OIF OPTICAL INTERNETWORKING FORUM

Recovery Amendment to E-NNI 2.0 OSPFv2-based Routing

OIF-ENNI-OSPF-02.2

February 4, 2014

Implementation Agreement created and approved
by the Optical Internetworking Forum
www.oiforum.com

The OIF is an international non profit organization with over 100 member companies, including the world's leading carriers and vendors. Being an industry group uniting representatives of the data and optical worlds, OIF's purpose is to accelerate the deployment of interoperable, cost-effective and robust optical internetworks and their associated technologies. Optical internetworks are data networks composed of routers and data switches interconnected by optical networking elements.

With the goal of promoting worldwide compatibility of optical internetworking products, the OIF actively supports and extends the work of national and international standards bodies. Working relationships or formal liaisons have been established with CFP-MSA, COAST, Ethernet Alliance, Fibre Channel T11, IEEE 802.1, IEEE 802.3, IETF, InfiniBand, ITU-T SG13, ITU-T SG15, MEF, ONF, Rapid I/O, SAS T10, SFF Committee, TMF and TMOC.

For additional information contact:
The Optical Internetworking Forum, 48377 Fremont Blvd.,
Suite 117, Fremont, CA 94538
510-492-4040 ☞ info@oiforum.com

www.oiforum.com

IA xxxxxx

Working Group:　　　**Networking and Operations**

TITLE:　　　**Recovery Amendment to E-NNI 2.0 OSPFv2-based Routing**

SOURCE:

TECHNICAL EDITOR

Remi Theillaud
Marben Products
176 rue Jean Jaures
92800 Puteaux
France
Phone: +33 (1) 79 62 10 22
Email: remi.theillaud@marben-products.com

WORKING GROUP CHAIR

Evelyne Roch
Huawei Technologies
303 Terry Fox Drive Suite 400
Ottawa, ON 　　K2K 3J1
Canada
Phone: +1 613 595 1900 x1612
Email:evelyne.roch@huawei.com

---

1   <u>Table of Contents</u>

2 <u>List of Figures</u>

3 <u>List of Tables</u>

4 Introduction

As Automatically Switched Optical Networks (ASONs) are deployed into new and existing networks, it cannot be assumed that such networks will be homogeneous (e.g., with respect to transport technologies, vendors, or approach to management and control). This is true even within a single carrier's network. To support deployment of an optical control plane into a heterogeneous environment, it is essential to introduce and support the concept of control domains, and in particular, the specification of the signaling and routing information exchanged between such domains.

A control domain is an architectural construct from ITU-T Recommendation [G.8080] that provides for encapsulation and information hiding, and the characteristics of the control domain are the same as those of its constituent set of distributed architectural components. The E-NNI reference point is defined to exist between control domains. The nature of the information exchanged between control domains across the E-NNI reference point captures the common semantics of the information exchanged amongst its constituent components, while allowing for different representations inside each control domain. Control domains are generally derived from architectural component types that serve a particular purpose; e.g., signaling control domains or routing control domains. Typically, signaling and routing control domains are expected to be congruent within ASON networks. The E-NNI reference point becomes an E-NNI signaling and routing interface when instantiated by signaling and routing protocols.

Figure 1 illustrates a simple example of a control plane subdivided into routing control domains interconnected by routing E-NNI interfaces. This example shows different domains potentially utilizing different I-NNI routing protocols communicating across the E-NNI interfaces by using a common set of signaling and routing protocols.
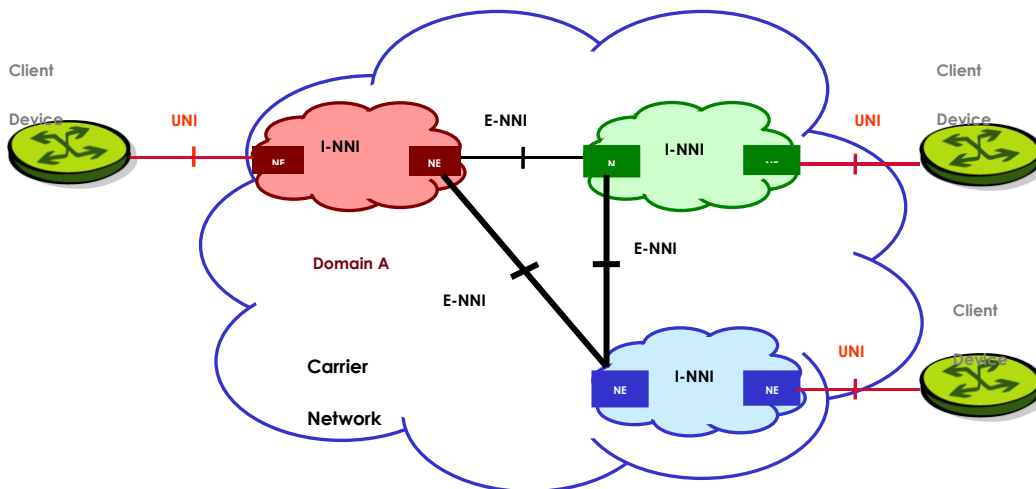
Figure 1: Example of Control Plane Configuration with Different Routing Control Domains

4.1 Problem Statement

The advent of the automatic switched transport network has necessitated the development of interoperable procedures for requesting and establishing dynamic connection services across heterogeneous, multi-domain networks. The development of such procedures requires the definition of:

- Control domains and associated reference points (E-NNI, I-NNI, UNI)

- Services offered by the transport network across control domains

- Routing protocols used to disseminate advertisements across E-NNI interfaces

This document addresses OSPF-based routing information exchange to support ASON routing architecture and requirements for the OIF E-NNI routing interface. Some of the requirements support interoperability and scalability in a multi-domain environment, diverse control plane characteristics within individual domains, and ASON-specific characteristics such as per-layer link capacity.

4.2 Scope

The scope of this implementation agreement is to define the E-NNI Routing Interface based on the [G.8080] routing architecture, with details as defined in [G.7715] and [G.7715.1], as applied to OSPFv2.

ITU-T has defined an ASON routing architecture and requirements for link-state protocols [G.7715.1], but did not specify how existing link state protocols, such as OSPF and ISIS, can fulfill such architecture and requirements.

The IETF CCAMP working group has defined OSPF-TE extensions (see [RFC4642] and [RFC5787]) to address the ITU-T ASON architecture, and ITU-T requirements captured in [RFC4258].

The base protocol used by this document is OSPFv2 [RFC2328] with extensions for Traffic Engineering [RFC3630] and GMPLS [RFC4202, RFC4203]. This document specifies the requirements on and use of OSPFv2-TE as an E-NNI routing protocol among multiple domains. This document relies as much as possible on IETF OSPFv2 protocol specifications (including extensions defined in [RFC5787]).

E-NNI Routing 2.0 provided support for SONET/SDH, OTN (ODUk) - [G.709] (2003-03), and Ethernet (EPL and EVPL services). The Multilayer amendment extensions [OIF-ENNI-OSPF-02.1] added advertisement of adaptation capabilities between those layers in support of the multilayer amendment [OIF-ENNI-ML-AM-01.0]. All extensions for the Multilayer amendment are in blue colored font to help the reader identify the changes. A summary of all changes can also be found in section 20.

This advertisement also defines VCAT layer advertisements required for transitional link advertisements. This allows path computation to determine a valid path across a VCAT layer instead of relying on signaling to determine whether VCAT capabilities are available once the path is already computed. Note that this amendment limits VCAT advertisements to transitional links, with VCAT layer routing for future study.

Backward compatibility with E-NNI Routing 2.0 is discussed in section 15.

This amendment specifies additions to E-NNI Routing 2.0 in support of the recovery amendment [OIF-ENNI-REC-AM-01.0]. All additions for the Recovery amendment are in green colored font to help the reader identify the changes.

The following areas are NOT covered within this document:

- Requirements for inter-carrier interfaces. The extensions in this document were defined within the framework of intra-carrier link state routing protocol requirements for ASON.

- Protocol extensions required to support multi-level hierarchy. This document only discusses the target architecture for multi-level hierarchy. Per the ITU-T G.8080 routing architecture with details as defined in [G.7715] and [G.7715.1], the routing infrastructure in ASON supports hierarchy using a link-state-based protocol at each routing level. The OSPF-TE operation at each routing level is

independent, i.e., it does not interfere with the operation of the routing protocol at other routing levels. However, some of the routing information at a given hierarchical level can be fed up to the next hierarchical level to be advertised in the parent routing area, and at the same time, the routing information at a higher level can be fed down to a lower level of hierarchy. Alternatively, routing information can be accessed by other means outside of routing protocol mechanisms. Collectively this provides a powerful mechanism for scaling of the routing protocol to large networks.

- Support for G.709ed3 (2009-12). It is anticipated that future IAs will cover the necessary extensions.

Private extensions have been defined in the forms of (sub-) TLVs to accommodate the requirements as defined in [G.8080], [G.7715], and [G.7715.1].

4.3   Relationship to Other Standards Bodies

This document, to the maximum extent possible, uses standards and specifications already available from other organizations. Specifically,

- The SDH/SONET service definitions are based on ITU-T specification [G.707] and ANSI specification [T1.105].

- The OTN (ODUk) service definitions are based on ITU-T specification [G.709] (2003-03).

- The Ethernet definitions are based on [IEEE802.3].

- The routing protocol requirements are based upon [G.7715] and [G.7715.1], and their normative specifications are based on IETF [RFC2328], [RFC3630], [RFC4203] and [RFC5787].

- The security and logging methods in this document are based on the OIF's profiles of IPsec and syslog as defined by the IETF (see section 6.4).

This version of the implementation agreement also documents private extensions, codepoints and formats of these extensions based on the E-NNI 1.0 Routing implementation agreement.

It is the intent of OIF to develop E-NNI protocols in close alignment with ITU-T Recommendations, and foundation IETF RFCs. As such, the OIF has aligned formats with IETF and ITU-T standard specifications where possible and will continue to pursue alignment with standards in its future work. As additional standard specifications become available that address functions included in this Implementation Agreement, additional revisions for further alignment with these standards will be considered.

4.4 Merits to OIF

The E-NNI OSPFv2 Routing 2.0 implementation agreement is a key step towards the implementation of an open inter-domain interface that allows offering dynamic setup and release of various services. This activity supports the overall mission of the OIF.

4.5 Working Groups

Networking and Operations Working Group

Carrier Working Group

Interoperability Working Group

OAM&P Working Group

4.6 Document Organization

This document is organized as follows:

- Section 1: Introduction and Scope of the Document

- Section 5: Terminology and Abbreviations

- Section 6: Basic Routing Components

- Sections 7 through14: ASON-based Routing Requirements and Extensions

- Section 15: Compatibility with E-NNI Routing 2.0

- Section 16: References

- Appendices

  - Section 17: Appendix I: E-NNI OSPF-based Routing with a Single Hierarchical Level

  - Section 18: Appendix II: Architecture for Operation with Multiple Hierarchical Levels

  - Section 19: Appendix III – Use of SNPP Aliases for Hierarchy

4.7   Keywords

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

5   Terminology and Abbreviations

5.1   Definitions

The following terms are used in this implementation agreement.

Control Domain   This terminology is adopted from ITU-T [G.8080]. A type of transport domain where the criterion for membership is the scope of a control plane component responsible for the transport resources within the transport domain.

Inter-domain Link   A link with endpoints in two different Routing Areas at a particular level of the routing hierarchy.

Intra-domain Link   A link with both endpoints within the same Routing Area at a particular level of the routing hierarchy.

Layer   This terminology is adopted from ITU-T [G.805]. A layer

(network) is a "topological component" that represents the complete set of access groups of the same type which may be associated for the purpose of transferring information.

Level

This terminology is adopted from ITU-T [G.8080]. A routing hierarchy describes the relationships between an RA and a containing RA or contained RAs. RAs at the same depth within the routing hierarchy are considered to be at the same routing level.

Node ID

This terminology is adopted from ITU-T [G.7715.1]. The Node ID identifies a node in the transport topology graph. A node may represent either an abstraction of a Routing Area or a subnetwork.

Protocol Controller

This terminology is adopted from ITU-T [G.8080]. The Protocol

Controller provides the function of mapping the parameters of the abstract interfaces of the control components into messages carried by a protocol to support interconnection via an interface.

RC ID

The RC ID is a unique value that identifies an RC instance. This identifier may be used by the database synchronization function for record ids.

RC PC ID

The RC PC ID is a unique value that identifies an RC Protocol Controller. As per [G.8080], the Protocol Controller takes the primitive interface supplied by one or more architectural components, and multiplexes this interface into a single instance of a protocol

RC PC SCN Address

The SCN Address where the RC attaches, via its Protocol Controller (PC), to the IP SCN. An RC may have multiple associated PCs that support the procedures and formats of specific protocols and attach to the SCN. The address referred to in this document is for the RC's OSPF PC.

Routing Area (RA)

This terminology is adopted from [G.8080]: A routing area is defined by a set of subnetworks, the SNPP links that interconnect

them, and the SNPPs representing the ends of the SNPP links exiting that routing area. A routing area may contain smaller routing areas interconnected by SNPP links. The limit of subdivision results in a routing area that contains a subnetwork.

Routing Controller (RC) This terminology is adopted from [G.7715]. The Routing Controller functional component provides the routing service interface and is responsible for coordination and dissemination of routing information.

Routing Control Domain This terminology is adopted from [G.8080]. A transport domain is a set of transport resources grouped according to some criteria established by operator policies. An RCD is a type of transport domain where the criterion for membership is assignment to an RC federation for the purposes of transport resource advertisement.

Signaling Control The packet network that carries control plane messages between

Network (SCN) Protocol Controllers

Signaling Controller Signaling Controller (see [G.8080])

TE Link This definition is per [RFC4203], which defines a TE link as a "logical" link that has TE properties. The TE link is logical in a sense that it represents a way to group or map the information about certain physical resources (and their properties) into the information used by Constrained SPF for path computation.

5.2   Abbreviations

The following abbreviations are used in this implementation agreement.

ASON	Automatically Switched Optical Networks
BN	Border Node
CC	Connection Controller
CD	Control Domain
CP	Connection Point
GMPLS	Generalized Multi-Protocol Label Switching
GRE	Generic Routing Encapsulation
E-NNI	External Network-Network Interface
ERO	Explicit Route Object
ID	Identifier
IETF	Internet Engineering Task Force
I-NNI	Internal Network-Network Interface
IP	Internet Protocol
IPsec	Internet Protocol Security
ISCD	Interface Switching Capability Descriptor
ITU-T	International Telecommunications Union - Telecommunications
L1VPN	Level 1 Virtual Private Network
LRM	Link Resource Manager
LSA	Link State Advertisement
NNI	Network-Network Interface
OSPF	Open Shortest Path First
PC	Protocol Controller

PCE	Path Computation Element
RA	Routing Area
RC	Routing Controller
RCD	Routing Control Domain
RP	Routing Performer
SCN	Signaling Communications Network
SCSI	Switching Capability Specific Information
SN	Subnetwork
SNP	Subnetwork Point
SNPP	Subnetwork Point Pool
SPF	Shortest Path First
SRG	Shared Risk Group
SRLG	Shared Risk Link Group
TE	Traffic Engineering
TCP	Termination Connection Point
TLV	Type/Length/Value
TNA	Transport Network Assigned Name
TTL	Time To Live
UNI	User-Network Interface
UNI-C	Client side of a UNI
UNI-N	Network side of a UNI
VLAN	Virtual Local Area Network

6 Basic Components for OSPFv2-Based E-NNI Routing

This routing implementation agreement is based on [RFC3630] but with a hierarchical operational model per [G.7715] for ASON networks as defined per G.8080. This implementation agreement uses the base OSPFv2 protocol as defined in [RFC2328], in [RFC3630] and in [RFC4203], although some additional requirements for optical transport networks are defined in the following Sections.

It should be noted that this implementation agreement does not include the use of OSPF for the maintenance of SCN topology, and as a result does not include the use of OSPF types 1-5 LSAs, path computation for IP routing or area border routers.

6.1 Basic Assumptions

This implementation agreement conforms to the routing architecture as specified for ASON in [G.7715]. It assumes that the network can be organized into a hierarchy of Routing Areas, as defined in [G.7715].

This implementation agreement implements the routing elements defined in ITU-T Recommendation [G.7715.1] for Link State Routing Protocols, using OSPFv2 as the basis. It makes use of work done in IETF on TE extensions to OSPFv2 [RFC3630], GMPLS extensions to OSPFv2 [RFC4203], and ASON routing extensions to OSPF [RFC5787], but identifies additional requirements and potential extensions as needed for ASON.

The hierarchical organization of Routing Areas used in this implementation agreement (as per [G.8080]) is orthogonal to the OSPFv2 multi-area operation defined for IP networks in [RFC2328]. Applicability of future GMPLS multi-area operations is for further study.

The purpose of this routing implementation agreement is to re-use OSPFv2-TE in networks with architecture as defined by G.8080, but it is not aimed at providing IP layer datagram routing. In addition it assumes that an IP-based control communications network or SCN, compliant with [G.7712], is in place to support communications between the various control entities.

6.2   Transport and Traffic Considerations for Routing Messages

It should be noted that sending of extraneous or invalid routing information, e.g., zero-length advertisements, should be prevented to reduce the overall traffic and processing load due to the routing protocol. Extraneous or invalid routing information SHOULD NOT be recorded in the routing database, and SHOULD NOT cause failure of the routing controller. If logging is used, this information SHOULD be logged with a higher (i.e., lower numbered) SEVERITY than Informational.

Unlike traditional IP networks where OSPF routers are physically interconnected to create adjacencies, RCs in an ASON network are most likely not topologically adjacent within the control plane, and not always one IP-hop away in the SCN topology

A number of methods are available to create one-hop adjacencies between OSPF instances in nodes that are not topologically adjacent, including a variety of tunneling methods (esp. GRE, IP-in-IP and IPsec tunnel mode), use of VLANs at layer 2, and use of OSPF virtual links. A number of associated impacts or limitations have been identified: VLANs can only be applied within SCNs consisting of a single ethernet broadcast domain; virtual links are an optional capability and currently are restricted to being part of an OSPF backbone area, which is a different topology than assumed for the E-NNI.

This implementation agreement uses the point-to-multipoint method defined below. Tunneling as described below is an alternative method; selection of a particular tunneling type is for further study.

6.2.1   Point-to-Multipoint Method

As in OSPF point-to-multipoint, all adjacencies between RCs are configured, and the OSPF Hello is not used for discovery purposes. OSPF adjacencies are allowed to be created between RCs more than one hop apart by allowing the IP TTL to be greater than 1 (as is done for OSPF virtual links).

It should be noted that the OSPF instance used for the E-NNI is providing Optical Network routing and not IP layer routing for the SCN.   As a result, the same OSPF adjacency type used for the E-NNI is independent of the actual interfaces used to connect to the SCN.

For point-to-multipoint adjacencies to operate across a multi-hop IP SCN, the IP header TTL field for Optical E-NNI Routing OSPF packets MUST be set to a value greater than 1 and SHOULD be set to a value of 255. Further, the Network mask included in OSPF Hello packets MUST be set to 0x00000000 to allow adjacencies with nodes that are not immediate neighbors.   Note: this configuration deviates from the typical configuration of OSPF for IP routing.

6.2.2   Tunneling Method

Tunneling is a commonly used technique between non-adjacent nodes, but tunneling introduces direct SCN links between non-adjacent RCs that could potentially be used for any application traffic, if the creation of the tunnel generates an entry in the node's IP forwarding table. In this case, to avoid unintended traffic routing and potential traffic looping, additional management is required to ensure that the tunnels are used only for E-NNI messages. Tunneling requires establishing appropriate tunnels between RCs, and then turning these tunnels into interfaces for Optical E-NNI related OSPF-TE instances only.

6.3   Considerations for Hierarchy and Topology Abstraction

Hierarchical routing can be used to enable the network to scale and to provide isolation between different network domains. Topology abstraction can be used to reduce the amount of information carried by the inter-domain routing protocol. When a hierarchy is created and topology abstraction is used, the externally advertised topology can be a transformed view of the actual internal topology of a contained Routing Area. This transformed view is intended specifically to provide information for computation of paths crossing the Routing Area, represented by advertisements of links and associated costs. This can impact routing performance, depending on the conditions within the Routing Area and the use of tools that provide additional routing information, e.g., a Path Computation Element as discussed below. If the available bandwidth in a domain is large compared to the average service request, node level abstraction will also have little negative impact on computed path quality.

Advertisement of an abstracted topology of a multi-node domain MUST support a valid representation of connectivity within that domain to support correct path computation, i.e., if multiple border nodes are advertised for a domain, some topological component MUST also be advertised to indicate when there is connectivity between these border

nodes. This reduces failure of path computation across the domain. In general, path computation should not have to infer from the control identifiers in use (such as the RC identifier) the data plane topology.

6.3.1 Multi-level Hierarchy

[G.7715] and [G.7715.1] specify that routing protocols for ASON support multiple levels of hierarchy, although they do not define specific mechanisms to support multiple hierarchical levels of RAs. In particular, if RCs bound to adjacent levels of the RA hierarchy were allowed to redistribute routing information in both directions between adjacent levels of the hierarchy without any additional mechanisms, they would not be able to determine looping of routing information.

It is necessary to have a means by which routing protocol LSAs indicate that particular routing information has been learned from a higher level RC when propagated to a lower level RC. Any downward RC from this level, which receives an LSA with this information would omit the information in this LSA and thus not re-introduce this information back into a higher level RC.

The complete procedures for supporting multi-level hierarchy are not covered in this document, but will be specified in a future OIF amendment addressing both signaling and routing aspects of a multi-level hierarchy, pending completion of associated standards in IETF and ITU-T. Initial work evaluating ASON requirements against existing routing protocol can be found in [RFC4642], and potential solutions being discussed in IETF can be found in [RFC5787].

6.3.2 Topology Abstraction

6.3.2.1 Topology Abstraction Concept

6.3.2.1.1 Separation of Routing Advertisement from Routing Advertiser

One differentiating characteristic of the E-NNI routing model is the separation of routing advertisement from routing advertiser. This separation is allowed by the ability to specify both endpoints of an advertised link separately from the identity of the advertising entity.

In conventional IP routing, this separation is not possible, because the Link ID is used both as a local identifier to manage the information about the link being advertised and the address of the advertising router. This means that advertisements of a link can only be generated by a node at each end of the link, and not by a physically separate routing controller.

6.3.2.1.2   Range of Abstraction

The protocol allows for a wide range of summarization of a domain topology. At one extreme, it is possible to advertise the full topology of the domain with no summarization, so that at the E-NNI level other routing controllers include all domain nodes and physical links in their topology database and compute paths based on a full knowledge of link resource availability within the domain.

At the other extreme, it is possible to render the entire internal topology of the domain as opaque, showing only the links into the domain and none of the domain's internal structure of nodes and links. This is commonly called an "abstract node" model, and is discussed in greater detail below.

6.3.2.1.3   Basic Routing Elements (Links and Nodes)

Types of nodal topology elements:

- border nodes – nodes that support an E-NNI interface

- interior nodes – nodes that do not support an E-NNI interface

- abstract nodes – nodes with no physical counterpart

Types of link topology elements:

- physical links – including E-NNI links

- abstract links – links with no physical counterpart

6.3.2.2   Topology Abstraction Types

6.3.2.2.1   Abstract Node Model

In this model depicted in Figure 2, the domain is advertised as a single node. As a result, no internal domain topology is visible to the outside, and E-NNI links appear from the advertisements to terminate on different ports of the same abstract node. This model supports advertisement of minimum information if desired for policy or scalability reasons.
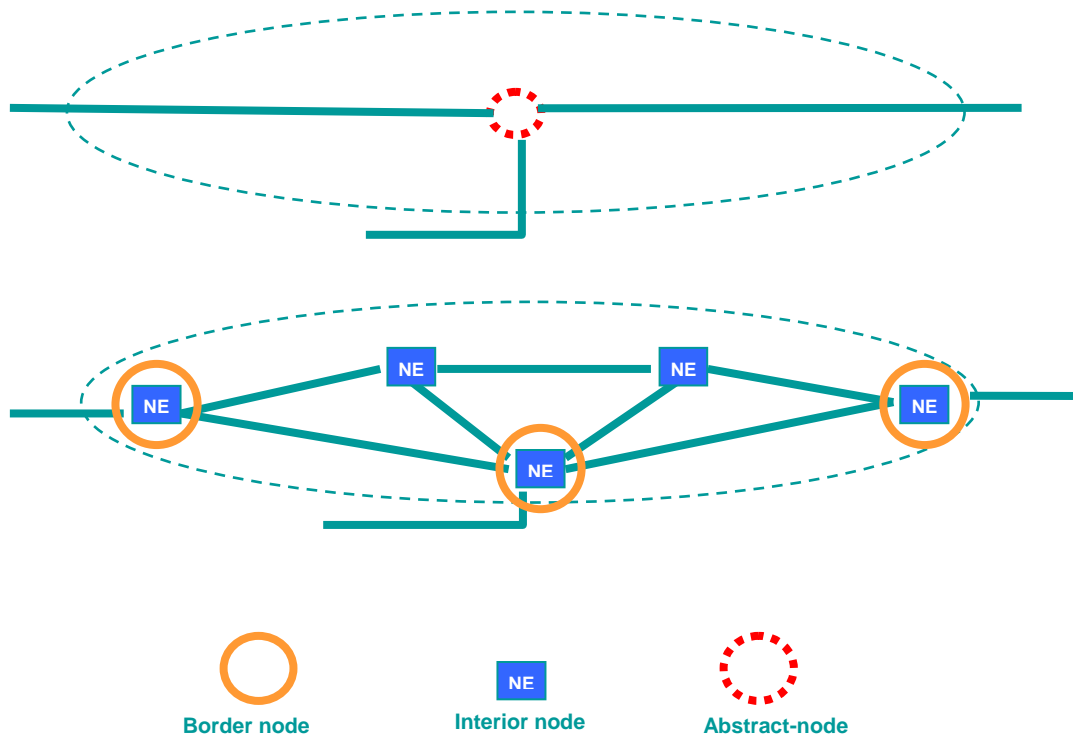
Figure 2: Abstract node model

6.3.2.2.2 Abstract Link Model

In this model depicted in Figure 3, the domain is advertised as a set of border nodes connected by a full or partial mesh of abstract links (full connectivity is being advertised when using a full mesh). Bandwidth and costs can be associated with each link to influence routing across the domain, but the links may not reflect the actual topology within the domain, only the connectivity supported. This model supports advertisement of additional information but at a cost of requiring $O(n^2)$ link advertisements, when using a full mesh, where n is the number of border nodes.
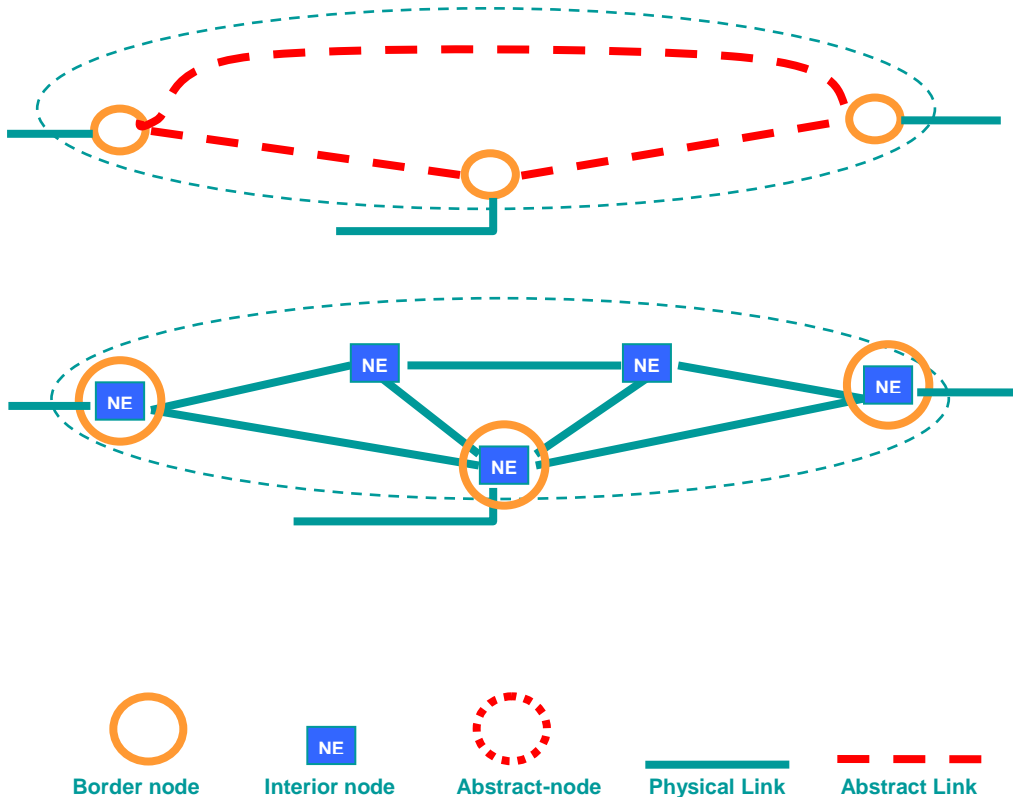
Figure 3: Abstract link model

6.3.2.2.3   More Complex Models

In more complex models, such as the one shown in Figure 4, a domain can be advertised with a combination of abstract links and abstract nodes, physical links and border nodes, to reveal a more complex topology. The insertion of abstract nodes, for example, into the advertised topology allows supported client TNAs to be associated with a virtualized node rather than having to advertise all interior nodes supporting UNI clients or having to advertise UNI clients all being attached to a border node.
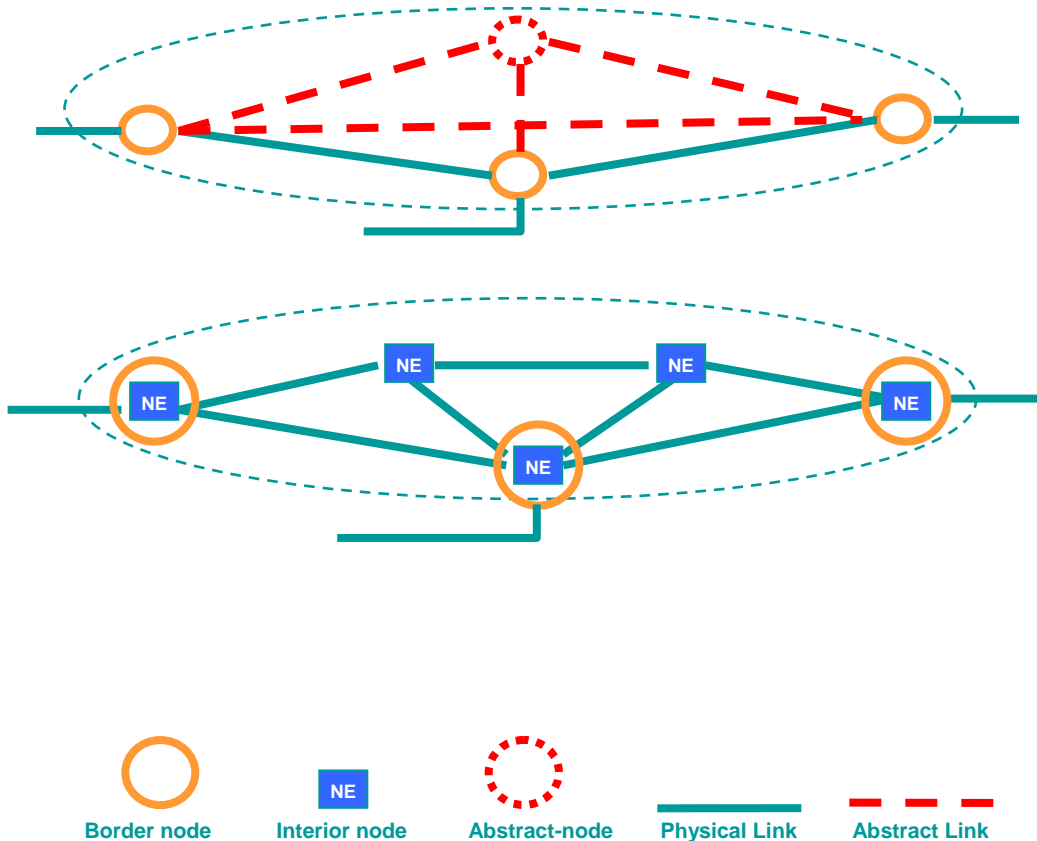
Figure 4: Combination of abstract links and abstract nodes

6.3.2.2.4 Relationship of Abstract to Real Topology Elements

Because the topology advertised through the routing protocol can be summarized or virtualized compared with the actual internal domain topology, one issue to consider with different models is the relationship of abstract and real topology elements. A number of possibilities can be supported, for example:

- One-to-one relationship: if the actual physical topology is advertised, or if abstract links and nodes are advertised with a one-to-one correspondence to the physical topology, it is possible for the advertised elements to reflect the status of the physical elements on a one-to-one basis. For example, if the physical link fails, then the corresponding abstract link can either no longer be advertised, or can be advertised as failed, i.e. zero bandwidth, to prevent it from being used in subsequent path computation. Note there is no advertisement of node status in the routing protocol, so node failure would not be advertised except as it impacts the status of links terminating on that node.

- Summarized relationship: links in an abstracted topology, even if they are not related one-for-one with physical links in the internal domain topology, can still reflect resource availability in some summarized or mapped way. Abstract links can, for example, reflect the up or down state of connectivity for some subset of physical links within the domain, or some bandwidth derived from the actual bandwidth of a subset of links, or some cost derived from the costs of a subset of links. This still allows the abstracted topology to convey more detailed information about the state of resource availability within the domain, suitable for making high level routing decisions.

- No relationship: an extremely simple topological model such as the abstract node model provides no flexibility to describe the internal state of links and nodes in the domain, so that there is no relationship between the advertised topology and the actual state of resources within the domain.

6.4 Security and Logging Considerations for Routing Messages

Security considerations for link state routing protocols are covered in the section titled "Link State Routing Protocols" of the Security Extension for UNI and E-NNI 2.0 [SecExt]. This section of [SecExt] recommends how implementations not using the Security Extension 2.0 SHOULD provide authentication of OSPF messages. It also states that implementations using the Security Extension 2.0 [SecExt] to protect signaling protocols MUST extend these mechanisms to OSPFv2-based routing as used in this IA. Such implementations SHOULD also provide the logging capabilities in [SysLog], in particular, the ability to log OSPFv2 messages with the PROT@26041 Structured Data Identifier. The PROT@26041 Structured Data item SHOULD contain the entire packet including network layer headers. Further formatting of this Structured Data item is NOT RECOMMENDED.

7 Opaque LSAs for E-NNI OSPFv2-Based Routing

7.1 Overview

[RFC3630] defines two types of top-level TLVs, i.e., the advertising router TLV and the link TLV. [RFC5787] uses a third top-level TLV, i.e., the node-attribute TLV, defined in [RFC5786].

Per [RFC3630] such top-level TLVs are included in a Type 1 TE LSA, flooded as a Type 10 opaque LSA.

This implementation agreement uses a vendor private LSA (see [RFC4940]). Its type is 252; it is a Type 10 opaque LSA too. The first four octets of this vendor private LSA is the OIF enterprise code: OIF has been assigned 26041. These four octets are then followed by one of the Router Address top-level TLV, Link top-level TLV or Node-attribute top-level TLV.

Type 1 TE LSAs are only used for backward compatibility with E-NNI Routing 1.0 (see section 14).

The format of the OIF vendor private LSA is shown below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            LS age             |    Options    |      10       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      252      |                  Instance                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Advertising Router                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        LS checksum            |              Length           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                OIF enterprise code:  26041                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Type             |              Length           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Value...                             |
//                                                            //
.                                                            .
.                                                            .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
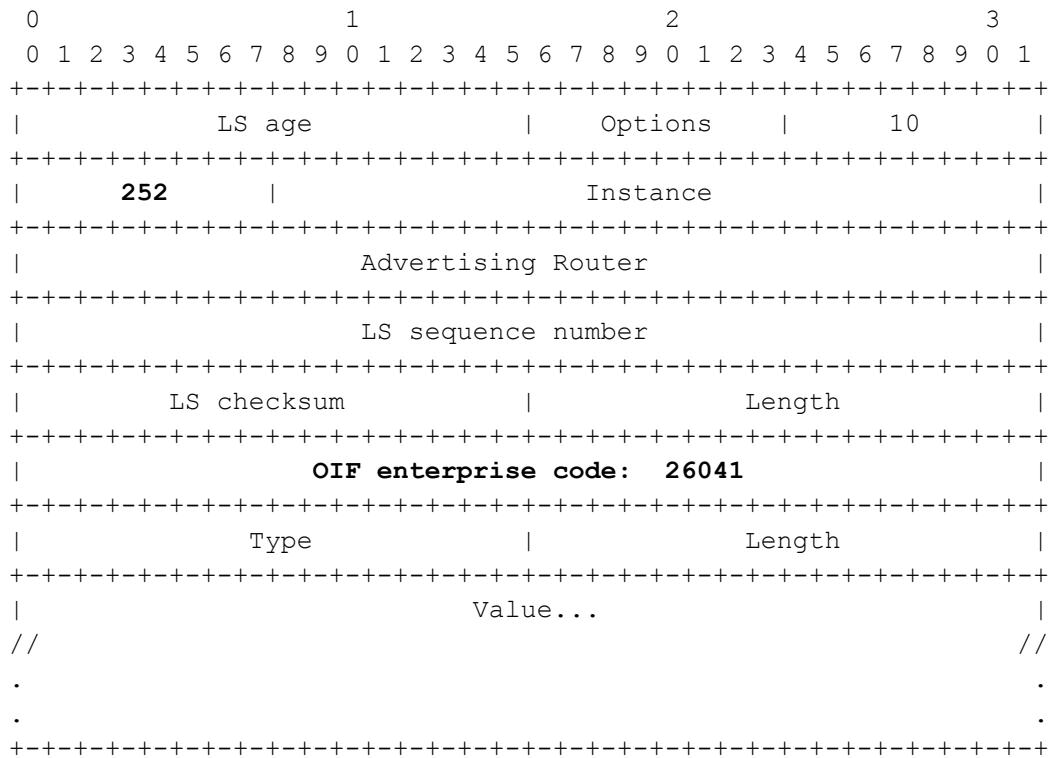
Advertisement of the Router Address TLV and (TE) Link TLV is *mandatory* for E-NNI routing. Advertisement of TNAs in Node-Attribute TLVs is dependent on the carrier network (the carrier may choose not to advertise TNAs if it uses a directory service to request the node a TNA is attached to). However, if received these MUST be stored and flooded to neighboring RCs.

Note: as per [RFC3630], the Router Address TLV appears in exactly one Traffic Engineering LSA originated by a RC. Only one Link TLV SHALL be carried in each Traffic Engineering LSA originated by a RC. The same rules apply to OIF private LSAs.

Per [RFC5786], only one node attribute TLV must be advertised in a Traffic Engineering LSA. [RFC5787] allows each RC to advertise multiple such LSAs. The same rules apply to OIF private LSAs.

Type value	Top Level TLV	Semantics	Reference	Scope	Mandatory /Optional
1	Router Address TLV	No standard semantic.	OSPFv2-TE [RFC3630]	Originated from any RC participating in the RA.	Mandatory for consistency with [RFC3630]
2	(TE) Link TLV	Point-to-point link	OSPFv2-TE [RFC3630]	Originated from any RC advertising one (or more) TE Link.	Mandatory for E-NNI routing.
5	Node-attribute TLV	Reachable TNAs	[RFC5786]	Originated from any RC advertising TNA Reachability.	Carrier dependent.

Table 1: OIF private opaque LSAs

7.2 Router Address TLV

The Router Address TLV carries a stable SCN address that belongs to the advertising OSPFv2-TE router. Note that this SCN address may not be reachable from where this LSA is received (implementations MUST NOT assume it is reachable; they must not use it for SCN IP traffic addressing).

7.3 Link TLV

The Link TLV is used to represent an inter-domain link or an intra-domain link. The Router ID field in the OSPFv2 packet header identifies the advertising OSPFv2 router. Because multiple Routing Controllers may be responsible for a routing domain, there may not be a one to one relationship between a node and a Routing Controller. A TE-Link LSA does not carry any Routing Controller identifiers.

The two endpoints of a TE-Link, in the transport plane, are identified by the Local and Remote TE Router ID sub-TLV. This is different from the way that [RFC3630] uses the OSPFv2 Router ID LSA header field and the Link ID sub-TLV, to identify the two

endpoints, and therefore does not support the separation between transport plane nodes and advertising routers (routing controllers). This separation has been addressed by [RFC5787] with the introduction of the Local and Remote TE Router ID sub-TLV, used in this implementation agreement.

For the purpose of E-NNI routing, the Link ID sub-TLV value SHOULD be set to 0.0.0.0, and MUST be ignored on receipt.

When included in an OIF private opaque LSA, a Link TLV contains the following information:

Type value	TLV	Semantics	Reference	Mandatory /Optional
1	Link type sub-TLV	Point-to-point link	[RFC3630]	M
2	Link ID sub-TLV	Should be set to 0.0.0.0	[RFC3630]	M
5	Link metric sub-TLV	Link cost	[RFC3630]	M (by default equal to 1)
9	Link resource class sub-TLV	Color	[RFC3630]	M (by default bit mask equal to 0…0)
11	Local/Remote Identifiers sub-TLV	Local interface ID and remote interface ID	[RFC4203]	M (if the Remote Identifier is unknown, it SHOULD be set to 0)
14	Link protection type sub-TLV	Link protection type	[RFC4203]	O (by default unprotected links)
15	Interface Switching Capability Descriptor sub-TLV (this implementation agreement defines its own TDM SCSI)	Describes the TE link bandwidth information	See Sections 11.1.2 (TDM) and 11.2 (packet)	M

16	SRLG sub-TLV	Shared risk link group	[RFC4203]	O (by default Link ID is the SRLG)
26	Local and Remote TE Router ID sub-TLV	Local endpoint (e.g., border node ID) and remote endpoint (e.g., remote border node ID)	See Section 9.2	M
32772	Link Attribute Scoping and Connection Type sub-TLV	Allows scoping of link attributes to a specific layer, and local connection type specification.	See Section 12.3	O
32776	Generic Multilayer TNA	Allows mapping of client layer SNPP into server layer TNA	See Section 14.2	O (MAY appear more than once)

Table-2: Inter-domain and Intra-domain Link information

Note 1: Per [RFC3630], Link Type and Link_ID sub-TLVs MUST appear exactly once. Per [RFC3630], Link metric and Link resource class may occur at most once, this implementation agreement states that both MUST appear exactly once. Per [RFC4203], the Link protection and SRLG sub-TLVs may occur at most once.

In addition,

- the Local and Remote TE Router ID sub-TLV MUST appear exactly once ;

- the Local and Remote Identifier sub-TLV MUST appear exactly once ;

- the Link Attribute Scoping and Connection Type sub-TLV and the TDM and Packet Interface Switching Capability Descriptor sub-TLV may appear multiple times, although at most once for a given layer.

Note 2: Setting of the Link ID to 0.0.0.0 as described in the table above deviates from the use of the Link ID as defined in [RFC3630].

Note 3: Multiple bits may be set in the Link Protection type sub-TLV, as noted in Section 13.

Note 4: OIF E-NNI 1.0/2.0 routing considers the optional advertisement of a "Link Protection Type" for E-NNI links, and states that if a Link State Advertisement doesn't carry this information, then the Link Protection Type is unprotected. However, [RFC4202] states that "The Link Protection Type is optional, and if a Link State Advertisement doesn't carry this information, then the Link Protection Type is unknown.". For the sake of interoperability it is recommended to advertise the corresponding "Link Protection Type" of E-NNI links.

The Local Identifier and the Remote Identifier are both part of the Link Local/Remote Identifiers sub-TLV (Type 11) defined in [RFC4203]. The format of this sub-TLV is defined as:

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Link Local Identifier                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Link Remote Identifier                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
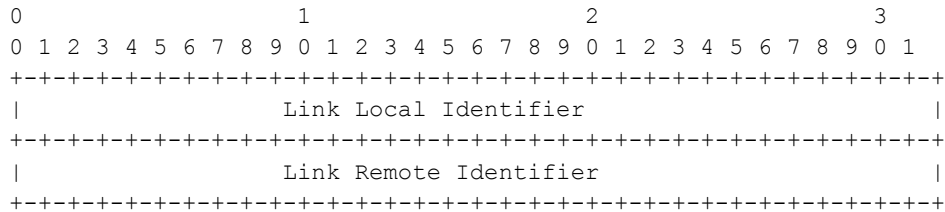
Note 4: under certain conditions the Link Remote Identifier MAY be coded 0 where the Identifier value is not known. When this is true, the link advertisement is not included in path calculation.

7.4   Node Attribute TLV

When included in an OIF private opaque LSA, a Node Attribute TLV (Type 5 per [IANA-OSPF-TE]) contains the following information:

Type value	TLV	Semantics	Reference	Optional/Mandatory
5	Local TE Router ID sub-TLV	Identifies the Node to which the TNAs are attached.	See Section 10.3.2	Mandatory

1	Node IPv4 Local Address sub-TLV	Specifies IPv4 TNAs	See Section 10.3.1	At least one of the Node IPv4/IPv6 Local Address or NSAP TNA sub-TLVs must be present.
2	Node IPv6 Local Address sub-TLV	Specifies IPv6 TNAs	See Section 10.3.1	
32772	NSAP TNA Address Sub-TLV	Specifies NSAP TNAs	See Section 10.4	

Table 3: Node Attribute Opaque LSA Information

8 Support of G.7715.1 Link Attributes – Link Identification

8.1 Link Identification with Full Separation of Node ID and RC/SC IDs

ASON has defined a number of different functional entities, each with its own identifier spaces. The identifier spaces used by ASON Routing are described in the Section 7.1 of [G.7715.1]:

Three categories of identifiers are used for ASON routing: transport plane names, control plane identifiers for components, and SCN addresses.

- Transport plane names describe [G.800]/[G.805] resources and are defined in G.8080/Y.1304.

  o SNPP names give a routing context to SNPs and are used by the control plane to identify transport plane resources. However, they do not identify control plane components but represent a (G.805) recursive subnetwork context for SNPs. Multiple SNPP name spaces may exist for the same resources.

  o UNI Transport Resource Identifiers are used to identify transport resources at a UNI reference point if they exist. They represent clients in (G.8080/Y.1304) access group containers and may be disseminated by RCs.

The OIF "TNA name" is an instantiation of the G.8080 "UNI Transport Resource Identifier" and both are used in this document to refer to the same thing.

- Control plane identifiers for G.8080/Y.1304 components may be instantiated differently from each other for a given ASON network. For example, one can have centralized routing with distributed signaling. Separate identifiers are thus used for:

    o   Routing Controllers (RCs)

    o   Connection Controllers (CCs)

- Additionally, components have Protocol Controllers (PCs) that are used for protocol-specific communication. These also have identifiers that are separate from the (abstract) components like RCs.

- SCN addresses enable control plane components to communicate with each other via the SCN as described in [G.7712].

Using these definitions, Table 4 reviews the different identifiers used in [RFC3630], and suggests a logical mapping to ASON identifiers. Note: the IETF has also defined a lexicography comparing GMPLS and ASON terminology [RFC4397], for general mapping of terminology.

Table 4: Identifier Table

Instance in OSPFv2-TE [RFC3630]	Description	OSPF-TE Address Space	G.8080 Architectural Name
9 Source and Destination IP Addresses	These addresses are used by the RC PCs to communicate with each other. They are also known as the RC PC IP addresses. They are located in the IP header of the OSPF packets	IPv4 Address space	RC PC SCN Address
Router Address	Traffic Engineering TopLevel TLV from [RFC3630].	IPv4 Address. ([RFC3630] states that this is a "reachable" IPv4 address.)	G.8080 distinguishes tranport plane node, Routing Controller and Signaling Controller, and therefore uses separate addresses (Transport Plane Node ID, RC and SC PC SCN addresses)
Router ID in OSPF packet Header	Used to identify the neighbor that generated the OSPF packet containing LSAs.	Router ID	Control Plane Name: RC PC ID
Advertising Router ID	Field contained in an LSA Header. For a given OSPF packet, this is likely to be different from the Router ID in the OSPF Header. For a TE Link TopLevel TLV, this field identifies the router at the near end of a link.	Router ID (see Note)	G.8080 distinguishes the entity advertising routing information, and the transport plane endpoints, and therefore uses two identifiers: RC ID and Transport Plane Node ID.

Link ID	SubTLV contained in a TE Link TopLevel TLV. For a given OSPF packet, this is likely to be different from the Router ID in the OSPF Header. For a TE Link TopLevel TLV, this identifies the router at the far end of a link.	Router ID (see Note)	G.8080 distinguishes the entity advertising routing information, and the transport plane endpoints, and therefore uses two identifiers: RC ID and Transport Plane Node ID.

Note: [RFC2328] defines the SPF algorithm used to traverse the topology shared by OSPF nodes in an area. This algorithm specifically uses Router ID as the Vertex ID when identifying a point-to-point link between two routers in the topology, as shown in Section 16.1 of [RFC2328]. This is further underscored in Section 16.1 Step 2 and Step 2b where Router-LSAs for vertex *V* (the near end of a link) and vertex W (the far end of a link) are retrieved using the Vertex ID. Because the Router IDs for the near end of a TE link and the far end of a TE link in [RFC3630] are located in the Advertising Router ID and the Link ID fields for the Link TLV, respectively (see [RFC3630] Sections 2.4.2 and 2.5.2), these fields are used in [RFC3630] to identify the link ends that exist in the TE topology.

By using different categories of identifiers for transport plane entities, Control plane entities and SCN addresses, it is possible in ASON to support a number of different function distributions including:

- 1:*N* relationship between an RC and Subnetworks

- *N*:1 relationship between RCs and a Subnetwork

This allows a separation between the transport plane entity being controlled, the control plane entity supporting it, and the SCN address where the control plane entity can be reached.

This allows for full separation of different control identifiers as required for ASON. However, this does not imply that different values are always used for each identifier. An implementation MAY use duplicate values (in mandatory fields) when full separation is not required; it MUST however accept TLVs from peer implementations that do support full separation.

Support for these distributions is considered useful for domain-to-domain networking and allows flexibility for support of E-NNI routing by domains with different characteristics. It helps support domains with differing characteristics and abstraction of domain topology and resource information as called for in OIF carrier requirements.

9.1 Local/Remote Node ID

To support the separation of control plane and transport plane identifiers as described above (and therefore 1:N or N:1 relationships) for unnumbered links, the capability is needed to identify unnumbered links uniquely when advertising. Therefore the transport plane Node IDs for local and remote link ends are advertised separately from the RC associated with the link.

Note: Since the Node ID extensions provide the Transport Plane Name for a Local Vertex and Remote Vertex on a link in the Transport Topology, the Advertising Router ID and Link ID fields are no longer used to identify the nodes at the ends of a link, reducing the role of the Advertising Router ID field to a part of the Database key used to name LSAs.

Furthermore, since the Node ID comes from the Transport Plane namespace, it is used as the identifier in an Explicit Route Object, removing the dependence on the Router Address TLV.

9.2 Protocol Extensions incorporated from IETF

The Local and Remote TE Router ID experimental sub-TLV (see [RFC5787]) is included in an inter-domain or intra-domain LSA to indicate the local and remote end points of a link. This is used to support separation of the control and data planes, as well as topology abstraction. This sub-TLV is mandatory in an E-NNI Link TLV. The format of this sub-TLV is defined as:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Type (see below)        |           Length (8)          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Local TE Router Identifier                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Remote TE Router Identifier                  |
```

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
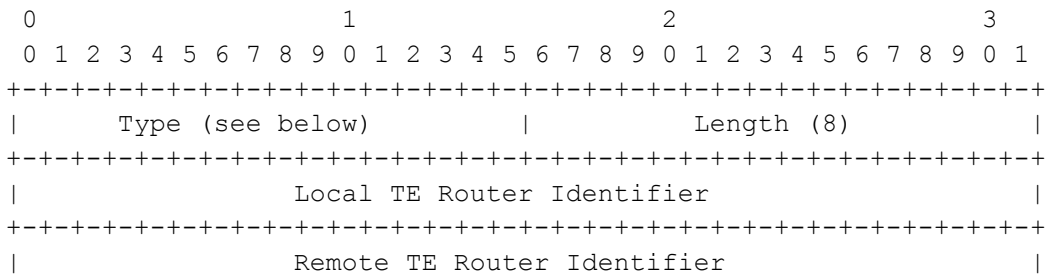
The Local TE Router Identifier field is set to the Local Node ID. The Remote TE Router Identifier field is set to the Remote Node ID.

This implementation agreement makes use of this sub-TLV (see section 7.3), and assigns code-point 26[1] to this sub-TLV Type, when it is encoded within the Link TLV of the OIF private LSA.

10 Support of G.7715.1 Node Attributes–Reachability Advertisement

10.1 Client reachability advertisement

In IP routing, it is expected that the way to calculate a route to an endpoint is for the endpoint to be announced in the routing protocol. However, end equipment in IPv4 networks is typically attached using Ethernet subnetworks advertised via Network LSAs or External LSAs. This makes separate end equipment advertisement unnecessary. Unfortunately, the optical network environment discussed in this document is outside the IPv4 network and does not have an analogous method for a router to advertise the UNI endpoints associated with a vertex in the area's topology. Consequently, a capability to advertise client reachability is needed, as is identified in [G.7715.1].

10.2 Reachability information and Node ID Advertisement

Within a single area, in a multi-domain environment, reachability information for connection endpoints can be exchanged.

Per [G.7715.1], reachability information may either be a set of UNI Transport Resource Identifiers (or TNAs), or a set of SNPP IDs/SNPP ID prefixes. Per [G.8080], UNI Transport Resource Identifiers and SNPP IDs/SNPP ID prefixes are not in the same namespace, and therefore routing protocol advertisements must provide a way to distinguish between them.

When an abstract topology is created in an upper area from a topology in a lower area that includes a Node with attached UNI Transport Resource Identifiers, the abstract

---

[1] 25 was the code-point proposed by [OSPF-ASON], before [RFC5787] made that sub-TLV experimental. IANA has later assigned 25 to the Interface Adjustment Capability Descriptor (IACD) sub-TLV. 26 has been requested from IANA for draft-ietf-ccamp-rfc5787bis.

topology associates the UNI Transport Resource Identifier with at least one Node in the abstract topology as allowed by policy.

10.3   Protocol Extensions incorporated from IETF

The Node Attribute Top-level TLV (see [RFC5787] and [RFC5786]) is used for advertising attached reachability information (e.g., TNAs). An LSA containing this new top-level TLV is only announced by the Routing Controller (RC) responsible for the Node in a topology to which the reachable endpoint is attached.

The Node Attribute TLV is advertised by a control domain routing controller towards RCs belonging to different control domains.

The Node Attribute TLV contains a list of TNAs attached to the same Node. The same TLV may also be used for other reachable endpoints, i.e. SNPP IDs/SNPP ID prefixes, per [G.7715.1], although this is not covered in this implementation agreement.

When a Node Attribute TLV advertises TNAs attached to the same node, it identifies this node using the Local TE Router ID sub-TLV. TNAs are specified using Node Local Address sub-TLVs (IPv4 and IPv6 TNAs) and NSAP TNA sub-TLVs (NSAP TNAs).

10.3.1   Node Local Address Sub-TLV

[RFC5786] defines two sub-TLVs that can be used for IPv4 and IPv6 TNAs: the Node IPv4 Local Address Sub-TLV and the Node IPv6 Local Address Sub-TLV. Each sub-TLV may specify multiple IPv4 (respectively IPv6) TNAs.

10.3.1.1 Node IPv4 Local Address Sub-TLV

The Node IPv4 Local Address sub-TLV has the following format:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               1               |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Prefix Len 1  |        IPv4 Prefix 1                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Prefix 1 cont. |                                               :
+-+-+-+-+-+-+-+-+                                               ~
:                           .                                   :
~                           .                +-+-+-+-+-+-+-+-+
:                           .                | Prefix Len n  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         IPv4 Prefix n                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
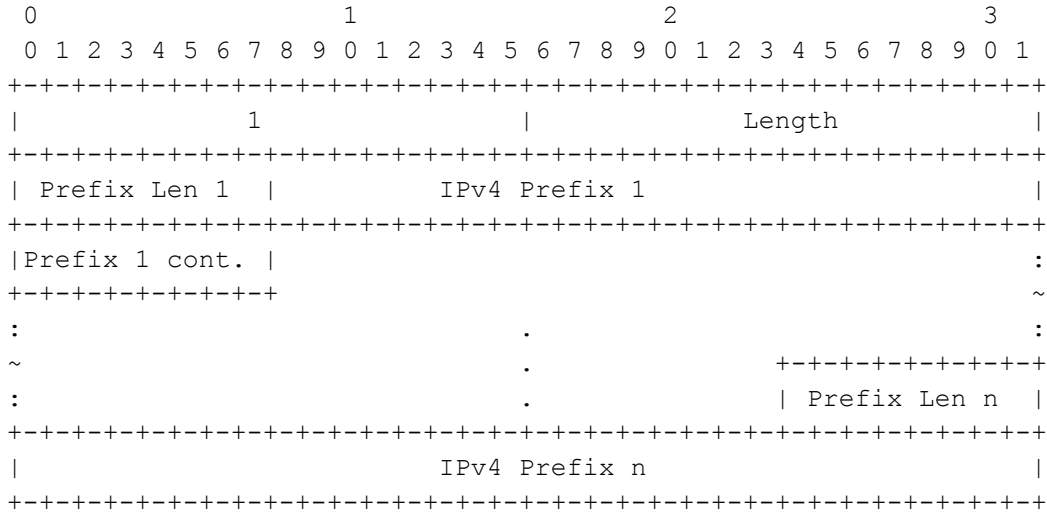
Per [RFC5786], each local IPv4 address is encoded as a <Prefix Length, Prefix> tuple. Prefix Length is encoded in 1 byte. It is the number of bits in the Address and can be at most 32. Prefix is an IPv4 address prefix and is encoded in 4 bytes with zero bits as necessary. The Node IPv4 Local Address sub-TLV length is in octets. It is the sum of the lengths of all n IPv4 Address encodings in the sub-TLV, where n is the number of local addresses included in the sub-TLV.

Per [RFC3630], the Node IPv4 Local Address sub-TLV is padded to four-octet alignment; padding is not included in the length field.

A Node Attribute TLV MUST NOT carry more than one Node IPv4 Local Address sub-TLV.

10.3.1.2 Node IPv6 Local Address Sub-TLV

The Node IPv6 Local Address sub-TLV has the following format:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               2               |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Prefix Len 1  | Prefix 1 Opt. | IPv6 Prefix 1                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   IPv6 Prefix 1 cont.                                         :
:                               .                               ~
~                               .
:                               .
:                                +-+-+-+-+-++-+-+-+-+-++-+-+-+-+-+
:                                | Prefix Len n  | Prefix n Opt. |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       IPv6   Prefix n                         :
|                                                               :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+--
```
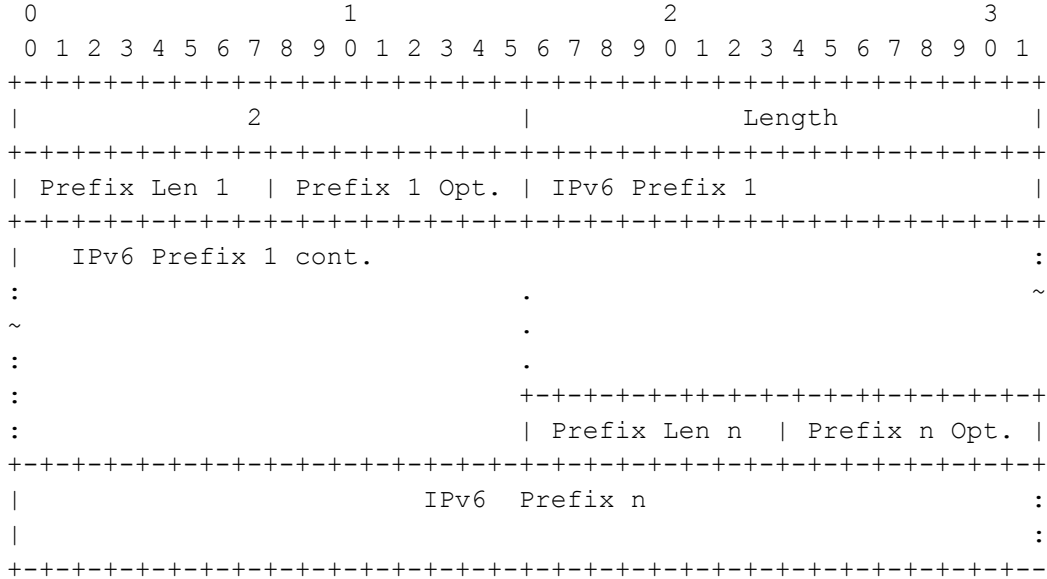
Per [RFC5786], each local IPv6 address is encoded using the procedures in [RFC5340]. Each IPv6 address MUST be represented by a combination of three fields: PrefixLength, PrefixOptions, and Address Prefix. PrefixLength is the length in bits of the prefix and is an 8-bit field.

PrefixOptions is an 8-bit field describing various capabilities associated with the prefix [RFC5340]. The originator of this sub-TLV must set the NU-bit, and leave all other bits unset (i.e., PrefixOptions must be set to 0x01). On receipt, the PrefixOptions should be ignored.

Address Prefix is an encoding of the prefix itself as an even multiple of 32-bit words[2], padding with zero bits as necessary. This encoding consumes((PrefixLength + 31) / 32) 32-bit words.

The Node IPv6 Local Address sub-TLV length is in octets. It is the sum of the lengths of all n IPv6 Address encodings in the sub-TLV, where n is the number of local addresses included in the sub-TLV.

Per [RFC3630], the Node IPv6 Local Address sub-TLV is padded to four-octet alignment; padding is not included in the length field.

A Node Attribute TLV MUST NOT carry more than one Node IPv6 Local Address sub-TLV.

---

[2] I.e., an exact multiple of 32-bit words.

10.3.2 Local TE Router ID Sub-TLV

The Local TE Router ID sub-TLV (see [RFC5787]) is included as part of the Node Attribute TLV and identifies the node hosting the advertised reachability information - TNAs. This sub-TLV is mandatory in a Node-Attribute TLV advertising TNAs. The format of this sub-TLV is defined as:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Type (see below)        |            Length (4)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Local TE Router Identifier                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
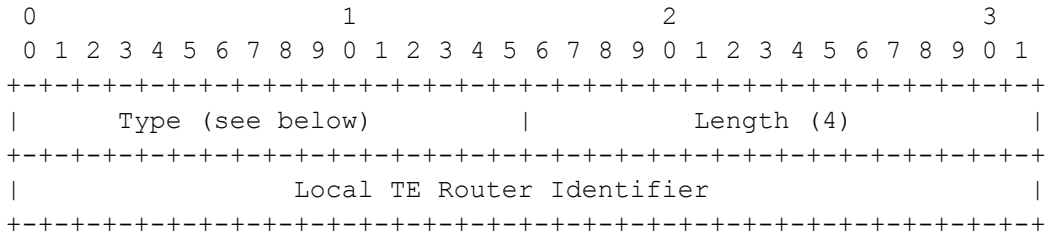
The Local TE Router Identifier field is set to the Node ID of the node to which the TNAs are attached.

This implementation agreement makes use of this sub-TLV (see section 7.4), and assigns code-point 5[3] to this sub-TLV Type, when it is encoded within the Node-attribute TLV of the OIF private LSA.

10.4 Standard Protocol Extensions

TNAs are advertised using sub-TLVs of the node attribute TLV ([RFC5786]).

[RFC5786] defines two sub-TLVs for IPv4 and IPv6 prefixes advertisement: the Node IPv4 Local Address Sub-TLV and the Node IPv6 Local Address Sub-TLV. However, [RFC5786] does not support the advertisement of NSAP prefixes.

10.4.1 NSAP TNA Sub-TLV

This implementation agreement defines a private OIF extension: a third sub-TLV (Type 32772) for NSAP TNA advertisement.

---

[3] 5 was the code-point proposed by [OSPF-ASON], before [RFC5787] made that sub-TLV experimental.

```
 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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |              32772            |              Length           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Prefix Len 1 |          NSAP Prefix 1                         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | NSAP Prefix 1 cont.                                          :
 +                                                              ~
 :                             .                                :
 ~                             .                +-+-+-+-+-+-+-+-+
 :                             .                | Prefix Len n  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 :                        NSAP Prefix n                          :
 :                                                              :
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
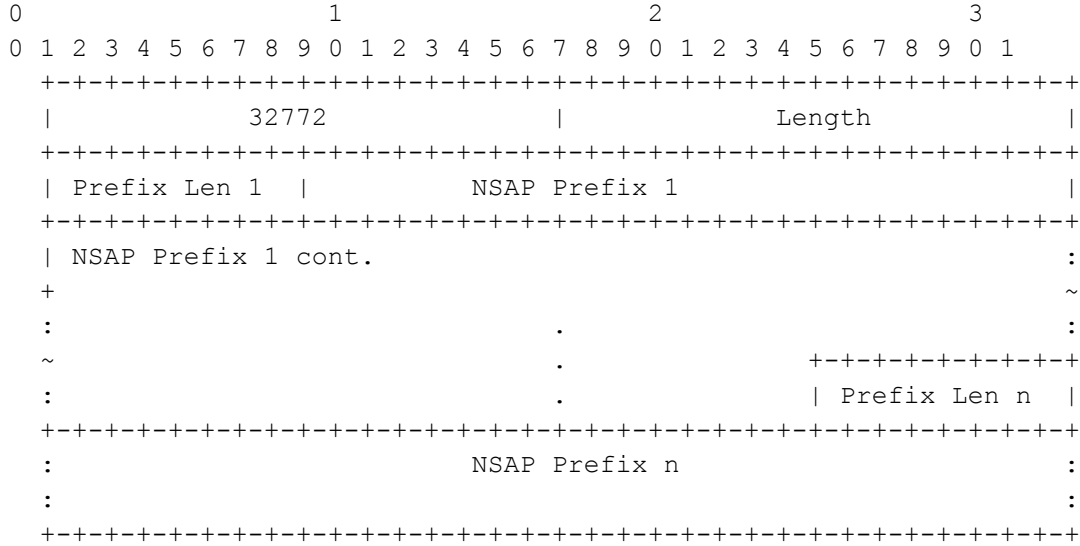
Each local NSAP prefix is encoded as a <Prefix Length, Prefix> tuple. Prefix Length is encoded in 1 byte. It is the number of bits in the prefix and can be at most 160. NSAP Prefix is an encoding of the prefix itself as an exact multiple of 32-bit words, padding with zero bits as necessary. This encoding consumes ((PrefixLength + 31) / 32) 32-bit words.

The NSAP TNA sub-TLV length is in octets. It is the sum of the lengths of all n NSAP prefix encodings in the sub-TLV, where n is the number of NSAP prefixes included in the sub-TLV.

Per [RFC3630], the NSAP TNA sub-TLV is padded to four-octet alignment; padding is not included in the length field.

A Node Attribute TLV MUST NOT carry more than one NSAP TNA sub-TLV.

11 <u>Support of G.7715.1 Link Attributes – Layer-specific Link Capacity</u>

11.1 TDM layer link capacity

11.1.1 Advertisement of Layer-specific link capacity

GMPLS Routing extensions to OSPF define an Interface Switching Capability Descriptor (ISCD) that delivers information about the (maximum and minimum) bandwidth per priority an LSP can make use of. In the ASON context, other representations are possible, e.g., in terms of a set of tuples <signal_type; number of unallocated timeslots>. The latter approach may require definition of additional signal types (from those

defined in [RFC4606]) to represent contiguous concatenation, i.e. STS-(3xN)c SPE / VC-4-Nc, N = 4, 16, 64, 256.

As [G.7715.1] specifies link capacity as a link characteristic specific to a particular layer network, a representation in the form of tuples of <signal_type; number of unallocated timeslots> is most closely consistent with ASON requirements and provides accurate and separable information on a fine grained, per layer network basis.

11.1.2   TDM Interface Switching Capability Descriptor sub-TLV

To provide this functionality, a new Switching Capability Specific Information (SCSI) has been defined for the IETF [RFC4203] Interface Switching Capability Descriptor (ISCD). It incorporates information about available connections at specific signal types. This provides an alternative accounting of resource availability, in particular taking into account the impact of time slot allocation on the availability of connections using contiguous concatenation.

The format for the TDM Interface Switching Capability Descriptor is given below.

This sub-TLV (Type 15) of the top-level Link TLV is dedicated to SONET/SDH and OTN (ODUk) bandwidth accounting.

It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (15)            |         Length = 4 + x        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap |   Encoding    |           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 0               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 1               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 2               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 3               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 4               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 5               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 6               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Max LSP Bandwidth at priority 7               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+-    Switching Capability-specific information: sub-TLVs    -+
|                            ...                                |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Reserved (16 bits):
> Set to zero when sent and ignored when received.

Switching Cap, Encoding, Max LSP Bandwidth at priority i:
> See sections 11.1.2.1 for SONET/SDH and OTN (ODUk) respectively.

Switching Capability-specific information:
> This document defines two sub-TLVs:

> 1. Type 1 sub-TLV provides bandwidth accounting using a 16 bits value.

> 2. Type 2 sub-TLV provides bandwidth accounting using a 32 bits value.

> The interpretation of the bandwidth value depends on the Signal Type. It may be provided as a number of fixed size containers, or as bytes per second.
> > • If the bandwidth value is provided as bytes per second, the Type 2 sub-TLV must be used, and the bandwidth value is encoded in a 4 octets field in the IEEE floating-point format.

- If the bandwidth value is not provided as bytes per second, then the sub-TLV originator is free to choose Type 1 or Type 2 sub-TLV (it is expected that it will choose Type 1 whenever possible, since it yields a more compact encoding).

Both Type 1 and Type 2 sub-TLVs may be used within the same SCSI. Exactly one Type 1 or Type 2 sub-TLV is encoded for a given signal type.

The choice of Type 1 or Type 2 sub-TLV is made by the advertisement originator. The receiver MUST accept both.

In case of link bundling, when component links are added or removed from the bundle, it may happen that the bundle TE-Link advertisement will switch from Type 1 to Type 2 sub-TLV (or vice-versa) for some signal types.

Type 1 sub-TLV has the following format:

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |        Type (1)               |          Length = 4 + x       |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Signal Type  |Bw Type| Flags |  Priority     |  Reserved     |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | BW at prio 0 (if supported)  |  BW at prio 1 (if supported)  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | BW at prio 2 (if supported)  |  BW at prio 3 (if supported)  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | BW at prio 4 (if supported)  |  BW at prio 5 (if supported)  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | BW at prio 6 (if supported)  |  BW at prio 7 (if supported)  |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type 2 sub-TLV has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Type (2)               |         Length = 4 + x        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Signal Type  |Bw Type| Flags |   Priority    |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bandwidth at priority 0 (if priority 0 supported)            +
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bandwidth at priority 1 (if priority 1 supported)            +
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bandwidth at priority 2 (if priority 2 supported)            +
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bandwidth at priority 3 (if priority 3 supported)            +
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bandwidth at priority 4 (if priority 4 supported)            +
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bandwidth at priority 5 (if priority 5 supported)            +
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bandwidth at priority 6 (if priority 6 supported)            +
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Bandwidth at priority 7 (if priority 7 supported)            +
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Signal Type

> This field identifies the particular container for which per-priority bandwidth is advertised. See the sub-sections below for SONET/SDH and OTN (ODUk) specifics.

Bw Type (4 bits)

> This field specifies which kind of bandwidth is advertised:
>
> 0:   Unreserved Bandwidth

Flags (4 bits)

> All bits in this bit-vector are reserved, and must be sent as 0 and ignored on reception.

Priority (8 bits)

> This field specifies which priorities are supported (a bandwidth value is advertised only for supported priorities). This field is a bitmap, each bit being associated to a priority: when set to 1 the priority is supported.
>
> - 0b1xxxxxxx : Priority 0
> - …
> - 0bxxxxxxx1 : Priority 7

If priorities are not used in a routing domain, then only one bandwidth value will be advertised: the one for the highest priority (priority 0, i.e., the Priority bitmap field is set to 0x80).

Reserved (8 bits):

Set to zero when sent and ignored when received.

11.1.2.1 SONET/SDH interfaces

Inherited from [RFC4203], the Switching Capability field, the Encoding field and the Max LSP Bandwidth fields MUST take the following values for SONET/SDH interfaces:

Switching Capability (8 bits): value 100 (TDM).

Encoding (8 bits): value 5 for SONET/SDH.

Max LSP Bandwidth at priority i: ignored when received, because per-signal type bandwidth values are provided in the SCSI.

Signal Type (8 bits): inherited from [RFC4606], the Signal Type field(s) MUST take one of the following values:

Value	Type	Bandwidth encoding
1	VT1.5 SPE / VC-11	The bandwidth is provided as a number of fixed size containers.
2	VT2 SPE / VC-12	
3	VT3 SPE	
4	VT6 SPE / VC-2	
5	STS-1 SPE / VC-3	
6	STS-3c SPE / VC-4	
21	STS-12c SPE/VC-4-4c	
22	STS-48c SPE/VC-4-16c	
23	STS-192c SPE/VC-4-64c	

Table 5: SONET/SDH signal types

As stated above, the LSA originator is free to choose either Type 1 or Type 2 sub-TLV for each signal type.

Unreserved Bandwidth values (Bw Type=0) must be advertised for all supported SONET/SDH signal types. The Unreserved Bandwidth specifies the number of identical unallocated timeslots per Signal Type and per Link. As such, the initial value(s) of this TLV indicates the total capacity in terms of number of timeslots per link. The signal type included in the BW announcement is specific to the layer link being reported and is not derived from some other signal type (e.g. STS-48c is not announced as 16 x STS-3c)

For instance on an OC-192/STM-64 interface either the number of STS-3c SPE/VC-4 unallocated timeslots is initially equal to 64, or the number of STS-48c SPE/VC-4-16c unallocated timeslots is equal to 4 or a combination of both type of signals depending on the interface capabilities. Once one of these components gets allocated for a given connection, the number of unallocated timeslots is decreased by the number of timeslots this connection implies.

The number of available timeslots per link is calculated independently for each signal type as resource usage on the link changes. For example, an OC-192/STM-64 interface with one STS-1/VC-3 timeslot in use would be advertised with the following unallocated timeslots (assuming that the link is able to support a full range of STS-192c and lower rate signals):

STS-192c/VC-4-64c	0
STS-48c/VC-4-16c	3
STS-12c/VC-4-4c	15
STS-3c/VC-4	63
STS-1/VC-3	191

For SONET/SDH interfaces, fragmentation of bandwidth caused by utilized timeslots can impact the usability of timeslots at higher rate signals and are accounted for in the number of unallocated timeslots advertised.

11.1.2.2 OTN (ODUk) interfaces

Inherited from [RFC4203], the Switching Capability field, the Encoding field and the Max LSP Bandwidth fields MUST take the following values for OTN (ODUk) interfaces:

Switching Capability (8 bits): value 100 (TDM).

Encoding (8 bits): value 12 for G.709 ODUk (Digital Path).

Max LSP Bandwidth at priority i: ignored when received, because per-signal type bandwidth values are provided in the SCSI.

Signal Type (8 bits): inherited from [RFC4328], the Signal Type field(s) MUST take one of the following values:

Value	Type	Bandwidth encoding
0	Not significant	
1	ODU1 (i.e., 2.5 Gbps)	The bandwidth is provided as a number of fixed size containers.
2	ODU2 (i.e., 10 Gbps)	
3	ODU3 (i.e., 40 Gbps)	
4	Reserved (for future use)	
5	Reserved (for future use)	

Table 6: OTN (ODUk) signal types

As stated above, the LSA originator is free to choose either Type 1 or Type 2 sub-TLV for each signal type.

Unreserved Bandwidth values (Bw Type=0) must be advertised for supported ODU1, ODU2 and ODU3 signal types. The Unreserved Bandwidth specifies the number of identical unallocated timeslots per Signal Type and per Link. As such, the initial value(s) of this TLV indicates the total capacity in terms of number of timeslots per link.

A future revision of this implementation agreement (or a future amendment) will address other OTN signal types, such as G.709 (2009-12) signal types.

11.2 Packet-based layer link capacity

For packet-based layer link advertisement, this implementation agreement uses [RFC4202]/[RFC4203] Interface Switching Capability Descriptor (ISCD), whose format is reproduced below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type (15)           |         Length = 4 + x        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap |   Encoding    |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Max LSP Bandwidth at priority 0                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Max LSP Bandwidth at priority 1                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Max LSP Bandwidth at priority 2                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Max LSP Bandwidth at priority 3                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Max LSP Bandwidth at priority 4                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Max LSP Bandwidth at priority 5                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Max LSP Bandwidth at priority 6                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               Max LSP Bandwidth at priority 7                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Switching Capability-specific information            |
|                        (variable)                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

OIF UNI [OIF-UNI-02.0] and E-NNI 2.0 [OIF-E-NNI-sig-02.0] signaling implementation agreement supports Ethernet services signaling using DCSC (for EPL services) and L2SC (for EVPL services) switching types. Although additional switching types are defined by IETF for packet-based layers, this implementation agreement uses only DCSC and L2SC.

Inherited from [RFC6002], the Switching Capability field and the Encoding field MUST take the following values for EPL services:

> Switching Capability (8 bits): value 125 (DCSC)

> Encoding (8 bits): value 2 for Ethernet

Inherited from [OIF-UNI-02.0] and [OIF-E-NNI-sig-02.0], the Switching Capability field and the Encoding field MUST take the following values for EVPL services:

> Switching Capability (8 bits): value 51 (L2SC)

> Encoding (8 bits): value 2 for Ethernet

For both switching types, no Switching Capability-specific information is defined.

If no priority is used (operator policy) then only the highest priority bandwidth (at priority 0) is used to specify the current link capacity. All other bandwidths (at priority 1 through 7) are set to 0.

If multiple priorities are used (operator policy) then all bandwidths associated with priorities not in use must be set to zero.

12 Support of G.7715.1 Link Attributes – Layer Scoped Attributes and Local Connection Type Supported

12.1 Scoping of Link Attributes to a Specific Layer

In addition to link capacity, Section 9.5.1 of [G.7715.1] describes the following link capabilities to be advertised on a per-layer basis:

- Link Weight: This attribute represents a vector of one or more metrics, each of which indicates the relative desirability of a particular link over another during path selection.

- Resource Class: This attribute corresponds to a set of administrative groups assigned by the operator to this link. A link may belong to zero, one or more administrative groups.

- Local Connection Type: This attribute identifies whether the local SNP represents a TCP, CP, or can be flexibly configured as either a TCP or a CP.

- Link Availability: This attribute represents a vector of one or more availability factors for the link or link end. Availability may be represented in different ways between domains and within domains. Within domains it may be used to represent a survivability capability of the link or link end. In addition, the availability factor may be used to represent a node survivability characteristic.

- Diversity Support: This attribute represents diversity information with respect to links, nodes and Shared Risk Groups (SRGs) that may be used during path computation.

- Local Client Adaptations Supported: This attribute represents the set of client layer adaptations supported by the TCP associated with the Local SNPP. This is only applicable when the local SNP represents a TCP or can be flexibly configured as either a TCP or CP.

Protocol extensions for all of these attributes, except Local Connection Type and Local Client Adaptations Supported, are already defined for TE-Links in [RFC3630] and [RFC4203]. Note that a single metric is supported for the Link Weight attribute.

However, since TE-Links address multiple layers, there is no method to scope an attribute to a specific layer. An extension that allows these attributes to be scoped to a layer is necessary.

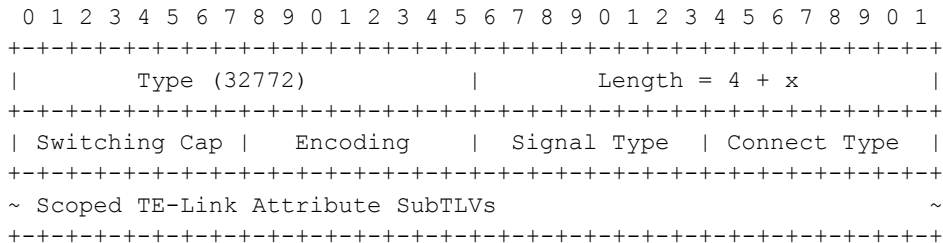If the connection type is not provided for a layer, then it defaults to TCP+CP.

12.2   Local Connection Type

Local Connection Type defines the type of G.805 entity that exists at the advertising end of the link. Since the entity at the far end may be flexibly configured, encoding of multiple entity types at the same time is necessary. To address this, a bit-vector is used to encode the connection types.

12.3   Link Attribute Scoping and Connection Type sub-TLV

This document proposes OIF private extensions to scope link attributes to a specific layer and to specify the local connection type.

This sub-TLV (Type 32772 and Length (4 + x) octets) is a sub-TLV of the top-level Link TLV. It is used for scoping TE-Link attributes to a specific layer and has the following format:

```
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Type (32772)        |          Length = 4 + x           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap |  Encoding   | Signal Type | Connect Type |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~ Scoped TE-Link Attribute SubTLVs                             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Note: x is the length of all the SubTLVs (including Type and Length fields) contained within the scoping subTLV.

Switching Cap, Encoding and Signal Type fields are encoded as defined in sections 11.1.2.1 (SONET/SDH), 11.1.2.2(OTN ODUk) and 11.2 (packet-based layer - Signal Type MUST be set to 0 in that case).

The following sub-TLVs can be encoded within a Link Attribute Scoping and Connection Type sub-TLV:

- Traffic Engineering Metric sub-TLV ;

- Administrative Group sub-TLV (Resource Class) ;

- Link Protection Type sub-TLV ;

- SRLG sub-TLV.

- Transitional Link (section 14.1)

A given sub-TLV may appear both as a sub-TLV of the top-level Link TLV, and as a sub-TLV of a Link Attribute Scoping and Connection Type sub-TLV. For a given layer, a sub-TLV encoded in a Link Attribute Scoping and Connection Type sub-TLV has precedence over the same sub-TLV encoded as a first-level sub-TLV of the top-level Link TLV.

The connection type is encoded using a bit vector:

- 0bxxxxxxx1 – Transit (i.e. CP)

- 0bxxxxxx1x – Trail Sink (i.e. TCP)

Other bits are in this bit-vector are reserved, and must be sent as 0 and ignored on reception.

13 Support of G.7715.1 Link Attributes – Link Availability

13.1 Link availability advertisement - I-NNI abstract links

A link may support multiple protection schemes. Moreover, a link may be an abstraction for multiple I-NNI links and nodes.

There are two possible approaches to deal with such I-NNI abstract links.

1. One approach generates as many advertisements as supported protection types. Each advertisement specifies the link capacity for a particular protection type.

2. The second approach advertises multiple protection types for each abstract link. The link capacity setting can either:

   a. Follow an "over-pessimistic" logic, where the advertised bandwidth is the one truly available for the highest protection type. However, such a link may then be excluded during path computation, while enough bandwidth was available for the (lower) requested protection;

   b. Follow an "over- optimistic" logic, where the advertised bandwidth is the one available for the lowest protection type. Crankback will be used if it turns out that not enough bandwidth is available for the requested protection type, at signaling time.

The first approach provides better accuracy at the cost of advertising multiple abstract links.

This document takes no stand about which approach should be deployed in a network. Both approaches must be implemented along with a configuration item that facilitates carrier configuration. The choice of deploying either approach should be left to the carrier according to its internal policy.

13.2   Link availability advertisement - E-NNI links

High-availability services over an E-NNI may be achieved:

- Either through 1+1 server-layer E-NNI links;

  In that case, the control plane will establish a single connection over the E-NNI link.

- Or through multiple unprotected E-NNI links;

  In that case, the control plane will establish two connections, a working connection and a protecting connection, each one over a different unprotected E-NNI link, using the E-NNI Recovery amendment signaling extensions.

From a path computation perspective, routing advertisements must specify that such high-availability services are supported over a particular E-NNI link.

- In case of 1+1 server-layer E-NNI link, the Link Protection type sub-TLV is set to Dedicated 1+1 (0x10) per RFC 4203;
- In case of multiple unprotected E-NNI links:
  - An abstract link should be advertised with a Link Protection type sub-TLV set to Dedicated 1+1 (0x10);
  - If unprotected services are to be supported over the E-NNI physical links, then additional individual E-NNI links should be advertised with a Link Protection type sub-TLV set to Unprotected (0x00);

Figure 5 provides an example physical topology supporting high-availability services, and the resulting routing advertisements: E-NNI links and I-NNI abstract links.

Physical Topology

E-NNI Routing Topology

Link advertised as "Dedicated 1+1"
Links advertised as "Unprotected"
Links advertised with both "Dedicated 1+1" and "Unprotected"
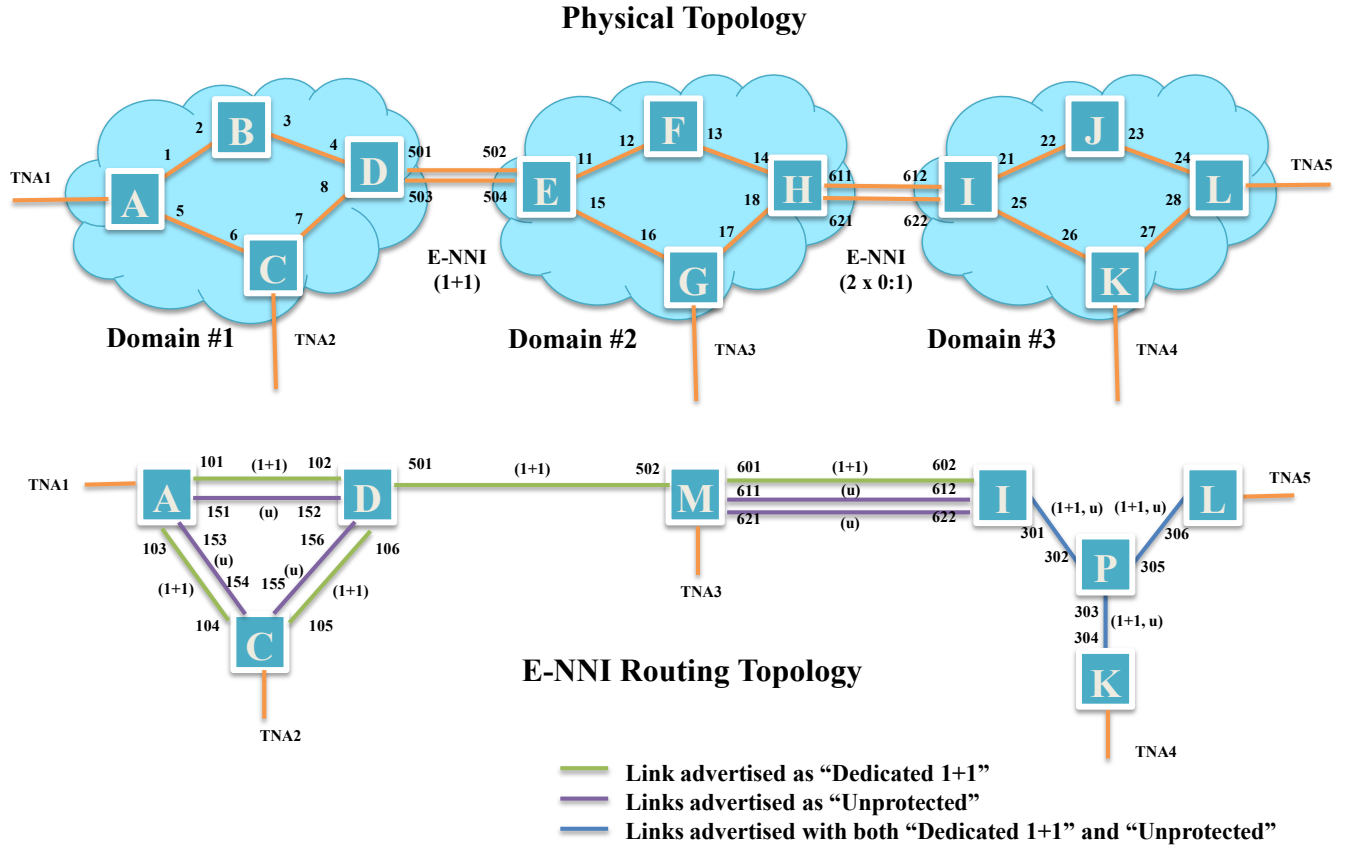
Figure 5 : Routing advertisements for high-availability services support

## 13.3   Standard Protocol Extensions

No protocol extension is proposed, since existing routing protocol specifications (Link Protection type sub-TLV in [RFC4203]) allow for either approach to be implemented by an I-NNI implementation as the bit vector format used to advertise the protection types on a TE-Link allow for one or multiple protection types to be advertised.

14 Multilayer Extensions

14.1   Transitional Links

Since Transitional Links connect between specific client and server layers, the lower and upper layers must be distinctly identified. As a result, the Transitional Link encoding utilizes the Link Attribute Scoping and Connection Type sub-TLV defined in section 12.3.

Within the server layer Link Attribute Scoping and Connection Type sub-TLV a transitional link sub-TLV is defined as follows:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (32777)          |             Length           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~             Transitional Link Attribute Sub-TLVs             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
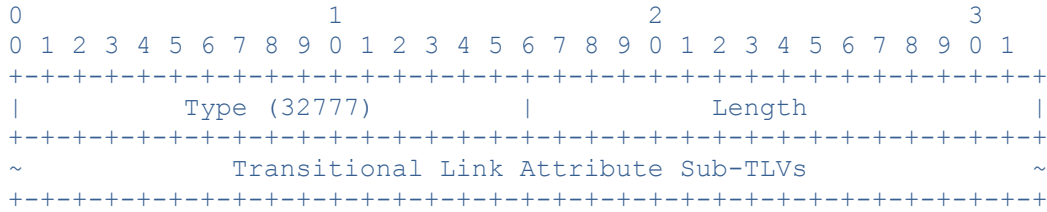
When included in a server layer Link Attribute Scoping and Connection Type sub-TLV, , a transitional link sub-TLV contains the following information:

Type value	TLV	Semantics	Reference	Mandatory /Optional
5	Link metric sub-TLV	Link cost	[RFC3630]	M (by default equal to 1) MUST be included once
9	Link resource class sub-TLV	Color	[RFC3630]	O (by default bit mask equal to 0…0) MUST not appear more than once
14	Link protection type sub-TLV	Link protection type	[RFC4203]	O (by default unprotected links) MUST not appear more than once
16	SRLG sub-TLV	Shared risk link group	[RFC4203]	O (by default Link ID is the SRLG) MUST not appear more than once
32769	Link Endpoint Identifier and Adaptation Stack sub-TLV	Stack of link endpoint identifiers and adaptation	Section 14.1	M (MUST be included once)

32770	Client:Server Capacity Ratio	Ratio of client to server layer capacity	Section 14.1	O (MUST not appear more than once)
32771	Inverse Multiplexing	Inverse multiplexing minimum and maximum constituents	Section 14.1	M for VCAT
32778	Layer-Scoped Multilayer TNA	Server layer TNA(s) for an adaptation	Section 14.2.2	Optional, MAY appear more than once

The transitional link attribute sub-TLVs apply to the transitional link, not to one adaptation in the stack. For example, cost represents the cost for the transitional link, not a particular adaptation cost.

The Transitional Link Attribute sub-TLVs reuses the sub-TLVs defined for Links, with the following sub-TLVs acceptable for use:

- Link Endpoint Identifier and Adaptation Stack sub-TLV

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Type (32769)          |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Server Interface ID for 1st adaptation (from server) in stack |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                 Intermediate layers                          ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap |   Encoding    |  Signal Type  |Adaptation-last|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Client Interface ID for last adaptation in stack      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Client Node ID for client layer                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+


Where Intermediate layers consists of 0 or more of the following:
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap |   Encoding    |  Signal Type  | Adaptation (N)|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Client Interface ID for the Nth adaptation       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Node ID for intermediate layer between Nth and N+1th adaptation|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Server Interface ID for the (N+1)th adaptation       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

  - This sub-TLV allows for multiple layers to be encoded. Since it is possible that a client adaptation may go through intermediate layers (e.g. Ethernet MAC over VCAT over VC4), the encoding allows for more than one higher layer to be named. The stack of intermediate layers described is from bottom to top (i.e.

the first Switching Type / Encoding / Signal Type describes the layer right above the announced server, the second then described the layer on top of that layer, etc.).

o The server interface ID for the bottom adaptation MAY be the same as the local link identifier encoded in the Local/Remote Identifiers sub-TLV for the top-level link TLV.

o For intermediate layers, the same node id is the client to the bottom adaptation and server to the top adaptation. The interface id before the node id represents the client layer identifier for the bottom adaptation and the one after the node id represents the server layer identifier for the top adaptation.

o Client interface identifier MUST be unique for the combination of client node id and adaptation.

o Server interface identifier MUST be unique for the combination of server node id and adaptation.

o The specific adaptation used (scoped by the server and client layers) is specified.

- Switching Cap / Encoding

  - VCAT:

    o Switching Capability (8 bits): value 100 (TDM)

    o Encoding (8 bits): value 0x85

  - SONET/SDH (excluding VCAT) and OTN:

    o As defined in section 11.1.2

  - Ethernet:

    o As defined in section 11.2

- Signal Type

  - Set to 0 for Ethernet and VCAT

  - As defined in section 11.1.2.1 for SONET/SDH

  - As defined in section 11.1.2.2 for OTN

- Adaptation (scoped by server layer)

  o SONET/SDH (excluding VCAT) server layer

  - GFP-F   0x01

  - GFP-T   0x02

  - VCAT (no LCAS) 0x03

- VCAT (LCAS)   0x04
  o   SONET/SDH VCAT server layer
    - GFP-F   0x01
    - GFP-T   0x02
  o   OTN ODUk (where k = 1,2,3) server layer
    - GFP-F   0x01
    - GFP-T   0x02
    - AMP   0x05
    - BMP   0x06

- Cost (Link Weight)

- Administrative Group (Resource Class)

- SRLG sub-TLV

- Link Protection type

- Client:Server Capacity Ratio

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type (32770)          |           Length = 4          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Client:Server Capacity Ratio                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

  o   The ratio of client to server layer capacity (e.g. 1:1, 24:1, 1:24) describes the relationship between the smallest allocatable unit in the client layer capacity and the corresponding number of client layer units that come from allocating the smallest unit in the server layer, allowing for path computation to be performed by a system unaware of the client/server adaptation technology. It is a 32-bit IEEE formatted floating point number. For inverse multiplexing, the smallest unit corresponds to one constituent and the Inverse Multiplexing sub-TLV indicates the minimum and maximum number of constituents that can be inverse multiplexed. For the Ethernet and VCAT layer, the smallest unit is bytes per second. Some example encodings for adaptations supported by this amendment are listed in section 21.

- Inverse Multiplexing

  o   Applicable and mandatory for VCAT encoding only. The following new Inverse Multiplexing sub-TLV is also defined to indicate the minimum and maximum number of constituents in inverse multiplexed groups. The Switching Cap, Encoding and Signal Type represent the constituent layer.

- o iMuxMin: The minimum number of constituents that can be inverse multiplexed together. For example, an implementation that supports full range of VCAT from VC4-1v to VC4-7v would set iMuxMin to 1.

- o iMuxMax: The maximum number of constituents that can be inverse multiplexed together. For example, an implementation that supports full range of VCAT from VC4-1v to VC4-7v would set iMuxMax to 7.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type (32771)         |            Length = 8         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap |   Encoding     |  Signal Type  |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            iMuxMin             |              iMuxMax          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- Multilayer TNA (as defined in section 14.2)

14.1.1 Example Transitional Link Encoding

An example use of the transitional link would be to encode the following capability
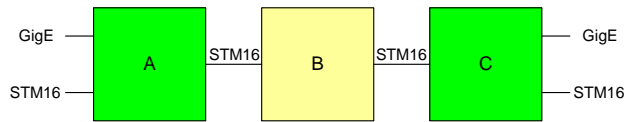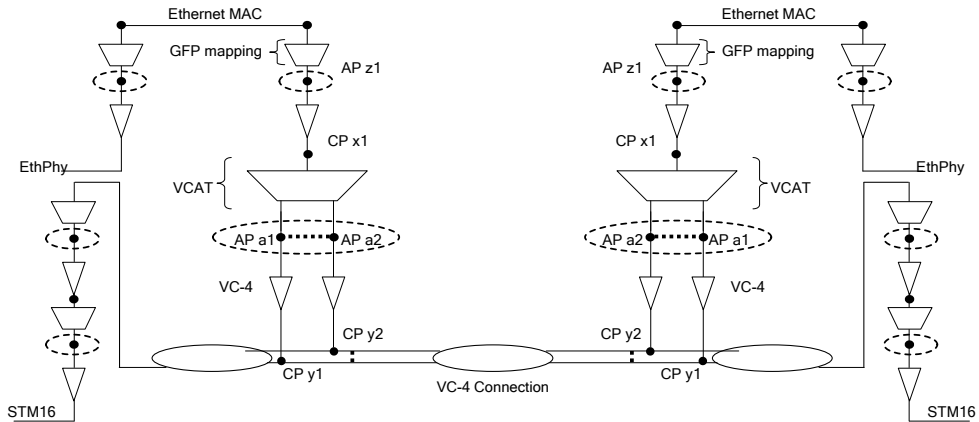

Figure 6. Box equipment view

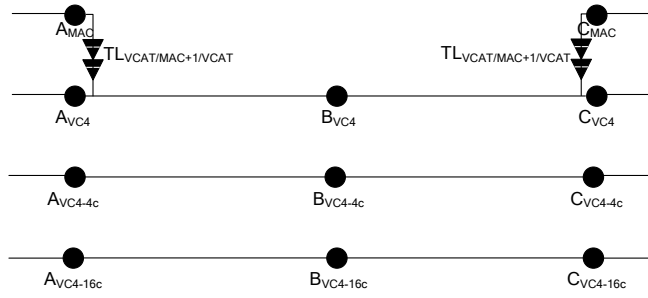Figure 7. G.805 drawing showing EthMac and VC4 layer relationship

Figure 8. Link-oriented Routing Topology showing EthMac/VCAT Transitional Link

The client adaptation sub-TLV is only present in the link advertisement generated on behalf of the node supporting the adaptation function, i.e. it is advertised by nodes A and C in Figure 7, but not advertised by node B.

The resulting bandwidth advertisement from A for the A-B link (when empty) using a Link-oriented encoding would be as described below. Note that this example only includes bandwidth-related TLVs. Many mandatory TLVs have been omitted for simplicity.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (2)             |          Length = 152         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (15)            |          Length = 72          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| TDM    (100) | SONET/SDH (5) |          Reserved (0)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Max LSP Bandwidth - Priority 0                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Max LSP Bandwidth - Priority 1                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Max LSP Bandwidth - Priority 2                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Max LSP Bandwidth - Priority 3                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Max LSP Bandwidth - Priority 4                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Max LSP Bandwidth - Priority 5                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Max LSP Bandwidth - Priority 6                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Max LSP Bandwidth - Priority 7                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (1)             |          Length =   8         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| VC4      (6)  |UnRs(0)|Rsvd(0)|1 0 0 0 0 0 0 0| Reserved (0)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Unallocated Timeslots (16)    | Padding (0)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (1)             |          Length =   8         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| VC4-4c (21)  |UnRs(0)|Rsvd(0)|1 0 0 0 0 0 0 0| Reserved (0)   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Unallocated Timeslots (4)     | Padding (0)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (1)             |          Length =   8         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| VC4-16c(22)  |UnRs(0)|Rsvd(0)|1 0 0 0 0 0 0 0| Reserved (0)   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Unallocated Timeslots (1)     | Padding (0)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|          Type (32772)         |          Length =  76         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

```
| TDM     (100) | SONET/SDH (5) | VC4      (6) |    CP+TCP      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Type (32777)       |         Length = 64           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Type (32769)       |         Length = 32           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Server Interface ID for VC4/VCAT adaptation           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| TDM     (100) | VCAT (0x85)   |Signal Type (0)|   VCAT (4)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Client Interface ID for VC4/VCAT adaptation           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Client Node ID for VCAT layer                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Server Interface ID for VCAT/Ethernet adaptation       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| L2SC    (51)  | Ethernet  (2) |Signal Type (0)|   GFP-F  (1)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Client Interface ID for VCAT/Ethernet adaptation        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Client Node ID for Ethernet layer                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Type (5)              |         Length = 4            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Cost  (10)                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Type (32771)          |         Length = 8            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| TDM (100)     |  SONET/SDH(5) | VC4 (6)       |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       iMuxMin  (1)            |       iMuxMax  (7)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Type (32770)          |         Length = 4            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Client:Server Capacity Ratio                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

14.2 Multilayer TNA

14.2.1 Generic Multilayer TNA

A Generic Multilayer TNA sub-TLV is defined as follows:

```
0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type (32776)              |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                       TNA Sub-TLVs                            :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
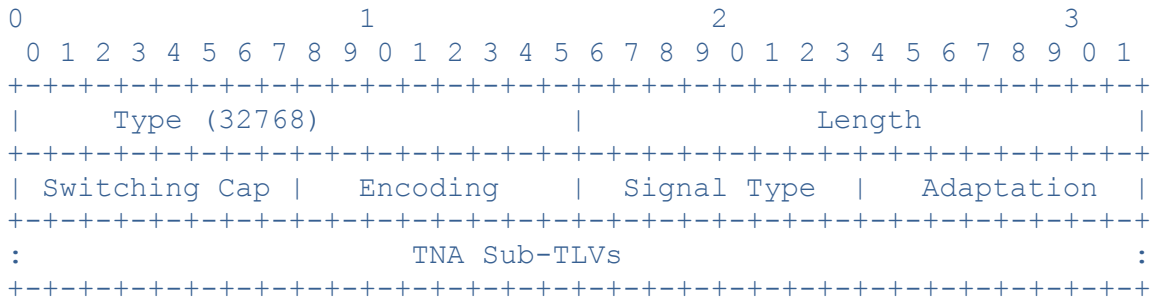
Where there may be one or more TNA Sub-TLVs.

14.2.2 Layer-Scoped Multilayer TNA

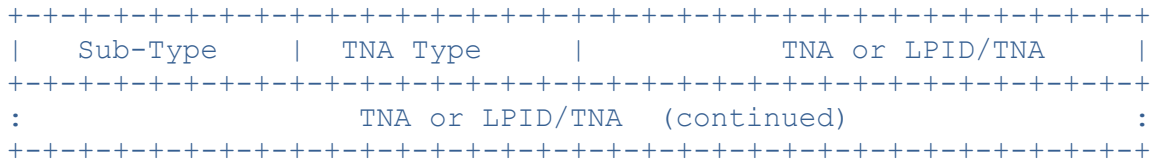A Layer-Scoped Multilayer TNA sub-TLV is defined as follows:

```
0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type (32768)              |              Length           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Switching Cap |    Encoding   |  Signal Type  |  Adaptation   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                         TNA Sub-TLVs                          :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Switching Cap, Encoding, Signal Type and Adaptation are set as in
section 14.1.

There may be one or more TNA Sub-TLVs.
```

14.2.3 Multilayer TNA sub-TLVs

TNA Sub-TLVs for the above two sub-TLVs are defined as follows:

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Sub-Type    |  TNA Type     |          TNA or LPID/TNA      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
:                   TNA or LPID/TNA   (continued)               :
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where Sub-Type is:

1- TNA

2- 32-bit Logical Port ID + TNA

TNA Type is:

1- IPv4 format

2- IPv6 format

3- NSAP format

TNA (for Sub-Type 1) contains a single TNA as a <Prefix Length, Prefix> tuple as described in section 10.3.1 (IPv4, IPv6) or section 10.4.1 (NSAP). This field is padded with zeroes for four-octet alignment.

LPID and TNA (for Sub-Type 2) contains a 32-bit logical port identifier followed by a TNA in the format specified above. This field is padded with zeroes for four-octet alignment.

14.2.4 Usage

In order to allow translation of client layer SNPP names into server layer TNAs and optional logical port identifier, TNAs are advertised as follows[4]:

- For pseudo-links: in the direction of the node to the pseudo-node, i.e. in the link TLV for the link where the local end is the advertising node and the remote end is the pseudo-node. The Generic Multilayer TNA is placed in the Link TLV.

- For transitional links: from the node that advertises the transitional link towards its adjacent neighbors. The Layer-Scoped Multilayer TNA is placed in the Transitional Link scope sub-TLVs. This includes a layer identifier to distinguish which Node ID/Interface ID this TNA is associated with.

15 Compatibility with OIF E-NNI Routing 2.0

In the ASON Architecture, the selection of routing protocol and the specific information elements carried by that protocol is specific to a routing area. For this reason, backward compatibility is only a consideration for an area where not all Routing Controllers support the multilayer Routing Extensions. Within such an area, the backward compatibility impact is different for each of the extensions proposed by this Amendment. They are as follows:

Pseudo-node Link advertisement

Links from a Node to a Pseudo-node (i.e. a pseudo-link) are advertised the same as a normal link with the additional optional Sub-TLV used to relate one or more server-layer TNA to the Pseudo Link link-end. All other characteristics of this link are identical to a normal link. As a result, the link can be included in paths developed by Routing Controllers that don't support the additional sub-TLV - the sub-TLV will be silently ignored. Ignoring this sub-TLV in path computation is of no consequence as the sub-TLV is providing translation information (i.e. acting as a directory service update mechanism) that is only needed when the client layer signaling process is requesting invocation of a server layer call.

The Client Layer Signaling Controller located at the point the Server Layer Call is initiated will need to be able to perform the translation of the Client-layer SNPP ID used for a pseudo-link into the Server Layer TNA. As this node, the Routing Controller will provide a directory service API to the Signaling Controller enabling lookup of a pseudo-link endpoint/Server-Layer TNA relationship.

---

[4] The same TNA or TNA/LPID tuple may be advertised several times, for multiple transitional links, pseudo-links or nodal TNAs within the same layer. Different LPIDs may be used to distinguish specific use.

Transitional Link advertisement

The Transitional Link adds adaptation information for the ends of links being advertised by routing. When this information is not understood by a Routing Controller (i.e. the RC does not support multilayer Routing Extensions), the Routing Controller ignores the transitional link advertisements for path computation purposes although it remains in the routing database and continues to be distributed to neighbors. It is impossible for the routing controller to compute a path that invokes server layer call(s). While client layer resources made available by server layer calls triggered by other client layer calls can be advertised and therefore used, the inability to invoke server layer calls may result in blocking of client layer call attempts.

16 <u>References</u>

16.1 ITU-T

[G.7041] ITU-T Recommendation G.7041/Y.1303 (2011), Amendment 1 (2012), *Generic Framing Procedure*

[G.707] ITU-T Recommendation G.707/Y.1322 (2003), *Network Node Interface for the Synchronous Digital Hierarchy (SDH)*

[G.709] ITU-T Recommendation G.709/Y.1331 (2003-03) *Interfaces for the Optical Transport Network (OTN)*

[G.783] ITU-T Recommendation G.783 (2006-03) *Characteristics of synchronous digital hierarchy (SDH) equipment functional blocks*

[G.7712] ITU-T Recommendation G.7712/Y.1703 (2003), *Architecture and Specification of Data Communication Network*

[G.7715] ITU-T Recommendation G.7715/Y.1706 (2002), *Architecture and Requirements for Routing in the Automatically Switched Optical Network*

[G.7715.1] ITU-T Recommendation G.7715.1/Y.1706.1 (2004), *ASON Routing Architecture and Requirements for Link State Protocols*

[G.7715.2] ITU-T Recommendation G.7715.2/Y.1706.2 (2007), *ASON routing architecture and requirements for remote route query*

[G.800] ITU-T Recommendation G.800 (2012), *Unified functional architecture of transport networks*

[G.805] ITU-T Recommendation G.805 (2000), *Generic Functional Architecture of Transport Networks*

[G.8080] ITU-T Recommendation G.8080/Y.1304 (2012), *Architecture for the Automatically Switched Optical Network (ASON)*

16.2   OIF

[OIF-UNI-02.0] OIF Implementation Agreement, "User Network Interface (UNI) 2.0 Signaling Specification Common Part," OIF-UNI-02.0-Common, February 2008.

[OIF-E-NNI-sig-02.0] OIF Implementation Agreement, "OIF E-NNI Signaling Specification", OIF-E-NNI-Sig-02.0, April 2009

[OIF-E-NNI-OSPF-01.0] OIF Implementation Agreement, "External Network-Network Interface (E-NNI) OSPF-based Routing - 1.0 (Intra-Carrier) Implementation Agreement", OIF-ENNI-OSPF-01.0, January 2007

[OIF-ENNI-ML-AM-01.0] OIF Implementation Agreement, "Multilayer Amendment to E-NNI 2.0 – Common Part", OIF-ENNI-ML-AM-01.0, April 2013.

[OIF-ENNI-OSPF-02.1] OIF Implementation Agreement, "Multilayer Amendment to E-NNI 2.0 – OSPFv2 based routing", OIF-ENNI-OSPF-02.1, April 2013.

[OIF-ENNI-REC-AM-01.0] OIF Implementation Agreement, "Recovery Amendment to E-NNI 2.0 – Common Part", OIF-ENNI-REC-AM-01.0

[SecAdd] Optical Internetworking Forum Implementation Agreement, "Addendum to the Security Extension for UNI and NNI," OIF-SEP-02.1, March 2006.

[SecExt] Optical Internetworking Forum Implementation Agreement, "Security Extension for UNI and E-NNI 2.0," OIF-SEP-03.0, November 2010.

[SysLog] Optical Internetworking Forum Implementation Agreement, "OIF Control Plane Logging and Auditing with Syslog version 1.1," OIF-SLG-01.1, July 2010.

16.3   IETF

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

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

[RFC4202] Kompella, K., Ed., and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005.

[RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-protocol Label Switching", RFC 4203, October 2005.

[RFC4328] D. Papadimitriou, Ed. "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 Optical Transport Networks Control", RFC 4328, January 2006.

[RFC4397] I. Bryskin, A. Farrel, A Lexicography for the Interpretation of Generalized Multiprotocol Label Switching (GMPLS) Terminology within the Context of the ITU-T's Automatically Switched Optical Network (ASON) Architecture. RFC 4397, February 2006.

[RFC4606] Mannie, E. and D. Papadimitriou, "Generalized Multi-Protocol Label Switching (GMPLS) Extensions for Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) Control", RFC 4606, August 2006.

[RFC4652] D.Papadimitriou (Ed.) et al. "Evaluation of existing Routing Protocols against ASON Routing Requirements", RFC 4652, October 2006..

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.

[RFC4657] Ash, J. and J. Le Roux, "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, September 2006.

[RFC4940] K. Kompella, B. Fenner, "IANA Considerations for OSPF", RFC 4940, June 2007.

[RFC5376] Bitar, N., Zhang, R., and K. Kumaki, "Inter-AS Requirements for the Path Computation Element Communication Protocol (PCECP)", RFC 5376, November 2008.

[RFC5441] JP. Vasseur, Ed., R. Zhang, N. Bitar, JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", RFC 5441, April 2009.

[RFC5786] R. Aggarwal, K. Kompella , "Advertising a Router's Local Addresses in OSPF TE Extensions" RFC5786, March 2010.

[RFC5787] D. Papadimitriou, "OSPFv2 Routing Protocol Extensions for ASON Routing", RFC5787, March 2010.

[OSPF-ASON] D. Papadimitriou, "OSPFv2 Routing Protocol Extensions for ASON Routing", draft-ietf-ccamp-gmpls-ason-routing-ospf-08.txt (Work in Progress).

[RFC6002] L. Berger, D. Fedyk, "Generalized MPLS (GMPLS) Data Channel Switching Capable (DCSC) and Channel Set Label Extensions" RFC6002, October 2010

[IANA-OSPF-TE] http://www.iana.org/assignments/ospf-traffic-eng-tlvs

16.4 ANSI

[T1.105] ANSI T1.105: *SONET Basic Description including Multiplex Structure, Rates and Formats*

16.5 IEEE

[IEEE802.3] IEEE 802.3-2005: IEEE Standard for Information technology - Telecommunications and information exchange between systems - IEEE standard for local and metropolitan area networks - Specific requirements – Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications.

17 Appendix I: E-NNI OSPF-based Routing with a Single Hierarchical Level

A prerequisite for hierarchical OSPF routing is that each control domain has at least one Routing Controller as defined in [G.7715]. This RC advertises topology associated with a Routing Area (with a specific RA ID), and has an RC ID and an SCN address for its OSPF Protocol Controller, to which all protocol messages will be addressed.

Via discovery or configuration, each RC finds out about its peer RCs within their common parent RA. Their RC IDs and corresponding SCN addresses are discovered or configured. Automated discovery of peer RCs is for further study.

If peer RCs are determined via configuration, a decision is made to establish a control adjacency with a particular neighbor RC for the purposes of routing information exchange.

17.1 Configuration

To bring up the hierarchy, there is a set of configuration parameters as described in the following sections.

The example of Figure 9 shows an optical network with three routing control domains. A single level of hierarchy of OSPF is configured as described in the following sections.
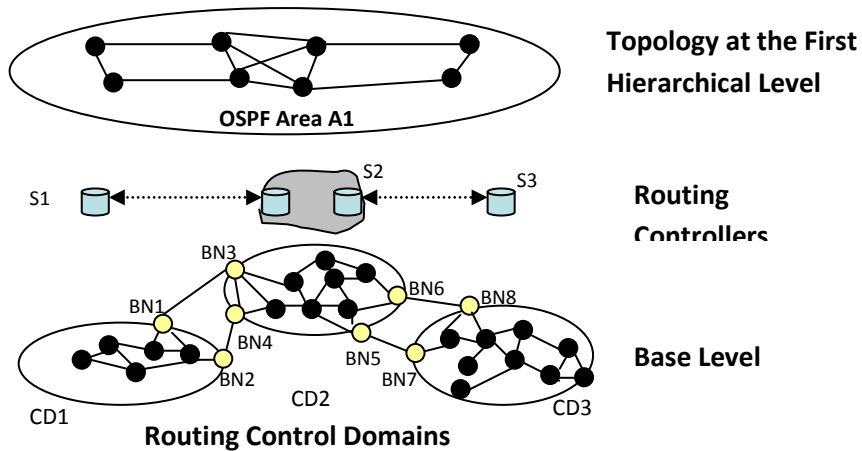
Figure 9: Example of Single Level OSPF-TE Operation.

17.1.1 Routing Controllers

Each routing control domain includes at least one routing controller. A routing controller is identified by its RC ID.

In the example of Figure 9, S2 is a federation of multiple RCs that advertises routing information for CD2 within Area A1.

17.1.2 Routing Controllers in Adjacent Routing Control Domains (per RC)

For each routing controller advertising for a given routing control domain, there exists at least one peer RC advertising for each adjacent control domain, and for each RC, the following information MUST be available:

1) The Routing Controller ID of the neighboring Routing Controller.

2) The SCN address of the neighboring Routing Controller.

For example, information about adjacent control domains provisioned in S2 is as follows (S2 being a federation of multiple RCs, note that each RC only needs to be provisioned with the information related to its neighboring domain RC, with which a routing adjacency exists):

Table 7 Adjacent Routing Controllers of S2 in Figure 9

Neighboring RCs	RC ID	SCN Address
S1	S1's Router ID	S1's SCN address
S3	S3's Router ID	S3's SCN address

17.1.3 Inter-Domain Links (per RC)

Information on inter-domain links can be configured on an RC. An inter-domain link reflects an inter-connection with an adjacent domain along with the traffic parameters in the outgoing direction, i.e., from the local node to the remote (adjacent) node.

Note that a link is identified within the scope of a node, not the scope of the advertising RC.

For example, there are four inter-domain links from the perspective of CD2 provisioned on S2 as follows (S2 being a federation of multiple RCs, note that each RC only needs to be provisioned with the inter-domain links between S2 and the neighboring domains for which this RC maintains a routing adjacency with a peer RC):

Table 8 Inter-domain Links Configured on RC S2 in Figure 9

Inter-domain links	Local border node	Remote border node
BN3-BN1	BN3	BN1
BN4-BN2	BN4	BN2
BN5-BN7	BN5	BN7
BN6-BN8	BN6	BN8

17.1.4 Intra-Domain Links (per RC)

An RC may advertise an intra-domain topology using a set of border nodes and abstract intra-domain links. See the abstract link model described in section 6.3.2.2.2.

If such a model is used, one or more intra-domain links can be configured on an RC or may be derived by the RC from internal domain routing information. An intra-domain link reflects some characteristics of traversing the domain, as reflected by advertised link traffic parameters on one direction, i.e. from the ingress node to the egress node.

Note that a link is identified within the scope of a node, not the scope of the advertising RC.

In the example of Figure 9, 12 intra-domain links are advertised by S2 to reflect characteristics of traversing CD2 from one border node to another as follows:

Table 9 Intra-domain links Configured on RC S2 in Figure 9

Intra-domain links	Local border node	Remote border node
BN3-BN4	BN3	BN4
BN4-BN3	BN4	BN3
BN3-BN5	BN3	BN5
BN5-BN3	BN5	BN3
BN3-BN6	BN3	BN6
BN6-BN3	BN6	BN3
BN4-BN5	BN4	BN5
BN5-BN4	BN5	BN4
BN4-BN6	BN4	BN6
BN6-BN4	BN6	BN4
BN5-BN6	BN5	BN6
BN6-BN5	BN6	BN5

Note the intra-domain links are abstract in nature, reflecting the aggregation of the topology in the RC. Also the number of intra-domain links that need to be provisioned is a local matter.

17.1.5 The Reachable TNA Names (per RC)

TNA addresses reachable within the CD2 and needing to be advertised by S2 are provisioned on S2 or derived from internal routing information.

Note that TNA names are associated to nodes, not to advertising RCs.

17.2   Operation

The purpose of the configuration as described in the section above is to start the first hierarchical level of an OSPF-TE based routing control domain. Each RC that has been configured starts to run as an OSPF-TE node at the first hierarchical level by exchanging OSPF-TE messages with the neighboring RCs. No routing adjacencies are created directly between neighboring border nodes unless they are also serving as RCs for their respective domains.

The RCs in the first hierarchical level form routing adjacencies in the control plane, and at the same time, each RC advertises the links that correspond to the inter-domain and intra-domain links for its associated domain. Also, each RC advertises the reachable TNA names for that domain.

In the example of Figure 9, RCs S1, S2 and S3 will form regular OSPF routing adjacencies in the control plane [Note: the detailed implementation of S2 as a federation of routing controllers is beyond the scope of this document]. At the same time, S2 will advertise the abstract links that correspond to Table 8 as part of the topology of the first hierarchical level area (OSPF area A1 in the Figure). The links that correspond to Table-3 specify the link attributes from CD2 to CD1 and from CD2 to CD3, and the corresponding link attributes from CD1 to CD2 and from CD3 to CD2 will be advertised by S1 and S3, respectively. The links that correspond to Table 9 expose routing information associated with CD2, and are useful during the routing path selections for connections that traverse CD2, i.e., which entry border node to use for ingress, and which exit border node to use for egress. In addition, S2 will advertise CD2's reachable TNA names throughout OSPF area A1.

Operation of the routing protocol (e.g., Database Synchronization and Link State Advertisement Flooding) otherwise follows procedures defined in [RFC2328] and in Section 3. Timers for generating link advertisements must be configurable by the operator to avoid mismatch at sending and receiving nodes.

17.3   Single Level Example

This appendix gives an E-NNI routing/signaling example for a single level of hierarchy.

The example uses a single RC node for each domain that advertises the domain topology. In practice, multiple RCs can be associated with a single domain.

The following are examined:

1) Control plane topology

2) Data plane topology

3) Connection path computation and ERO construction

4) Call progression at the domain boundary

17.3.1 The Control Domains

In a routing hierarchy, an RA is partitioned to create a lower level of RAs and interconnecting SNPP links. The internal structure of the RA is known "inside" the RA, but not from "outside". (That is, inside RA 1, the topology is known to include three child RAs interconnected by two SNPP links; from outside RA 1 this is opaque).
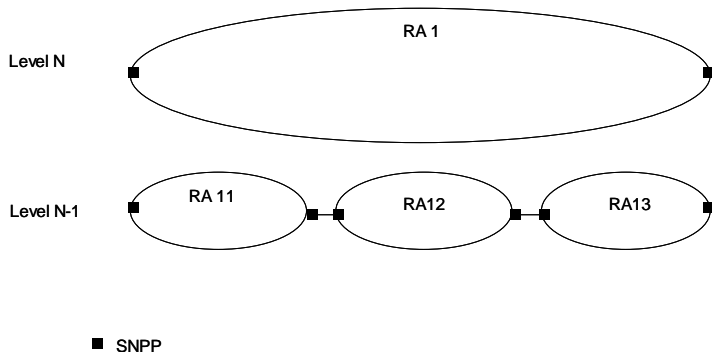

Figure 10: ASON Routing Hierarchy

Consider now two RCDs at a given hierarchical routing level with an SNPP link between them.
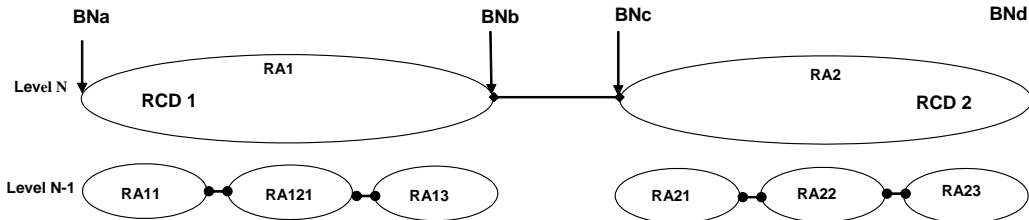
Figure 11: Routing Control Domains

There are several potential approaches to advertising costs of traversing an RCD. Two approaches are discussed below.

Abstract node: The representation of an RCD is as a single node with no internal structure. The topology seen in the E-NNI routing protocol at Level *N* includes two nodes (AN1 and AN2) and one (inter-RCD) link as below.
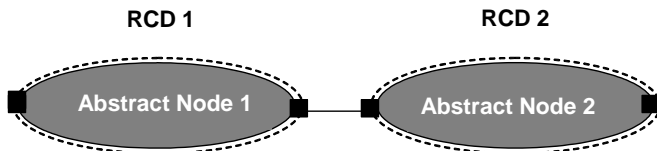
Figure 12: Abstract Node Representation

Abstract link: The representation of an RCD is in terms of its border nodes and intervening (intra-RCD) "abstract" SNPP links. The resulting topology seen in the E-NNI routing protocol at Level *N* includes 4 nodes (BNa, BNb, BNc and BNd), and three SNPP links.
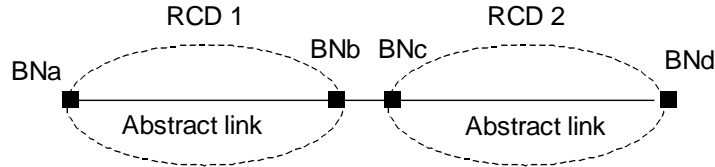

Figure 13: Abstract Link Representation

17.3.2   Single level topology example

Four routing control domains in the example for single level hierarchy are shown in Figure 14, i.e., CD1, CD2, CD3 and CD4.

In this example, the abstract link model is used for CD 1 and CD 2 and the abstract node model is used for CD 3 and CD 4.
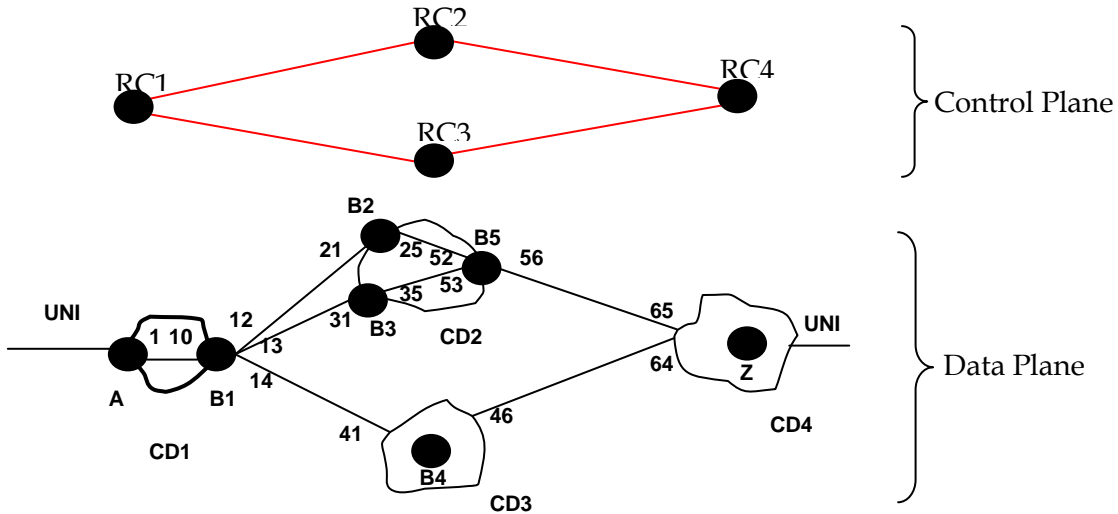

Figure 14: Example Topology

17.3.2.1   The Control Plane

Four OSPF nodes are shown in Figure 14, i.e., RC1, RC2, RC3 and RC4 that form the control adjacencies as shown in the red color. The four OSPF nodes represent the four control domains, respectively. Again, the example uses one RC per domain, but in practice, multiple RCs may be used for a particular domain.

17.3.2.2 Data Plane

The data plane and its topology are shown as in Figure 14. Note there are border nodes, inter-domain links and intra-domain links. Both border node ID (B1, B2, etc.) and link interface ID (13, 31, etc.) are marked in Figure 14.

17.3.2.3 Advertising Links from RC1

The following links are advertised by RC1:

- `B1->B2 (an inter-domain link)`

    - Advertising Router is RC1

    - Local and Remote TE Router ID sub-TLV contains B1 and B2

    - Local interface ID sub-TLV contains 12

    - Remote interface ID sub-TLV contains 21

    - Link ID sub-TLV set to 0.0.0.0

- `B1->B3 (an inter-domain link)`

    - Advertising Router is RC1

    - Local and Remote TE Router ID sub-TLV contains B1 and B3

    - Local interface ID sub-TLV contains 13

    - Remote interface ID sub-TLV contains 31

    - Link ID sub-TLV set to 0.0.0.0

- `B1->B4 (an inter-domain link)`

    - Advertising Router is RC1

    - Local and Remote TE Router ID sub-TLV contains B1 and B4

    - Local interface ID sub-TLV contains 14

    - Remote interface ID sub-TLV contains 41

    - Link ID sub-TLV set to 0.0.0.0

- A->B1 (an intra-domain link)

    - Advertising Router is RC1

    - Local and Remote TE Router ID sub-TLV contains A and B1

    - Local interface ID sub-TLV contains 1

    - Remote interface ID sub-TLV contains 10

    - Link ID sub-TLV set to 0.0.0.0

- B1->A (an intra-domain link)

  - Advertising Router is RC1

  - Local and Remote TE Router ID sub-TLV contains B1 and A

  - Local interface ID sub-TLV contains 10

  - Remote interface ID sub-TLV contains 1

  - Link ID sub-TLV set to 0.0.0.0

17.3.2.4   Advertising Links from RC2

The following links are advertised by RC2:

a) B2->B1 (an inter-domain link)

  - Advertising Router is RC2

  - Local and Remote TE Router ID sub-TLV contains B2 and B1

  - Local interface ID sub-TLV contains 21

  - Remote interface ID sub-TLV contains 12

  - Link ID sub-TLV set to 0.0.0.0

b) B3->B1 (an inter-domain link)

  - Advertising Router is RC2

  - Local and Remote TE Router ID sub-TLV contains B3 and B1

  - Local interface ID sub-TLV contains 31

  - Remote interface ID sub-TLV contains 13

  - Link ID sub-TLV set to 0.0.0.0

c) B5->Z (an inter-domain link)

  - Advertising Router is RC2

  - Local and Remote TE Router ID sub-TLV contains B5 and Z

  - Local interface ID sub-TLV contains 56

  - Remote interface ID sub-TLV contains 65

  - Link ID sub-TLV set to 0.0.0.0

d) B2->B5 (an intra-domain link)

  - Advertising Router is RC2

  - Local and Remote TE Router ID sub-TLV contains B2 and B5

  - Local interface ID sub-TLV contains 25

- Remote interface ID sub-TLV contains 52

- Link ID sub-TLV set to 0.0.0.0

e) B5->B2 (an intra-domain link)

- Advertising Router is RC2

- Local and Remote TE Router ID sub-TLV contains B5 and B2

- Local interface ID sub-TLV contains 52

- Remote interface ID sub-TLV contains 25

- Link ID sub-TLV set to 0.0.0.0

f) B3->B5 (an intra-domain link)

- Advertising Router is RC2

- Local and Remote TE Router ID sub-TLV contains B3 and B5

- Local interface ID sub-TLV contains 35

- Remote interface ID sub-TLV contains 53

- Link ID sub-TLV set to 0.0.0.0

g) B5->B3 (an intra-domain link)

- Advertising Router is RC2

- Local and Remote TE Router ID sub-TLV contains B5 and B3

- Local interface ID sub-TLV contains 53

- Remote interface ID sub-TLV contains 35

- Link ID sub-TLV set to 0.0.0.0

17.3.2.5 Advertisements from RC3

The following links are advertised by RC3:

a) B4->B1 (an inter-domain link)

- Advertising Router is RC3

- Local and Remote TE Router ID sub-TLV contains B4 and B1

- Local interface ID sub-TLV contains 41

- Remote interface ID sub-TLV contains 14

- Link ID sub-TLV set to 0.0.0.0

b) B4->Z (an inter-domain link)

- Advertising Router is RC3

- Local and Remote TE Router ID sub-TLV contains B4 and Z

- Local interface ID sub-TLV contains 46

- Remote interface ID sub-TLV contains 64

- Link ID sub-TLV set to 0.0.0.0

17.3.2.6 Advertisements from RC4

The following links are advertised by RC4:

```
a) Z->B5 (an inter-domain link)
```

- Advertising Router is RC4

- Local and Remote TE Router ID sub-TLV contains Z and B5

- Local interface ID sub-TLV contains 65

- Remote interface ID sub-TLV contains 56

- Link ID sub-TLV set to 0.0.0.0

```
b) Z->B4 (an inter-domain link)
```

- Advertising Router is RC4

- Local and Remote TE Router ID sub-TLV contains Z and B4

- Local interface ID sub-TLV contains 64

- Remote interface ID sub-TLV contains 46

- Link ID sub-TLV set to 0.0.0.0

17.3.2.7 Path Computation at the UNI-N and ERO

Suppose one wants to make a connection from A to Z in Figure 14: the source node A sees there are three possible routes, i.e.,

1) A->B1->B2->B5-> Z

2) A->B1->B3->B5-> Z

3) A->B1->B4-> Z

If the chosen route is 1) above, then the ERO built by A is:

  A:1 -> B1:12 -> B2:25 -> B5:56 -> Z

If the chosen route is 3) above, then the ERO built by A is:

  A:1 -> B1:14 -> B4:46 -> Z

17.3.2.8 Path Expansion

If internal topology exists within a CD and is not advertised externally, a mapping or expansion of a received ERO is needed to fit the actual internal topology of the CD. For example, if CD3 in the figure above consists of multiple nodes, the ERO entry {B4:46} is expanded internally to match the actual ingress border node and internal path to the destination TNA.

18 Appendix II: Architecture for Operation with Multiple Hierarchical Levels

This appendix does not provide any protocol details about how to achieve multiple hierarchical levels. It only provides procedural guidance.

The routing hierarchy as proposed by this document is achieved by stacking separate routing areas vertically. The following convention is used throughout that document: hierarchical levels are numbered (e.g., *N*, *N*+1, *N*-1…) in such a way that a level is assigned a higher number than the lower levels it contains. A lower level routing area (level *N*) is completely contained within a single higher level routing area (level *N*+1).

The requirements defined in this section are intended to be consistent with requirements for hierarchical routing defined in [G.7715.1], Section 8. In a given routing area, a single routing protocol runs independently, and at least one RC, selected either via provisioning or election, which represents that RA at the next higher level in the routing hierarchy. Usually some communication mechanisms exist between the RC at hierarchical level *N* and routing entities within the control domain it represents in order to exchange routing information in both directions, i.e., routing information feed-up and feed-down, but this is internal to the domain.

18.1 Configuration

Some configuration is required to build up the routing hierarchy. An operator chooses the hierarchical structure of the routing areas, that is, the containment hierarchy of routing areas, which is usually a reflection of the hierarchical organization of the operator's network.
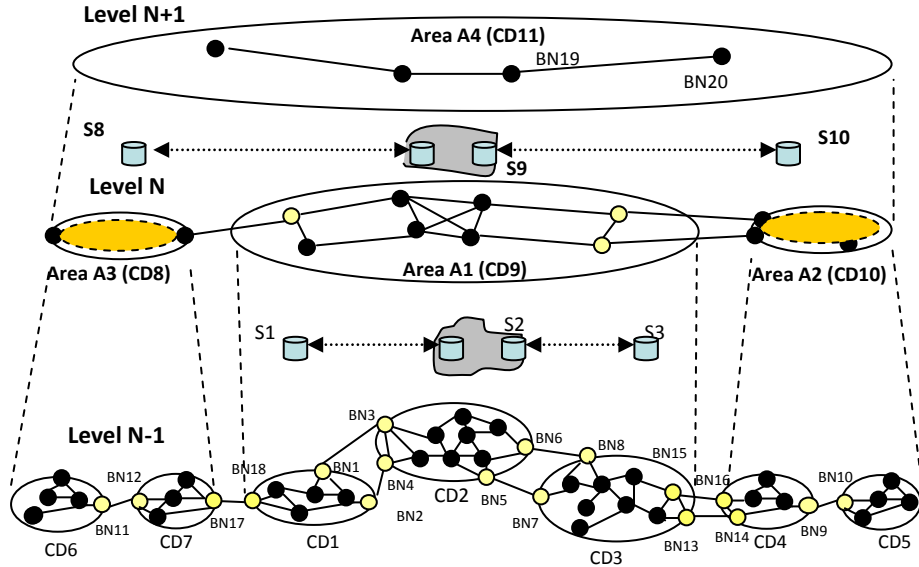
Figure 15: An Example of a Multi-level Hierarchy.

For example, in Figure 15, the network is arranged into two levels of routing hierarchy. In level N of the hierarchy there are three distinct routing areas: A1, A2, and A3. No routing messages are exchanged by routing controllers within these areas and routing controllers within these areas only can find routes across their respective areas. All three of these routing areas are hierarchically contained in a fourth routing area which operates up a level in the hierarchy. In area A4, the RCs advertise routing information for routing control domains encompassing the areas A1, A2 and A3 (S8, S9, and S10 respectively).

18.1.1   Routing Controllers and Routing Areas

For a Routing Area RA^N that is at hierarchical level N, there is at least one Routing Controller RC^{N+1} at hierarchical level $N+1$, up to the highest level of the hierarchy.

With the hierarchical routing model as proposed by this document, the operation at each level of hierarchy associated with a single Routing Area is independent. However, the routing information obtained as a result of executing link state routing at a given hierarchy level can feed up (except at the highest hierarchical level) and feed down (except at the lowest hierarchical level), or alternatively be configured on the Routing Controller for that level.

The reason for information feed-up is so that the routing information associated with one Routing Area can be advertised to others and can be used for routing decisions for the setup of connections that cross optical control domains. The feed-up of routing information is performed level-by-level on a given node. For scaling purposes, it is desirable that feed-up be accomplished together with aggregation and summarization.

The routing information fed up from level N is advertised by the Routing Controller RC^{N+1} at the Level N+1 with the advertiser identified as RC^{N+1}. Therefore Routing Controllers at higher levels of the hierarchy do not need to learn about the identifiers (Routing Controller ID, Routing Area ID, etc.) at lower levels. Another reason for the information feed-up is to reduce the configuration burden, i.e., some components especially in the data plane can be automatically aggregated by RCs at lower levels.

The reason for information feed-down is so that the routing information associated with other Routing Areas is available in the "local" Routing Area and the routing for connections across or to the remote Routing Areas can be calculated by nodes in the "local" Routing Area in a distributed fashion. The information provided from the LSAs originated by RCs in Routing Areas at higher-levels, extends in both control plane and data plane to Routing Areas beyond the local one, with aggregation and summarization, and the information can be directly used in the call/connection routing procedure.

Note the feed-down of routing information is optional; see the following sections for details.

18.1.2   Routing Controllers in Adjacent RCDs (per RC)

The provisioning of information concerning RCs in adjacent domains is exactly the same as described in the Section 17.1.2.

For example in Figure 3, suppose S9 is a federation of RCs representing Area A1, and S10 and S8 are RCs for areas A3 and A2, respectively. The Router ID and SCN address of S8 and S10 are configured on S9, and the Router IDs and SCN addresses of S9 are configured on S8 and S10, as well.

18.1.3   Inter-Domain Links (per RC)

The configuration for inter-domain links may be different from that described in Section 17.1.3 due to additional aggregation of the border nodes and inter-domain links at higher levels of the routing hierarchy.

Note that a link is identified within the scope of a border node, not the scope of the advertising RC.

For example when interconnecting the control domains CD9 and CD10 where there exist two physical links BN13-BN14 and BN15-BN16, the border nodes and links are aggregated at the higher level so that the following configuration is applied to S9 and S10:

Table 10 Inter-Domain Links Configured on S9 in Figure-3

Inter-domain	Local border	Remote border

links	node	node
BN19-BN20	BN19	BN20

Table 11 Inter-Domain Links Configured on S10 in Figure-3

Inter-domain links	Local border node	Remote border node
BN20-BN19	BN20	BN19

18.2 Operation

18.2.1 Adjacency in the Control Plane

Given a routing area RA^N at hierarchical level N, there is a correspondent routing controller RC^{N+1} in RA^{N+1} at hierarchical level $N+1$. In the Figure 3, the domains CD8, CD9 and CD10 are represented by RCs S8, S9 and S10, respectively in area A4 (corresponding to CD11), and form routing adjacencies for exchange of OSPF routing information for area A4.

18.2.2 Topology Aggregation and Feed-Up for Advertising

Topology of the control domain from Level 1 up can be aggregated by RCs and advertised at the next level of hierarchy automatically or through configuration. Note: alternatives to topology aggregation may be defined in future versions of this document.

18.2.2.1 Inter-Domain Links

As described in Section 17.1.2, information on the inter-domain links can be configured on the routing controllers at the level of the routing hierarchy containing both link endpoints.

18.2.2.2 Intra-Domain Links

Intra-domain links can be configured as described in Section 17.1.4 in a hierarchical routing network with multi-level hierarchies, but they can also be discovered and originated automatically. When at least two border nodes are advertised externally for a routing area, the intra-domain topology can be aggregated by computing virtual intra-domain links. The intra-domain links, once aggregated, can be advertised by the RCs that belong to the control domain in the next higher level of hierarchy.

18.2.3 TNA Name Summarization and Feed-Up for Advertising

In a routing area that is at the hierarchical level N ($N \geq 1$), each node in that area can advertise one or more TNA names throughout that area. The RC in that area can

summarize on all these reachable TNA names before advertising TNA reachability at the next higher level of hierarchy.

Note that TNA names are associated to nodes, not to advertising RCs.

18.2.4 Routing Information Feed Down from Level *N* to *N*-1

Routing information that is recorded at the nodes at hierarchical level *N* ($N \geq 2$) can be fed down to the nodes at level *N*-1 with a standardized mechanism such as the one described below. Note: alternatives to feed down that reduce the information storage requirements for lower level RCs may be defined in future versions of this document.

The feed-down of the routing information can be performed by the RC at Level *N* by passing the routing information down to an associated RC or RCs at Level *N*-1, which then in turn advertise the routing information throughout the routing area where the RC belongs.

The routing information that can be fed down includes the following:

1) LSA that contains inter-domain links.

2) LSA that contains intra-domain links.

3) LSA that contains reachable TNA names.

The LSAs at Level *N* that have been fed down may be advertised by the RC at Level *N*-1 as is. The same information can also be further fed down to Level *N*-2, etc., in the same manner.

The purpose of the routing information feed-down is to distribute the traffic engineering information across the control domains to all nodes at the lower hierarchy levels, so that the path selection for end-to-end connections can be accomplished in distributed manner.

19 Appendix III – Use of SNPP Aliases for Hierarchy

19.1 Introduction

The OIF E-NNI Routing project has had a requirement to support multiple hierarchically organized areas for quite some time. While the OIF E-NNI Routing Interoperability demonstrations held in 2003, 2004, 2005, 2007 and 2009 did not test this feature, implementations of hierarchical routing were developed and tested. This appendix describes one method developed and tested in all five interoperability events that requires translation of SNPPs at one hierarchical level into SNPPs at another hierarchical

level. It should be noted that the abstraction model used in this example is only one of many possible abstraction models that can be useful in E-NNI routing.

19.2   Area Hierarchy and Abstract Topologies

[G.7715] states, "routing areas may be hierarchically contained, with a separate routing performer associated with each routing area in the hierarchy." Since the routing performer for this area only has visibility to the topology of its area, it has no specific knowledge of the topology of areas that contain it, or any of the areas it contains. However, the routing performer will still show the contained area along with the SNPP links that connect the contained area to other sub-networks and areas.

At the same time, the lower level RA has visibility to the ends of the links that are used to connect the abstract node to other nodes/areas in the upper level area. Visibility to these ends is necessary so that route computations can be performed across the lower level routing area.

Figure 16 below (also Figure 7 of [G.7715]) illustrates such a topology. The Areas are represented by the shaded circles, link ends are represented by solid dots, and links are represented by arcs. Note the correspondence between the links shown in the upper area topology and the link ends in the lower level topology.
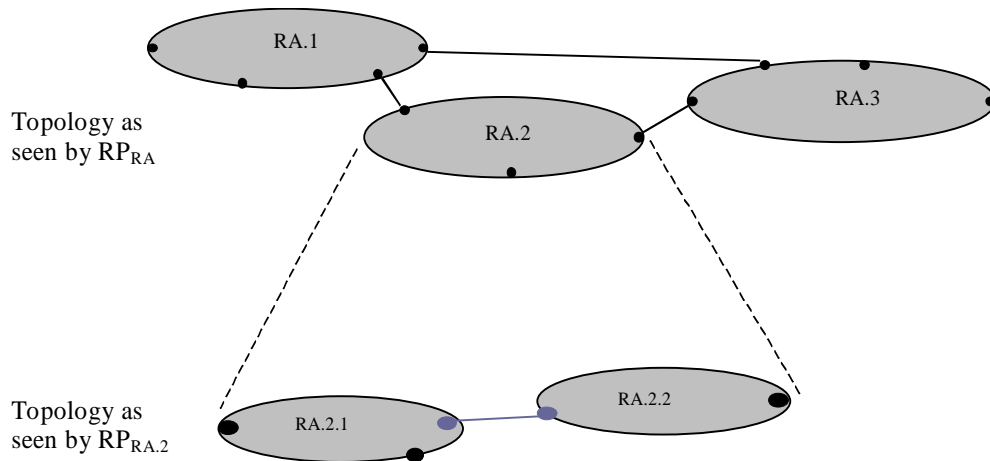

Figure 16: Topology Views as Seen by RP Associated with Hierarchical Routing Areas (Figure 7 in [G.7715])

Since the contained routing area is represented as a single node in the containing area, it is actually an abstraction of the contained area's topology. Therefore, this is called an "Abstract Node".

19.2.1   SNPP links Terminating on Abstract Nodes

The definition of an area in [G.8080] requires that links be wholly contained within an area. Consequently, a link does not exist in any area other than the lowest area that

contains both endpoints of a link. The example illustrated in Figure A3-2 shows links that are contained within area RA11, as well a link that is contained within area RA1. As shown, the Routing Controller for RA1 located on SN3 has visibility to the link in RA1, while the Routing Controller in RA11 located on SN3 has visibility limited to the link-end.
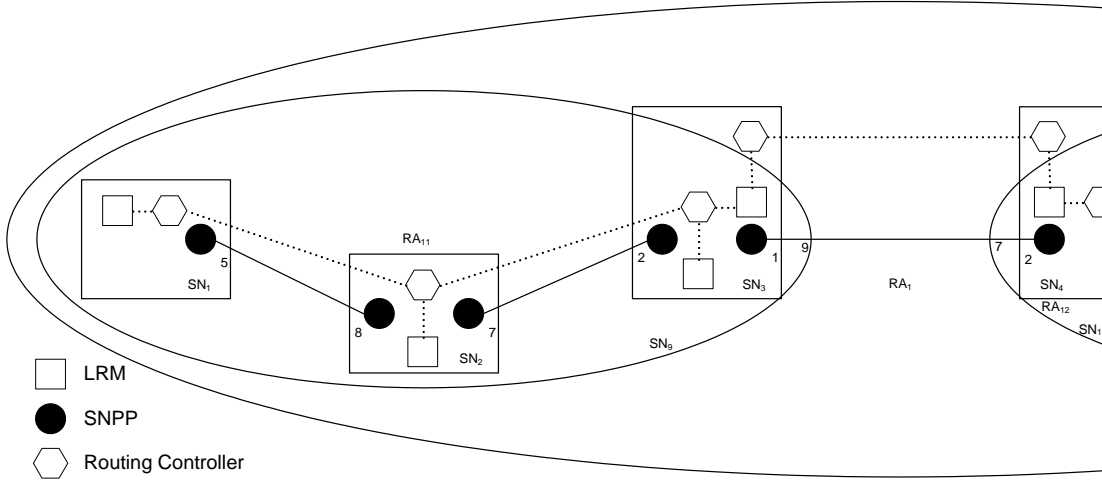

Figure 17: Hierarchical Routing Controller Relationships

According to [G.8080] two separate SNPP names exist for the link end in SN_3 that is connected to SN_4:

RA=<RA_1, RA_{11}> SN=SN_3 LC=1 (in the RA_{11} context)

and

RA=< RA_1> SN= SN_9 LC=9 (in the RA_1 context)

How this interacts with the process of Hierarchical Routing is described below.

19.3 Hierarchical Routing Example

 As an example of how to apply this representation, Hierarchical Routing can be accomplished by performing path calculations in successively higher areas. As stated in [G.7715]:

"1)   The child RC shall first be consulted to develop a path to the destination. If the child RC knows the destination, the path developed by the child RC shall be used. This path shall have the highest preference."

"2)   When the child RC does not know the destination, the parent RC shall be requested to develop a path to the destination. If the parent RC is able to develop a path, the first link end of the path returned will identify the SNPP used to exit the child routing area. The child RC will next be consulted for a route to the SNPP. The path that

is returned by the child RC is then pre-pended to the path that is returned from the parent RC. This path shall have the lowest preference."

So to compute a path from SN_1 in RA_{11} to SN_4 in RA_{12}, the child RC in RA_{11} will first evaluate the destination to see if it is contained within RA_{11}. Since it is not, the child RC will ask an RC in the parent RA (RA_1) to develop a route to SN_4 in RA_{12}. Again, the RC in parent area RA_1 will evaluate the destination to see if it is contained within RA_1. Since the prefix for SN_4 and/or its TNAs are advertised within RA_1 by the RC for SN_{10}, the RC can compute a path from RA_{11} to RA_{12}. The resulting path through the parent RA (RA_1) specifies the near link end for the link which connects SN_9 to SN_{10}, specifically RA=< RA_1>, SN= SN_9, LC=9. This can then translated into the child RA's SNPP name for the visible link end, specifically RA=<RA_1, RA_{11}>, SN=SN_3, LC=1. The translated name can then be used by the RC in the child area to compute a path across the child RA. These paths are then concatenated, providing the end-to-end path.

This interaction between child and parent RC recurses, allowing any number of hierarchical areas to exist between the lowest level child area and the root of the hierarchy.

19.4   Information Necessary for This Example

To perform hierarchical routing as described, a method is necessary to translate the SNPP name used in the parent RA to the SNPP name in the child RA. To accomplish this, a routing announcement is generated by SN_3 in the child RA containing the following information:

Field	# included
Child SNPP name	1
Parent SNPP name	1

This announcement is made into the child RA instead of the parent RA to maintain the requirement for hiding the specifics of the child RA.

Communications between the Child RC and Parent RC can be local to a system or can occur across a Remote Path Computation query interface.

19.5   Scalability

This approach scales linearly with the number of links in the Parent RA that terminate on this RA.

19.6 Versatility

Since [G.8080] defines the use of SNPP aliases for not just hierarchical routing, but also for L1VPN style functionality, the translation information defined above can also be used to facilitate L1VPN services.

20 Appendix IV – Summary of Multilayer Extensions

The following additions summarize the multilayer changes:

- Add the Multilayer TNA to the Link TLV in section 7.3 Table-2.

- Add the Transitional Link to the Link Attribute Scoping and Connection type sub-TLV in section 12.3.

- Define multilayer extensions for transitional links and server-layer TNAs in section 14.

- Discuss compatibility with OIF E-NNI Routing 2.0 in section 15.

21 Appendix V – Client:Server Capacity Ratio

The following table shows example values for client:server capacity ratio for most combinations of client layer, server layer and adaptation.

Client	Server	Adaptation	Units for Client:Server Ratio	Client:Server Ratio Value	IEEE format	Source
Ethernet	ODU1	GFP-F	Bps:u	276,480,000:1 Bps:u	0x4D83D600	Converted from 2 488 320.000 kbit/s [G.709 Table7-3] with GFP-F factor (1:1.125)
Ethernet	ODU2	GFP-F	Bps:u	1,110,586,329.11:1	0x4E846468	Converted from 9 995 276.962 kbit/s [G.709 Table7-3] with GFP-F factor (1:1.125)
Ethernet	ODU3	GFP-F	Bps:u	4,461,168,813.55:1	0x4F84F404	Converted from 40 150 519.322 kbit/s [G.709 Table 7-3] with GFP-F factor (1:1.125)
Ethernet	VC4 (add CCAT note)	GFP-F	Bps:u	16,640,000:1	0x4B7DE800	Converted from 149,760 kbit/s [G.707/Table6-1] with GFP-F factor (1:1.125)
Ethernet	VC3	GFP-F	Bps:u	5,376,000:1	0x4AA41000	Converted from 48,384 kbit/s [G.707/Table6-1] with

						GFP-F factor (1:1.125)
Ethernet	VC4-nV (STS-3c-nV)	GFP-F	Bps:Bps	1:1.125	0x3F638E39	worst case scenario: 8 bytes overhead per GFP frame for 64 bytes frames
Ethernet	VC4-nV (STS-3c-nV)	GFP-T	Bps:Bps	1: 1.334393656	0x3F3FD8F1	Worst case scenario based on EthMAC/Phy expansion (1.3125 based on min packet size of 64 bytes and preamble and inter-packet gap of 20 bytes [IEEE802.3]) * EthPHY/GFP-T expansion of((8 bytes overhead + 7695 bytes per GFP-T frame)/7695 bytes per GFP-T frame) [G.7041] Note that this is limited to 1GigE and n=7.
Ethernet	VC3-nV (STS-1-nV)	GFP-F	Bps:Bps	1:1.125	0x3F638E39	worst case scenario: 8 bytes overhead per GFP frame for 64 bytes frames
Ethernet	VC3-nV (STS-1-nV)	GFP-T	Bps:Bps	1: 1.334393656	0x3F3FD8F1	Worst case scenario based on EthMAC/Phy expansion (1.3125 based on min packet size of 64 bytes and preamble and inter-packet gap of 20 bytes [IEEE802.3]) * EthPHY/GFP-T expansion of((8 bytes overhead + 7695 bytes per GFP-T frame)/7695 bytes per GFP-T frame) [G.7041]. Note that this is limited to 1GigE and n=21.
VC4-nV/STS-3c	VC4/STS-3c	VCAT (LCAS)	Bps:u	18,720,000:1	0x4B8ED280	Converted from 149,760 kbit/s [G.707/Table6-1]
VC4-nV/STS-3c	VC4/STS-3c	VCAT (no LCAS)	Bps:u	18,720,000:1	0x4B8ED280	Converted from 149,760 kbit/s [G.707/Table6-1]
VC3-nV/STS-	VC3/STS-1	VCAT (LCAS)	Bps:u	6,048,000:1	0x4AB89200	Converted from 48,384 kbit/s

1						[G.707/Table6-1]
VC3-nV/STS-1	VC3/STS-1	VCAT (no LCAS)	Bps:u	6,048,000:1	0x4AB89200	Converted from 48,384 kbit/s [G.707/Table6-1]
Ethernet	VC4-nv/VC4	GFP-F/VCAT(LCAS)	Bps:u	16,640,000:1	0x4B7DE800	Ethernet over VC4-nV ratio * VC4-nV/VC4 ratio
Ethernet	VC4-nv/VC4	GFP-F/VCAT(no LCAS)	Bps:u	16,640,000:1	0x4B7DE800	Ethernet over VC4-nV ratio * VC4-nV/VC4 ratio
Ethernet	VC4-nv/VC4	GFP-T/VCAT(LCAS)	Bps:u	14,028,843.67:1	0x4B56102C	Ethernet over VC4-nv ratio *VC4-nv/VC4 ratio. Note that this is limited to 1GigE and n=7.
Ethernet	VC4-nv/VC4	GFP-T/VCAT(no LCAS)	Bps:u	14,028,843.67:1	0x4B56102C	Ethernet over VC4-nv ratio *VC4-nv/VC4 ratio. Note that this is limited to 1GigE and n=7.
Ethernet	VC3-nv/VC3	GFP-F/VCAT(LCAS)	Bps:u	5,376,000:1	0x4AA41000	Ethernet over VC3-nV ratio * VC3-nv/VC3 ratio
Ethernet	VC3-nv/VC3	GFP-F/VCAT(no LCAS)	Bps:u	5,376,000:1	0x4AA41000	Ethernet over VC3-nV ratio * VC3-nv/VC3 ratio
Ethernet	VC3-nv/VC3	GFP-T/VCAT(LCAS)	Bps:u	4,532,395.65:1	0x4A8A5157	Ethernet over VC3-nV ratio * VC3-nv/VC3 ratio. Note that this is limited to 1 GigE and n=21.
Ethernet	VC3-nv/VC3	GFP-T/VCAT(no LCAS)	Bps:u	4,532,395.65:1	0x4A8A5157	Ethernet over VC3-nV ratio * VC3-nv/VC3 ratio. Note that this is limited to 1 GigE and n=21.
VC3	ODU1	AMP (RS16)	u:u	48:1	0x42400000	
VC3	ODU2	AMP (RS64)	u:u	192:1	0x43400000	
VC3	ODU3	AMP (RS256)	u:u	768:1	0x44400000	
VC4	ODU1	AMP (RS16)	u:u	16:1	0x41800000	
VC4	ODU2	AMP (RS64)	u:u	64:1	0x42800000	
VC4	ODU3	AMP (RS256)	u:u	256:1	0x43800000	

VC3	ODU1	BMP (RS16)	u:u	48:1	0x42400000	
VC3	ODU2	BMP (RS64)	u:u	192:1	0x43400000	
VC3	ODU3	BMP (RS256)	u:u	768:1	0x44400000	
VC4	ODU1	BMP (RS16)	u:u	16:1	0x41800000	
VC4	ODU2	BMP (RS64)	u:u	64:1	0x42800000	
VC4	ODU3	BMP (RS256)	u:u	256:1	0x43800000	

22 Appendix VI: List of companies belonging to OIF when document is approved

Acacia Communications
ADVA Optical Networking
Agilent Technologies R & D
Alcatel-Lucent
Altera
AMSS
Amphenol
Anritsu
Applied Communication
Sciences
AT&T
Avago
Broadcom
Brocade
Centellax
China Telecom
Ciena
Cisco
ClariPhy
Coriant
Cortina Systems
CPqD
Department of Defense
Deutsche Telekom
Emcore
Ericsson
FCI USA LLC
Fiberhome Technologies Group
Finisar
Fujikura
Fujitsu
Furukawa Electric Japan
Google
Hewlett Packard
Hitachi
Hittite Microwave
Huawei Technologies
IBM
Infinera
Inphi
Intel

JDSU
Juniper Networks
Kaiam
Kandou
KDDI R & R Laboratories
LeCroy
LSI
Luxtera
M/A-COM Technology Solutions
Marben Products
Mellanox
Metaswitch
Mindspeed
Mitsubishi Electric
Molex
MoSys
MultiPhy
NEC
NeoPhotonics
NTT
Oclaro
Optelian
Orange
PETRA
PMC Sierra
QLogic
Ranovus
Semtech
Skorpios
Sumitomo Electric
Sumitomo Osaka Cement
TE Connectivity
Tektronix
Tellabs
TELUS Communications
TeraXion
Texas Instruments
Time Warner Cable
TriQuint Semiconductor
u2t Photonics AG
US Conec
Verizon
Xilinx
Xtera Communications
Yamaichi Electronics