Internet Engineering Task Force
Internet-Draft
Intended status: Experimental
Expires: September 22, 2016

H. Chen
R. Li
Huawei Technologies
G. Cauchie

A. Retana
Cisco Systems, Inc.
N. So
Tata Communications
F. Xu
Verizon
V. Liu
China Mobile
M. Toy
Comcast
L. Liu
Fujitsu

March 21, 2016

OSPF TE Topology-Transparent Zone
draft-chen-ospf-te-ttz-02.txt

Abstract

A topology-transparent zone is virtualized as the edges of the zone fully connected. This document proposes extensions to OSPF protocols to support Traffic Engineering (TE) topology-transparent zone.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

The number of routers in a network becomes larger and larger as the Internet traffic keeps growing. Through splitting the 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 a network into more areas is a very challenging and time consuming since it is involved in significant network architecture changes.

Secondly, the services carried by the network may be interrupted while the network is being split into more areas.

Furthermore, it is complex for a TE LSP crossing areas to be setup. In one option, a TE path crossing areas is computed by using collaborating PCEs [RFC5441] through PCEP[RFC5440], which is not easy to configure by operators since the manual configuration of the sequence of domains is required. Especially, the current PCE standard method may not guarantee that the path found is optimal.

Topology-transparent zone (TTZ) resolves these issues. This document proposes extensions to OSPF protocols to support Traffic Engineering (TE) topology-transparent 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. 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 fully 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.

The figure below illustrates an area containing a TTZ: TTZ 600.

```
                   TTZ 600
                   \
                    \ ^~^~^~^~^~^~^~^~^~^~^~^~
                     (                        )
    ===[R15]========(==[R61]------------[R63]==)======[R29]===
        ||         (   |    \          /    |   )       ||
        ||         (   |     \        /     |   )       ||
        ||         (   |      \      /      |   )       ||
        ||         (   |    ___\    /       |   )       ||
        ||         (   |   /   [R71]        |   )       ||
        ||         (   | [R73] /    \       |   )       ||
        ||         (   |      /      \      |   )       ||
        ||         (   |     /        \     |   )       ||
        ||         (   |    /          \    |   )       ||
    ===[R17]========(==[R65]------------[R67]==)======[R31]===
         \\          (//                    \\)       //
          ||         //v~v~v~v~v~v~v~v~v~v~v~\\      ||
          ||        //                        \\     ||
          ||       //                          \\    ||
           \\     //                            \\  //
       ======[R23]==============================[R25]=====
             //                                     \\
            //                                       \\
```

Figure 1: An Example of TTZ

The area comprises routers R15, R17, R23, R25, R29 and R31. It also contains TTZ 600, which comprises routers R61, R63, R65, R67, R71 and R73, and the links connecting them.

There are two types of routers in a TTZ: TTZ internal routers and TTZ edge routers. A TTZ internal router is a router inside the TTZ and its adjacent routers are in 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 hides the information inside the TTZ from the outside. It does not directly distribute any internal information about the TTZ to a router outside of the TTZ.

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

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. The cost of the "link" from one edge router to another edge router is the cost of the shortest path between these two routers. The bandwidth of the "link" is the maximum bandwidth of a path between the two routers.

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.

4. Extensions to OSPF Protocols

There are a couple of ways in which OSPF protocols are extened to support OSPF TE TTZ. One way is to add a TTZ ID TLV into an existing TE LSA. Another way is to put the contents of an existing TE LSA into another opaque LSA with a TTZ ID TLV.

4.1. Add TTZ ID TLV into Existing TE LSA

A TTZ ID TLV below, which is defined in OSPF TTZ, is added into an existing OSPF TE LSA.

```
       0                   1                   2                   3
       0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |       TTZ-ID-TLV-type         |          Length = 4           |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                            TTZ ID                             |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

4.1.1. Updating TE LSAs for a TTZ Router

When a TTZ router receives a CLI command triggering TTZ information distribution for migration or an LSA containing a TTZ Options TLV with T = 1, it knows that distributing TTZ information is started.

A TTZ router adds a TTZ ID TLV into each of its existing TE LSAs and floods the LSAs to its neighbors after distributing TTZ information starts.

When a router inside the TTZ receives a TE LSA containing a TTZ ID TLV from a neighboring router in the TTZ, it stores the TE link state for the TE TTZ and floods the link state to the other neighboring routers.

4.1.2. Originating TE LSAs for a TTZ Edge Router

When a TTZ router receives a CLI command activating migration to TTZ or an LSA containing a TTZ Options TLV with M = 1, it knows that the migration to TTZ is initiated.

A TTZ edge router originates a TE LSA for a P2P TE "link" to each of the other TTZ edge routers after the migration to TTZ starts. The metric of the link is the metric of the shortest path between two edge routers within the TTZ. The bandwidth of the link is the maximum bandwidth of a path between the two TTZ edge routers within the TTZ.

The edge router of a TTZ does not distribute any TE LSA with a TTZ ID TLV containing the ID of the TTZ to a router outside of the TTZ after the migration to TTZ starts.

4.2. Put an Existing TE LSA in Another LSA

The TE LSAs about a TTZ describes the TE TTZ topology. These LSAs can be contained and distributed in opaque LSAs within the TTZ. These opaque LSAs are called TTZ opaque LSAs or TTZ LSAs for short.

The following is a general form of a TTZ LSA, which is defined in

OSPF TTZ. It has an LS type = 10 and TTZ-LSA-Type, and contains a number of TLVs.

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |            LS age             |     Options   | LS Type = 10  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |  TTZ-LSA-Type |                     Opaque ID                 |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                      Advertising Router                       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                      LS Sequence Number                       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |         LS checksum           |             Length            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                                                               |
   ~                              TLVs                             ~
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where a new value (TTZ-TE-LSA-Type) for TTZ TE LSA is introduced for TTZ-LSA-Type.

4.2.1. Originating TTZ TE LSAs for a TTZ Router

After distributing TTZ information starts, a router in a TTZ originates a TTZ TE LSA for each of its existing TE LSAs and floods the TTZ TE LSA to its neighbors. For an existing TE LSA, the TTZ TE LSA contains the TLVs in the existing TE LSA and a TTZ ID TLV with the ID of the TTZ.

When a router inside the TTZ receives a TTZ TE LSA from a neighboring router in the TTZ, it stores the TE link state for the TTZ and floods the link state to the other neighboring routers.

4.2.2. Originating TE LSAs for a TTZ Edge Router

A TTZ edge router originates a TE LSA for a P2P TE "link" to each of the other TTZ edge routers after the migration to TTZ starts. The metric of the link is the metric of the shortest path between two edge routers within the TTZ. The bandwidth of the link is the maximum bandwidth of a path between the two TTZ edge routers within the TTZ.

The edge router of a TTZ does not distribute any TTZ TE LSA to a router outside of the TTZ after the migration to TTZ starts.

4.2.3. Flushing Out TE LSAs for a TTZ Router

A TTZ router SHOULD flush out its existing TE LSAs after their corresponding TTZ TE LSAs are originated and the migration to TTZ is done for a short given time such as one minute.

4.3. Comparison of Two Ways

The first way seems simple, in which the existing TE LSAs are used. In addition, it may use less memory. However, it is hard to flush out the TE LSAs with TTZ ID TLVs after migration to TTZ.

The second way uses two separated sets of LSAs. One set is for the normal TE topology of a TTZ; the other set is for the TTZ TE topology of the TTZ. Thus it is easy to flush out the normal TE LSAs of the TTZ after migration to TTZ. Moreover, it is cleaner.

It seems that the second way is preferred since it is cleaner.

5. Computation of TE Path

The computation of a TE path on a router outside of a TTZ is the same as before. On a router in a TTZ, the computation of a TE path has the same procedure flow as before, with one exception. A router in a TTZ MUST ignore the TE links in the TE LSAs generated by the edge routers of the TTZ for virtualizing the TE TTZ.

A TE path on a router inside the TTZ is computed through using the TE link state database (LSDB) containing the TE topology of the TTZ and the TE topology outside of the TTZ.

6. Summarizing TE Information in TTZ

The Traffic Engineering (TE) information about a TTZ may be summarized to the outside of the TTZ as the edges of the TTZ fully connected. The TE link (virtual) between two edges may have the maximum bandwidth of a path between them. The procedure below illustrates a way to find the bandwidth for the TE link.

```
L1: candidate-list = {{root, MaxBW}}; result-tree = { }.
    Where for edge Ei, root = Ei; MaxBW is a maximum number.

L2: While candidate-list != { } do

L3:    Select node with maximum bandwidth from candidate-list
       as working node k;
       remove it from candidate-list; add it into result-tree.

L4:    Suppose that BWk is the bandwidth of working node k
       (i.e., BWk is the maximum bandwidth from root to node k).
       For each node x connected to node k and not in result-tree,
        find the bandwidth BWx of node x as follows:
        BWx = min{BWk, BWk-x}, where
          BWk-x is the bandwidth of the link from node k to node x.
        If node x is not in candidate-list,
        then add {x, BWx} into candidate-list;
        otherwise (i.e., {x, BWx0} is in candidate-list),
        if BWx > BWx0,
        then replace {x, BWx0} in candidate-list with {x, BWx}.

L5: end-while

L6: Maximum bandwidth from Ei to every other edge node Ej is found.
```

Figure 2: Find Bandwidth of TE Link between Two Edges

Note that we should have solutions for summarizing SRLGs and link colors for a TTZ, which are challenging.

7. Security Considerations

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

8. IANA Considerations

TBD

9. Contributors

Veerendranatha Reddy Vallem
Huawei Technologies
Banglore
India
Email: veerendranatharv@huawei.com

William McCall
cisco Systems, Inc.
Bellevue, WA
USA
wimccall@cisco.com

Anil Kumar S N
Huawei Technologies
Banglore
India
Email: anil.sn@huawei.com

10. Acknowledgement

The author would like to thank Hannes Gredler, Igor Bryskin, Acee Lindem, Abhay Roy, Wenhu Lu, and Dean Cheng for their valuable comments.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <http://www.rfc-editor.org/info/rfc2328>.

[RFC2370] Coltun, R., "The OSPF Opaque LSA Option", RFC 2370, DOI 10.17487/RFC2370, July 1998, <http://www.rfc-editor.org/info/rfc2370>.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <http://www.rfc-editor.org/info/rfc3630>.

11.2. Informative References

[RFC5441] Vasseur, JP., Ed., 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, DOI 10.17487/RFC5441, April 2009, <http://www.rfc-editor.org/info/rfc5441>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <http://www.rfc-editor.org/info/rfc5440>.

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

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

Email: aretana@cisco.com

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

Email: ning.so@tatacommunications.com

Fengman Xu
Verizon
2400 N. Glenville Dr
Richardson, TX 75082
USA

Email: fengman.xu@verizon.com

Vic Liu
China Mobile
No.32 Xuanwumen West Street, Xicheng District
Beijing,   100053
China

Email: liuzhiheng@chinamobile.com

Mehmet Toy
Comcast
1800 Bishops Gate Blvd.
Mount Laurel, NJ 08054
USA

Email: mehmet_toy@cable.comcast.com

Lei Liu
Fujitsu
USA

Email: lliu@us.fujitsu.com