edge routers: R61, R63, R65 and R67. These four TTZ edge routers are fully connected.

In addition, a router outside of the TTZ sees TTZ edge routers having normal connections to the routers outside of the TTZ. For example, router R15 sees four TTZ edge routers R61, R63, R65 and R67, which have the normal connections to R15, R29, R17 and R23, R25 and R31 respectively.

5. Changes to OSPF Protocols

There are a number of ways to extend the existing OSPF protocol to support TTZ. This section describes a couple of them.

- One way is to use one bit to indicate that a link is a TTZ link in a router LSA.

- Another option is to have a TLV in a Router Information LSA containing TTZ ID and flags to indicate that the router is a TTZ edge router or a TTZ internal router.

5.1. One Bit to Indicate an Internal TTZ Link

A router LSA contains the description of a number of router links. The existing format of a router LSA is illustrated as follows:

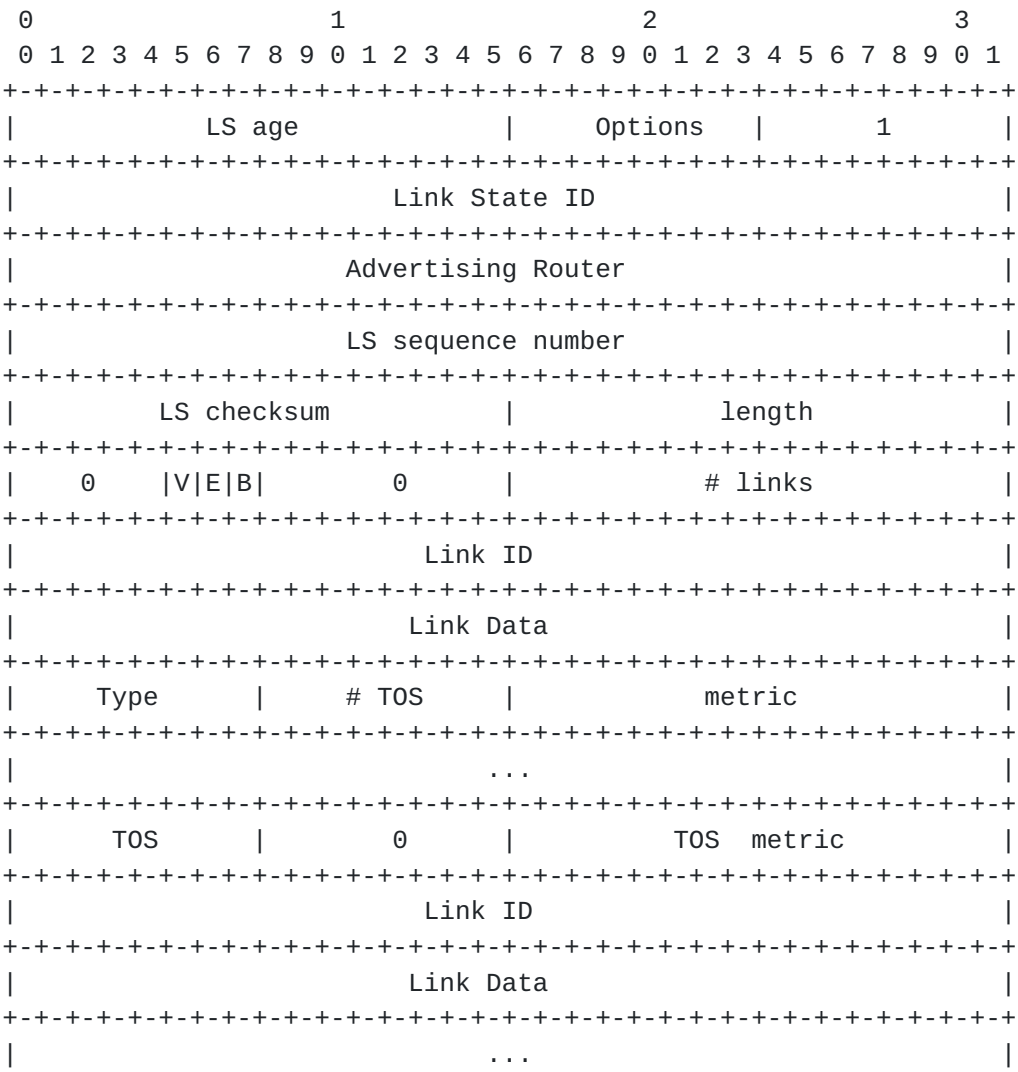
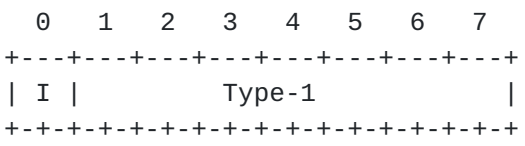


Figure 2: Format of Router LSA

For a router link, the value of an eight bit Type field indicates the kind of the link. The value of the Type field may be 1, 2, 3 or 4,

which indicates that the kind of the link is a point-to-point connection to another router, a connection to a transit network, a connection to a stub network, or a virtual link respectively.

The existing eight bit Type field for a router link may be split into two fields as follows:



I bit flag:

1: This indicates that the router link is an internal link to a router inside the TTZ.

0: This indicates that the router link is an external link.

Type-1:

The kind of the link.

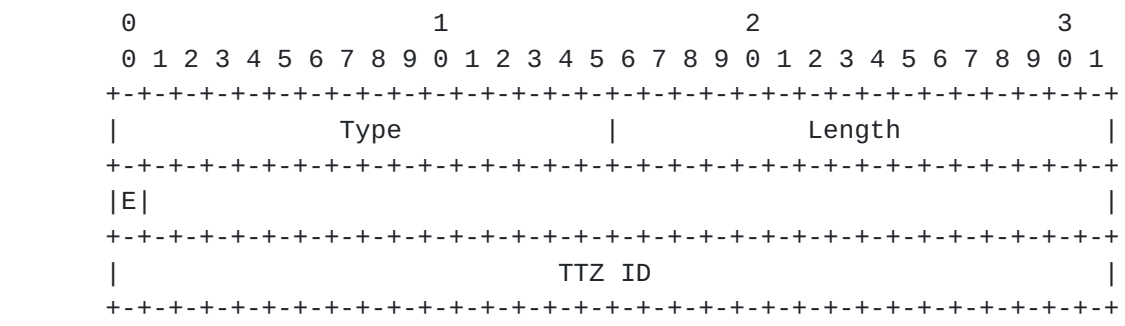
Figure 3: Bit to Indicate Internal TTZ Link

For a link inside a TTZ, the value of I bit flag is set to one, indicating that this link is an internal TTZ link. For a link connecting to a router outside of a TTZ from a TTZ edge router, the value of I bit flag is set to zero, indicating that this link is an external TTZ link.

The value of Type-1 field may have value 1, 2, 3, or 4, which indicates that the kind of a link being described is a point-to-point connection to another router, a connection to a transit network, a connection to a stub network, or a virtual link respectively.

5.2. A TTZ TLV in Router Information LSA

A new TLV is proposed in Router Information LSA for TTZ in the figure below to indicate a router's TTZ capability and the TTZ to which the router belongs.



Type A 16-bit field set to a number to be determined by IANA.

Length A 16-bit field indicating the length of the value, which is 8.

Value E bit set to 1 indicating a TTZ edge router
0 indicating a TTZ internal router

TTZ ID gives the TTZ to which the router belongs

Figure 4: TTZ TLV in Router Information LSA

The TTZ TLV follows the Router Informational Capabilities TLV in a Router Information LSA. Every edge router of a TTZ generates a Router Information LSA containing a TTZ TLV with E bit set to 1 and the ID of the TTZ. Each internal router of the TTZ originates a Router Information LSA containing a TTZ TLV with E bit set to 0 and the ID of the TTZ.

A TTZ is defined by all the routers with the same TTZ ID and all the TTZ links. For a TTZ edge router, its links connected to other TTZ routers belong to the TTZ. For a TTZ internal router, all its links belong to the TTZ.

6. Constructing Router LSA

Two types of router LSAs are generated by an edge router of a TTZ. The first type describes the links connecting to it; in fact, this LSA is the "normal" LSA that would be constructed if the TTZ feature is not used. This LSA is also generated by a router inside the TTZ. The second is generated to virtualize the TTZ as a group of edge routers connected.

The first type of LSA comprises both the router links connecting the

routers inside the TTZ and the router links connecting to the routers outside of the TTZ. For each of the router links in the LSA, it can be represented in one of the ways described in the previous section.

The second router LSA generated by an edge router of the TTZ comprises two groups of links in general.

The first group of links are the router links connecting the routers outside of the TTZ from this TTZ edge router. These router links are normal router links. There is a router link for every adjacency between this TTZ edge router and a router outside of the TTZ.

The second group of links are the "virtual" router links. For each of the other TTZ edge routers, there is a "virtual" router link to it from this TTZ edge router. The cost of the router link from this TTZ router to one of the other TTZ edge routers may be the cost of the shortest path from this TTZ edge router to it.

In addition, the LSA may contain a third group of links, which are stub links for other destinations inside the TTZ.

7. Establishing Adjacencies

A router in a TTZ forms an adjacency with another router in the TTZ in the same way as a normal router when these two routers have a connection.

For an edge router in a TTZ, it also forms an adjacency with any router outside of the TTZ that has a connection with the edge router.

When the edge router synchronizes its link state database with the router outside of the TTZ, it sends the router outside of the TTZ the information about all the LSAs except for the LSAs belonging to the TTZ that are hidden from any router outside of the TTZ.

At the end of the link state database synchronization, the edge router originates its own router LSA for virtualizing the TTZ and sends this LSA to the router outside of the TTZ.

From the point of view of the router outside of the TTZ, it sees the other end as a normal router and forms the adjacency in the same way as a normal router. It is not aware of anything about its neighboring TTZ. From the LSAs related to the TTZ edge router in the other end, it knows that the TTZ edge router is connected to each of the other TTZ edge routers and some routers outside of the TTZ.

8. Distribution of LSAs

LSAs can be divided into three classes according to their distributions. The first class of LSAs is distributed within a TTZ. The second is distributed through a TTZ. The third is distributed only to the edge routers of the TTZ.

8.1. Distribution of LSAs within TTZ

Any LSA about a link state in a TTZ is distributed within the TTZ. It will not be distributed to any router outside of the TTZ.

For example, any router LSA generated for a router in a TTZ is distributed within the TTZ. It will not be distributed to any router outside of the TTZ.

Any network LSA generated for a broadcast or NBMA network inside a TTZ is distributed within the TTZ. It will not be distributed to any router outside of the TTZ.

Any opaque LSA generated for a TTZ internal TE link is distributed within the TTZ. It will not be distributed to any router outside of the TTZ.

8.2. Distribution of LSAs through TTZ

Any LSA about a link state outside of a TTZ received by an edge router of the TTZ is distributed through the TTZ.

For example, when an edge router of a TTZ receives an LSA for a link state outside of the TTZ from a router outside of the TTZ, it floods it to its neighboring routers both inside the TTZ and outside of the TTZ. This LSA may be any LSA such as a router LSA and an opaque LSA that is distributed in a domain.

The routers in the TTZ continue to flood the LSA. When another edge router of the TTZ receives the LSA, it floods the LSA to its neighboring routers both outside of the TTZ and inside the TTZ.

8.3. Distribution of LSAs only to Edges of TTZ

In the case that a TTZ is virtualized as a group of edge routers of the TTZ connected, every edge router of the TTZ generates a router LSA for the TTZ. This LSA is distributed to the routers outside of the TTZ but not to the TTZ internal routers.

When an edge router of the TTZ receives a router LSA originated by another edge router of the TTZ for virtualizing the TTZ, it floods

the LSA to its neighboring routers outside of the TTZ but not to the TTZ internal routers.

9. Computation of Routing Table

The computation of the routing table on a router is the same as that described in RFC 2328, with one exception. An edge router of a TTZ MUST ignore the router LSAs generated by the edge routers of the TTZ for virtualizing the TTZ.

10. Smooth Migration to TTZ

This section describes the mechanisms which allow users to make a smooth migration to the TTZ with minimum interruption to the network.

For a group of routers and a number of links connecting the routers in an area, making them to work as a TTZ eventually with minimum interruption to the network may take a few of steps or stages.

At first, users configure the TTZ feature on every router in the TTZ. In this stage, the router has dual roles. One role is to function as a normal router. The other is to generate and distribute some TTZ information among the routers in the TTZ.

Secondly, users may allow every router in the TTZ to work as a TTZ router after they determine that every router in the TTZ is ready for transferring to work as a TTZ router eventually. For a router in the TTZ, users may allow it to work as a TTZ router after it has received all the necessary information from all the routers in the TTZ. This information may be displayed on a router through a CLI command.

And then users may activate the TTZ. There are a few of ways to activate the TTZ. One way is to activate it on a TTZ router through a CLI command such as activate TTZ directly. Another is through a TTZ activate timer, which activates the TTZ once the timer expires.

After a TTZ router is requested to activate the TTZ, it transfers to work as a TTZ router. When a TTZ router receives a LSA for virtualizing the TTZ and it is allowed to work as a TTZ router, it also transfers to work as a TTZ router. Thus, after every router in a TTZ is allowed to work as a TTZ router, activating the TTZ on one TTZ router will make every router in the TTZ transfer to work as a TTZ router.

For an edge router of the TTZ, transferring to work as a TTZ router comprises flushing its LSA originated for its link state as a normal

router as needed, generating a router LSA to virtualize the TTZ and flooding this LSA to all its neighboring routers except for the TTZ internal routers.

11. Security Considerations

The mechanism described in this document does not raise any new security issues for the OSPF protocols.

12. IANA Considerations

IANA is asked to assign a TLV in the OSPF Router Informational LSA, as described in Section 5.

13. Acknowledgement

The author would like to thank Acee Lindem, Dean Cheng, Lin Han and Yang Yu for their valuable comments on this draft.

14. References

14.1. Normative References

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

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC4970] Lindem, A., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 4970, July 2007.

[RFC2740] Coltun, R., Ferguson, D., and J. Moy, "OSPF for IPv6", RFC 2740, December 1999.

14.2. Informative References

[RFC5441] Vasseur, JP., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, April 2009.

[RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440,

March 2009.

Authors' Addresses

Huaimo Chen
Huawei Technologies
Boston, MA
USA

Email: huaimo.chen@huawei.com

Renwei Li
Huawei Technologies
2330 Central expressway
Santa Clara, CA
USA

Email: renwei.li@huawei.com

Gregory Cauchie
FRANCE

Email: greg.cauchie@gmail.com

Ning So
Tata Communications
2613 Fairbourne Cir.
Plano, TX 75082
USA

Email: ning.so@tatacommunications.com

Lei Liu
UC Davis
CA
USA

Email: liulei.kddi@gmail.com

Alvaro Retana
Cisco Systems, Inc.
7025 Kit Creek Rd.
Raleigh, NC 27709
USA

Email: aretana@cisco.com