Network Working Group P. Srisuresh
INTERNET-DRAFT Kuokoa Networks
Expires as of March 16, 2003 P. Joseph
Force10 Networks
September 16, 2002

TE LSAs to extend OSPF for Traffic Engineering
<draft-srisuresh-ospf-te-03.txt>

Status of this Memo

Abstract

OSPF is a link state routing protocol used for IP-network topology discovery and collection and dissemination of link access metrics. The resulting Link State Database (LSDB) is used to compute IP address forwarding table based on shortest-path criteria. Traffic Engineering extensions(OSPF-TE) outlined in this document are built on the native OSPF foundation, utilizing new LSAs, designed specifically for TE. OSPF-TE sets out to discover TE network topology and perform collection and dissemination of TE metrics within the TE network. This results in the generation of an independent TE-LSDB, that would permit computation of TE circuit paths. Unlike the native OSPF link metrics, TE metrics can be rapidly changing and varied across different elements of the network. TE circuit

paths are computed using varied TE criteria, often different from the shortest-path, to route traffic around congestion paths. Principal motivations to designing the OSPF-TE over [OPQLSA-TE] and transition path for vendors currently using [OPQLSA-TE] to adapt the OSPF-TE are outlined in separate sections within the document. OSPF-TE provides a single unified mechanism for traffic engineering across packet and non-packet networks, and may be adapted for a peer networking model.

Table of Contents

1. Introduction

There is substantial industry experience with deploying OSPF link state routing protocol. That makes OSPF a good candidate to adapt for traffic engineering purposes. The dynamic discovery of network topology, link access metrics, flooding algorithm and the hierarchical organization of areas can all be used effectively in creating and tearing traffic links on demand. The intent of OSPF-TE is to discover TE network topology and the TE metrics of the nodes and links in the network.

The objective of traffic engineering is to set up circuit path(s) across a pair of nodes or links, as the case may be, so as to forward traffic of a certain forwarding equivalency class. Circuit emulation in a packet network is accomplished by each MPLS intermediary node performing label swapping. Whereas, circuit emulation in a TDM or Fiber cross-connect network is accomplished by configuring the switch fabric in each intermediary node to do the appropriate switching (TDM, fiber or Lamda) for the duration of the circuit.

The objective of this document is not how to set up traffic circuits, but rather provide the necessary TE parameters for the nodes and links that constitute the TE topology. Unlike the native OSPF, OSPF-TE will be used to build circuit paths, meeting certain TE criteria. The only requirement is that end-nodes and/or end-links of a circuit be identifiable with an IP address.

The approach suggested in this document is different from the Opaque-LSA-based approach outlined in [OPQLSA-TE]. Section 4 describes the motivations behind designing OSPF-TE. Section 6 outlines a strategy to transition Opaque-LSA based implementations to adapt the OSPF-TE outlined here.

2. Traffic engineering overview

A traffic engineered circuit may be identified by the tuple of (Forwarding Equivalency Class, TE parameters for the circuit, Origin Node/Link, Destination node/Link).

The Forwarding Equivalency Class(FEC) may be constituted of a number of criteria such as (a) Traffic arriving on a specific interface, (b) Traffic meeting a certain classification criteria (ex: based on fields in the IP and transport headers), (c) Traffic in a certain priority class, (d) Traffic arriving on a specific set of TDM (STS) circuits on an interface, (e) Traffic arriving on a certain wave-length of an interface, (f) Traffic arriving at a certain time of day, and so on. A FEC may be constituted as a combination of one or more of the above criteria. Discerning traffic based on the FEC criteria is a mandatory requirement on Label Edge Routers (LERs). Traffic content is transparent to the Intermediate Label Switched Routers (LSRs), once a circuit is formed. LSRs are simply responsible for keeping the circuit in-tact for the lifetime of the circuit(s). As such, this document will not address FEC or the associated signaling to setup circuits. [MPLS-TE] and [GMPLS-TE] address the FEC criteria. Whereas, [RSVP-TE] and [CR-LDP] address different types of signaling protocols.

This document is concerned with the collection of TE parameters for all the nodes and links within an autonomous system. TE parameters for a node may include a) ability to perform traffic prioritization, b) ability to provision bandwidth on interfaces, c) support for zero or more CSPF algorithms, d) support for a specific TE-Circuit switch type, e) support for a certain type of automatic protection switching and so forth. TE parameters for a link may include a) available bandwidth, b) reliability of the link, c) color assigned to the link, d) cost of bandwidth usage on the link, and e) membership to a Shared Risk Link Group (SRLG) and so forth.

Only the unicast paths circuit paths are considered here. Multicast variations are currently considered out of scope for this document. The requirement is that the originating as well as the terminating entities of a TE path are identifiable by their IP address.

3. Terminology

Definitions for majority of the terms used in this document with regard to OSPF protocol may be found in [OSPF-V2]. MPLS and traffic engineering terms may be found in [MPLS-ARCH]. RSVP-TE and CR-LDP

signaling specific terms may be found in [RSVP-TE] and [CR-LDP] respectively.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALLNOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

Below are definitions for the terms used within this document.

3.1. OSPF-TE node

This is a router that supports the OSPF-TE described in this document. At least one of the attached links for the node supports IP packet termination and runs the OSPF-TE protocol.

An OSPF-TE node supports native OSPF as well as the OSPF-TE.

3.2. Native OSPF node

A native OSPF node is an OSPF router that does not support the TE extensions described in this document or does not have a TE link attached to it. A Native OSPF node forwards IP traffic, using the shortest-path forwarding algorithm.

A native OSPF node may be enhanced to be an OSPF-TE node. An autonomous system (AS) could be constituted of a combination

of native-OSPF and OSPF-TE nodes.

3.3. TE nodes vs. native(non-TE) nodes

A TE-Node is an intermediate or edge node taking part in the traffic engineered (TE) network. A TE-circuit is constituted of a series of TE nodes connected to each other through TE links. In a SONET/TDM network or a photonic cross-connect network, a TE node is not required to support OSPF-TE. An external OSPF-TE node may represent the TE node for protocol processing.

A native (or non-TE) node is an IP router capable of IP packet forwarding, does not have TE link attachments and does not take part in a TE network.

3.4. TE links vs. native(non-TE) links

A TE Link is a network attachment that supports traffic engineering. A TE-circuit is constituted of a series of TE nodes connected to each other through TE links.

A native (or non-TE) link is one that is used for IP packet traversal. A link may be configured to be pure TE link or native link or a both.

3.5. Packet-TE network vs. non-packet-TE network

Packet-TE network is one in which TE-circuit emulation is accomplished by each MPLS intermediary node performing label swapping on the packet data.

Non-packet-TE network, such as SONET/TDM or Fiber cross-connect network is one in which TE-circuit emulation is accomplished by configuring the switch fabric in each intermediary node to do the appropriate switching (TDM, fiber or Lamda) for the duration of the circuit.

In either case, OSPF-TE can only be enabled on interfaces supporting IP packet termination. Interfaces supporting OSPF and/or OSPF-TE constitute the OSPF control network. The OSPF control network can be independent of the packet or non-packet data network.

3.6. TE topology vs. non-TE topology

A TE topology is constituted of a set of contiguous TE nodes and TE links. Associated with each TE node and link is a set of TE criteria that may be supported at any given time. A TE topology

allows circuits to be overlayed statically or dynamically based on a specific TE criteria.

A non-TE topology specifically refers to the network that does not support TE. Control protocols such as OSPF may be run on the non-TE topology. IP forwarding table used to forward IP packets on this network is built based on the control protocol specific algorithm, such as OSPF shortest-path criteria.

3.7. TLV

A TLV stands for an object in the form of Tag-Length-Value. All TLVs are assumed to be of the following format, unless specified otherwise. The Tag and length are 16 bits wide each. The length includes the 4 bytes required for Tag and Length specification.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Tag               |     Length (4 or more)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Value ....                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          ....                                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

3.8. Router-TE TLVs

TLVs used to describe the TE capabilities of a TE-node.

3.9. Link-TE TLVs

TLVs used to describe the TE capabilities of a TE-link.

4. Motivations to designing the OSPF-TE using TE-LSAs

The motivation behind designing the OSPF-TE using TE-LSAs is that the approach is clean, extendible, scalable, unified, efficient, and SLA enforceable. The approach also provides the right framework for future OSPF extensibility. Each of these motivations is explained in detail in the following subsections.

The last subsection lists network scenarios that benefit from the TE-LSA design.

4.1. Clean design - TE-LSDB, independent of the native LSDB

OSPF-TE using TE LSAs provides a clean separation of Link State Data Bases (LSDB) between native (SPF-based) and TE networks. The OSPF-TE dynamically discovers TE network topology and the associated TE metrics of the nodes and links in the TE network. OSPF-TE design is based on the tried and tested OSPF paradigm. As such, it inherits all the benefits of the OSPF, present and future.

With OSPF-TE, native OSPF nodes will keep just the native OSPF link state database. The OSPF-TE nodes will keep the native as well as the TE LSDB. In the case, where the network is used only for Traffic engineering purposes, the native-LSDB describes the control topology.

In the Opaque-LSA-based TE scheme([OPQLSA-TE]), the TE-LSDB built using opaque LSAs refers the native LSDB to build the TE topology. Further, the LSDB has no knowledge of the TE capabilities of the routers. Point-to-point links are the only type of links that can form a TE network. It is apparent that [OPQLSA-TE] is a new protocol in itself within the constraints of an Opaque-LSA and is not tailored to benefit from the tried and tested native-OSPF.

4.2. Extendible design - based on the OSPF foundation

TE LSAs are extendible, just as the native OSPF on which OSPF-TE is founded. [OPQLSA-TE], on the other hand, is not extendible and is constrained by the Opaque LSA on which it is founded.

For example, Opaque LSAs are not suited to advertising summary LSAs along a restricted flooding scope (as with TE-Summary network LSA). Opaque LSAs are also not suited to advertising incremental TLV changes. A change in any TLV of a TE-link will mandate the entire Opaque-LSA (with all the TLVs included) to be transmitted. TE-incremental-link-update LSA, on the other hand, is capable of advertising just the delta TLVs. Opaque LSAs are also not extendible to support advertisement of TLVs for non-members of the OSPF control network. This is a necessity for certain non-packet networks such as a SONET/TDM network. In a SONET/TDM network, data-path topology often differs from its OSPF control network counterpart. TE-Router-Proxy-LSA (section 9.7) permits advertising LSAs for non-members via a proxy node that is a member of the control network.

Lastly, the expressibility of data in an Opaque LSA is strictly restricted to being in the form of TLVs and sub-TLVs, some mandatorily required and some positionally dependent in the TLV sequence for interpretation.

4.3. Scalable design - TE-AS may be sub-divided into areas

OSPF-TE using TE LSAs inherits the hierarchical area organization used within native-OSPF. Without revealing the nodes and characteristics of the attached links within a TE-area, the TE-Summary network LSA (refer section 9.4) advertises the reachability of TE networks within an area to the areas outside in the same AS.

Providing area level abstraction and having the abstraction be distinct for TE and native topologies is a necessity for inter-area communication. When the topologies are separate, the area border routers can advertise different summary LSAs to TE and non-TE routers. For example, a native Area Border router (ABR) simply announces the shortest path network summary LSAs (LSA type 3) for nodes outside the area. A TE-ABR, on the other hand, would use TE-summary network LSA to advertise network Reachability information - not aggregated path metric as required for a native OSPF LSDB. Clearly, the data content and flooding scope should be different for the TE nodes. The flooding boundary for TE-summary LSAs would be (AS - OriginatingArea - StubAreas - NSSAs).

Opaque-LSA-based TE scheme([OPQLSA-TE]) is restricted for use within an area and is not applicable for flooding across areas. As-wide scope Opaque LSAs (Type 11 LSAs) will be unable to restrict flooding in its own originating area.

4.4. Unified design - for packet and non-packet networks

OSPF-TE uses the same set of TE LSAs for disseminating TE characteristics - irrespective of whether the network is a packet network or a non-packet network or a combination of both. Only the TLVs used to describe the characteristics will vary. Each TE node will be required to advertise its own TE capabilities and that of the attached TE links.

In a peer networking TE model, the TE nodes are heterogeneous and have different TE characteristics. As such, the signaling protocols will need the TE characteristics of all nodes and attached links so they can signal the nodes to formulate TE circuits across heterogeneous nodes. The underlying control protocol must be capable of providing a unified LSDB for all nodes in the network. OSPF-TE clearly meets this requirement.

Opaque-LSA-based TE scheme([OPQLSA-TE]) is limited in scope for

packet networks. Extensions ([OPQLSA-GMPLS]) are underway to support GMPLS links within opaque LSAs. However, neither

[OPQLSA-TE] nor [OPQLSA-GMPLS] is sufficient by itself or when combined for use within a peer networking model with heterogeneous nodes. Neither of the Opaque LSA based extensions have provision to distinguish between the various nodes and link attachments that are different from point-to-point connections. SONET specific ring topologies and packet network specific LAN and other mesh topologies are not permitted.

4.5. Efficient design - in LSA content and flooding reach

OSPF-TE is capable of identifying the boundaries of a TE topology and limits the flooding of TE LSAs to only the TE-nodes. Nodes that do not have TE link attachments are not bombarded with TE specific LSAs. This is a useful characteristic for networks supporting native and TE traffic in the same connected network.

The more frequent and wider the flooding scope, the larger the number of retransmissions and acknowledgements. The same information (needed or not) may reach a router through multiple links. Even if the router did not forward the information past the node, it would still have to send acknowledgements across all the various links on which the LSAs tried to converge. Clearly, it is not desirable to flood LSAs to nodes that do not require it. This can be a considerable impediment to non-TE nodes in a network that is constituted of native and TE nodes.

Opaque-LSA-based TE scheme([OPQLSA-TE]) makes no distinction between TE and native OSPF nodes as far as LSA flooding is concerned. It is possible for the native OSPF nodes to silently ignore the unsupported Opaque LSAs or add knobs within implementation to decide whether or not a certain opaque LSA mandates dijkstra SPF recomputation. In any case, unintended LSAs are disruptive and can be the cause of a large number of acknowledgements and retransmissions.

TE metrics in a network could be rapidly changing. Only a subset of the metrics may be prone to rapid change, while others remain largely unchanged. Changes must be communicated at the earliest throughout the network to ensure that the TE-LSDB is up-to-date.

TE-Incremental-Link-update LSA (section 9.2) permits advertising only a subset of the link metrics and not the entire router-LSA all over. [OPQLSA-TE] does not have provision to advertise just the TLVs that changed. A change in any TLV of a TE-link will mandate the entire LSA to be transmitted. This is clearly a serious shortcoming in the protocol.

4.6. SLA enforceable TE network can coexist with IP network

OSPF-TE is designed to draw distinction between links that support TE traffic and links that support native best-effort IP traffic. This flexibility to configure links as appropriate for a service, permits enforceability of service level agreements (SLAs). A link, configured to support TE traffic alone will not permit native IP traffic on the link.

Best-effort IP transit network and constraint based TE network have different SLA requirements and hence different billing models. Keeping the two networks physically isolated will enable SLA enforceability, but can be expensive. Combining the two networks into a single physically connected network will bring economies of scale, if the SLA enforceability can be retained. When the links of a TE-network LSDB do not overlap the links of a native LSDB, such a virtual isolation of networks and hence SLA enforceability becomes possible.

Opaque-LSA-based TE scheme([OPQLSA-TE]) is inherently not capable of having two virtual networks in a single physically connected network. All point-to-point links in a packet network are subject to best-effort IP traffic, irrespective of whether a link is usable for TE traffic or not. In order to ensure that TE links are not cannibalized by best-effort traffic, the network provider will simply have to restrict best-effort traffic from entering the network. This is a severe limitation and is a direct result of not having LSDB isolation. When TE and native topologies are not separated (as is the case with Opaque-LSAs), a native OSPF node could be utilizing a TE link as its least cost link, thereby stressing the TE link and rendering the TE link ineffective for TE purposes.

4.7. Right Framework for future OSPF extensibility

OSPF-TE design provides the right framework for future OSPF extensions based on independent service provider needs. The framework essentially calls for building independent service specific LSDBs. Each such LSDB will consist of service specific metrics of all resources within the service-specific topology. The TE-LSDB permits TLV scalability as well as the ability to perform fast searches through the database. Just as the TE-LSDB may be used for MPLS TE application, a different type of LSDB may be used for a different type of application across the same physically connected IP network. E.g., one can derive QOS based IP forwarding on an IP network.

Limiting flooding scope of service specific LSAs within the service specific topology eliminates LSA contamination between virtual service networks of a single physically connected

network. Using service specific LSAs provides flexibility in LSA content and flooding scope.

Opaque-LSA-based TE scheme([OPQLSA-TE]) works best when a single type of service is assumed for a single physically connected network. As such, multiple disparate networks can function running various flavors of OSPF. [OSPF-v2] for native best-effort IP networks, [OPQLSA-TE] for packet networks and [OPQLSA-GMPLS] for non-packet networks.

4.8. Network scenarios benefiting from the OSPF-TE design

Many real-world scenarios are better served by the new-TE-LSAs scheme. Here are a few examples.

4.8.1. IP providers transitioning to TE services

Providers needing to support MPLS based TE in their IP network may choose to transition gradually. Perhaps, add new TE links or convert existing links into TE links within an area first and progressively advance to offer in the entire AS.

Not all routers will support TE extensions at the same time during the migration process. Use of TE specific LSAs and their flooding to OSPF-TE only nodes will allow the vendor to

introduce MPLS TE without destabilizing the existing network. As such, the native OSPF-LSDB will remain undisturbed while newer TE links are added to network.

4.8.2. Providers offering Best-effort-IP & TE services

Providers choosing to offer both best-effort-IP and TE based packet services simultaneously on the same physically connected network will benefit from the OSPF-TE design. By maintaining independent LSDBs for each type of service, TE links are not cannibalized by the non-TE routers for SPF forwarding. Unlike the [OPQLSA-TE] scheme, OSPF-TE provides the framework for SLA enforcement.

4.8.3. Multi-area networks

The OSPF-TE design parallels the tried and tested native-OSPF design. Unlike [OPQLSA-TE], OSPF-TE scales naturally to multi-area networks.

4.8.4. Non-packet and Peer-networking models

OSPF-TE is the only scheme that can support the following

network attachments For a non-Packet TE network.
   (a) "Positional-Ring" type network LSA and
   (b) Router Proxying - allowing a router to advertise on behalf
       of other nodes (that are not Packet/OSPF capable).

Opaque LSA based extensions ([OPQLSA-TE], [OPQLSA-GMPLS]) are not suited to distinguish the heterogeneous nodes in a peer-connected network. Opaque-LSA based extensions are also not suited to support link attachments that are different from point-to-point connections.

5. OSPF-TE solution, assumptions and limitations

5.1. OSPF-TE Solution

The OSPF-TE uses the options flag as a means to determine the TE topology. New TE LSAs are designed to generate an independent TE-service tailored LSDB. Sections 8.0 and 9.0 describe the TE extensions in detail. Changes required of the OSPF data

structures in order to support OSPF-TE are described in section 11.0. The OSPF-TE design is based on the tried and tested OSPF paradigm. With TE-LSDB, you have the advantages of retaining the scalability of TLV's and the ability to run fast searches through the database.

With the new TE-LSA scheme, an OSPF-TE node will have two types of Link state databases (LSDB). A native LSDB that describes the native control topology and a TE-LSDB that describes the TE topology. Shortest-Path-First algorithm will be used to forward IP packets along the native control network. OSPF neighbors data structure will be used for flooding along the control topology.

The TE-LSDB is constituted only of TE nodes and TE links. A variety of CSPF algorithms may be used to dynamically setup TE circuit paths along the TE network. A new TE-neighbors data structure will be used to flood TE LSAs along the TE-only topology. Clearly, the the TE nodes will need the control (non-TE) network for OSPF communication. The control network may also be used for pinging OSPF-TE nodes and performing any debug and monitoring tasks on the nodes. However, the ability to make distinction between TE and non-TE topologies, allows the bandwidth on TE links to be strictly SLA enforceable, even as a TE link is packet-capable. The actual characteristics of the TE-link are irrelevant from the OPSF-TE perspective. As such, that allows for packet and non-packet networks to operate in peer mode.

Consider the following network where some of the routers and links are TE enabled and others are native OSPF routers and links. All nodes in the network belong to the same OSPF area.


```
   +---+
   |   |--------------------------------------+
   |RT6|\\                                    |
   +---+  \\                                  |
    ||      \\                                |
    ||        \\                              |
    ||          \\                            |
    ||          +---+                         |
    ||          |   |----------------+        |
    ||          |RT1|\\              |        |
    ||          +---+  \\            |        |
```

```
                     ||          //|      \\          |        |
                     ||        //  |        \\        |        |
                     ||      //    |          \\      |        |
                    +---+  //      |            \\  +---+      |
                    |RT2|//        |              \\|RT3|------+
                    |   |----------|----------------|   |
                    +---+          |                +---+
                                   |                  |
                                   |                  |
                                   |                  |
                                 +---+              +---+
                                 |RT5|--------------|RT4|
                                 +---+              +---+
         Legend:
              --   Native(non-TE) network link
              |    Native(non-TE) network link
              \\   TE network link
              ||   TE network link
```

Figure 6: A (TE + native) OSPF network topology

In the above network, TE and native OSPF Link State Data bases (LSDB) would have been synchronized within the area along the following nodes.

Native OSPF LSDB nodes	TE-LSDB nodes
RT1, RT2, RT3. RT4, RT5, RT6	RT1, RT2, RT3, RT6

Nodes such as RT1 will have two LSDBs, a native LSDB and a TE-LSDB to reach native and TE networks. The TE LSA updates will not impact non-TE nodes RT4 and RT5.

5.2. Assumptions

OSPF-TE design makes the following assumptions.

1. An OSPF-TE node with links in an OSPF area will need to establish router adjacency with at least one other neighboring OSPF-TE node in order for the router's database to be synchronized with other routers in the area. Failing this, the

OSPF router will not be in the TE calculations of other TE routers in the area. Refer [FLOOD-OPT] for flooding optimizations.

2. Unlike the native OSPF, OSPF-TE must be capable of advertising link state of interfaces that are not capable of handling IP packet data. As such, the OSPF-TE protocol cannot be required to synchronize its link-state database with neighbors across all its links. It is sufficient to synchronize link-state database in an area, across a subset of the IP termination links. Yet, the TE LSDB (LSA database) should be synchronized across all OSPF-TE nodes within an area.

   All nodes and interfaces described by the TE LSAs will be present in the TE topology database (a.k.a. TE LSDB).

3. A link in a packet network can be a TE-link or a native-IP link or both. There may be different ways by which to use a link for TE and non-TE traffic. For example, a link may be used for both types of traffic and satisfy the TE SLA requirements, so long as the link is under-subscribed for TE (say, 50% of the link capacity is being used). Once the TE capacity requirement exceeds the set mark (say, the 50% mark), the link may be removed from the non-TE topology.

4. This document does not require any changes to the existing OSPF LSDB implementation. Rather, it suggests the use of another database, the TE-LSDB, comprised of the TE LSAs, for TE purposes.

5. As a general rule, all nodes and links that may be party to a TE circuit should be uniquely identifiable by an IP address. As for router ID, a separate loopback IP address for the router, independent of the links attached, is recommended.

6. The assumption about to be stated is principally meant for non-packet networks such as a SONET TDM network. In non-packet networks, each IP subnet on a TE-configurable network MUST have a minimum of one node with an interface running OSPF-TE protocol. For example, a SONET/SDH TDM ring must have a minimum of one node (say, a Gateway Network Element) running the OSPF protocol in order to enable TE configuration on all nodes within the ring.

   An OSPF-TE node may advertise more than itself and the links

it is directly attached to. It may also advertise other TE participants and their links, on their behalf.

5.3. Limitations

Below are the limitations of the OSPF-TE.

1. Disjoint TE topologies would have the same problem as an OSPF-TE node not forming adjacencies with any other node. The disjoint nodes will not be included in the TE topology of the rest of the TE routers. It will be the responsibility of the network administrator(s) to ensure connectedness of the TE network.

   For example, two routers that are physically connected to the same link (or broadcast network) need not be router adjacent via the Hello protocol, if the link is not IP packet terminated.

6. Transition strategy for implementations using Opaque LSAs

Below is a strategy to transition implementations using opaque LSAs to adapt the new TE LSA scheme in a gradual fashion.

1. Restrict the use of Opaque-LSAs to within an area.

2. Fold in the TE option flag to construct the TE and non-TE topologies in an area, even if the topologies cannot be used for flooding within the area.

3. Use TE-Summary LSAs and TE-AS-external-LSAs for inter-area Communication. Make use of the TE-topology within area to summarize the TE networks in the area and advertise the same to all TE-routers in the backbone. The TE-ABRs on the backbone area will in-turn advertise these summaries again within their connected areas.

7. The OSPF Options field

A new TE flag is introduced within the options field to identify TE extensions to the OSPF. This bit will be used to distinguish between routers that support Traffic engineering extensions and those that do not. The OSPF options field is present in OSPF Hello packets, Database Description packets and all link state advertisements. The TE bit, however, is a requirement only for

the Hello packets. Use of TE-bit is optional in Database

Description packets or LSAs.

Below is a description of the TE-Bit. Refer [OSPF-V2], [OSPF-NSSA] and [OPAQUE] for a description of the remaining bits in options field.

```
 --------------------------------------
 |TE | O | DC | EA | N/P | MC | E | * |
 --------------------------------------
 The OSPF options field - TE support
```

TE-Bit: This bit is set to indicate support for Traffic Engineering extensions to the OSPF. The Hello protocol which is used for establishing router adjacency and bidirectionality of the link will use the TE-bit to build adjacencies between two nodes that are either both TE-compliant or not. Two routers will not become TE-neighbors unless they agree on the state of the TE-bit. TE-compliant OSPF extensions are advertised only to the TE-compliant routers. All other routers may simply ignore the advertisements.

There is however a caveat with the above use of the last remaining reserved bit in the options field. OSPF v2 will have no more reserved bits left for future option extensions. If it is deemed necessary to leave this bit as is, we could use OPAQUE-9 LSA (Local scope) along each interface to communicate the support for OSPF-TE.

8. Bringing up TE adjacencies; TE vs. Non-TE topologies

OSPF creates adjacencies between neighboring routers for the purpose of exchanging routing information. In the following subsections, we describe the use of Hello protocol to establish TE capability compliance between neighboring routers of an area. Further, the capability is used as the basis to build a TE vs. non-TE network topology.

8.1. The Hello Protocol

The Hello Protocol is primarily responsible for dynamically establishing and maintaining neighbor adjacencies. In a TE network, it may not be required or possible for all links and neighbors to establish adjacency using this protocol.

The Hello protocol will use the TE-bit to establish Traffic Engineering capability (or not) between two OSPF routers.

For NBMA and broadcast networks, this protocol is responsible for electing the designated router and the backup designated router. For a TDM ring network, the designated and backup designated routers may either be preselected (ex: GNE, backup-GNE) or determined via the same Hello protocol. In any case, routers supporting the TE option shall be given a higher precedence for becoming a designated router over those that do not support TE.

8.2. Flooding and the Synchronization of Databases

In OSPF, adjacent routers within an area must synchronize their databases.   However, as observed in [FLOOD-OPT], the requirement may be stated more concisely that all routers in an area must converge on the same link state database. To do that, it suffices to send single copies of LSAs to the neighboring routers in an area, rather than send one copy on each of the connected interfaces. [FLOOD-OPT] describes in detail how to minimize flooding (Initial LSDB synchronization as well as the asynchronous LSA updates) within an area.

With the OSPF-TE described here, a TE-only network topology is constructed based on the TE option flag in the Hello packet. Subsequent to that, TE LSA flooding in an area is limited to TE-only routers in the area, and do not impact non-TE routers in the area. A network may be constituted of a combination of a TE topology and a non-TE (control) topology. Standard IP packet forwarding and routing protocols are possible along the control topology.

In the case where some of the neighbors are TE compliant and others are not, the designated router will exchange different sets of LSAs with its neighbors. TE LSAs are exchanged only with the TE neighbors. Native LSAs do not include description

for TE links. As such, native LSAs are exchanged with all neighbors (TE and non-TE alike) over a shared non-TE link.

Flooding optimization in a TE network is essential for two reasons. First, the control traffic for a TE network is likely to be much higher than that of a non-TE network. Flooding optimizations help to minimize the announcements and the associated retransmissions and acknowledgements on the network. Secondly, the TE nodes need to converge at the earliest to keep up with TE state changes occurring throughout the TE network.

This process of flooding along a TE topology cannot be folded into the Opaque-LSA based TE scheme ([OPQLSA-TE]), because Opaque LSAs (say, LSA #10) have a pre-determined flooding scope. Even as a TE topology is available from the use of

TE option flag, the TE topology is not usable for flooding unless a new TE LSA is devised, whose boundaries can be set to span the TE-only routers in an area.

NOTE, a new All-SPF-TE Multicast address may be used for the exchange of TE compliant database descriptors.

8.3. The Designated Router

The Designated Router is elected by the Hello Protocol on broadcast and NBMA networks. In general, when a router's interface to a network first becomes functional, it checks to see whether there is currently a Designated Router for the network. If there is, it accepts that Designated Router, regardless of its Router Priority, so long as the current designated router is TE compliant. Otherwise, the router itself becomes Designated Router if it has the highest Router Priority on the network and is TE compliant.

Clearly, TE-compliance must be implemented on the most robust routers only, as they become most likely candidates to take on additional role as designated router.

Alternatively, there can be two sets of designated routers, one for the TE compliant routers and another for the native OSPF routers (non-TE compliant).

8.4. The Backup Designated Router

The Backup Designated Router is also elected by the Hello Protocol. Each Hello Packet has a field that specifies the Backup Designated Router for the network. Once again, TE-compliance must be weighed in conjunction with router priority in determining the backup designated router. Alternatively, there can be two sets of backup designated routers, one for the TE compliant routers and another for the native OSPF routers (non-TE compliant).

8.5. The graph of adjacencies

An adjacency is bound to the network that the two routers have in common. If two routers have multiple networks in common, they may have multiple adjacencies between them. The adjacency may be split into two different types - Adjacency between TE-compliant routers and adjacency between non-TE compliant routers. A router may choose to support one or both types of adjacency.

Two graphs are possible, depending on whether a Designated Router is elected for the network. On physical point-to-point

networks, Point-to-MultiPoint networks and virtual links, neighboring routers become adjacent whenever they can communicate directly. The adjacency can only be one of (a) TE-compliant or (b) non-TE compliant. In contrast, on broadcast and NBMA networks the Designated Router and the Backup Designated Router may maintain two sets of adjacency. However, the remaining routers will participate in either TE-compliant adjacency or non-TE-compliant adjacency, but not both. In the Broadcast network below, you will notice that routers RT7 and RT3 are chosen as the designated and backup routers respectively. Within the network, Routers RT3, RT4 and RT7 are TE-compliant. RT5 and RT6 are not. So, you will notice the adjacency variation with RT4 vs. RT5 or RT6.

```
      +---+          +---+
      |RT1|----------|RT2|            o---------------o
      +---+    N1    +---+           RT1             RT2
```

```
                                                 RT7
                                                  o::::::::::
            +----+  +----+  +----+               /|:         :
            |RT7|   |RT3|   |RT4|               / | :        :
            +----+  +----+  +----+             /  |  :       :
              |       |       |               /   |   :      :
         +------------------------+       RT5o RT6o    oRT4 :
                  |       |     N2            *   *    ;     :
                +----+  +----+                 *  *   ;      :
                |RT5|   |RT6|                   * * ;        :
                +----+  +----+                   **;         :
                                                  o::::::::::
                                                 RT3
```

Figure 6: The graph of adjacencies with TE-compliant routers.

9. TE LSAs

The native OSPF protocol, as of now, has a total of 11 LSA types. LSA types 1 through 5 are defined in [OSPF-v2]. LSA types 6, 7 and 8 are defined in [MOSPF], [NSSA] and [BGP-OSPF] respectively. Lastly, LSA types 9 through 11 are defined in [OPAQUE].

Each of the LSA types have a unique flooding scope defined. Opaque LSA types 9 through 11 are general purpose LSAs, with flooding scope set to link-local, area-local and AS-wide (except

stub areas) respectively. As will become apparent from this document, the general purpose content format and the coarse flooding scope of Opaque LSAs are not suitable for disseminating TE data.

In the following subsections, we define new LSAs for Traffic engineering use. The Values for the new TE LSA types are assigned such that the high bit of the LS-type octet is set to 1. The new TE LSAs are largely modeled after the existing LSAs for content format and have a custom suited flooding scope. Flooding optimizations discussed in previous sections shall be used to disseminate TE LSAs along the TE-restricted topology.

A TE-router LSA is defined to advertise TE characteristics of the router and all the TE-links attached to the TE-router. TE-Link-Update LSA is defined to advertise individual link specific TE updates. Flooding scope for both these LSAs is the TE topology within the area to which the links belong. I.e., only those OSPF nodes within the area with TE links will receive these TE LSAs.

TE-Summary network and router LSAs are defined to advertise the reachability of area-specific TE networks and Area border routers(along with router TE characteristics) to external areas. Flooding Scope of the TE-Summary LSAs is the TE topology in the entire AS less the non-backbone area for which the the advertising router is an ABR. Just as with native OSPF summary LSAs, the TE-summary LSAs do not reveal the topological details of an area to external areas. But, the two summary LSAs do differ in some respects. The flooding scope of TE summary LSAs is different. As for content, TE summary network LSAs simply describe reachability without summarization of network access costs. And, unlike the native summary router LSA, TE-summary router LSA content includes TE capabilities of the advertising TE router.

TE-AS-external LSA and TE-Circuit-Path LSA are defined to advertise AS external network reachability and pre-established TE circuits respectively. While flooding scope for both these LSAs can be the TE-topology in the entire AS, flooding scope for the pre-established TE circuit LSA may optionally be restricted to just the TE topology within an area.

Lastly, the new TE LSAs are defined so as to permit peer operation of packet networks and non-packet networks alike. As such, a new TE-Router-Proxy LSA is defined to allow advertisement of a TE router, that is not OSPF capable, by an OSPF-TE node as a proxy.

9.1. TE-Router LSA (0x81)

The TE-router LSA (0x81) is modeled after the router LSA with the same flooding scope as the router-LSA, except that the scope is

restricted to TE-only nodes within the area. The TE-router LSA describes the TE metrics of the router as well as the TE-links attached to the router. Below is the format of the TE-router LSA. Unless specified explicitly otherwise, the fields carry the same meaning as they do in a router LSA. Only the differences are explained below. Router-TE flags, Router-TE TLVs, Link-TE options, and Link-TE TLVs are each independently described in a separate sub-section.

```
 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   |      0x81     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Link State ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Advertising Router                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    0    |V|E|B|      0        |       Router-TE flags         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Router-TE flags (contd.)     |       Router-TE TLVs          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     ....                                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     ....      |          # of TE links        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Link ID                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Link Data                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |       0       |    Link-TE flags              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Link-TE flags (contd.)      |  Zero or more Link-TE TLVs    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Link ID                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Link Data                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              ...                              |
```

Option

In TE-capable router nodes, the TE-bit may be set to 1.

The following fields are used to describe each router link (i.e., interface). Each router link is typed (see the below Type field). The Type field indicates the kind of link being described.

Type

A new link type "Positional-Ring Type" (value 5) is defined. This is essentially a connection to a TDM-Ring. TDM ring network is different from LAN/NBMA transit network in that, nodes on the TDM ring do not necessarily have a terminating path between themselves. Secondly, the order of links is important in determining the circuit path. Third, the protection switching and the number of fibers from a node going into a ring are determined by the ring characteristics. I.e., 2-fiber vs 4-fiber ring and UPSR vs BLSR protected ring.

Type	Description
1	Point-to-point connection to another router
2	Connection to a transit network
3	Connection to a stub network
4	Virtual link
5	Positional-Ring Type.

Link ID

Identifies the object that this router link connects to. Value depends on the link's Type. For a positional-ring type, the Link ID shall be IP Network/Subnet number, just as with a broadcast transit network. The following table summarizes the updated Link ID values.

Type	Link ID
1	Neighboring router's Router ID
2	IP address of Designated Router
3	IP network/subnet number
4	Neighboring router's Router ID
5	IP network/subnet number

Link Data

This depends on the link's Type field. For type-5 links, this specifies the router interface's IP address.

9.1.1. Router-TE flags - TE capabilities of the router

Below is an initial set of definitions. More may be standardized if necessary. The TLVs are not expanded in the current rev. Will be done in the follow-on revs. The field imposes a restriction of no more than 32 flags to describe the TE capabilities of a router-TE.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|L|P|T|L|F|                                           |S|S|S|C|
|S|E|S|D|S|S|                                           |T|E|I|S|
|R|R|C|M|C|C|                                           |A|L|G|P|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|<---- Boolean TE flags ------->|<- TE flags pointing to TLVs ->|
```

Bit LSR
When set, the router is considered to have LSR capability.

Bit LER
When set, the router is considered to have LER capability. All MPLS border routers will be required to have the LER capability. When the E bit is also set, that indicates an AS Boundary router with LER capability. When the B bit is also set, that indicates an area border router with LER capability.

Bit PSC
Indicates the node is Packet Switch Capable.

Bit TDM
Indicates the node is TDM circuit switch capable.

Bit LSC
Indicates the node is Lamda switch Capable.

Bit FSC
Indicates the node is Fiber (can also be a non-fiber link type) switch capable.

Bit STA

Label Stack Depth limit TLV follows. This is applicable only when the PSC flag is set.

Bit SEL
TE Selection Criteria TLV, supported by the router, follows.

Bit SIG
MPLS Signaling protocol support TLV follows.

BIT CSPF
CSPF algorithm support TLV follows.

9.1.2. Router-TE TLVs

The following Router-TE TLVs are defined.

TE-selection-Criteria TLV (Tag ID = 1)

The values can be a series of resources that may be used as the criteria for traffic engineering (typically with the aid of a signaling protocol such as RSVP-TE or CR-LDP or LDP).

- Bandwidth based LSPs (1)
- Priority based LSPs (2)
- Backup LSP (3)
- Link cost (4)

Bandwidth criteria is often used in conjunction with Packet Switch Capable nodes. The unit of bandwidth permitted to be configured may however vary from vendor to vendor. Bandwidth criteria may also be used in conjunction with TDM nodes. Once again, the granularity of bandwidth allocation may vary from vendor to vendor.

Priority based traffic switching is relevant only to Packet Switch Capable nodes. Nodes supporting this criteria will be able to interpret the EXP bits on the MPLS header to prioritize the traffic across the same LSP.

Backup criteria refers to whether or not the node is capable

of finding automatic protection path in the case the originally selected link fails. Such a local recovery is specific to the node and may not need to be notified to the upstream node.

MPLS-Signaling protocol TLV (Tag ID = 3)
The value can be 2 bytes long, listing a combination of RSVP-TE, CR-LDP and LDP.

Constraint-SPF algorithms-Support TLV (Tag ID = 4)
List all the CSPF algorithms supported. Support for CSPF algorithms on a node is an indication that the node may be requested for all or partial circuit path selection during

circuit setup time. This can be beneficial in knowing whether or not the node is capable of expanding loose routes (in an MPLS signaling request) into an LSP. Further, the CSPF algorithm support on an intermediate node can be beneficial when the node terminates one or more of the hierarchical LSP tunnels.

Label Stack Depth TLV (Tag ID = 5)
Applicable only for PSC-Type traffic. A default value of 1 is assumed. This indicates the depth of label stack the node is capable of processing on an ingress interface.

9.1.3. Link-TE options - TE capabilities of a TE-link

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|T|N|P|T|L|F|D|                                         |S|L|B|C|
|E|T|K|D|S|S|B|                                         |R|U|W|O|
| |E|T|M|C|C|S|                                         |L|G|A|L|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|<---- Boolean TE flags ------->|<- TE flags pointing to TLVs ->|
```

TE - Indicates whether TE is permitted on the link. A link can be denied for TE use by setting the flag to 0.

NTE - Indicates whether non-TE traffic is permitted on the TE link. This flag is relevant only when the TE flag is set.

PKT   - Indicates whether or not the link is capable of
           packet termination.

TDM, LSC, FSC bits
         - Same as defined for router TE options.

DBS   - Indicates whether or not Database synchronization
           is permitted on this link.

SRLG Bit - Shared Risk Link Group TLV follows.

LUG bit - Link usage cost metric TLV follows.

BWA bit - Data Link bandwidth TLV follows.

COL bit - Data link Color TLV follows.

9.1.4. Link-TE TLVs

SRLG-TLV
This describes the list of Shared Risk Link Groups the link belongs to. Use 2 bytes to list each SRLG.

BWA-TLV
This indicates the maximum bandwidth, available bandwidth, reserved bandwidth for later use etc. This TLV may also describe the Data link Layer protocols supported and the Data link MTU size.

LUG-TLV
This indicates the link usage cost - Bandwidth unit, Unit usage cost, LSP setup cost, minimum and maximum durations permitted for setting up the TLV etc., including any time of day constraints.

COLOR-TLV
This is similar to the SRLG TLV, in that an autonomous system may choose to issue colors to link based on a certain criteria. This TLV can be used to specify the

color assigned to the link within the scope of the AS.

9.2. TE-incremental-link-Update LSA (0x8d)

A significant difference between a non-TE OSPF network and a TE OSPF network is that the latter is subject to dynamic circuit pinning and is more likely to undergo state updates. Specifically, some links might undergo changes more frequently than others. Advertising the entire TE-router LSA in response to a change in any single link could be repetitive. Flooding the network with TE-router LSAs at the aggregated speed of all the dynamic changes is simply not desirable. The TE-incremental-link-update LSA advertises only the incremental link updates.

The TE-incremental-link-Update LSA will be advertised as frequently as the link state is changed. The TE-link sequence is largely the advertisement of a sub-portion of router LSA. The sequence number on this will be incremented with the TE-router LSA's sequence as the basis. When an updated TE-router LSA is advertised within 30 minutes of the previous advertisement, the updated TE-router LSA will assume a sequence no. that is larger than the most frequently updated of its links.

Below is the format of the TE-incremental-link-update 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            LS age             |     Options   |      0x8d     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Link State ID (same as Link ID)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Advertising Router                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Link Data                            |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |       0       |    Link-TE options            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Link-TE options           | Zero or more Link-TE TLVs     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     # TOS     |                            metric             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              ...                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      TOS      |       0       |          TOS  metric          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Link State ID

This would be exactly the same as would have been specified as as Link ID for a link within the router-LSA.

Link Data

This specifies the router ID the link belongs to. In majority of cases, this would be same as the advertising router. This choice for Link Data is primarily to facilitate proxy advertisement for incremental link updates.

Say, a router-proxy-LSa was used to advertise the TE-router-LSA of a SONET/TDM node. Say, the proxy router is now required to advertise incremental-link-update for the same SONET/TDM node. Specifying the actual router-ID the link in the incremental-link-update-LSA belongs to helps receiving nodes in finding the exact match for the LSA in their database.

The tuple of (LS Type, LSA ID, Advertising router) uniquely identify the LSA and replace LSAs of the same tuple with an older sequence number. However, there is an exception to this rule in the context of TE-link-update LSA. TE-Link update LSA will initially assume the sequence number of the TE-router LSA it belongs to. Further, when a

new TE-router LSA update with a larger sequence number is advertised, the newer sequence number is assumed by al the link LSAs.

9.3. TE-Circuit-paths LSA (0x8C)

TE-Circuit-paths LSA may be used to advertise the availability of pre-established TE circuit path(s) originating from any router

in the network. The flooding scope may be Area wide or AS wide.

```
 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   |      0x84     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Link State ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Advertising Router                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      LS sequence number                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      0    |G|E|B|D|S|T|CktType| Circuit Duration (Optional)   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Circuit Duration cont...                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Circuit Duartion cont..       | Circuit Setup time (Optional) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Circuit Setup time cont...                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Circuit Setup time cont..     |Circuit Teardown time(Optional)|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Circuit Teardown time cont...                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Circuit Teardown time cont..  |  No. of TE circuit paths      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Circuit-TE ID                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Circuit-TE Data                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |        0      |    Circuit-TE flags           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Circuit-TE flags (contd.)   |  Zero or more Circuit-TE TLVs |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Circuit-TE ID                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Circuit-TE Data                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
|                              ...                              |
```

Link State ID
    The ID of the far-end router or the far-end Link-ID to which the
    TE circuit path(s) is being advertised.

TE-circuit-path(s) flags

```
    Bit G - When set, the flooding scope is set to be AS wide.
            Otherwise, the flooding scope is set to be area wide.

    Bit E - When set, the advertised Link-State ID is an AS boundary
            router (E is for external). The advertising router and
            the Link State ID belong to the same area.

    Bit B - When set, the advertised Link state ID is an Area border
            router (B is for Border)

    Bit D - When set, this indicates that the duration of circuit
            path validity follows.

    Bit S - When set, this indicates that Setup-time of the circuit
            path follows.

    Bit T - When set, this indicates that teardown-time of the
            circuit path follows.
```

CktType
    This 4-bit field specifies the Circuit type of the Forward
    Equivalency Class (FC).

```
             0x01 - Origin is Router, Destination is Router.
             0x02 - Origin is Link,   Destination is Link.
             0x04 - Origin is Router, Destination is Link.
             0x08 - Origin is Link,   Destination is Router.
```

Circuit Duration (Optional)
    This 64-bit number specifies the seconds from the time of the
    LSA advertisement for which the adversited pre-established
    TE circuit path will be valid. This field is specified only
    when the D-bit is set in the TE-circuit-path flags.

Circuit Setup time (Optional)
    This 64-bit number specifies the time at which the TE-circuit
    path may be setup. This field is specified only when the
    S-bit is set in the TE-circuit-path flags. The setup time is
    specified as the number of seconds from the start of January
    1 1970 UTC.

Circuit Teardown time (Optional)
This 64-bit number specifies the time at which the TE-circuit path may be torn down. This field is specified only when the T-bit is set in the TE-circuit-path flags. The teardown time is specified as the number of seconds from the start of January 1 1970 UTC.

No. of TE Circuit paths
This indicates the number of pre-established TE circuit paths between the advertising router and the router specified in the link state ID.

Circuit-TE ID
This is the ID of the far-end router for a given TE-circuit path segment.

Circuit-TE Data
This is the virtual link identifier on the near-end router for a given TE-circuit path segment. This can be a private interface or handle the near-end router uses to identify the virtual link.

The sequence of (circuit-TE ID, Circuit-TE Data) list the end-point nodes and links in the LSA as a series.

Circuit-TE flags
This lists the Zero or more TE-link TLVs that all member elements of the LSP meet.

9.4. TE-Summary LSAs

TE-Summary-LSAs are the Type 0x83 and 0x84 LSAs. These LSAs are originated by area border routers. TE-Summary-network-LSA (0x83) describes the reachability of TE networks in a non-backbone area, advertised by the Area Border Router. Type 0x84 summary-LSA describes the reachability of Area Border Routers and AS border routers and their TE capabilities.

One of the benefits of having multiple areas within an AS is that frequent TE advertisements within the area do not impact outside the area. Only the TE abstractions as befitting the external areas are advertised.

9.4.1. TE-Summary Network LSA (0x83)

TE-summary network LSA may be used to advertise reachability of TE-networks accessible to areas external to the originating

area. The content and the flooding scope of a TE-Summary LSA is different from that of a native summary LSA.

The scope of flooding for a TE-summary network is AS wide, with the exception of the originating area and the stub areas. The area border router for each non-backbone area is responsible for advertising the reachability of backbone networks into the area.

Unlike a native-summary network LSA, TE-summary network LSA does not advertise summary costs to reach networks within an area. This is because, TE parameters are not necessarily additive or comparative. The parameters can be varied in their expression. A TE-summary network LSA will not be know to summarize a network whose links do not fall under an SRLG (Shared-Risk Link Group). This is way, the TE-summary LSA merely advertises the reachable of TE networks within an area. The specific circuit paths can be computed by the BDRs. On the other hand, if there are specific circuit paths to advertise, that can be done independently using TE-Circuit-path LSA (refer: section 9.3)

```
 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   |     0x83      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Link State ID  (IP Network Number)           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Advertising Router (Area Border Router)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
   |                       Network Mask                            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                       Area-ID                                 |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

9.4.2. TE-Summary router LSA (0x84)

TE-summary router LSA may be used to advertise the availability of Area Border Routers (ABRs) and AS Border Routers (ASBRs) that are TE capable. The TE-summary router LSAs are originated by the Area Border Routers. The scope of flooding for the TE-summary router LSA

is the non-backbone area the advertising ABR belongs to.

```
    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   |      0x84     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                     Link State ID                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                     Advertising Router (ABR)                  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                     LS sequence number                        |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |         LS checksum           |             length            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |    0      |E|B|      0        |         No. of Areas          |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                     Area-ID                                   |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                       ...                                     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                   Router-TE flags                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                   Router-TE TLVs                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                     ....                                      |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Link State ID
The ID of the Area border router or the AS border router whose TE capability is being advertised.

Advertising Router
The ABR that advertises its TE capabilities (and the OSPF areas it belongs to) or the TE capabilities of an ASBR within one of the areas the ABR is a border router of.

No. of Areas
Specifies the number of OSPF areas the link state ID belongs to.

Area-ID
Specifies the OSPF area(s) the link state ID belongs to. When the link state ID is same as the advertising router ID, this lists all the areas the ABR belongs to. In the case the link state ID is an ASBR, this simply lists the area the ASBR belongs to. The advertising router is assumed to be the ABR from the same area the ASBR is located in.

Summary-router-TE flags

Bit E - When set, the advertised Link-State ID is an AS boundary router (E is for external). The advertising router and the Link State ID belong to the same area.

Bit B - When set, the advertised Link state ID is an Area border router (B is for Border)

Router-TE flags,
Router-TE TLVs (TE capabilities of the link-state-ID router)

TE Flags and TE TLVs are as applicable to the ABR/ASBR specified in the link state ID. The semantics is same as specified in the Router-TE LSA.

9.5. TE-AS-external LSAs (0x85)

TE-AS-external-LSAs are the Type 0x85 LSAs. This is modeled after AS-external LSA format and flooding scope. These LSAs are originated by AS boundary routers with TE extensions (say, a BGP node which can

communicate MPLS labels across to external ASes), and describe networks and pre-established TE links external to the AS. The flooding scope of this LSA is similar to that of an AS-external LSA. I.e., AS wide, with the exception of stub areas.

```
     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   |      0x85     |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                        Link State ID                          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                     Advertising Router                        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                     LS sequence number                        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |         LS checksum           |             length            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                         Network Mask                          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                       Forwarding address                      |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                      External Route Tag                       |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  #  of Virtual TE links       |                 0             |
```



```
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                      Link-TE flags                            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                      Link-TE TLVs                             |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                              ...                              |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                      TE-Forwarding address                    |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                      External Route TE Tag                    |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                              ...                              |
```

Network Mask
    The IP address mask for the advertised TE destination. For

        example, this can be used to specify access to a specific
        TE-node or TE-link with an mask of 0xffffffff. This can also
        be used to specify access to an aggregated set of destinations
        using a different mask, ex: 0xff000000.

    Link-TE flags,
    Link-TE TLVs
        The TE attributes of this route. These fields are optional and
        are provided only when one or more pre-established circuits can
        be specified with the advertisement. Without these fields,
        the LSA will simply state TE reachability info.

    Forwarding address
        Data traffic for the advertised destination will be forwarded to
        this address. If the Forwarding address is set to 0.0.0.0, data
        traffic will be forwarded instead to the LSA's originator (i.e.,
        the responsible AS boundary router).

    External Route Tag
        A 32-bit field attached to each external route. This is not
        used by the OSPF protocol itself. It may be used to communicate
        information between AS boundary routers; the precise nature of
        such information is outside the scope of this specification.

9.6. Changes to Network LSA

Network-LSA is the Type 2 LSA. With the exception of the following, no additional changes will be required to this LSA for TE compatibility. The LSA format and flooding scope remains unchanged.

A network-LSA is originated for each broadcast, NBMA and Positional-Ring type network in the area which supports two or more routers. The TE option is also required to be set while

propagating the TDM network LSA.

9.6.1. Positional-Ring type network LSA - New Network type for TDM-ring.

- Ring ID: (Network Address/Mask)
- No. of elements in the ring (a.k.a. ring neighbors)
- Ring Bandwidth
- Ring Protection (UPSR/BLSR)
- ID of individual nodes (Interface IP address)

- Ring type (2-Fiber vs. 4-Fiber, SONET vs. SDH)

Network LSA will be required for SONET RING. Unlike the broadcast type, the sequence in which the NEs are placed on a RING-network is pertinent. The nodes in the ting must be described clock wise, assuming the GNE as the starting element.

9.7. TE-Router-Proxy LSA (0x8e)

This is a variation to the TE-router LSA in that the TE-router LSA is not advertised by the network element, but rather by a trusted TE-router Proxy. This is typically the scenario in a non-packet TE network, where some of the nodes do not have OSPF functionality and count on a helper node to do the advertisement for them. One such example would be the SONET/SDH ADM nodes in a TDM ring. The nodes may principally depend upon the GNE (Gateway Network Element) to do the advertisement for them. TE-router-Proxy LSA shall not be used to advertise Area Border Routers and/or AS border Routers.

```
 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   |      0x8e     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Link State ID  (Router ID of the TE Network Element)     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Advertising Router                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                0              |       Router-TE flags         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Router-TE flags (contd.)     |       Router-TE TLVs          |
+---------------------------------------------------------------+
|                     ....                                      |
+---------------------------------------------------------------+
|                     ....      |      # of TE links            |
```



```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
   |                          Link ID                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         Link Data                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |     Type      |       0       |    Link-TE options            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |   Link-TE flags               |  Zero or more Link-TE TLVs    |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                          Link ID                              |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                         Link Data                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                              ...                              |
```

9.8. Others

We may also introduce a new TE-NSSA LSA, similar to the native-NSSA LSA. TE-NSSA will help ensure that not all external TE routes are flooded into the NSSA area. A TE capable router can become the NSSA translator. All parameters and contents of TE-NSSA LSAs are transferred as is.

10. Abstract topology representation with TE support

Below, we assume a TE network that is composed of three OSPF areas, namely Area-1, Area-2 and Area-3, attached together through the backbone area. The following figure is an inter-area topology abstraction from the perspective of routers in Area-1. The abstraction is similar, but not the same, as that of the non-TE abstraction. As such, the authors claim the model is easy to understand and emulate. The abstraction illustrates reachability of TE networks and nodes in areas external to the local area and ASes external to the local AS. The abstraction also illustrates pre-established TE links that may be advertised by ABRs and ASBRs.

Area-1 an has a single border router, ABR-A1 and no ASBRs. Area-2 has an Area border router ABR-A2 and an AS border router ASBR-S1. Area-3 has two Area border routers ABR-A2 and ABR-A3; and an AS border router ASBR-S2. There may be any number of Pre-engineered TE links amongst ABRs and ASBRs. The following example assumes a single TE-link between ABR-A1 and ABR-A2; between ABR-A1 and ABR-A3; between ABR-A2 to ASBR-S1; and between ABR-A3 to ASBR-S2. All Area border routers and AS border routers are assumed to be represented by their TE capabilities.

```
                           +-------+
                           |Area-1 |
                           +-------+
    +-------------+            |
    |Reachable TE |        +------+
    |networks in  |--------|ABR-A1|
    |backbone area|        +------+
    +-------------+         | | |
              +-------------+ | +-------------------+
              |                 |                   |
    +-----------------+       |             +-----------------+
    |Pre-engineered TE|    +----------+     |Pre-engineered TE|
    |circuit path(s)  |    | Backbone |     |circuit path(s)  |
    |to ABR-A2        |    | Area     |     |to ABR-A3        |
    +-----------------+    +----------+     +-----------------+
              |               |   |                 |
              +----------+    |   |                 |
                         |    |   +--------------+  |
    +-----------+        |    |                  |  |     +-----------+
    |Reachable  |   +------------+              +------+  |Reachable  |
    |TE networks|---|    ABR-A2  |              |ABR-A3|--|TE networks|
    |in Area A2 |   +------------+              +------+  |in Area A3 |
    +-----------+   /   | | |                     | |     +-----------+
                  /     | | +-------------------+ | +----------+
                /       | +-------------+       | |            |
    +-----------+ +--------------+      |       | |    +--------------+
    |Reachable  | |Pre-engineered|      |       | |    |Pre-engineered|
    |TE networks| |TE Ckt path(s)|  +------+  +------+ |TE Ckt path(s)|
    |in Area A3 | |to ASBR-S1    |  |Area-2|  |Area-3| |to ASBR-S2    |
    +-----------+ +--------------+  +------+  +------+ +--------------+
                             |      /              |        /
    +-------------+          |     /               |       /
    |AS external  |      +---------+          +-------------+
    |TE-network   |------| ASBR-S1 |          |  ASBR-S2    |
    |reachability |      +---------+          +-------------+
    |from ASBR-S1 |         |                   |        |
    +-------------+         |                   |        |
            +-----------------+   +-------------+   +-----------------+
            |Pre-engineered TE|   |AS External  |   |Pre-engineered TE|
            |circuit path(s)  |   |TE-Network   |   |circuit path(s)  |
            |reachable from   |   |reachability |   |reachable from   |
            |ASBR-S1          |   |from ASBR-S2 |   |ASBR-S2          |
            +-----------------+   +-------------+   +-----------------+
```

Figure 9: Inter-Area Abstraction as viewed by Area-1 TE-routers

11. Changes to Data structures in OSPF-TE nodes

11.1. Changes to Router data structure

The router with TE extensions must be able to include all the TE capabilities (as specified in section 7.1) in the router data structure. Further, routers providing proxy service to other TE routers must also track the router and associated interface data structures for all the TE client nodes for which the proxy service is being provided. Presumably, the interaction between the Proxy server and the proxy clients is out-of-band.

11.2. Two set of Neighbors

Two sets of neighbor data structures will need to be maintained. TE-neighbors set is used to advertise TE LSAs. Only the TE-nodes will be members of the TE-neighbor set. Native neighbors set will be used to advertise native LSAs. All neighboring nodes supporting non-TE links can be part of this set. As for flooding optimizations based on neighbors set, readers may refer [FLOOD-OPT].

11.3. Changes to Interface data structure

The following new fields are introduced to the interface data structure. These changes are in addition to the changes specified in [FLOOD-OPT].

TePermitted
: If the value of the flag is TRUE, the interface is permissible to be advertised as a TE-enabled interface.

NonTePermitted
: If the value of the flag is TRUE, the interface permits non-TE traffic on the interface. Specifically, this is applicable to packet networks, where data links may permit both TE and non-TE packets. For FSC and LSC TE networks, this flag will be set to FALSE. For Packet networks that do not permit non-TE traffic on TE links also, this flag is set to TRUE.

PktTerminated
If the value of the flag is TRUE, the interface terminates Packet data and hence may be used for IP and OSPF data exchange.

AdjacencySychRequired
If the value of the flag is TRUE, the interface may be used to synchronize the LSDB across all adjacent neighbors. This is TRUE by default to all PktTerminated interfaces that are enabled for OSPF. However, it is possible to set this to FALSE

for some of the interfaces.

TE-TLVs
Each interface may potentially have a maximum of 16 TLVS that describe the link characteristics.

The following existing fields in Interface data structure will take on additional values to support TE extensions.

Type
The OSPF interface type can also be of type "Positional-RING". The Positional-ring type is different from other types (such as broadcast and NBMA) in that the exact location of the nodes on the ring is relevant, even as they are all on the same ring. SONET ADM ring is a good example of this. Complete ring positional-ring description may be provided by the GNE on a ring as a TE-network LSA for the ring.

List of Neighbors
The list may be statically defined for an interface, without requiring the use of Hello protocol.

12. IANA Considerations

12.1. TE-compliant-SPF routers Multicast address allocation

12.2. New TE-LSA Types

12.3. New TLVs (Router-TE and Link-TE TLVs)

12.3.1. TE-selection-Criteria TLV (Tag ID = 1)

- Bandwidth based LSPs (1)
- Priority based LSPs (2)
- Backup LSP (3)
- Link cost (4)

12.3.2. MPLS-Signaling protocol TLV (Tag ID = 3)
- RSVP-TE signaling
- LDP signaling
- CR-LDP signaling

12.3.3. Constraint-SPF algorithms-Support TLV (Tag ID = 4)
- CSPF Algorithm Codes.

12.3.4. SRLG-TLV (Tag ID = 0x81)
- SRLG group IDs

12.3.5. BW-TLV (Tag ID = 0x82)

12.3.6 CO-TLV (Tag ID = 0x83)

13. Acknowledgements

The authors wish to thank Vishwas Manral, Chitti Babu, Riyad Hartani and Tricci So for their valuable comments and feedback on the draft.

14. Security Considerations

This memo does not create any new security issues for the OSPF protocol. Security considerations for the base OSPF protocol are covered in [OSPF-v2]. As a general rule, a TE network is likely to generate significantly more control traffic than a native OSPF network. The excess traffic is almost directly proportional to the rate at which TE circuits are setup and torn down within an autonomous system. It is important to ensure that TE database synchronizations happen quickly when compared to the aggregate circuit setup an tear-down rates.

REFERENCES

[IETF-STD] Bradner, S., " The Internet Standards Process -- Revision 3", RFC 1602, IETF, October 1996.

[RFC 1700] J. Reynolds and J. Postel, "Assigned Numbers", RFC 1700

[MPLS-TE] Awduche, D., et al, "Requirements for Traffic Engineering Over MPLS," RFC 2702, September 1999.

[GMPLS-TE] P.A. Smith et. al, "Generalized MPLS - Signaling Functional Description", work in progress, draft-ietf-mpls-generalized-signaling-03.txt

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

[CR-LDP] Jamoussi, B. et al, "Constraint-Based LSP Setup using LDP", draft-ietf-mpls-cr-ldp-06.txt, Work in Progress.

[OSPF-v2] Moy, J., "OSPF Version 2", RFC 2328, April 1998.

[MOSPF] Moy, J., "Multicast Extensions to OSPF", RFC 1584, March 1994.

[NSSA] Coltun, R., V. Fuller and P. Murphy, "The OSPF NSSA Option", draft-ietf-ospf-nssa-update-10.txt, Work in Progress.

[OPAQUE] Coltun, R., "The OSPF Opaque LSA Option," RFC 2370, July 1998.

[FLOOD-OPT] Zinin, A. and M. Shand, "Flooding Optimizations in link-state routing protocols", work in progress, <draft-ietf-ospf-isis-flood-opt-01.txt>

[OPQLSA-TE] Katz, D., D. Yeung and K. Kompella, "Traffic Engineering Extensions to OSPF", work in progress,

<draft-katz-yeung-ospf-traffic-06.txt>

[OPQLSA-GMPLS] Kompella, K., Y. Rekhter, A. Banerjee, J. Drake, G. Bernstein, D. Fedyk, E. Mannie, D. Saha and V. Sharma, "OSPF Extensions in Support of Generalized MPLS", <draft-ietf-ccamp-ospf-gmpls-extensions-01.txt>, work in progress.

Authors' Addresses

Pyda Srisuresh
Kuokoa Networks, Inc.
475 Potrero Avenue
Sunnyvale, CA 94085
U.S.A.
EMail: srisuresh@yahoo.com

Paul Joseph
Force10 Networks
1440 McCarthy Boulevard
Milpitas, CA 95035
U.S.A.
EMail: pjoseph@Force10Networks.com