Network Working Group
INTERNET-DRAFT
Expires as of January 20, 2002

P. Srisuresh
Kuokoa Networks
P. Joseph
Jasmine Networks
July, 2001

# TE LSAs to extend OSPF for Traffic Engineering
<draft-srisuresh-ospf-te-01.txt>

## Status of this Memo



## Abstract

OSPF is a well established link state routing protocol used for topology discovery and computing forwarding table based on shortest-Path criteria. Traffic Engineering extensions (OSPF-TE) will use criteria different from shortest-path so as to route traffic around congestion paths and meet varying Service Level agreements. OSPF-TE may also be used by non-IP networks such as photonic and TDM (SONET/SDH) circuit switch networks for light-path or TDM circuit setup between two end-points. The approach outlined in this document differs from that of [OPQLSA-TE]. The document does not suggest the use of Opaque LSAs to add TE extensions to OSPF. Rather, new TE LSAs, modeled after existing LSAs and flooding scope are proposed to overcome the scaling limitations of the approach outlined in [OPQLSA-TE]. The

document draws a distinction between TE and non-TE topologies and restricts flooding of TE LSAs into non-TE topology. The document covers OSPF extensions for packet and non-packet networks alike, providing a unified extension mechanism for all networks. As such, this approach improves interoperability between peer network elements. Lastly, the document specifies a transition path for vendors currently using opaque LSAs to transition to using new TE LSAs outlined here.

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, flooding algorithm and the hierarchical organization of areas can all be used effectively in creating and tearing traffic links on demand. The intent of the document is to build an abstract view of the topology in conjunction with the traffic engineering characteristics of the nodes and links involved.

The connectivity topology may remain relatively stable, except when

the existing links or networking nodes go down or flap or new nodes and links are added to 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 traditional OSPF, the TE extended OSPF 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. For non-IP networks (such as TDM or photonic cross connect networks), Mapping IP addresses to a well-known name can be done through a DNS-like mechanism.

The approach suggested in this document is different from the Opaque-LSA-based approach outlined in [OPQLSA-TE]. Section 11 describes the motivations behind conceiving this approach and why the authors claim the benefits of the approach significantly substantial over the opaque LSA based approach. Section 12 outlines a strategy to transition from Opaque-LSA based deployments to the new-TE-LSA approach outlined here.

## 2. Traffic Engineering

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.

As for TE parameters for the circuit, this refers to the TE parameters for all the nodes and links constituting a circuit. Typically, TE parameters for a node in a TE circuit may include the following.

* Traffic prioritization ability,





* Ability to provision bandwidth on interfaces,
* Support of CSPF algorithms,
* TE-Circuit switch type,
* Automatic protection switching.

TE parameters for the link include:

* Bandwidth availability,
* reliability of the link,
* color assigned to the link
* cost of bandwidth usage on the link.
* membership to a Shared Risk Link Group and so on.

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 (or) TE-compliant OSPF node

This is a router that supports the OSPF TE extensions described in this document and at least one of the attached links support TE extensions. Further, this requires that at least one of the attached links support Packet termination and run the OSPF-TE protocol.

An OSPF-TE node supports native OSPF as well as the TE extensions outlined here.

## 3.2. Native OSPF router

A native OSPF router 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. An autonomous system (AS) could be





constituted of a combination of native-OSPF and OSPF-TE nodes.

A native OSPF router, when enhanced to include the extensions described in this document can become a OSPF-TE node.

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

A TE-Node is an intermediate or edge node taking part in the traffic engineered (TE) network. Specifically, a TE circuit is constituted of a series of TE nodes connected to each other via the TE links.

A non-TE node or a native node is a node that does not have any TE links attached to it and does not take part in a TE network. Specifically, native OSPF nodes that do not take part in a TE

network fall under this category.

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

A TE Link is a network attachment that supports traffic engineering. Specifically, a TE circuit can only be setup using a combination of TE nodes and TE links connected to each other.

Non-TE link or a native link is one that supports IP packet communication, but does not support traffic engineering on the link. For example, native OSPF protocol and least-cost criteria may be used to determine reachability of subnets in a network constituted of native OSPF nodes and native OSPF links.

## 3.5. Packet interface vs. non-packet interface

Interfaces on an OSPF-TE node may be characterized as those that terminate (i.e., send and receive) IP packet data and those that do not. Both types of links can be part of a traffic engineered network. In contrast, a native OSPF router does not support non-packet interfaces.

Needless to say, the OSPF protocol and its TE extensions can only be enabled on interfaces supporting IP packet termination. While the OSPF protocol can be run only on interfaces terminating IP packets - the protocol can advertise link state information of non-packet interfaces attached to it - thereby allowing for traffic engineering over non-packet links. For example - control interfaces can advertise link state information of the SONET interfaces on a SONET Add-Drop Multiplexer.

## 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 TE 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.

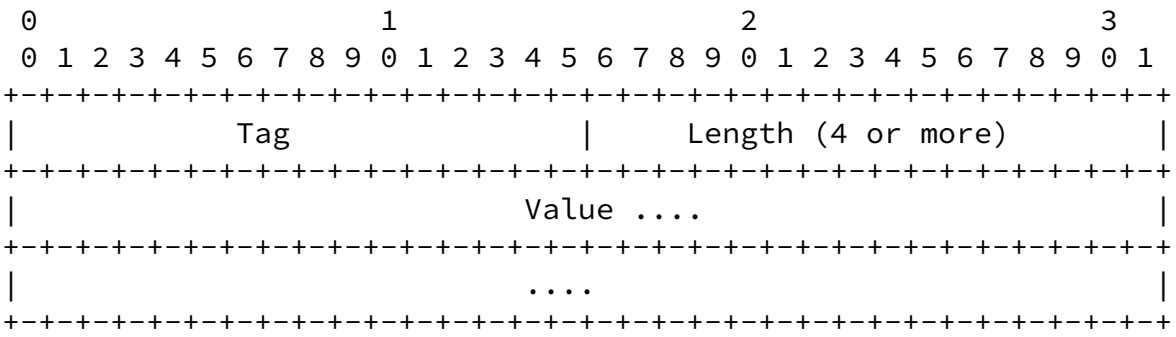
A non-TE topology specifically refers to the network that does not support TE. Control protocols such as OSPF may be run on the non-TE

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, strictly stands for an object in the form of Tag-Length-Value. However, this term is also used in the document, at times, to simply refer a Traffic Engineering attribute of a TE-node or TE-link.

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.




## 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. Motivation and Implicit assumptions for the TE extensions

The motivation behind the OSPF-TE described in this document is to dynamically discover the TE-network topology, devise a scalable flooding methodology and benefit from the hierarchical area





organization and other techniques of the native OSPF. The result

would be the ability to build an abstract view of a network topology with all the traffic engineering characteristics.

With traditional OSPF, the goal is to build a forwarding table to reach various subnets in the IP network with least-cost as the basis. However, the goal of OSPF-TE is to determine a circuit path (that can be pinned-down for a desired duration) meeting a certain set of traffic engineering criteria. Further, the circuit path could consist entirely of nodes and links that do not carry IP traffic.

The following assumptions are made throughout the document for the discussion of OSPF-TE.

1. Interfaces on an OSPF-TE node may be characterized as those that can terminate (i.e., send and receive) IP packet data and those that wont. Both types of links can be part of a traffic engineered network. Needless to say, the OSPF-TE protocol can only be enabled on interfaces that support IP packet data termination. As such, the control network over which TE LSAs are exchanged may be constituted of a combination of non-TE links and TE links that also permit non-TE packet traffic.

2. Unlike traditional OSPF, OSPF-TE protocol must be capable of advertising link state of interfaces that are not capable of handling 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 links - say, the packet terminating interfaces. Yet, the TE LSDB (LSA database) should be synchronized across all OSPF-TE nodes within an area.

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

3. 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 [OSPF-FL1] for flooding optimizations.

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

4. Each IP subnet on a TE-configurable network MUST have a minimum of one node with an interface running OSPF-TE protocol. This is despite the fact that all nodes on the subnet may take part in Traffic Engineering. (Example: SONET/SDH TDM ring with a single Gateway Network Element, a.k.a. GNE running the OSPF protocol, yet all other nodes in the ring are also full members of a TE circuit).

   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.

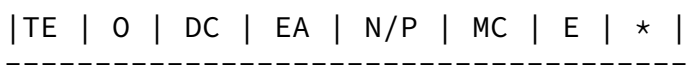
5. As a general rule, all nodes and links that may be party to a TE circuit should be uniquely identifiable by an IP address. As for router ID, a separate loopback IP address for the router, independent of the links attached, is recommended.

6. 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. TE nodes may have 2 types of link state databases - a native OSPF LSDB and a TE-LSDB. A native OSPF LSDB, constituted of native links and nodes attached to these links (i.e., non-TE as well as TE nodes), will use shortest-path criteria to forward IP packets over native links. The TE-LSDB, constituted only of TE nodes and TE links, may be used to setup TE circuit paths along the TE topology.

## 5. The OSPF Options field

A new TE flag is introduced to identify TE extensions to the OSPF. With this, the OSPF v2 will have no more reserved bits left for future option extensions. 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. See [OSPF-V2], [OSPF-NSSA] and [OPAQUE] for a description of the bits in options field. Only the TE-Bit is described in this section.






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.

## 6. 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.

### 6.1. 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 donot support TE.

## 6.2. Flooding and the Synchronization of Databases

In OSPF, adjacent routers within an area must synchronize their databases. However, as observed in [OSPF-FL1], 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. [OSPF-FL1] 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.

## 6.3. 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).

## 6.4. 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).

## 6.5. The graph of adjacencies

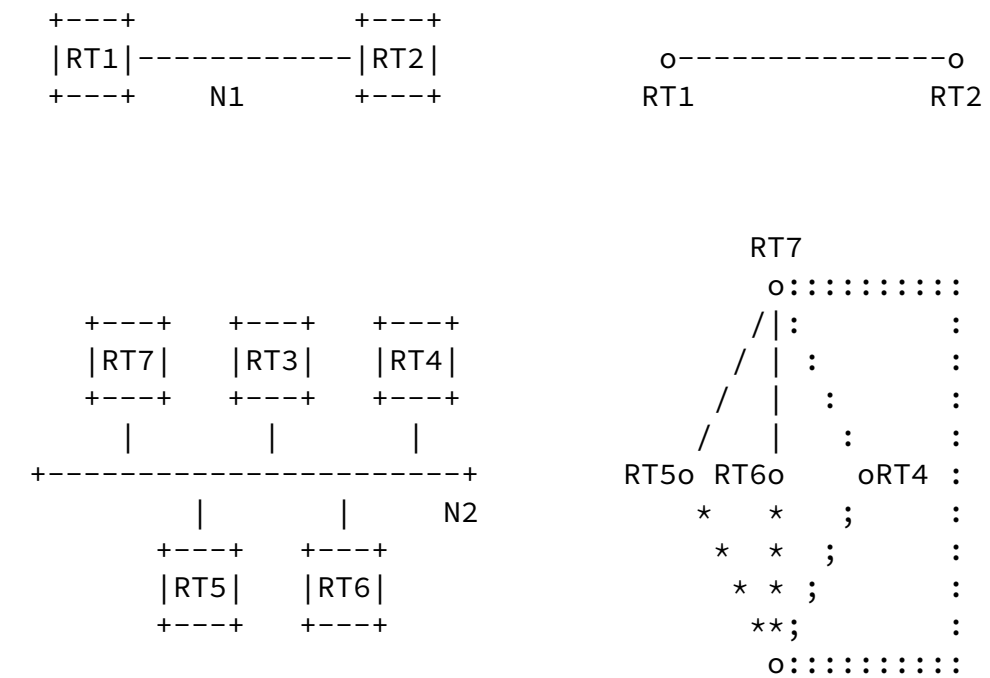
An adjacency is bound to the network that the two routers have in common. If two routers have multiple networks in common, they may have multiple adjacencies between them. The adjacency

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.







RT3

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

## 7. 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-engineered TE circuits respectively. While flooding scope for both these LSAs can be the TE-topology in the entire AS, flooding scope for the pre-engineered 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.

## 7.1. TE-Router LSA

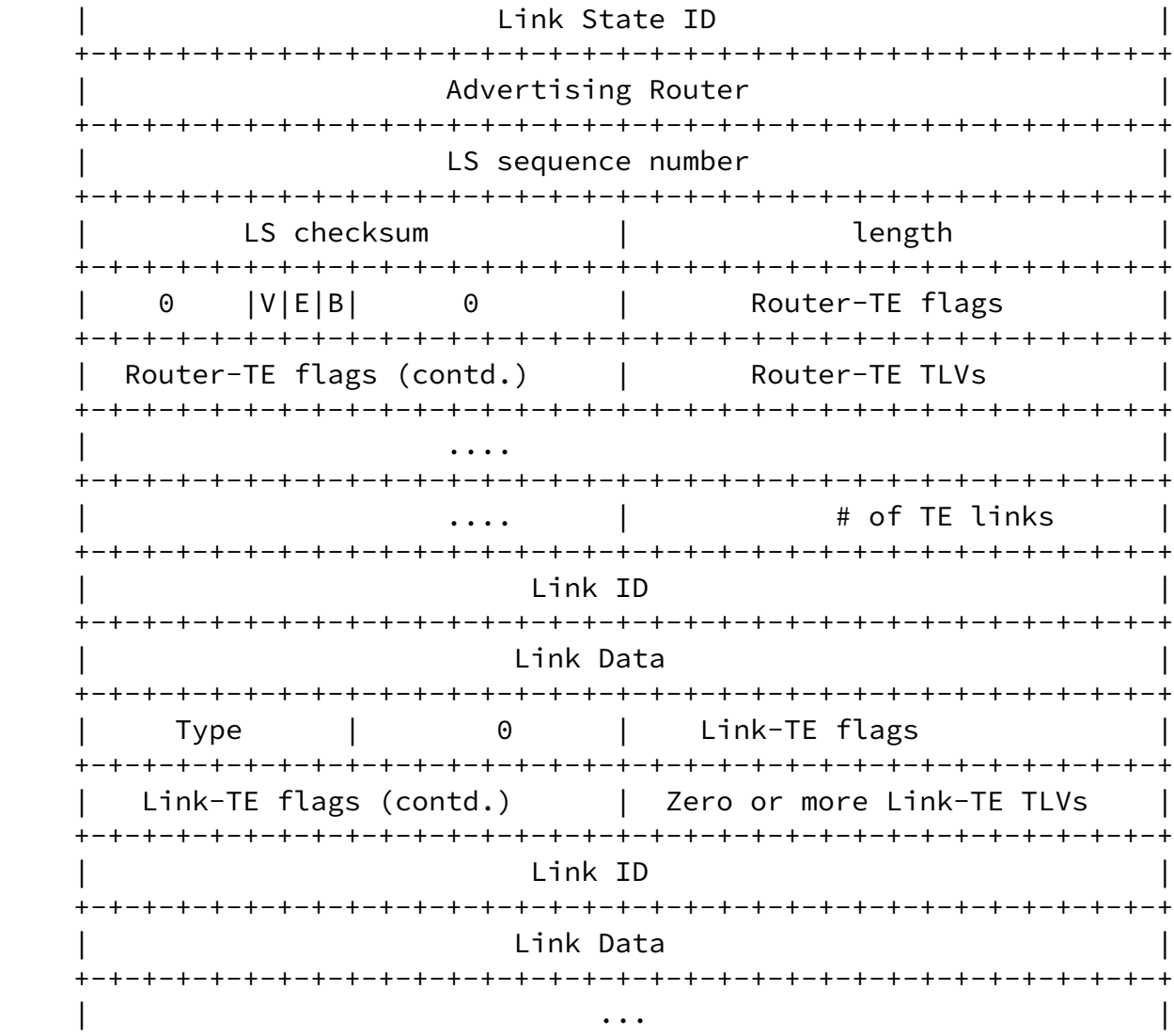
Router LSAs are Type 1 LSAs. The TE-router LSA is modeled after the router LSA with the same flooding scope as the router-LSA, except that the scope is further restricted to TE-only nodes within the area. A value of 0x81 is assigned to TE-router LSA. The TE-router LSA describes the router-TE metrics as well as the link-TE metrics of 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.

```
     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      |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

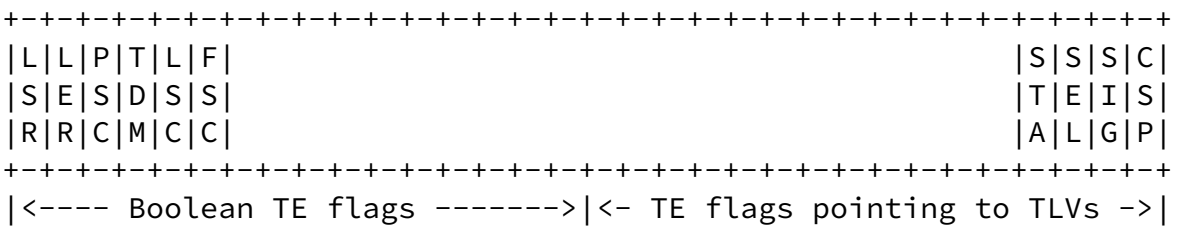



Option

In TE-capable router nodes, the TE-compliance bit is set to 1.

Router-TE flags field (TE capabilities of the router node)

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.




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.

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.

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 donot 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. |

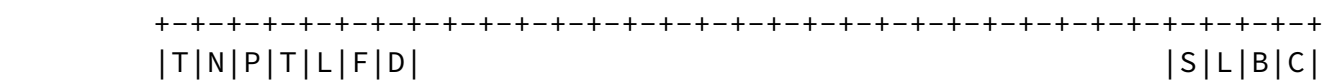
Link ID
Identifies the object that this router link connects to. Value depends on the link's Type. For a positional-ring type, the Link ID shall be IP Network/Subnet number, just as with a broadcast transit network. The following table summarizes the updated Link ID values.

| 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 |

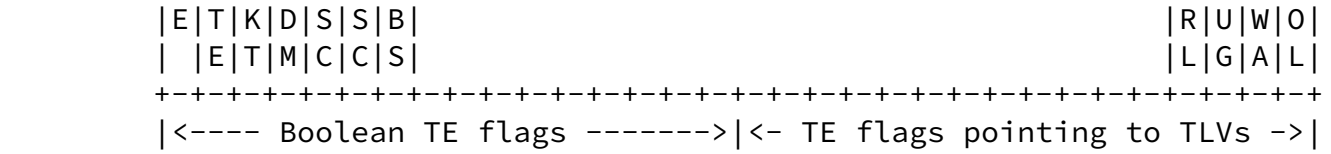
Link Data
This depends on the link's Type field. For type-5 links, this specifies the router interface's IP address.

Link-TE options (TE capabilities of a link)








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.

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.

## 7.2. 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.

### 7.2.1. 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 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.

## 7.3. 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 donot impact outside the area. Only the TE abstractions as befitting the external areas are advertised.

### 7.3.1. TE-Summary Network LSA (0x83)

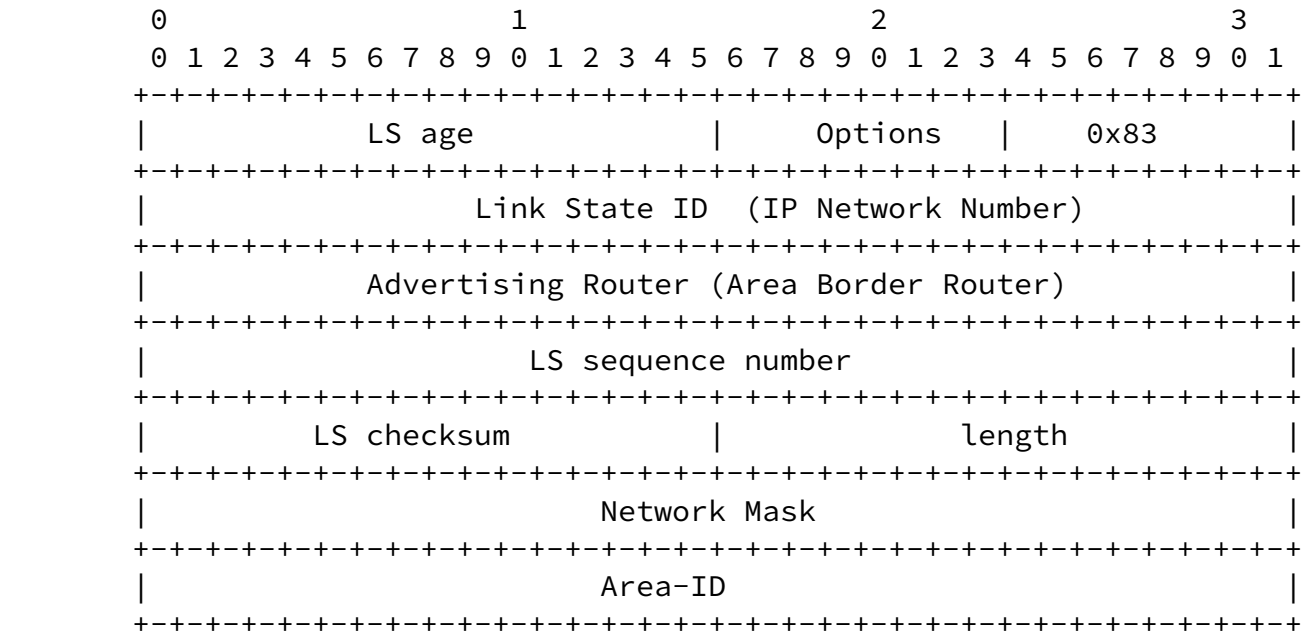
TE-summary network LSA may be used to advertise reachability of TE-networks accessible to areas external to the originating





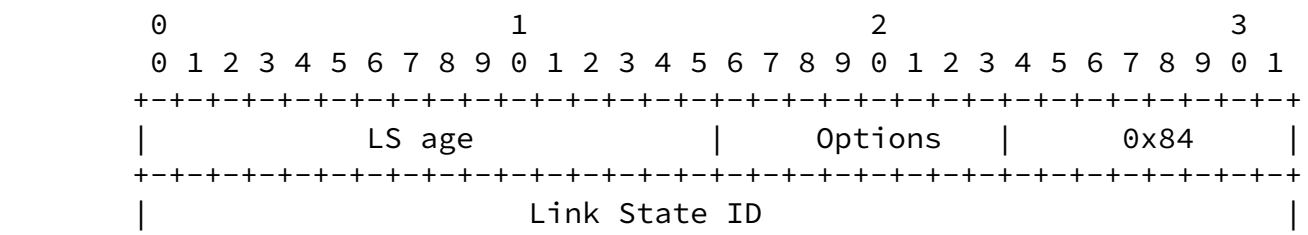
area. The scope of flooding is AS wide, with the exception of the originating area and the stub areas. For example, the TE-summary network LSA advertised by the border router of a non-backbone area is readvertised to all other areas, not just the backbone area. The area border router for each non-backbone area is responsible for advertising the reachability of backbone networks into the area.

The flooding scope of TE-summary network LSA is unlike that of the summary network LSA (type 0x03), which simply uses this as an inter-area exchange of network accessibility and limits the flooding scope to just the backbone area.



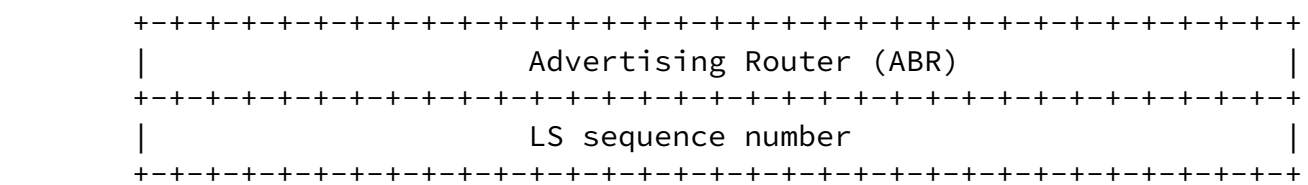

### 7.3.2. TE-Summary router LSA (0x84)

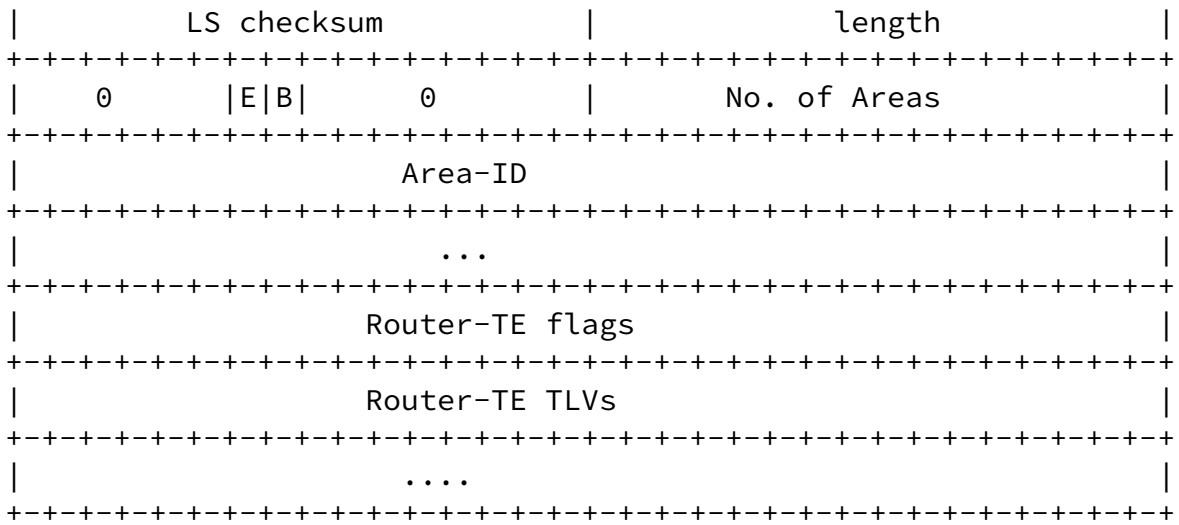
TE-summary router LSA may be used to advertise the availability of Area Border Routers (ABRs) and AS Border Routers (ASBRs) that are TE capable. The TE-summary router LSAs are originated by the Area Border Routers. The scope of flooding for the TE-summary router LSA is the entire AS, with the exception of the non-backbone areas the advertising ABRs belong to.










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

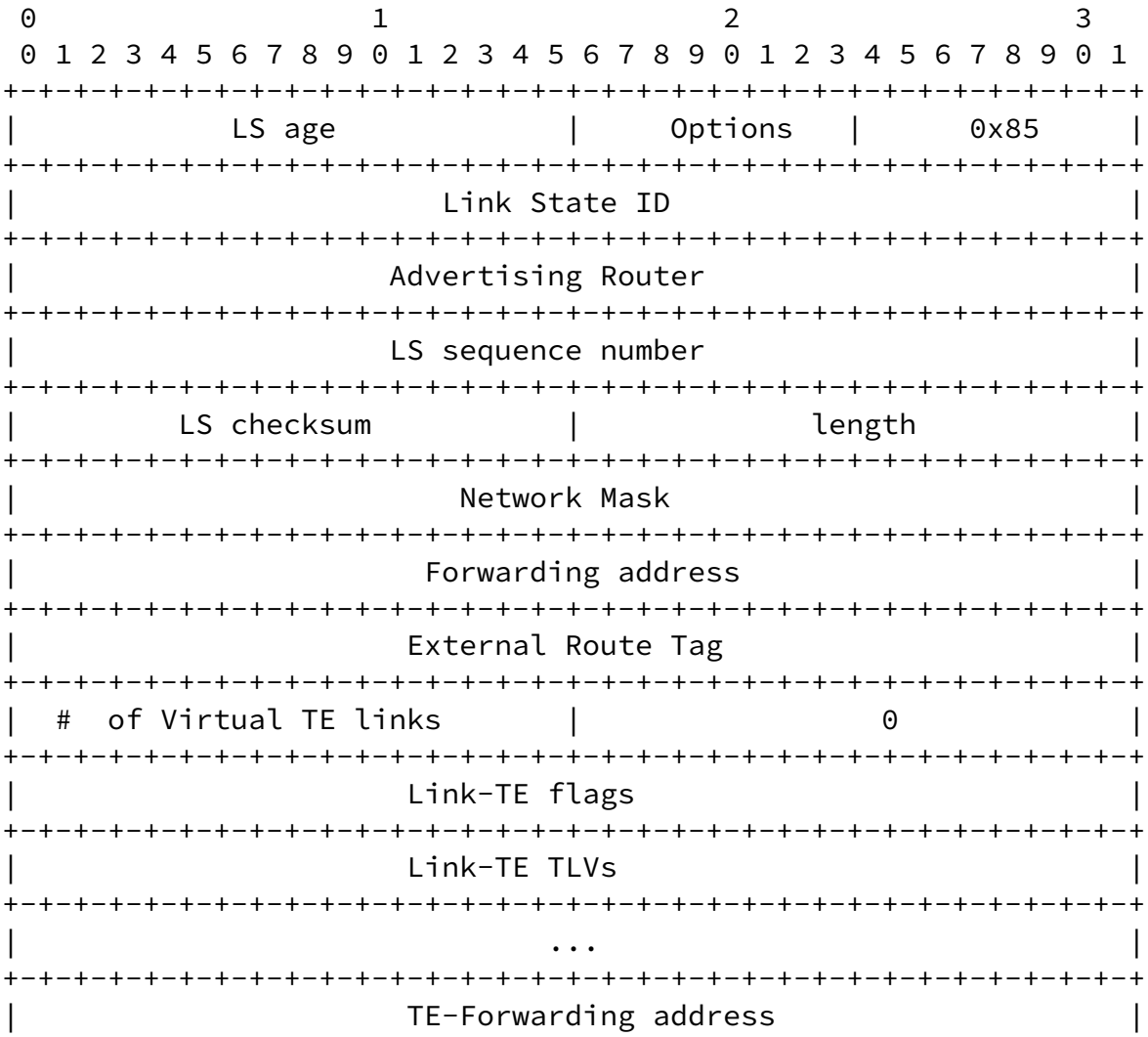
border router (B is for Border)

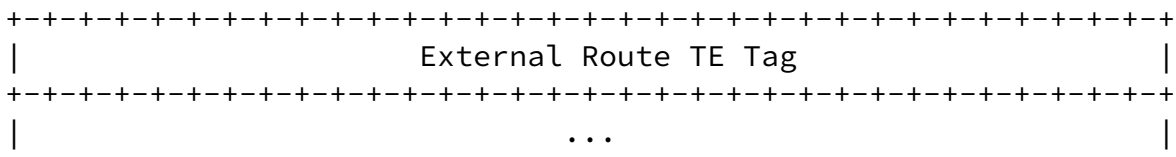
Router-TE flags,
Router-TE TLVs  (TE capabilities of the link-state-ID router)

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.

## 7.4. 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-engineered 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.






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.

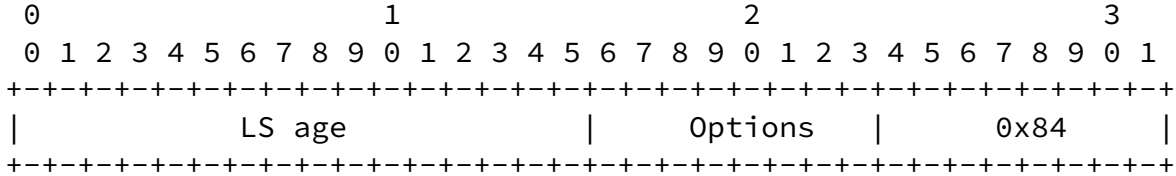
Link-TE flags,
Link-TE TLVs
The TE attributes of this route. These fields are optional and are provided only when one or more pre-engineered circuits can be specified with the advertisement. Without these fields, the LSA will simply state TE reachability info.

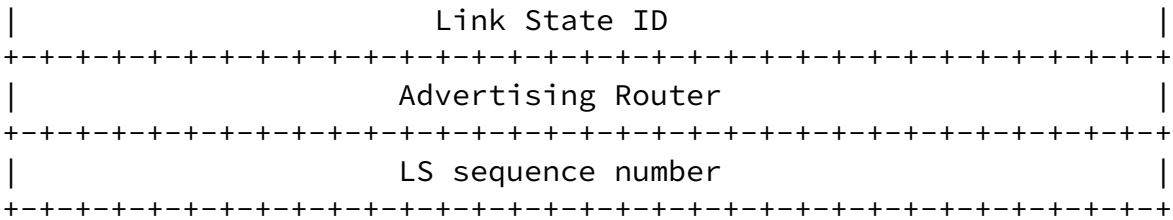
Forwarding address
Data traffic for the advertised destination will be forwarded to this address. If the Forwarding address is set to 0.0.0.0, data traffic will be forwarded instead to the LSA's originator (i.e., the responsible AS boundary router).

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.

## 7.5. TE-Circuit-paths LSA (0x8C)

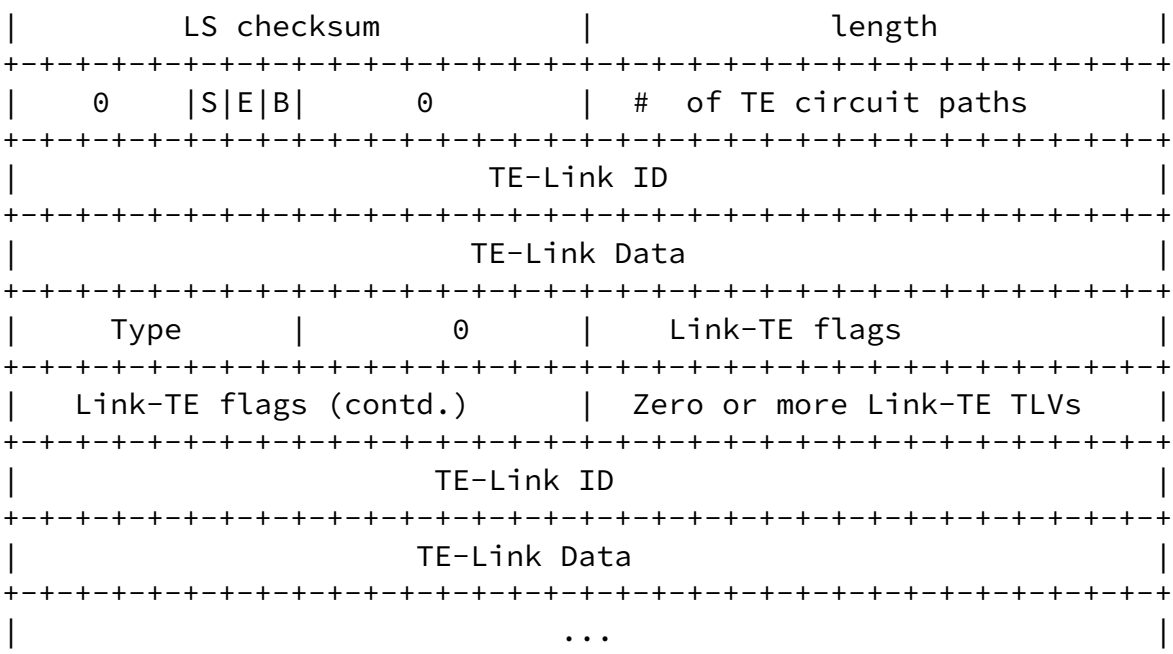
TE-Circuit-paths LSA may be used to advertise the availability of pre-engineered TE circuit path(s) originating from any router in the network. The flooding scope may be Area wide or AS wide.










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

TE-circuit-path(s) flags

Bit S - 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)

No. of Virtual TE Links
This indicates the number of pre-engineered TE links between the advertising router and the router specified in the link state ID.

TE-Link ID
This is the ID by which to identify the virtual link on the advertising router. This can be any private interface index or handle that the advertising router uses to identify the pre-engineered TE virtual link to the ABR/ASBR.

TE-Link Data
This specifies the IP address of the physical interface on the advertising router.



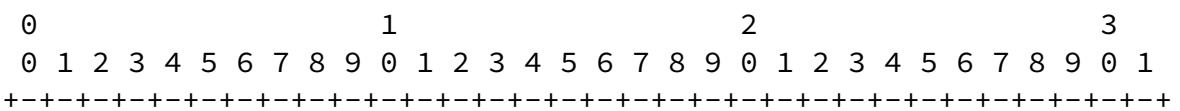


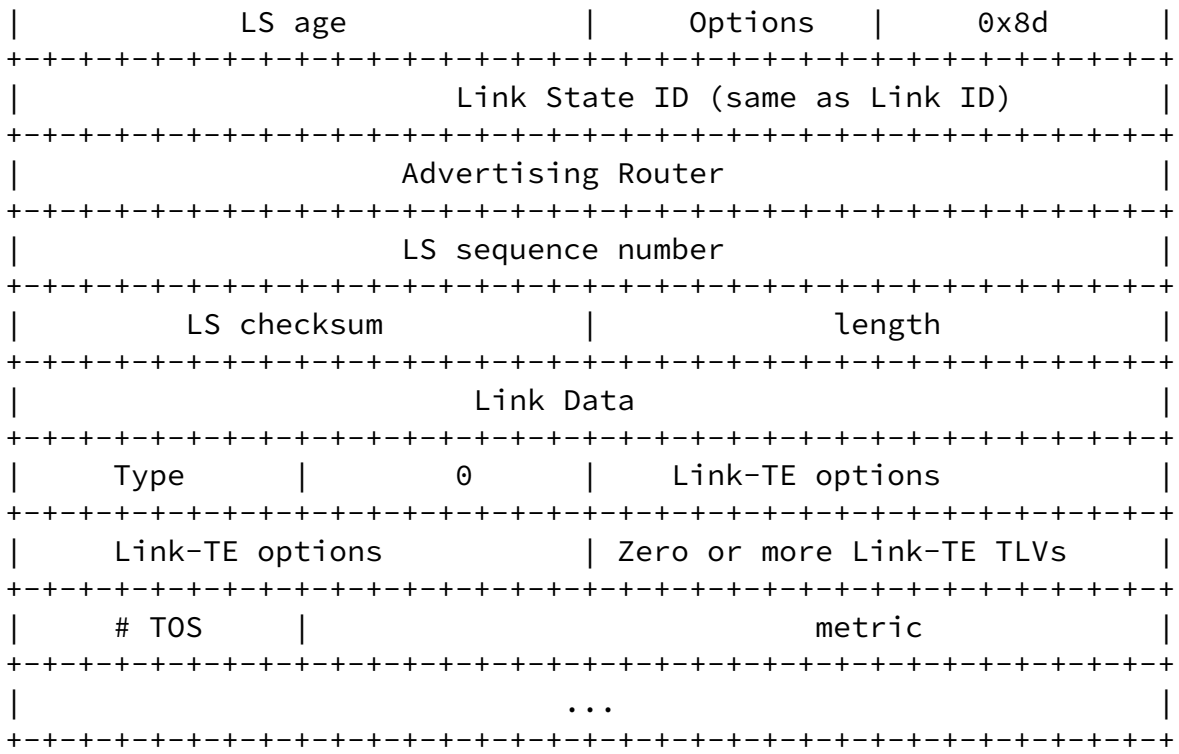
## 7.6. TE-Link-Update LSA (0x8d)

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 more changes and 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. Hence, the new TE-link-update LSA, that advertises link specific updates alone.

The TE-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.

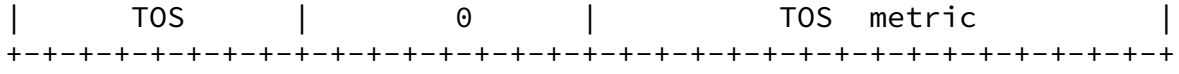
Below is the format of the TE-link update LSA.









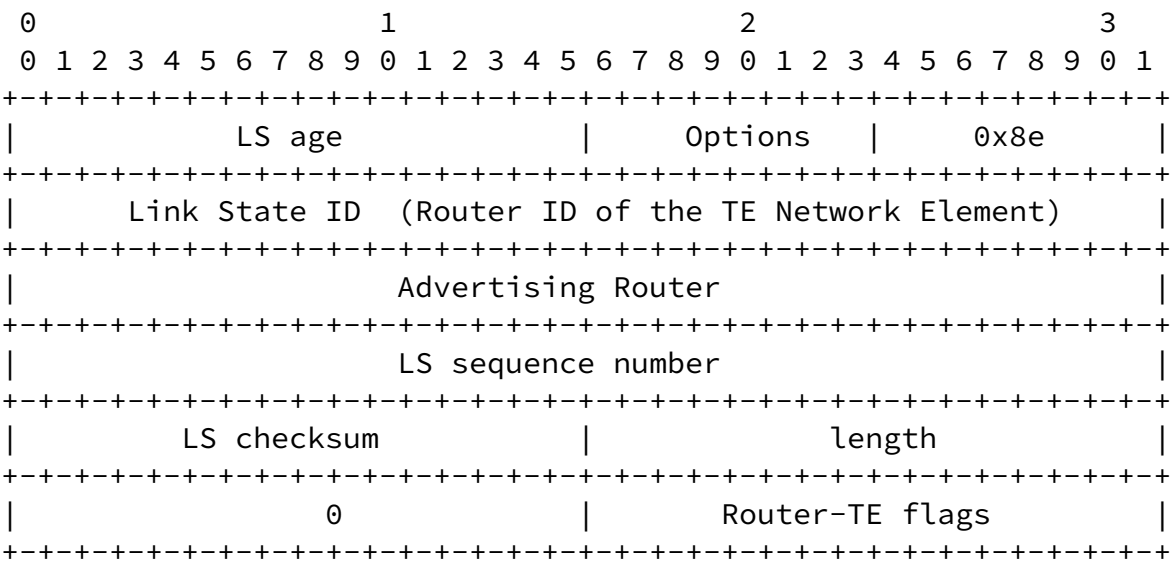

Link State ID
This would be exactly the same as would have been specified as as Link ID for a link within the router-LSA.

Link Data
This specifies the router ID the link belongs to. In majority of cases, this would be same as the advertising router.

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.

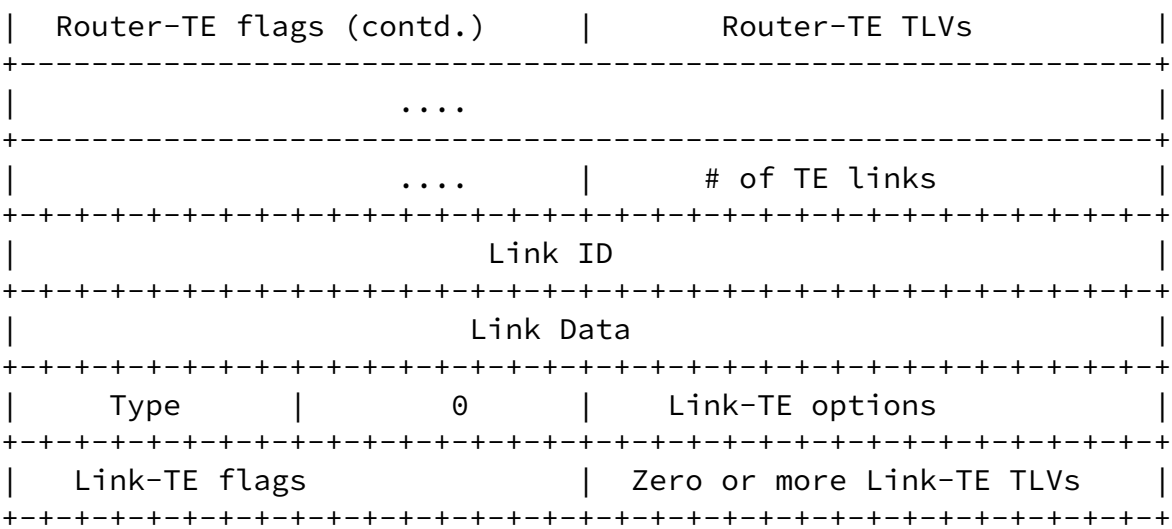
## 7.7. TE-Router-Proxy LSA (0x8e)

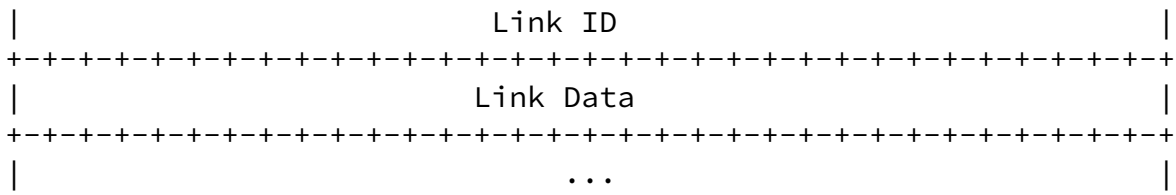
This is a variation to the TE-router LSA in that the TE-router LSA is not advertised by the network element, but rather by a trusted TE-router Proxy. This is typically the scenario in a non-packet TE network, where some of the nodes donot have OSPF functionality and count on a helper node to do the advertisement 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.










## 8. Link State Databases

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 control (non-TE) 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.

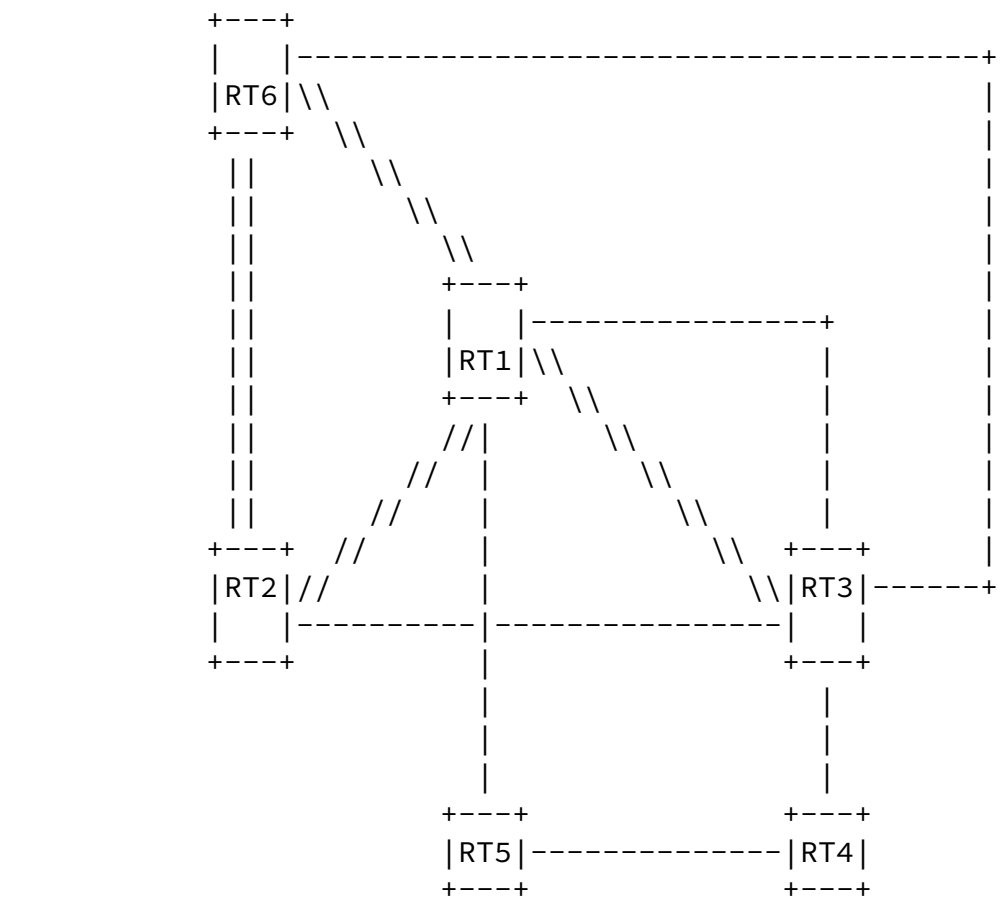
The TE-LSDB is constituted only of TE nodes and TE links. A variety of CSPF algorithms may be used to dynamically setup TE circuit paths along the TE network. A new TE-neighbors data structure will be used to flood TE LSAs along the TE-only topology. Clearly, the the TE nodes will need the control (non-TE) 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.




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 |

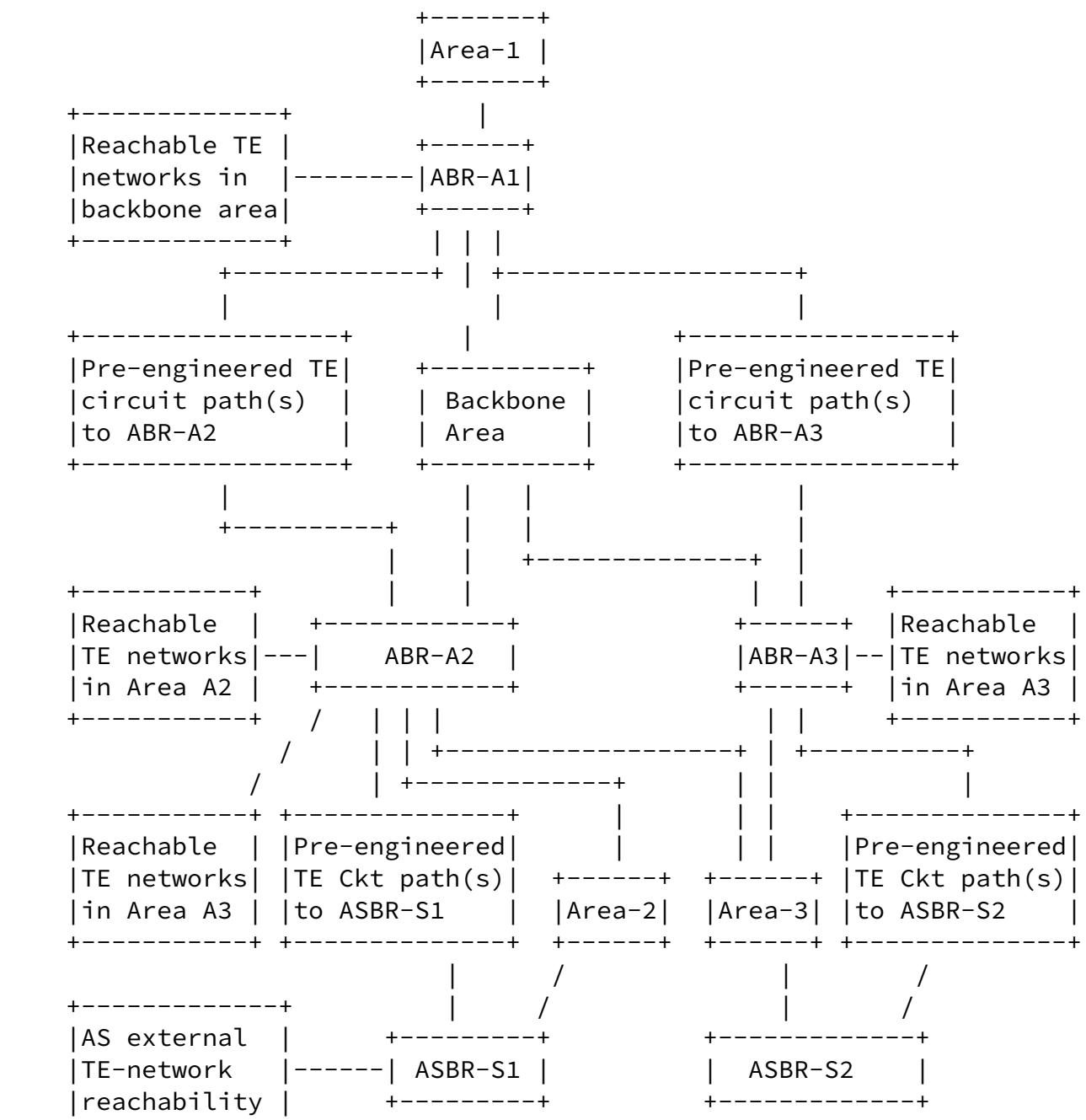
Nodes such as RT1 will have two LSDBs, a native LSDB and a TE-LSDB to reach native and TE networks. The TE LSA updates will not impact non-TE nodes RT4 and RT5.

## 9. 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-engineered 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 backbone area
ABR-A1
Pre-engineered TE circuit path(s) to ABR-A2
Backbone Area
Pre-engineered TE circuit path(s) to ABR-A3
Reachable TE networks in Area A2
ABR-A2
ABR-A3
Reachable TE networks in Area A3
Reachable TE networks in Area A3
Pre-engineered TE Ckt path(s) to ASBR-S1
Area-2
Area-3
Pre-engineered TE Ckt path(s) to ASBR-S2
AS external TE-network reachability
ASBR-S1
ASBR-S2

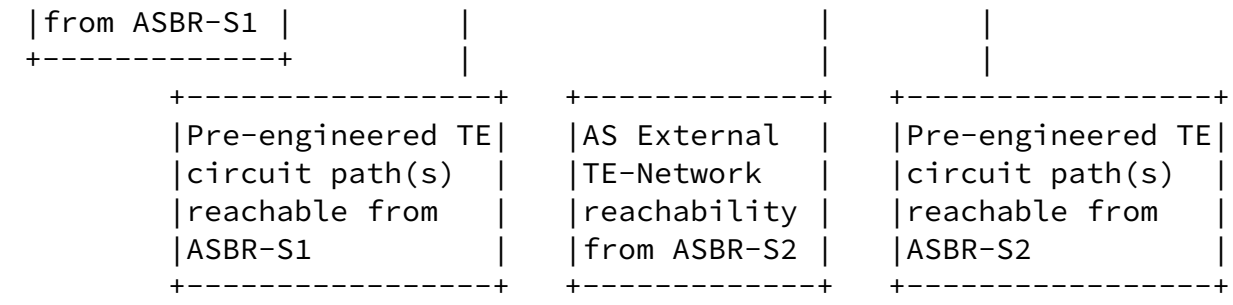



Figure 9: Inter-Area Abstraction as viewed by Area-1 TE-routers





## 10. Changes to Data structures in OSPF-TE nodes

### 10.1. 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.

### 10.2. 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 [OSPF-FL1].

### 10.3. 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 [OSPF-FL1].

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 donot 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.

## 11. Motivations to this approach

Use of TE LSAs bring substantial benefits over using Opaque LSAs as described below. These benefits cannot be retrofitted into Opaque LSAs due to fundamental scalability limitations of the Opaque-LSA approach.

The primary motivation behind the TE-LSA model is that the approach is clean (clean separation of LSDB between TE vs non-TE networks), scalable (across more than one OSPF area), unified (for packet and non-packet networks alike), efficient (efficient flooding algorithm) and SLA enforceable. The model proposed also provides the right framework for future enhancements.

### 11.1. TE flooding isolated to TE-only nodes

A TE network can generate a large number of LSA updates due to the many state changes the TE links undergo dynamically. For example, bandwidth assignment on a TE link for a specific circuit path setup will mandate that the change in bandwidth availability be communicated network wide. While such frequent link state updates is reasonable for an OSPF-TE node, neither the frequency nor the content of TE link state is desirable for native OSPF nodes. This can be a considerable interruption to non-TE nodes in a network that is constituted of multiple types of nodes and links





(ex: A network constituted of packet routing nodes/links and SONET network ADMs/links, A packet-network where the ratio of TE nodes to non-TE nodes is quite considerable).

The wider the flooding scope (and number of TE nodes), 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 multiple links on which the LSAs tried to converge. By restricting the flooding of TE LSAs to TE-only nodes within a TE topology, we obviate any TE based processing for non-TE nodes.

The flooding topology for opaque LSAs makes no distinction between TE and native OSPF nodes. In a network where the TE and native

nodes coexist, a native OSPF router would be bombarded with opaque LSAs. It is possible for the native OSPF nodes to silently ignore the unsupported Opaque LSAs (during network migration) or add knobs within implementation to decide whether or not a certain opaque LSA mandates dijkstra SPF recomputation. But, the latter can be tricky and will need non-trivial amounts of Opaque LSA processing to make the determination. In the case where routers donot validate the need to recompute, routers might end up recomputing for all new Opaque LSA advertisements. Clearly, that would be a considerable computational demand and can be cause for instability on the OSPF routers.

### 11.2. Clean separation between native and TE LSDBs

Most vendors wishing to support MPLS based TE in their network tend to migrate gradually to support the TE extensions. 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. As such, the TE network cannot be assumed to exist independently without native OSPF network even in the long term.

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.

With the new TE-LSA scheme, native OSPF nodes will keep just the native OSPF link state database. The OSPF-TE nodes will keep native as well as the TE LSDB. The native LSDB describes the control (non-TE) topology. Shortest-Path-First algorithm will be used to forward IP packets along this network. OSPF neighbors





data structure will be used for flooding along the control topology.

In the Opaque-LSA-based TE scheme, the TE-LSDB built using opaque LSAs will be required to refer the native LSDB to build the TE topology. Even with that, there is way to know the TE capabilities of the routers. The Opaque-LSA approach does not deal with TE capabilities for a router. Opaque LSAs are flooded to all nodes.

Some nodes that happen to support the TE extensions will have a hit and accept the opaque LSAs. Others that donot support will have a miss and simply drop the received Opaque LSAs. This type of hit-and-miss approach is not only disruptive, but also blind to SLA requirements on TE links.

### 11.3. Scalability across a hierarchical Area topology

Use of TE LSAs for inter-area communication is clearly superior to using Opaque LSAs with AS wide scoping. Without revealing the TE nodes and characteristics of the attached links, an Opaque LSA (type 11) simply does not disseminate reachability of TE networks and nodes outside the area. Stated differently, Use of opaque LSA can, work at best, for a single area AS.

Providing area level abstraction and having this abstraction be distinct for TE and native topologies is a necessity in inter-area communication. When the topologies are separate, the area border routers can advertise different summary LSAs for 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, could 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-LSAs are suitable neither for content nor for flooding scope in the context of inter-area communication. The flooding boundaries of Opaque LSAs make the approach suitable at best to single-area topologies. For example, Opaque LSAs cannot support the flooding scope of TE-summary-networks. Opaque LSAs (AS-wide scope) will be unable to restrict flooding in its own originating area. Opaque LSAs are also not adequate to establish TE peering relationship with neighbors.

### 11.4. Usable across packet and non-packet TE networks

In a peer networking TE model, you are likely to want different





types of TE information flooded by various nodes, as they are

heterogenous and will remain that way. The TE LSA based approach offers a single set of LSAs that may uniformly be used across packet and non-packet nodes and links. Once a link is declared as TE, the TE properties advertised of the link can be link specific, but all advertisements would use the same LSA format.

Implementations reusing the opaque LSA with GMPLS extensions are burden for the routers that do not need it. Clear separation (as proposed here) between TE and native LSAs and having independent flooding scopes for native and TE state information will be extremely useful in inheriting the right set of LSAs for the right application (i.e, TE vs native).

### 11.5. SLA enforceable network modeling

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 effectively rendering the TE link ineffective for TE purposes. Separating the two topologies (as advocated by this document with new TE LSAs and TE option flag) ensure that the SLA objectives on TE links are properly met.

### 11.6. Framework for future extensibility

The approach outlined provides a framework for future extensibility based on service provider needs.

There may be many types of information that should not be disseminated along the Opaque LSA flooding boundaries. Take for example, the TE-summary network LSA. This LSA does not follow the scope of an area or an AS, but something in between. As a general rule, the proposed framework can be extended to define newer TE LSAs with a suitable flooding scope.

Having a clean framework which argues for having different link state databases for different applications on the same network will provide the right forum for future extensibility. Just as the TE LSDB may be used for MPLS TE application, a different type of LSDB may be used for yet another type of application (such as QOS based IP forwarding) using the same IP network.

lastly, an opaque LSA is restricted in the format in which it can express different types of data. Everything should be expressible in the form of a TLV. Summary-TE-networks-from each Area, TE-ABR routers, TE-ASBR routers, TE-AS-External-networks, TE-Router Capabilities, TE-link updates, Pre-engineered-TE-Links - All of

these data have to be engineered to be expressible in a TLV form with one or more sub-TLVs. Some of the TLVs will be required to be mandatory. Some would be expected to appear in a pre-specified order and some are expected to appear just once in the LSA. TLVs should not be a panacea for all kinds of TE data. TLVs are generally more difficult to process and debug than fixed format messages.

Opaque LSAs demand more processing to assimilate into topology abstraction. A single Opaque LSA type is bent in many ways (using a variety of TLVs) to update the native OSPF topology abstraction nodes. Not a framework that could be easily extended for future applications.

### 11.7. Real-world scenarios benefiting from this approach

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

1. Multi-area network.

2. Single-Area networks - The TE links are not cannibalized by the non-TE routers for SPF forwarding.

3. Credible SLA enforcement in a (TE + non-TE) packet network. Ability to restrict flooding to some links (say, non-TE links) ensures the service provider is able to devote the entire bandwidth of a TE-link for TE circuit purposes. This makes SLA enforcement credible.

4. For a non-Packet TE network, the Opaque-LSA-based-TE scheme is not adequate to represent
   (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).

## 12. Transition strategy for implementations using Opaque LSAs

Below is a strategy to transition current implementations to adapt the new TE LSA scheme in a gradual fashion. Implementations using Opaque-LSAs can take the following steps to accomplish this. Once the OSPF-TE is completely transitioned to using the new TE LSAs as described here, the TE network can reap the full benefits of the scheme. Amongst other things, packet and non-packet networks may be combined with ease into a unified network. As such, the MPLS

traffic engineering can be based on either of the overlayed or peer models espoused in [GMPLS-TE].





1. Restrict the use of Opaque-LSAs for 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 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.

4. Replace Opaque LSAs with TE LSAs within the area as well.

## 13. IANA Considerations

### 13.1. TE-compliant-SPF routers Multicast address allocation

### 13.2. New TE-LSA Types

### 13.3. New TLVs (Router-TE and Link-TE TLVs)

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

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

#### 13.3.2. MPLS-Signaling protocol TLV (Tag ID = 3)

- RSVP-TE signaling
- LDP signaling
- CR-LDP signaling

#### 13.3.3. Constraint-SPF algorithms-Support TLV (Tag ID = 4)

- CSPF Algorithm Codes.

## 14. Acknowledgements

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

## 15. 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 sychronizations 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", draft-ietf-mpls-generalized-signaling-03.txt, work in progress.

[RSVP-TE] Awduche, D.O., L. Berger, Der-Hwa Gan, T. Li, V. Srinivasan and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", Work in progress, draft-ietf-mpls-rsvp-lsp-tunnel-08.txt

[CR-LDP] Jamoussi, B. et. al, "Constraint-Based LSP Setup using LDP", draft-ietf-mpls-cr-ldp-05.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.

[OSPF-FL1] Zinin, A. and M. Shand, "Flooding Optimizations in link-state routing protocols", work in progress, <draft-ietf-ospf-isis-flood-opt-01.txt>

[OSPF-FL2] Moy, J., "Flooding over a subset topology", <draft-ietf-ospf-subset-flood-00.txt>, work in progress.

[OPQLSA-TE] Katz, D., D. Yeung and K. Kompella, "Traffic Engineering Extensions to OSPF", work in progress, <draft-katz-yeung-ospf-traffic-05.txt>





Authors' Addresses

Pyda Srisuresh
Kuokoa Networks, Inc.
2901 Tasman Dr., Suite 202
Santa Clara, CA 95054
U.S.A.
EMail: srisuresh@yahoo.com

Paul Joseph
Jasmine Networks
3061 Zanker Road, Suite B
San Jose, CA 95134
U.S.A.
EMail: pjoseph@jasminenetworks.com