Internet Engineering Task Force H. Chen
Internet-Draft Huawei Technologies
Intended status: Standards Track M. Toy
Expires: January 4, 2016 Comcast
V. Liu
China Mobile
L. Liu
Fijitsu
July 3, 2015

Extensions to OSPF for Temporal LSP
draft-chen-ospf-tts-00.txt

Abstract

This document specifies extensions to OSPF for distributing Traffic Engineering (TE) information on a link in a sequence of time intervals.

Table of Contents

1. Introduction

Once an existing multiprotocol label switching (MPLS) traffic engineering (TE) label switched path (LSP) is set up, it is assumed to carry traffic forever until it is down. When an MPLS TE LSP tunnel is up, it is assumed that the LSP consumes its reserved network resources forever even though the LSP may only use network resources during some period of time. As a result, the network resources are not used efficiently. Moreover, a tunnel service can not be reserved or booked in advance for a period of time.

This document specifies extensions to OSPF for supporting the setup of an MPLS TE LSP in a period of time called a time interval or a sequence of time intervals. It is assumed that the LSP carries traffic during this time interval or each of these time intervals. Thus the network resources are efficiently used. More importantly, some new services can be provided. For example, a consumer can book a tunnel service in advance for a time interval or a sequence of time intervals. Tunnel services may be scheduled.

2. Terminology

A Time Interval: a time period from time Ta to time Tb.

LSP: Label Switched Path. An LSP is a P2P (point-to-point) LSP or a P2MP (point-to-multipoiint) LSP.

LSP in a time interval: LSP that carries traffic in the time interval.

LSP in a sequence of time intervals: LSP that carries traffic in each of the time intervals.

Temporal LSP: LSP in a time interval or LSP in a sequence of time intervals.

TEDB: Traffic Engineering Database.

This document uses terminologies defined in RFC2328 and RFC3630.

3. 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.

4. Representation of TE Information

The existing Open Shortest Path First (OSPF) Traffic Engineering (TE) distributes an unreserved bandwidth Bi at each of eight priority levels for a link at one point of time, for example, at the current time.

```
  Bandwidth
    ^
    |
  Bi|_____________________________________________________
    |
    |
   -+-------------------------------------------------------> Time
    |
```

This means that the link has bandwidth Bi at a priority level from now to forever until there is a change to it. This TE information on the link is stored in TEDB.

Thus, a temporal LSP (i.e., an LSP in a time interval) cannot be set up using the information in the TEDB and the bandwidth cannot be reserved in advance for the LSP in the time interval.

To support temporal LSPs, we should extend OSPF to distribute TE information on a link in a series of time intervals.

4.1. TE Information in Absolute Time

Suppose that the amount of the unreserved bandwidth at a priority level on a link is Bj in a time interval from time Tj to Tk (k = j+1), where j = 0, 1, 2, ... The unreserved bandwidth on the link can be represented as

    [T0, B0], [T1, B1], [T2, B2], [T3, B3], ...

This is an absolute time representation of bandwidths on a link. Time Tj (j = 0, 1, 2, ...) MUST be a synchronized time among all network nodes.


```
     Bandwidth
       ^
       |                                    B3
       |          B1                        -----------
       |          ----------
       |B0                                             B4
       |----------         B2                          ---------
       |                   ----------------
       |
      -+-------------------------------------------------------> Time
       |T0        T1        T2              T3         T4
```

If an LSP is deleted or down at time T and uses bandwidth B, then for every time interval/period (after time T) during which bandwidth B is reserved for the LSP on a link attached to a network node, the network node adds B to the link for that interval/period.

If an LSP is set up at time T and uses bandwidth B, then for every time interval/period (after time T) during which bandwidth B is reserved for the LSP on a link attached to a network node, the network node subtracts bandwidth B from the link for that interval/

period.

If there are significant changes to the bandwidths on a link attached to a network node, the network node distributes the bandwidths on the link to other network nodes. That is that a updated [T0, B0], [T1, B1], [T2, B2], [T3, B3], etc., are distributed to other network nodes in the network. Each of the other network nodes can construct or determine the bandwidth for a series of time intervals/periods on a link after receiving the information.

4.2. TE Information in Relative Time

Alternatively, a relative time representation of bandwidths on a link can be used. For example, the amount of the unreserved bandwidth at a priority level on a link is Bj during a series of time intervals/ periods can be expressed as

```
[P0, B0], [P1, B1], [P2, B2], [P3, B3], ..., where
Pj = Tk - Tj, k = (j+1) and j = 0, 1, 2, 3, ....
```

In this representation, every time Tj (j = 0, 1, 2, ...) can be a local time. A timer may expire after every unit of time (e.g., every second) and trigger --P0, which decrements P0. When P0 = 0, P1 becomes P0, P2 becomes p1, and so on.

If there are significant changes to the bandwidths on a link attached to a node, the node distributes the bandwidths on the link to other nodes. That is that a updated [P0, B0], [P1, B1], [P2, B2], [P3, B3], ..., are distributed to other network nodes in the network. On each of the other network nodes, a timer may expire for every unit of time (e.g., every second) and trigger --P0, which decrements P0. When P0 = 0, P1 becomes P0, P2 becomes p1, and so on.

An advantage of using relative time representation is that the times or clocks on all the network nodes can be different.

5. Extensions to OSPF

This section describes the extensions to OSPF for supporting the

setup of temporal LSPs.

5.1. TE LSA

An opaque LSA of type 10 is originated by a network node to distribute TE information such as the bandwidth of a link that is attached to the network node.

```
 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   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       1       |                  Opaque ID                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Advertising Router                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                              TLVs                             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The opaque LSA comprises a link-state (LS) age field, an options field, an LS type field, an opaque identifier (ID) field, an advertising router field, an LS sequence number field, an LS checksum field, a length field, and one or more TLVs.

The LS age field indicates the time since the LSA was originated in seconds. The options field indicates the optional capabilities

supported by the described portion of the routing domain. The LS type field indicates the type of the LSA. The opaque ID field is a number used to maintain multiple opaque LSAs. The advertising router field indicates the Router ID of the router that originated the LSA. The LS sequence number field is used to detect old or duplicate LSAs. Successive instances of an LSA are given successive LS sequence numbers. The LS checksum field indicates the Fletcher checksum of the complete contents of the LSA, including the LSA header but

excluding the LS age field. The length field indicates the length of the LSA in bytes.

5.2. TTS Link TLV

In addition to existing router address TLV and link TLV, TLVs fields may comprise a new temporal tunnel service (TTS) link TLV.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Type (5)           |       Length (variable)       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Reserved (0)         |         Segment-Number        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            sub TLVs                           |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The TTS link TLV comprises a type field, a length field, a reserved field, a segment-number field, and a sub TLVs field.

The type field may comprise a value assigned by the Internet Assigned Number Authority (IANA) to indicate that the TLV is a TTS link TLV. For example, the type field may comprise a value of 5. The length field may indicate the length of the values in the TTS link TLV in bytes.

The segment-number indicates a segment of the TTS link TLV. The information on a link may be too big to fit into one TTS link TLV. In this case, the information may be split into a few of segments, each of which is put into a TTS link TLV and identified by a segment number.

The sub TLV field comprises a link type sub-TLV and a link ID sub-TLV. It may further comprise a local address sub-TLV, a remote address sub-TLV, a TE metric sub-TLV, a maximum bandwidth sub-TLV, a maximum reservable bandwidth sub-TLV, an unreserved bandwidth sub-TLV, an administrator group sub-TLV, a relative TTS unreserved

bandwidth sub-TLV, an absolute TTS unreserved bandwidth sub-TLV, and

any other suitable sub-TLVs.

The format of an absolute TTS unreserved bandwidth sub-TLV is shown below.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Type (21)          |      Length (36+36*n)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              T0                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             B0[8]                             |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              T1                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             B1[8]                             |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                                                               ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              Tn                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Bn[8]                             |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

It comprises a type field, a length field, absolute time fields, and unreserved bandwidth fields.

The type field may comprise a value assigned by the IANA to indicate that the sub-TLV is an absolute TTS unreserved bandwidth sub-TLV. For example, the type field may comprise a value of 21.

The length field may indicate the length of the values in the absolute TTS unreserved bandwidth sub-TLV in bytes.

The absolute time fields and the unreserved bandwidth fields may be in pairs such as

    [T0, B0[8]], [T1, B1[8]], ..., [Tn, Bn[8]],

where T0, T1, ..., Tn are the times synchronized among all the nodes and Bj[8] (j=0, 1,..., n) are the amount of unreserved bandwidth at eight priority levels in the time interval/period from Tj to Tk (k=j+1).

The format of a relative TTS unreserved bandwidth sub-TLV is illustrated as follows.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type (22)         |      Length (36+36*n)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              P0                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             B0[8]                             |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              P1                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             B1[8]                             |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                                                               ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              Pn                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Bn[8]                             |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

It comprises a type field, a length field, relative time fields, and unreserved bandwidth fields.

The type field may comprise a value assigned by the IANA to indicate that the sub-TLV is a relative TTS unreserved bandwidth sub-TLV. For example, the type field may comprise a value of 22.

The length field may indicate the length of the values in the relative TTS unreserved bandwidth sub-TLV in bytes.

The relative time fields and the unreserved bandwidth fields may be in pairs such as

[P0, B0[8]], [P1, B1[8]], ... , [Pn, Bn[8]],

where Pj (j=0, 1,..., n) is the time period during which the unreserved bandwidth is Bj[8], containing the amount of unreserved bandwidth at eight priority levels.

6. Security Considerations

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

7. IANA Considerations

This section specifies requests for IANA allocation.

8. Acknowledgement

The author would like to thank people for their valuable comments on this draft.

9. References

9.1. 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.

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, July 2008.

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

9.2. Informative References

[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.

Authors' Addresses

Huaimo Chen
Huawei Technologies
Boston, MA
US

Email: huaimo.chen@huawei.com

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

Email: mehmet_toy@cable.comcast.com

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

Email: liuzhiheng@chinamobile.com

Lei Liu
Fijitsu
USA

Email: lliu@us.fujitsu.com