

Network Working Group
INTERNET-DRAFT
Expires as of June 8, 2003

P. Srisuresh
Kuokoa Networks
P. Joseph
Force10 Networks
December 8, 2002

OSPF-TE: An experimental extension to OSPF for Traffic Engineering
<[draft-srisuresh-ospf-te-04.txt](#)>

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>.

Abstract

This document defines OSPF-TE, an experimental traffic engineering (TE) extension to the link-state routing protocol OSPF. New TE LSAs are designed to disseminate TE metrics within an autonomous System (AS) - intra-area as well as inter-area. An Autonomous System may consist of TE and non-TE nodes. Non-TE nodes are unaffected by the distribution of TE LSAs. A stand-alone TE Link State Database (TE-LSDB), separate from the native OSPF LSDB, is generated for the computation of TE circuit paths. OSPF-TE is also extendible to non-packet networks such as SONET/TDM and optical networks. A transition path is provided for those currently using [[OPQLSA-TE](#)] and wish to adapt OSPF-TE.

Table of Contents

1.	Introduction	3
2.	Principles of traffic engineering	3
3.	Terminology	5
3.1.	TE node	5
3.2.	TE link	5
3.3.	TE circuit path	5
3.4.	OSPF-TE node	6
3.5.	TE control network	6
3.6.	TE network (TE topology)	6
3.7.	Packet-TE network	6
3.8.	Non-packet-TE network	6
3.9.	Native (non-TE) node	7
3.10.	Native (non-TE) link	7
3.11.	Non-TE network (Non-TE topology)	7
3.12.	Peer network (combination network)	7
3.13.	LSP	7
3.14.	LSA	7
3.14.	LSDB	7
3.15.	CSPF	7
3.16.	TLV	8
3.17.	Router-TE TLVs	8
3.18.	Link-TE TLVs	8
4.	Motivations behind the design of OSPF-TE	8
4.1.	Scalable design	9
4.2.	Coexistent design	9
4.3.	Efficient in flooding reach	9
4.4.	Ability to reserve TE-exclusive links	10
4.5.	Extendible design	10
4.6.	Unified for packet and non-packet networks	11
4.7.	Networks benefiting from the OSPF-TE design	11
5.	OSPF-TE solution overview	12
5.1.	OSPF-TE Solution	12
5.2.	Assumptions	13
6.	Opaque LSAs to OSPF-TE transition strategy	14
7.	OSPF-TE router adjacency - TE topology discovery	14
7.1.	The OSPF Options field	15
7.2.	The Hello Protocol	15
7.3.	Flooding and the Synchronization of Databases	16
7.4.	The Designated Router	16
7.5.	The Backup Designated Router	16
7.6.	The graph of adjacencies	17

8.	TE LSAs - Packet network	18
8.1.	TE-Router LSA (0x81)	19
8.2.	TE-incremental-link-Update LSA (0x8d)	26
8.3.	TE-Circuit-paths LSA (0x8C)	27
8.4.	TE-Summary LSAs	30

8.5.	TE-AS-external LSAs (0x85)	33
9.	TE LSAs - Non-packet network	34
9.1.	TE-Router LSA (0x81)	34
9.2.	Changes to Network LSA	36
9.3.	TE-Router-Proxy LSA (0x8e)	36
10.	Abstract topology representation with TE support	37
11.	Changes to Data structures in OSPF-TE routers	40
11.1.	Changes to Router data structure	40
11.2.	Two set of Neighbors	40
11.3.	Changes to Interface data structure	40
12.	IANA Considerations	41
12.1.	TE LSA type values	41
12.2.	TE TLV tag values	42
13.	Acknowledgements	42
14.	Security Considerations	42
15.	Normative References	44
16.	Informative References	44

[1.](#) Introduction

This document defines OSPF-TE, an experimental traffic engineering (TE) extension to the link-state routing protocol OSPF. The objective of OSPF-TE is to discover TE network topology and disseminate TE metrics within an autonomous system(AS). A stand-alone TE Link State Database (TE-LSDB), different from the native OSPF LSDB, is created to facilitate computation of TE circuit paths. Algorithms to compute TE circuit paths is however not the objective of this document.

OSPF-TE is different from the Opaque-LSA-based design outlined in [\[OPQLSA-TE\]](#). [Section 4](#) describes the motivations behind the design of OSPF-TE. [Section 6](#) outlines a strategy to transition Opaque-LSA based implementations to adapt OSPF-TE.

Those interested in TE extensions for the packet networks only may skip [section 9.0](#).

[2.](#) Principles of traffic engineering

The objective of traffic engineering is to set up circuit path(s) between a pair of nodes or links and to forward traffic of a certain forwarding equivalency class through the circuit path. Only the unicast circuit paths are considered here. Multicast variations are out of scope for this document.

A traffic engineered circuit path may be identified by the tuple of (Forwarding Equivalency Class, TE parameters for the

circuit, Origin Node/Link, Destination node/Link).

Forwarding Equivalency Class (FEC) is a grouping of traffic that is forwarded in the same manner by a node. A FEC may be classified based on a number of criteria as follows.

- a) Traffic arriving on a specific interface,
- b) Traffic arriving at a certain time of day,
- c) Traffic meeting a certain classification criteria (ex: based on a match of the fields in the IP and transport headers),
- d) Traffic in a certain priority class,
- e) Traffic arriving on a specific set of TDM (STS) circuits on an interface,
- f) Traffic arriving on a certain wavelength of an interface

Discerning traffic based on the FEC criteria is mandatory for Label Edge Routers (LERs). The intermediate Label Switched Routers (LSRs) are transparent to the traffic content. LSRs are merely responsible for keeping the circuit in-tact for the circuit lifetime. This document will not address defining FEC criteria, or the mapping of a FEC to circuit, or the associated signaling to set up circuits. [[MPLS-TE](#)] and [[GMPLS-TE](#)] address the FEC criteria. [[RSVP-TE](#)] and [[CR-LDP](#)] address signaling protocols to set up circuits.

This document is concerned with the collection of TE metrics for all the TE enforceable nodes and links within an autonomous system. TE metrics for a node may include the following.

- a) Ability to perform traffic prioritization,

- b) Ability to provision bandwidth on interfaces,
- c) Support for Constrained Shortest Path First (CSPF) algorithms,
- d) Support for certain TE-Circuit switch type,
- e) Support for a certain type of automatic protection switching

TE metrics for a link may include the following.

- a) Available bandwidth,
- b) Reliability of the link,
- c) Color assigned to the link,
- d) Cost of bandwidth usage on the link,
- e) Membership to a Shared Risk Link Group (SRLG)

A number of CSPF algorithms may be used to dynamically set up TE circuit paths in a TE network.

As for origin node/link and destination node/link, the originating and the terminating entities of a TE circuit path are identifiable

by their IP addresses.

3. Terminology

Definitions of terms used in the context of the 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", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[IETF-STD](#)].

Below are definitions for the terms used within this document.

3.1. TE node

TE-Node is a node in the traffic engineered (TE) network. A TE-node has a minimum of one TE-link attached to it. Associated with each TE node is a set of supported TE metrics. A TE node may also participate in a native IP network.

In a SONET/TDM or photonic cross-connect network, a TE node is not required to be an OSPF-TE router. An external OSPF-TE router may act as proxy for the TE nodes that cannot be routers themselves.

[3.2.](#) TE link

TE Link is a network attachment point to a TE-node and is intended for traffic engineering use. Associated with each TE link is a set of supported TE metrics. A TE link may also optionally carry native IP traffic.

Of the various links attached to a TE-node, only the links that take part in a traffic engineered network are called the TE links.

[3.3.](#) TE circuit path

A TE circuit path is a uni-directional data path, defined by a list of TE nodes connected to each other through TE links. A TE circuit path is also often referred merely as a circuit path or a circuit.

For the purposes of OSPF-TE, the originating and terminating entities of a TE circuit path must be identifiable by their

IP addresses. As a general rule, all nodes and links party to a Traffic Engineered network should be uniquely identifiable by an IP address.

[3.4.](#) OSPF-TE node

An OSPF-TE node is a TE node that runs the OSPF routing protocol and the OSPF-TE extensions described in this document.

An autonomous system (AS) may be constituted of a combination of native and OSPF-TE nodes.

[3.5.](#) TE Control network

The IP network used by the OSPF-TE nodes for OSPF-TE

communication is referred as the TE control network or simply the control network. The control network can be independent of the TE data network.

[3.6.](#) TE network (TE topology)

A TE network is a network of connected TE-nodes and TE-links for the purpose of setting up one or more TE circuit paths. The terms TE network, TE data network and TE topology are used synonymously throughout the document.

[3.7.](#) Packet-TE network

A packet-TE network is a TE network in which the nodes switch MPLS packets. An MPLS packet is defined in [[MPLS-TE](#)] as a packet with an MPLS header, followed by data octets. The intermediary node(s) of a circuit path in a packet-TE network perform MPLS label swapping to emulate the circuit.

Unless specified otherwise, the term packet network is used throughout the document to refer a packet-TE network.

[3.8.](#) Non-packet-TE network

A non-packet-TE network is TE-network in which the nodes switch non-packet entities such as an STS time slot, a Lambda wavelength or simply an interface.

SONET/TDM and Fiber cross-connect networks are examples of non-packet-TE networks. Circuit emulation in these networks is accomplished by the switch fabric in the intermediary nodes (based on TDM time slot, fiber interface or Lambda).

Unless specified otherwise, the term non-packet network is used throughout the document to refer a non-packet-TE network.

[3.9.](#) Native (non-TE) node

A native or non-TE node is an OSPF router capable of IP packet forwarding and does not take part in a TE network. A native

OSPF node forwards IP traffic using the shortest-path forwarding algorithm and does not run the OSPF-TE extensions.

[3.10.](#) Native (non-TE) link

A native (or non-TE) link is a network attachment to a TE or non-TE node used for IP packet traversal.

[3.11.](#) non-TE network (Non-TE topology)

A non-TE network refers to an OSPF network that does not support TE. Non-TE network, native-OSPF network and non-TE topology are used synonymously throughout the document.

[3.12.](#) Peer network (combination network)

A peer network is a network that is constituted of packet and non-packet networks combined. In a peer network, a TE node could potentially support TE links for the packet as well as non-packet data.

OSPF-TE is usable within a packet network or a non-packet network or a peer network, which is a combination of the two.

[3.13.](#) LSP

LSP stands for "Label Switched Path". LSP is a TE circuit path in a packet network. The terms LSP and TE circuit path are used synonymously in the context of packet networks.

[3.14.](#) LSA

LSA stands for OSPF "Link State Advertisement".

[3.15.](#) LSDB

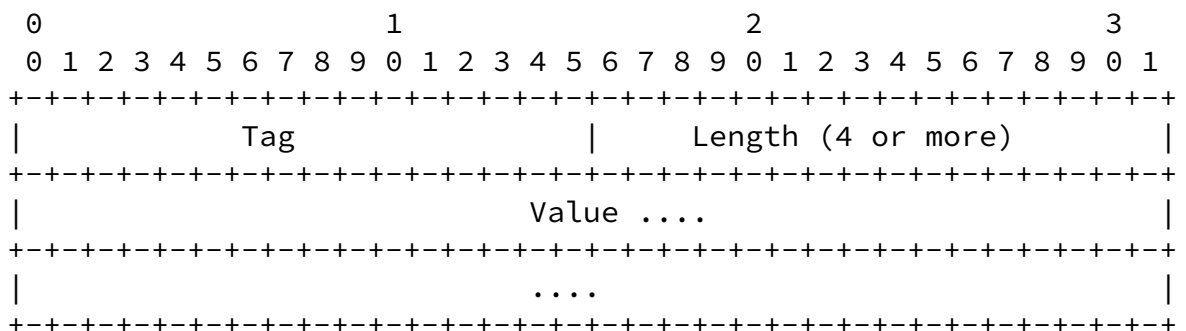
LSDB stands for "LSA Database". LSDB is a representation of the topology of a network. A native LSDB, constituted of native OSPF LSAs, represents the topology of a native IP network. TE-LSDB, on the other hand, is constituted of TE LSAs and is a representation

[3.16.](#) CSPF

CSPF stands for "Constrained Shortest Path First". Given a TE LSDB and a set of constraints that must be satisfied to form a circuit path, there may be several CSPF algorithms to obtain a TE circuit path that meets the criteria.

[3.17.](#) 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 octets required for Tag and Length specification. All TLVs described in this document are padded to 32-bit alignment. Any padding required for alignment will not be a part of the length field, however. TLVs are used to describe traffic engineering characteristics of the TE nodes, TE links and TE circuit paths.



[3.18.](#) Router-TE TLVs

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

[3.19.](#) Link-TE TLVs

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

[4.](#) Motivations behind the design of OSPF-TE

There are several motivations that lead to the design of OSPF-TE. OSPF-TE is scalable, coexistent and efficient in flooding reach. The motivations are explained in detail in the following subsections. Also listed in the last subsection are network scenarios that benefit from the OSPF-TE design.

[4.1.](#) Scalable design

Area level abstraction provides the scaling necessary for a large autonomous system (AS). OSPF-TE allows for independent area abstractions for the TE and native topologies. The TE and native area border routers will advertise different summary LSAs to TE and non-TE routers. Readers may refer [section 10](#) for a topological view of the AS from an OSPF-TE node in an area.

[4.2.](#) Coexistent design

OSPF-TE regards an AS as constituted of a TE and non-TE networks coexisting within the same bounds. OSPF-TE dynamically discovers TE topology and the associated TE metrics of the nodes and links within, just as the native OSPF does in a non-TE network. An independent TE-LSDB, representative of the TE topology is generated as a result. A stand-alone TE-LSDB allows for speedy searches through the database.

In [[OPQLSA-TE](#)], the TE-LSDB is derived from the combination of opaque LSAs and native LSDB. The TE-LSDB derived has no knowledge of the TE capabilities of the routers in the network.

[4.3.](#) Efficient in 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. Non-TE nodes are not bombarded with TE LSAs. This is a useful characteristic for networks supporting native and TE traffic in the same connected network.

A subset of the TE metrics may be prone to rapid change, while others remain largely unchanged. Changes in TE metrics must be communicated at the earliest throughout the network to ensure that the TE-LSDB is up-to-date within the network. 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 set up and torn down within the TE network. The TE database synchronization should occur much quicker compared to the aggregate circuit set up and tear-down rates.

TE-Incremental-Link-update LSA ([section 8.2](#)) permits advertising a subset of the link metrics.

The more frequent and wider the flooding frequency, 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. It is undesirable to flood non-TE nodes with TE information.

[OPQLSA-TE] uses Opaque LSAs for advertising TE information. Opaque LSAs reach all nodes within the network - TE-nodes and non-TE nodes alike. [\[OPQLSA-TE\]](#) also 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.

[4.4.](#) Ability to reserve TE-exclusive links

OSPF-TE is designed to draw distinction between TE-links and non-TE links. A TE link, configured to support TE traffic alone, will not permit best-effort IP traffic on the link. This permits TE enforceability on the TE links.

When links of a TE-topology do not overlap the links of a native IP network, OSPF-TE allows for virtual isolation of the two networks. Best-effort IP transit network and constraint based TE network often have different service requirements. 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 service enforceability can be retained.

[OPQLSA-TE] does not support the ability to isolate best-effort IP traffic from TE traffic on a link. All links are subject to best-effort IP traffic. An OSPF router could potentially select a TE link to be its least cost link and inundate the link with best-effort IP traffic, thereby rendering the link unusable for TE purposes.

[4.5.](#) Extendible design

OSPF-TE design is based on the tried and tested OSPF paradigm, and inherits all the benefits of the OSPF, present and future.

TE-LSAs are extendible, just as the native OSPF on which OSPF-TE is founded.

[[OPQLSA-TE](#)], on the other hand, is constrained by the semantics of the Opaque LSA on which it is founded. The content within an Opaque LSA is restricted to being in the form of TLVs and sub-TLVs, some of which are mandatory and some of which are positionally dependent in the TLV sequence for proper interpretation. Opaque LSAs are also restrictive when the flooding

scope for the content is required to be different from the scope of the opaque LSA itself.

[4.6.](#) Unified for packet and non-packet networks

OSPF-TE is usable within a packet network or a non-packet network or a combination peer network.

Signaling protocols such as RSVP and LDP work the same across packet and non-packet networks. Signaling protocols merely need the TE characteristics of nodes and links so they can signal the nodes to formulate TE circuit paths. In a peer network, the underlying control protocol must be capable of providing a unified LSDB for all TE nodes (nodes with packet-TE links as well as non-packet-TE links) in the network. OSPF-TE meets this requirement.

[OPQLSA-TE] is limited in scope for packet networks. An independent [[OPQLSA-GMPLS](#)] is required to support GMPLS links in a non-packet network. Neither of the Opaque LSA based extensions have provision to distinguish between node types.

[4.7.](#) Networks benefiting from the OSPF-TE design

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

[4.7.1.](#) IP providers transitioning to provide 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. The native OSPF-LSDB will remain undisturbed while newer TE links are added to the network.

[4.7.2.](#) 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.

[4.7.3.](#) Large TE networks

The OSPF-TE design is advantageous in large TE networks that require the AS to be sub-divided into multiple areas.

[4.7.4.](#) Non-packet networks and Peer networks

OSPF-TE is also the right choice for vendors opting for a stable, well-founded protocol for their non-IP TE networks. OSPF-TE is uniquely qualified to support the following network attachments in non-Packet TE networks.

- (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).

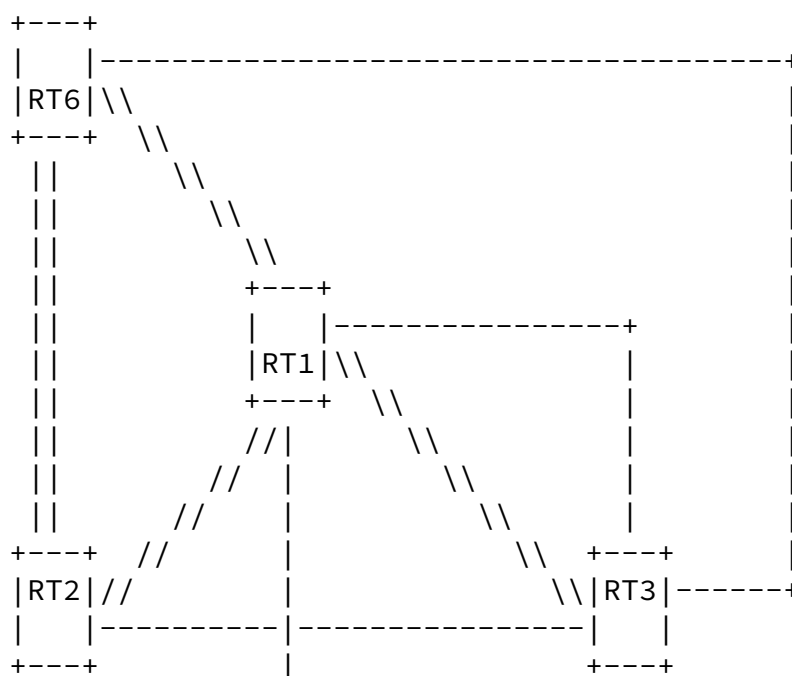
[5.](#) OSPF-TE solution overview

[5.1.](#) OSPF-TE Solution

A new TE flag is introduced within the OSPF options field to enable discovery of TE topology. [Section 8.0](#) describes the semantics of the TE flag. TE LSAs are designed for use by the OSPF-TE nodes. [Section 9.0](#) describes the TE LSAs in detail.

Changes required of the OSPF data structures to support OSPF-TE are described in [section 11.0](#). A new TE-neighbors data structure will be used to flood TE LSAs along TE-topology.

An OSPF-TE node will have the native LSDB and the TE-LSDB, A native OSPF node will have just the native LSDB. Consider the following OSPF area constituted of OSPF-TE and native OSPF routers. Nodes RT1, RT2, RT3 and RT6 are OSPF-TE routers with TE and non-TE link attachments. Nodes RT4 and RT5 are native OSPF routers with no TE links. When the LSA database is synchronized, all nodes will share the same native LSDB OSPF-TE nodes alone will have the additional TE-LSDB.



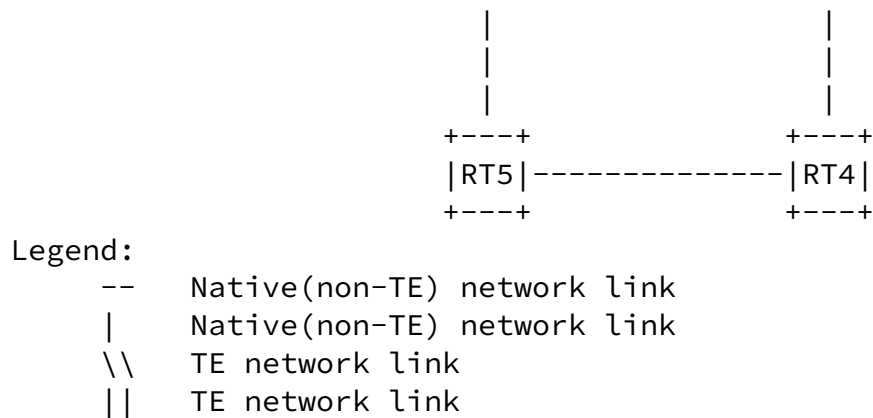


Figure 6: A (TE + native) OSPF network topology

5.2. Assumptions

OSPF-TE is an extension to the native OSPF protocol and does not mandate changes to the existing OSPF. OSPF-TE design makes the following assumptions.

1. An OSPF-TE node will need to establish router adjacency with at least one other OSPF-TE node in the area in order for the router's TE-database to be synchronized within the area. Failing this, the OSPF router will not be in the TE calculations of other TE routers in the area.

It is the responsibility of the network administrator(s) to ensure connectedness of the TE network. Otherwise, there can be disjoint TE topologies within a network.

2. OSPF-TE nodes must advertise the link state of its TE-links. TE-links are not obligated to support native IP traffic. Hence, an OSPF-TE node cannot be required to synchronize its link-state database with neighbors on all its links. The only requirement is to have the TE LSDB synchronized across all OSPF-TE nodes in the area. Refer [[FLOOD-OPT](#)] for flooding optimizations.
3. A link in a packet network may be designated as a TE-link or a native-IP link or both. For example, a link may be used for both TE and non-TE traffic, so long as the link is

under-subscribed in bandwidth for TE traffic - say, 50% of the link capacity is set aside for TE traffic.

4. Non-packet TE sub-topologies MUST have a minimum of one node running OSPF-TE protocol. For example, a SONET/SDH TDM ring must have a minimum of one Gateway Network Element(GNE) running OSPF-TE. The OSPF-TE node will advertise on behalf of all the in the ring.

6. Opaque LSAs to OSPF-TE transition strategy

Below is a strategy to transition implementations using opaque LSAs to adapt the OSPF-TE 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 topologies area-wise. By doing this, the TE topology for the AS will be available at area level abstraction.
3. Use TE-Summary LSAs and TE-AS-external-LSAs for inter-area Communication. Make use of the TE-topology within an 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 within their connected areas.

7. OSPF-TE router adjacency - TE topology discovery

OSPF creates adjacencies between neighboring routers for the purpose of exchanging routing information. In the following subsections, we describe modifications to the OSPF options field and the use of Hello protocol to establish TE capability compliance between neighboring routers in an area. The capability is used as the basis to build TE topology.

7.1. The OSPF Options field

A new TE flag is introduced within the options field by this draft

to identify TE extensions to the OSPF. This bit will be used to distinguish routers that support OSPF-TE. 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 the 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 will use the TE-bit to establish OSPF-TE adjacency. 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 deemed necessary to leave this bit as is, the OPAQUE-9 LSA (local scope) can be used on each interface to communicate the support for OSPF-TE.

7.2. The Hello Protocol

The Hello Protocol is primarily responsible for dynamically establishing and maintaining neighbor adjacencies. In a TE network, it is not required for all links and neighbors to establish adjacency using this protocol. The Hello protocol will use the TE-bit to establish traffic engineering capability 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.

If deemed necessary to leave the TE-bit unused in the options field, the OSPF-TE routers could use OPAQUE-9 LSA (local scope) to communicate TE capability between two OSPF routers.

[7.3.](#) The Designated Router

The Designated Router is elected by the Hello Protocol on broadcast and NBMA networks. In general, when a router's non-TE link first becomes functional, it checks to see whether there is currently a Designated Router for the network. If there is one, 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.

TE-compliance (I.e., OSPF-TE) must be implemented on the most robust routers, as they become likely candidates to take on the 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).

[7.4.](#) 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 electing 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).

[7.5.](#) Flooding and the Synchronization of Databases

In OSPF, adjacent routers within an area must synchronize their databases. However, as observed in [[FLOOD-OPT](#)], a more concise requirement of OSPF is that all routers in an area must converge on the same link state database. It is sufficient to send a

single copy of the LSAs to the neighboring routers in an area

than send one copy on each connected interface. [[FLOOD-OPT](#)] describes in detail how to minimize flooding (Initial LSDB synchronization as well as the asynchronous LSA updates) within an area.

In the case where some of the neighbors are TE compliant and others are not, the designated OSPF-TE router will exchange different sets of LSAs with its neighbors. TE LSAs are exchanged only with the TE neighbors. Native LSAs are exchanged with all neighbors (TE and non-TE alike).

A new OSPFIGP-TE multicast address 224.0.0.24 may be used for the exchange of TE compliant database descriptors. Flooding optimization in a TE network is essential as the control traffic for a TE network is likely to be higher than that of a non-TE network. Flooding optimization will help minimize LSA announcements and the associated retransmissions and acknowledgements on the network.

[7.6.](#) The graph of adjacencies

If two routers have multiple networks in common, they may have multiple adjacencies between them. The adjacency may be one of two types - native OSPF adjacency and TE adjacency. OSPF-TE routers will form both types of adjacency.

Two types of adjacency 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 be one of (a) TE-compliant or (b) native. In contrast, on broadcast and NBMA networks the designated router and the backup designated router may maintain two sets of adjacency. The remaining routers will form either TE-compliant or native adjacency. In the Broadcast network below, routers RT7 and RT3 are chosen as the designated and backup routers respectively. Routers RT3, RT4 and RT7 are TE-compliant. RT5 and RT6 are not. So, RT4 will have TE and native adjacencies with the designated and backup routers. RT5 and RT6 will only have native adjacency with the

designated and backup routers.

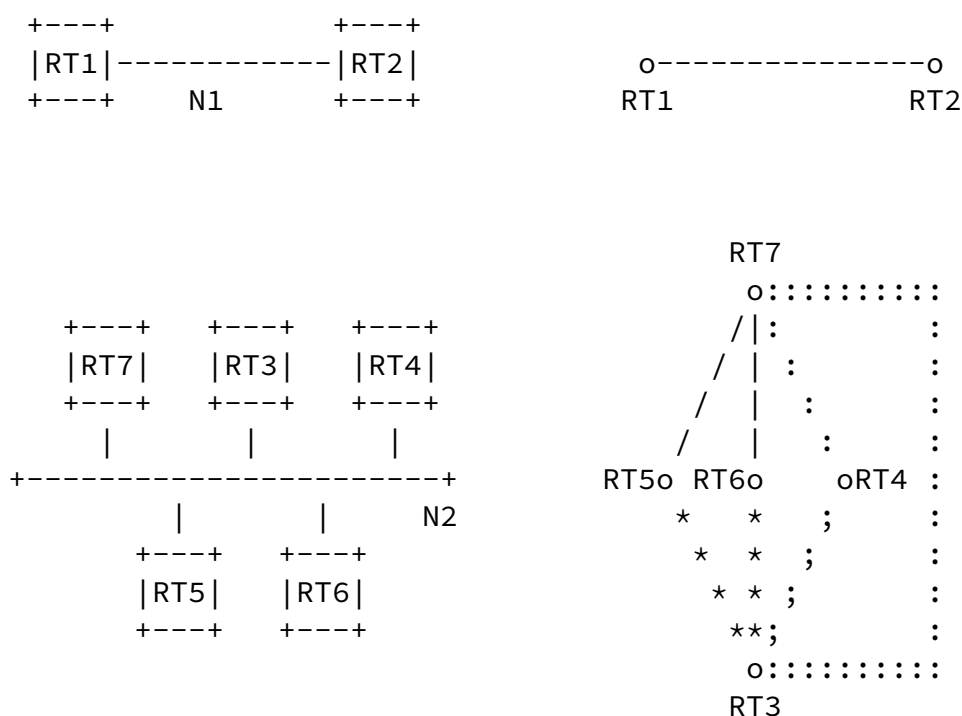


Figure 6: The graph of adjacencies with TE-compliant routers.

8. TE LSAs - Packet network

The OSPFv2 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. LSA types 9 through 11 are defined in [OPAQUE].

Each LSA type has a unique flooding scope. 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.

In the following subsections, we define new LSAs for traffic engineering (TE) use. The Values for the new TE LSA types are assigned such that the high bit of the LSA-type octet is set to 1. The new TE LSAs are largely modeled after the existing LSAs for content format and have a unique flooding scope.

TE-router LSA is defined to advertise TE characteristics of an OSPF-TE router and all the TE-links attached to the router. TE-incremental-Link-Update LSA is defined to advertise incremental updates to the metrics of a TE link. Flooding scope for both these LSAs is restricted to the TE nodes in the area.

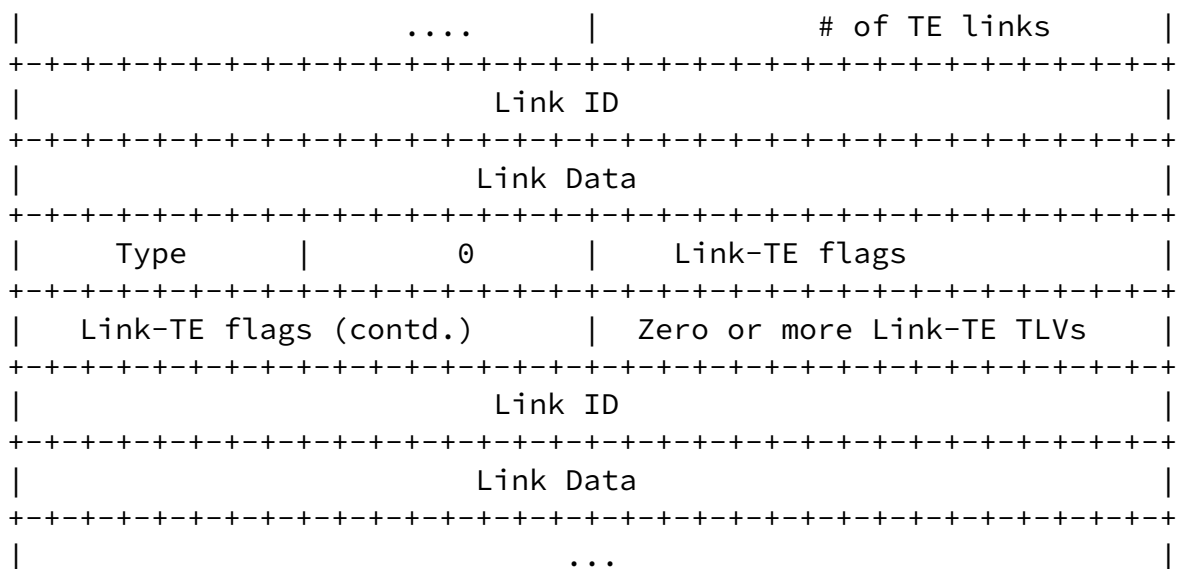
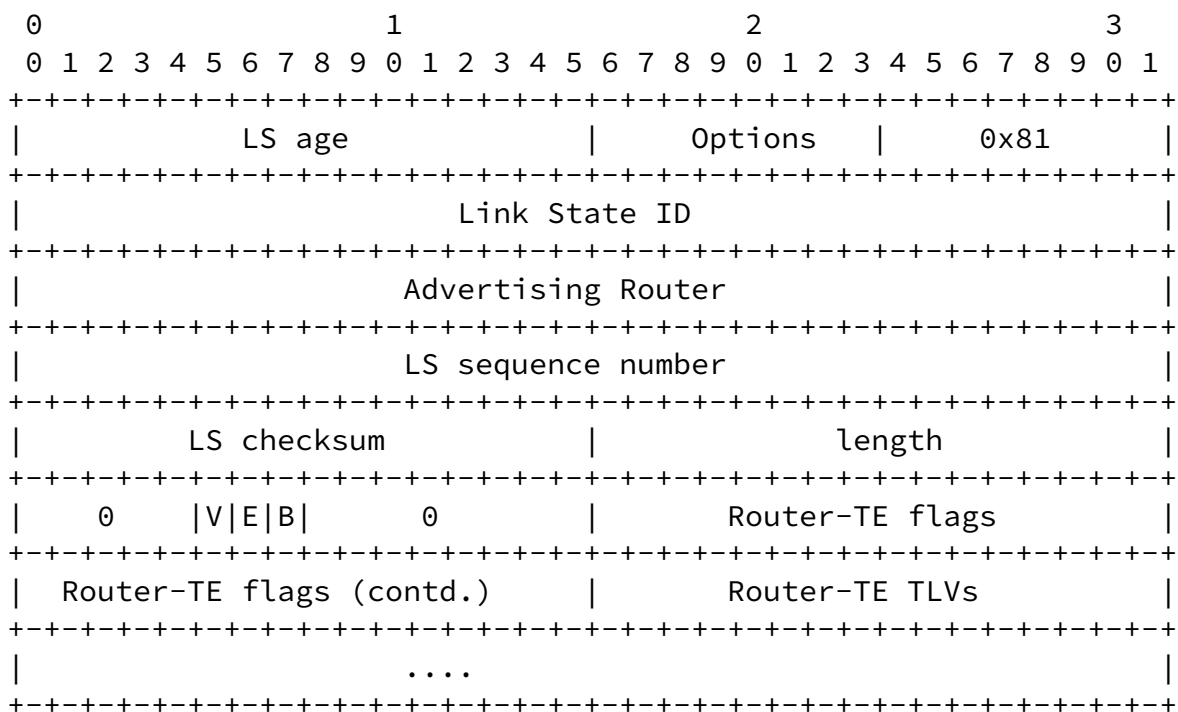
TE-Summary network and router LSAs are defined to advertise the reachability of area-specific TE networks and Area Border Routers (along with router TE characteristics) to external areas. Flooding Scope of the TE-Summary LSAs is the TE topology in the entire AS less the non-backbone area for which the 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.

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

[8.1.](#) TE-Router LSA (0x81)

The TE-router LSA (0x81) is modeled after the router LSA and has the same flooding scope as the router-LSA. However, the scope is restricted to only the OSPF-TE 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 described in the following sub-sections.



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

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

The following flags are used to describe the TE capabilities of an OSPF-TE router. The remaining bits of the 32-bit word are reserved for future use.

```
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|L|L|P| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|S|E|S| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|R|R|C| | | | | | | | | | | | | | | | | | | | | | | | | | | |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|<--- 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 LSP

MPLS Label switch TLV TE-NODE-TLV-MPLS-SWITCHING follows. This is applicable only when the PSC flag is set.

Bit SIG

MPLS Signaling protocol support TLV TE-NODE-TLV-MPLS-SIG-PROTOCOLS follows.

BIT CSPF

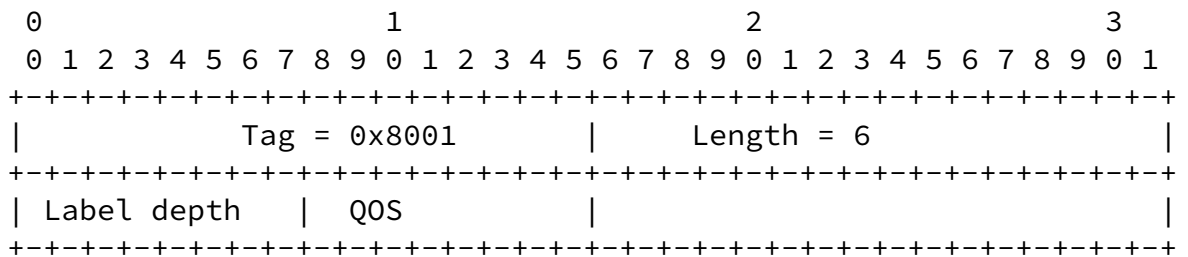
CSPF algorithm support TLV TE-NODE-TLV-CSPF-ALG follows.

8.1.2. Router-TE TLVs

The following Router-TE TLVs are defined.

8.1.2.4. TE-NODE-TLV-MPLS-SWITCHING

MPLS switching TLV is applicable only for packet switched nodes. The TLV specifies the MPLS packet switching capabilities of the TE node.

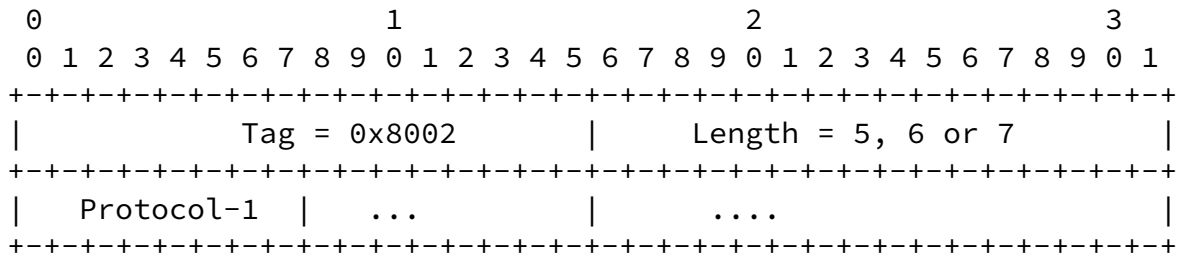


'Label depth' is the depth of label stack the node is capable of processing on its ingress interfaces. An octet is used to represent label depth. A default value of 1 is assumed when the TLV is not listed.

'QOS' is a single octet field that may be assigned '1' or '0'. Nodes supporting QOS are able to interpret the EXP bits in the MPLS header to prioritize multiple classes of traffic through the same LSP.

8.1.2.2. TE-NODE-TLV-MPLS-SIG-PROTOCOLS

MPLS signaling protocols TLV lists all the signaling protocol supported by the node. An octet is used to list each signaling protocol supported.

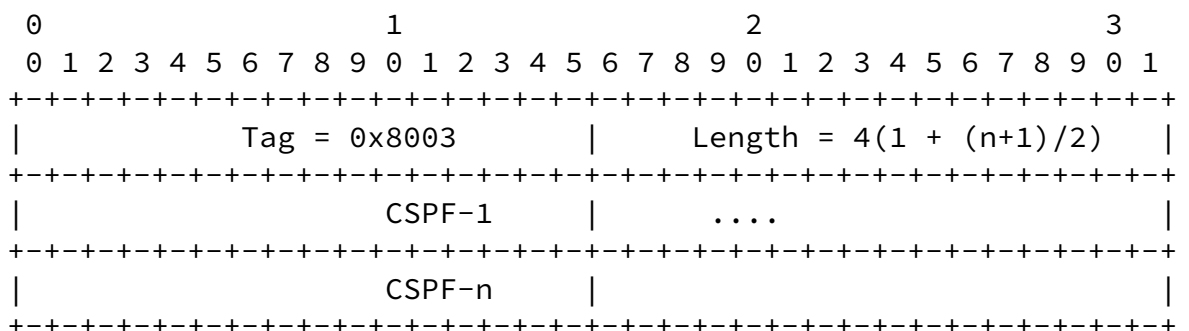


RSVP-TE protocol is represented as 1, CR-LDP as 2 and LDP as 3.
These are the only permitted signaling protocols at this time.

[8.1.2.3](#). TE-NODE-TLV-CSPF-ALGORITHMS

The CSPF algorithms TLV lists all the CSPF algorithm codes supported. Support for CSPF algorithms makes the node eligible to compute complete or partial circuit paths. Support for CSPF algorithms can also be beneficial in knowing whether or not a node is capable of expanding loose routes (in an MPLS signaling request) into a detailed circuit path.

Two octets are used to list each CSPF algorithm code. The algorithm codes may be vendor defined and unique within an Autonomous System. If the node supports 'n' CSPF algorithms, the Length would be $(4 + 4 * ((n+1)/2))$ octets.



8.1.3. Link-TE flags - TE capabilities of a link

The following flags are used to describe the TE capabilities of a link. The remaining bits of the 32-bit word are reserved for future use.

```

+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|T|N|P| | |D|                                     |S|L|B|C|
|E|T|K| | |B|                                     |R|U|W|O|
| |E|T| | |S|                                     |L|G| |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 IP packet processing.

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

SRLG Bit - Shared Risk Link Group TLV TE-LINK-TLV-SRLG follows.

LUG bit - Link usage cost metric TLV TE-LINK-TLV-LUG follows.

BW bit - Link bandwidth TLV TE-LINK-TLV-BANDWIDTH follows.

COL bit - Link Color TLV TE-LINK-TLV-COLOR follows.

8.1.4. Link-TE TLVs

8.1.4.1. TE-LINK-TLV-SRLG

The SRLG describes the list of Shared Risk Link Groups (SRLG) the link belongs to. Two octets are used to list each SRLG. If the link belongs to 'n' SRLGs, the Length would be $(4 + 4 * ((n+1)/2))$ octets.

[illegible]

```

|          SRLG-1          |          ....          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          SRLG-n          |          |          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

[8.1.4.2](#). TE-LINK-TLV-BANDWIDTH

The bandwidth TLV specifies maximum bandwidth, bandwidth available for TE use and reserved bandwidth 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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Tag = 0x0002          |          Length = 16          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Maximum Bandwidth          |          |          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Bandwidth available for TE use          |          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|          Reserved Bandwidth          |          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Bandwidth is expressed in units of 32 bytes/sec (256 bits/sec). A 32-bit field for bandwidth would permit specification not exceeding 1 tera-bits/sec.

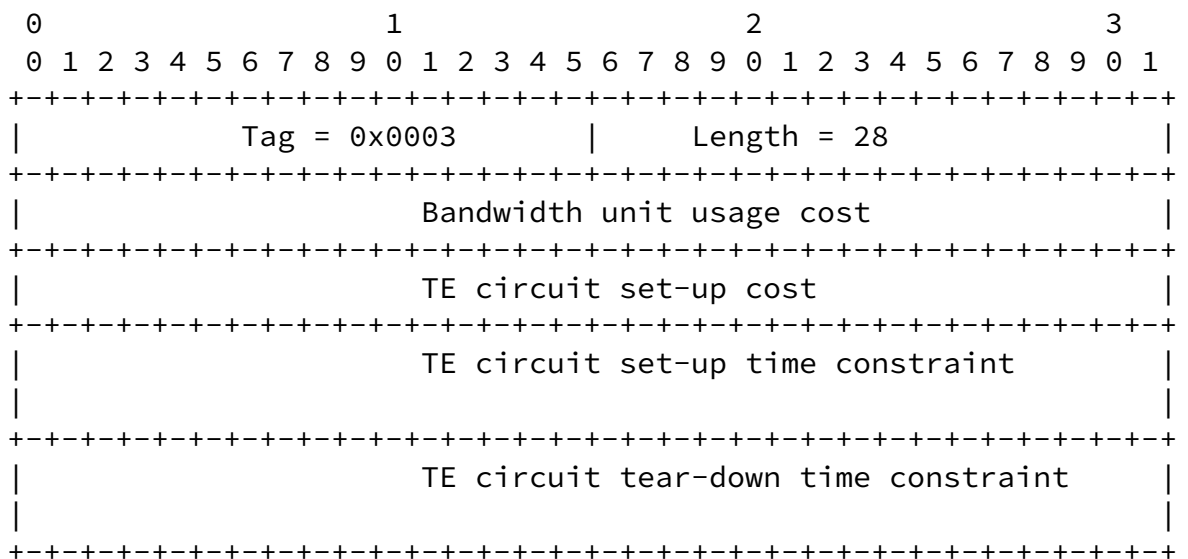
'Maximum bandwidth' is be the maximum link capacity expressed in bandwidth units.

'Bandwidth available for TE use' is the maximum reservable bandwidth on the link for use by all the TE circuit paths traversing the link. The link is oversubscribed when this field is more than the 'Maximum Bandwidth'. When the field is less than the 'Maximum Bandwidth', the remaining bandwidth on the link may likely be used for non-TE traffic.

'Reserved Bandwidth' is the bandwidth that is currently subscribed from of the link. 'Reserved Bandwidth' must be less than the 'Bandwidth available for TE use'. New TE circuit paths are able to claim no more than the difference between the two bandwidths for reservation.

[8.1.4.3](#). TE-LINK-TLV-LUG

The link usage cost TLV specifies Bandwidth unit usage cost, TE circuit set-up cost, and any time constraints for setup and teardown of TE circuits on the link.



Circuit Setup time constraint

This 64-bit number specifies the time at or after which a TE-circuit path may be set up on the link. The set-up time constraint is specified as the number of seconds from the start of January 1, 1970 UTC. A reserved value of 0 implies no circuit setup time constraint.

Circuit Teardown time constraint

This 64-bit number specifies the time at or before which all TE-circuit paths using the link must be torn down. The teardown time constraint is specified as the number of seconds from the start of January 1 1970 UTC. A reserved value of 0 implies no circuit teardown time constraint.

No. of TE Circuit paths

This specifies the number of pre-engineered TE circuit paths between the advertising router and the router specified in the

link state ID.

[8.1.4.4. TE-LINK-TLV-COLOR](#)

The color TLV is similar to the SRLG TLV, in that an Autonomous System may choose to issue colors to a TE-link meeting certain criteria. The color TLV can be used to specify one or more colors assigned to the link as follows. Two octets are used to list each color. If the link belongs to 'n' number of colors, the Length would be $(4 + 4 * ((n+1)/2))$ octets.

```

      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 = 0x0004                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Color-1                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Color-n                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               ....                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

[8.2. TE-incremental-link-Update LSA \(0x8d\)](#)

A significant difference between a non-TE OSPF network and a TE OSPF network is that the latter may be subject to frequent real-time circuit pinning and is likely to undergo TE-state updates. Some links might undergo changes more frequently than others. Flooding the network with TE-router LSAs at the aggregated speed of all link metric changes is simply not desirable. A smaller in size, TE-incremental-link-update LSA is designed to advertise only the incremental link updates.

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 all the link LSAs.

8.3. TE-Circuit-path LSA (0x8C)

TE-Circuit-path LSA may be used to advertise the availability of pre-engineered 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 Duration 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 pre-engineered 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 set up. This field is specified only when the S-bit is set in the TE-circuit-path flags. The set-up 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 specifies the number of pre-engineered 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.

[8.4.](#) 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 befitting the external areas are advertised.

[8.4.1.](#) 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. For example, a TE-summary network LSA will not summarize a

network whose links do not fall under an SRLG (Shared-Risk Link Group). This way, the TE-summary LSA merely advertises the reachability of TE networks within an area. The specific circuit paths can be computed by the BDRs. Pre-engineered circuit paths are advertised using TE-Circuit-path LSA (refer [section 8.3](#)).

```

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           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

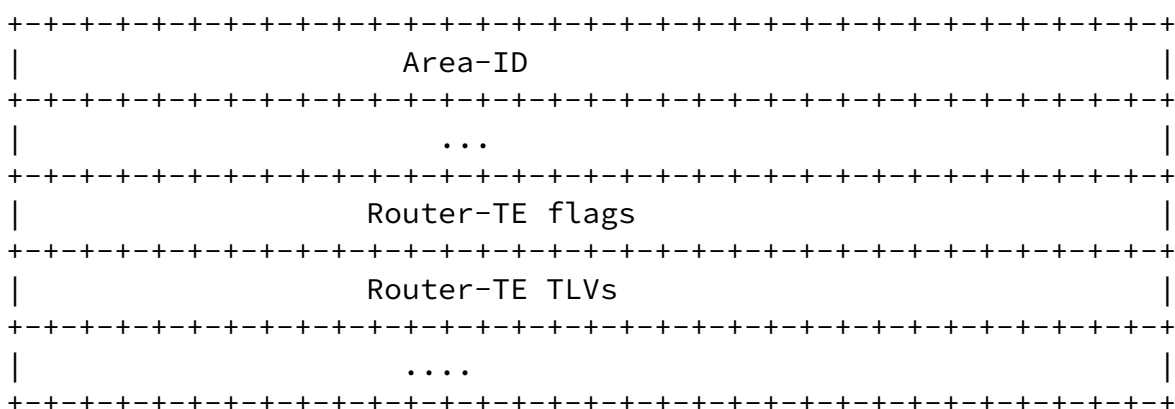
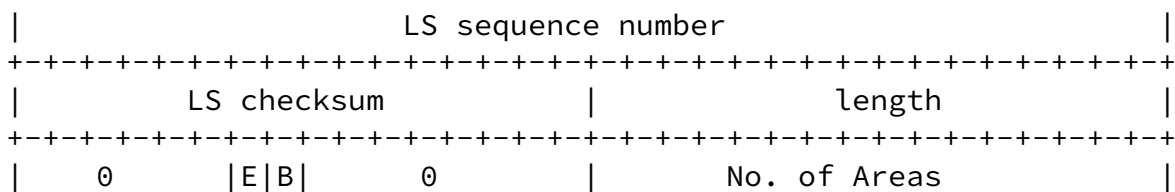
[8.4.2](#). TE-Summary router LSA (0x84)

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)           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

**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, the Area-ID lists all the areas the ABR belongs to. In the case the link state ID is an ASBR, the Area-ID 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.

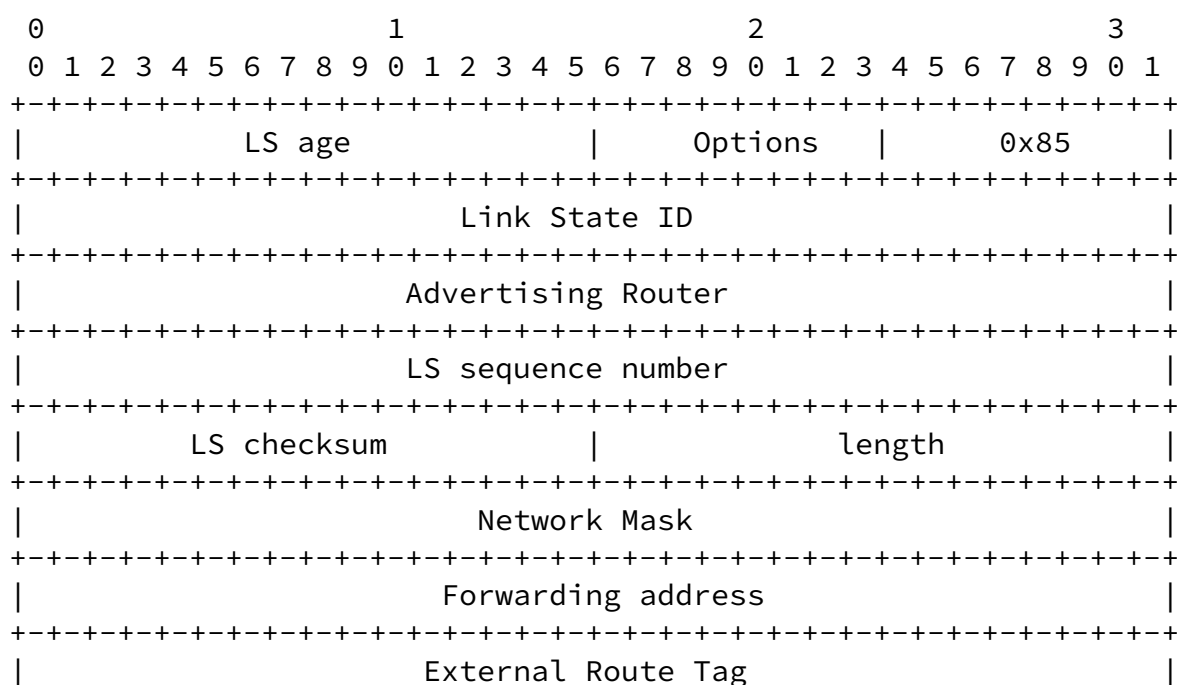
Bit B - When set, the advertised Link state ID is an Area border router (B is for Border)

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.

8.5. 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. TE-AS-external LSAs are originated by AS boundary routers with TE extensions, and describe the TE networks and pre-engineered circuit paths external to the AS. As with AS-external LSA, the flooding scope of the TE-AS-external LSA is AS wide, with the exception of stub areas.



```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| # 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-engineered 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.](#) TE LSAs - Non-packet network

A non-packet network would use all the TE LSAs described in the previous section for a packet network, albeit with some variations. These variations are described in the following subsections.

TE-Router-Proxy LSA is defined to allow proxy advertisement for non-packet TE-nodes by an OSPF-TE router.

[9.1.](#) TE-Router LSA (0x81)

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 the case 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

Flags specific to non-packet TE-nodes are described below.

```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|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 TDM

Indicates the node is TDM circuit switch capable.

Bit LSC

Indicates the node is Lambda switch Capable.

Bit FSC

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

[9.1.2.](#) 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 ->|

```

TDM, LSC, FSC bits

- Same as defined for router TE options.

9.2. 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.

9.2.1. 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 is required for SONET RING. Unlike the broadcast type, the sequence in which the Network Elements (NEs) are placed on a RING-network is pertinent. The nodes in the ring must be described clock wise, assuming the Gateway Network Element (GNE) as the starting element.

9.3. 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								
+-----+																																							

[10](#). Abstract topology representation with TE support

Below, we consider a TE network composed of three OSPF areas - Area-1, Area-2 and Area-3, attached together through the backbone area. Area-1 has a single area border router, ABR-A1 and no ASBRs. Area-2 has an area border router ABR-A2 and an AS border

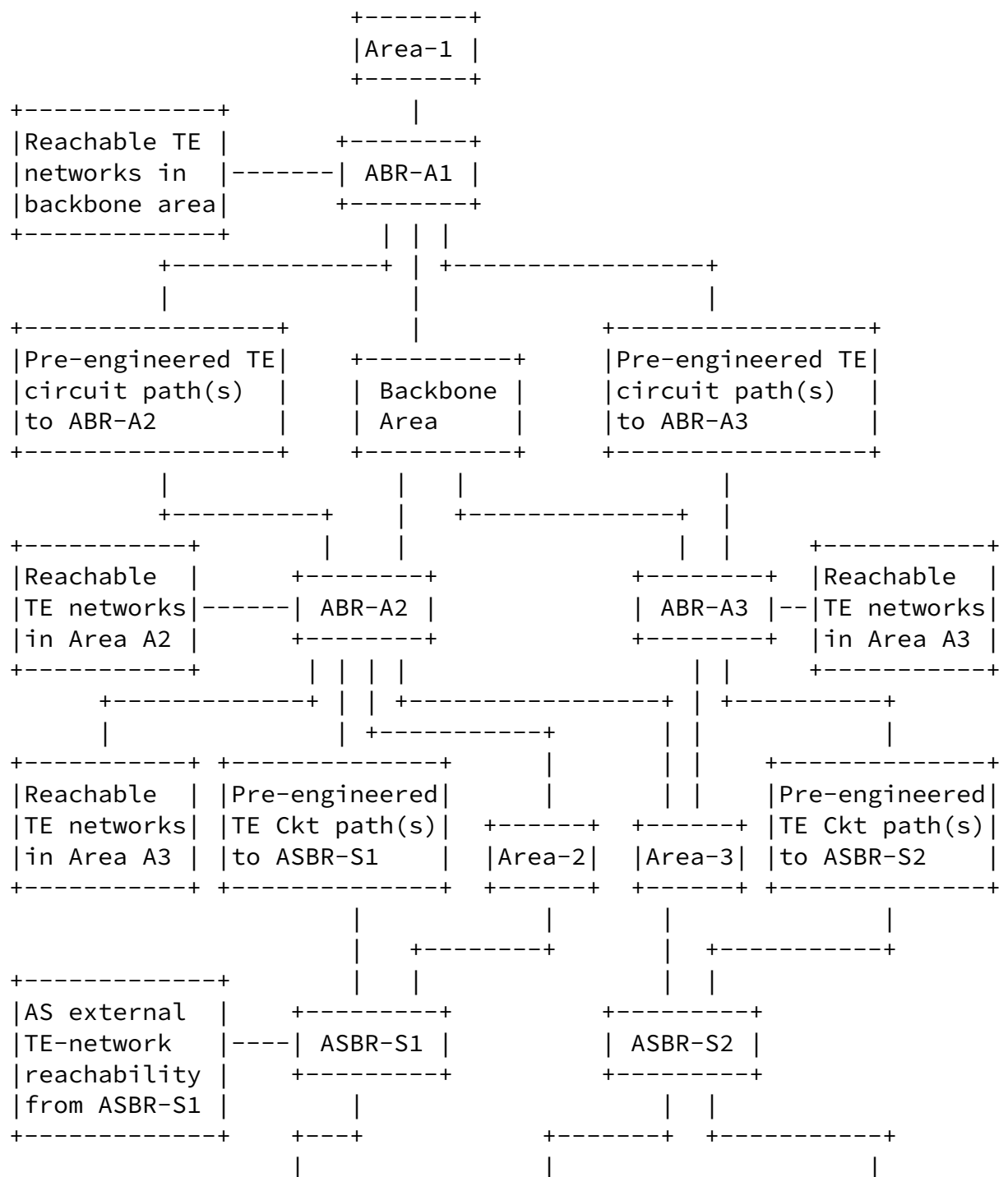
router ASBR-S1. Area-3 has two area border routers ABR-A2 and ABR-A3 and an AS border router ASBR-S2. The following network also assumes a pre-engineered TE circuit path between ABR-A1 and ABR-A2; between ABR-A1 and ABR-A3; between ABR-A2 and ASBR-S1; and between ABR-A3 and ASBR-S2.

The following figure is an inter-area topology abstraction from the perspective of routers in Area-1. The abstraction illustrates reachability of TE networks and nodes within area to the external areas in the same AS and to the external ASes. The abstraction also illustrates pre-engineered TE circuit paths advertised by ABRs and ASBRs.

Internet-Draft

OSPF TE extensions

December 2002



+-----+	+-----+	+-----+
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 sets of Neighbors

Two sets of neighbor data structures are required. 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 are 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 may 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 IP packets. For FSC and LSC TE networks, this flag is set to FALSE.

IpTerminated

If the value of the flag is TRUE, the interface processes IP Packet data and hence may be used for 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 IpTerminated 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 though 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.

This document proposes that TE LSA types and TE TLVs be maintained by the IANA. The document also proposes an OSPFIGP-TE multicast address be assigned by the IANA for the exchange of TE database descriptors.

OSPF-TE multicast address is suggested a value of 224.0.0.24 so as not to conflict with the recognized multicast address definitions, as defined in <http://www.iana.org/assignments/multicast-addresses>

The following sub-section explains the criteria to be used by the IANA to assign TE LSA types and TE TLVs.

12.1. TE LSA type values

LSA type is an 8-bit field required by each LSA. TE LSA types will have the high bit set to 1. TE LSAs can range from 0x80 through 0xFF. The following values are defined in sections 8.0 and 9.0. The remaining values are available for assignment by the IANA with IETF Consensus [Ref 11].

TE LSA Type	Value
-------------	-------

TE-Router LSA	0x81
TE-Summary Network LSA	0x83
TE-Summary router LSA	0x84
TE-AS-external LSAs	0x85
TE-Circuit-paths LSA	0x8C
TE-incremental-link-Update LSA	0x8d
TE-Router-Proxy LSA	0x8e

12.2. TE TLV tag values

TLV type is a 16-bit field required by each TE TLV. TLV type shall be unique across the router and link TLVs. A TLV type can range from 0x0001 through 0xFFFF. TLV type 0 is reserved and unassigned. The following TLV types are defined in sections 8.0 and 9.0. The remaining values are available for assignment

by the IANA with IETF Consensus [Ref 11].

TE TLV Tag	Reference Section	Value

TE-LINK-TLV-SRLG	Section 8.1.4.1	0x0001
TE-LINK-TLV-BWA	Section 8.1.4.2	0x0002
TE-LINK-TLV-LUG	Section 8.1.4.3	0x0003
TE-LINK-TLV-COLOR	Section 8.1.4.4	0x0004
TE-NODE-TLV-MPLS-SWITCHING	Section 8.1.2.1	0x8001
TE-NODE-TLV-MPLS-SIG-PROTOCOLS	Section 8.1.2.2	0x8002
TE-NODE-TLV-CSPF-ALG	Section 8.1.2.3	0x8003

[13.](#) Acknowledgements

The authors wish to specially thank Chitti Babu and his team for verifying portions of the specification for a packet network. The authors also wish to thank Vishwas Manral, Riyad Hartani and Tricci So for their valuable comments and feedback on the draft.

[14.](#) Security Considerations

Security considerations for the base OSPF protocol are covered in [[OSPF-v2](#)] and [[SEC-OSPF](#)]. This memo does not create any new security issues for the OSPF protocol. Security measures applied to the native OSPF (refer [[SEC-OSPF](#)]) are directly

applicable to the TE LSAs described in the document. Discussed below are the security considerations in processing TE LSAs.

Secure communication between OSPF-TE nodes has a number of components. Authorization, authentication, integrity and confidentiality. Authorization refers to whether a particular OSPF-TE node is authorized to receive or propagate the TE LSAs to its neighbors. Failing the authorization process might indicate a resource theft attempt or unauthorized resource advertisement. In either case, the OSPF-TE nodes should take proper measures to audit/log such attempts so as to alert the

administrator to take necessary action. OSPF-TE nodes may refuse to communicate with the neighboring nodes that fail to prompt the required credentials.

Authentication refers to confirming the identity of an originator for the datagrams received from the originator. Lack of strong credentials for authentication of OSPF-TE LSAs can seriously jeopardize the TE service rendered by the network. A consequence of not authenticating a neighbor would be that an attacker could spoof the identity of a "legitimate" OSPF-TE node and manipulate the state, and the TE database including the topology and metrics collected. This could potentially lead to denial-of-service on the TE network. Another consequence of not authenticating is that an attacker could pose as OSPF-TE neighbor and respond in a manner that would divert TE data to the attacker.

Integrity is required to ensure that an OSPF-TE message has not been accidentally or maliciously altered or destroyed. The result of a lack of data integrity enforcement in an untrusted environment could be that an imposter will alter the messages sent by a legitimate adjacent neighbor and bring the OSPF-TE on a node and the whole network to a halt or cause a denial of service for the TE circuit paths effected by the alteration.

Confidentiality of MIDCOM messages ensure that the TE LSAs are accessible only to the authorized entities. When OSPF-TE is deployed in an untrusted environment, lack of confidentiality will allow an intruder to perform traffic flow analysis and snoop the TE control network to monitor the traffic metrics and the rate at which circuit paths are being setup and torn-down. The intruder could cannibalize a lesser secure OSPF-TE node and destroy or compromise the state and TE-LDSB on the node. Needless to say, the least secure OSPF-TE will become the achilles heel and make the TE network vulnerable to security attacks.

15. Normative References

- [RFC 1700] J. Reynolds and J. Postel, "Assigned Numbers", [RFC 1700](#)
- [RFC 2434] Narten, T. and H. Alvestrand, "Guidelines for writing an IANA Considerations Section in RFCs", [BCP 26](#), [RFC 2434](#), October 1998.
- [MPLS-TE] Awduche, D., et al, "Requirements for Traffic Engineering Over MPLS," [RFC 2702](#), September 1999.
- [OSPF-v2] Moy, J., "OSPF Version 2", [RFC 2328](#), April 1998.
- [SEC-OSPF] Murphy, S., Badger, M., and B. Wellington, "OSPF with Digital Signatures", [RFC 2154](#), June 1997
- [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](#)>

[15](#). Informative References

- [GMPLS-TE] P.A. Smith et. al, "Generalized MPLS - Signaling Functional Description", work in progress, [draft-ietf-mpls-generalized-signaling-09.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.
- [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-11.txt](#), Work in Progress.
- [OPAQUE] Coltun, R., "The OSPF Opaque LSA Option," [RFC 2370](#), July 1998.

[OPQLSA-TE] Katz, D., D. Yeung and K. Kompella, "Traffic Engineering Extensions to OSPF", work in progress, <[draft-katz-yeung-ospf-traffic-09.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-09.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

