CCAMP Working Group
Internet Draft
Expires: October 2005

Kohei Shiomoto (NTT)
Rajiv Papneja (ISOCORE)
Richard Rabbat (Fujitsu)

April 2005

Addressing and Messaging in Generalized Multi-Protocol Label Switching (GMPLS) Networks

draft-shiomoto-ccamp-gmpls-addressing-01.txt

Status of this Memo

Abstract

This document explains and clarifies addressing and messaging in Generalized Multi-Protocol Label Switching (GMPLS) networks. The aim of this document is to facilitate and ensure better interworking of GMPLS-capable Label Switching Routers (LSR) based on experience gained in deployments and interoperability testing and proper interpretation of published RFCs.

The document recommends a proper approach for the interpretation and choice of address and identifier fields within GMPLS protocols and references specific control plane usage models. It also examines

some common GMPLS Resource Reservation Protocol-Traffic Engineering (RSVP-TE) signaling message processing issues and recommends solutions.

Table of Contents

Changes from version -00 to version -01:

- Used conventions of section 2 throughout draft

- Removed dedicated sections for IS-IS: discussed in unified section on IGP

- Removed dedicated sections for IPv6: text now addresses v4 and v6

- Cleaned up all sections

- Separated references into informational and normative sections

1. Introduction

This document describes explains and clarifies addressing and messaging in networks that use GMPLS [RFC3945] as their control plane. For the purposes of this document it is assumed that there is a one-to-one correspondence between control plane and data plane entities. That is, each data plane switch has a unique control plane presence responsible for participating in the GMPLS protocols, and that each such control plane presence is responsible for a single data plane switch. The combination of control plane and data plane entities is referred to as a Label Switching Router (LSR). Various more complex deployment scenarios can be constructed, but these are out of scope of this document.

2. Conventions Used in This Document

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

3. Terminology

Note that the term 'Router ID' is used in two contexts within GMPLS. It may refer to an identifier for a participant in a routing protocol, or it may be an identifier for an LSR that participates in TE routing. These could be considered as the control plane and data plane contexts.

In this document, the contexts are distinguished by the following definitions.

Loopback address - A loopback address is a stable IP address of the advertising router that is always reachable if there is any IP connectivity to it [RFC3630]. Thus, for example, an IPv4 127/24 address is excluded from this definition.

TE Router ID - A stable IP address of an LSR that is always reachable if there is any IP connectivity to the LSR, e.g., a loopback address. The most important requirement is that the address does not become unusable if an interface on the LSR is down [RFC3477].

Router ID - The OSPF protocol version 2 [RFC2328] defines the Router ID to be a 32-bit network unique number assigned to each router running OSPF. IS-IS [RFC1195] includes a similar concept in the System ID. This document describes both concepts as the "Router ID" of the router running the routing protocol. The Router ID is not required to be a reachable IP address, although an operator MAY set it to a reachable IP address on the same system.

TE link û "A TE link is a representation in the IS-IS/OSPF Link State advertisements and in the link state database of certain physical resources, and their properties, between two GMPLS nodes." [RFC3945]

Data plane node û A vertex on the TE graph. It is a data plane switch or router. Data plane nodes are connected by TE links which are constructed from physical data links. A data plane node is controlled through some combination of management and control plane actions. A data plane node may be under full or partial control of a control plane node.

Control plane node - A GMPLS protocol speaker. It may be part of a data plane switch or may be a separate computer. Control plane nodes are connected by control channels which are logical connectionless or connection-oriented paths in the control plane. A control plane node is responsible for controlling zero, one or more data plane nodes.

Interface ID û The Interface ID is defined in [RFC3477] and in section 9.1 of [RFC3471].

4. Numbered Addressing

When numbered addressing is used, addresses are assigned to each node and link in both control and data planes in GMPLS networks. A TE Router ID is defined to identify the LSR for TE purposes. It is a requirement stated in [RFC3477] that the TE Router ID MUST be a reachable address in the control plane.

The reason why the TE Router ID must be a reachable IP address is because in GMPLS, control and data plane names /addresses are not completely separated. An Explicit Route Object (ERO) signaled as a part of a Label Switched Path (LSP) setup message contains an LSP path specified in data plane addresses, namely TE Router IDs and TE link IDs. The message needs to be forwarded as IP/RSVP packet between LSRs that manage data plane nodes along the path. Hence, each LSR along the path needs to resolve the next hop data plane address into the next hop control plane address before the message could be forwarded to the next hop. Generally speaking there is a need for a module/protocol that discovers and manages control plane/ data plane address bindings for the address spaces to be completely separated. In this case, the TE Router ID could be just a network unique number. Mandating that TE Router ID be a reachable IP address eliminates the need of the mentioned above module û the next hop data plane TE Router ID could be used as a destination for IP packets encapsulating the LSP setup (RSVP Path) message. Note that every TE link ID could always be resolved to the link originating TE Router ID.

An IP address MAY also be assigned to each physical interface connected to the control plane network. Both numbered and unnumbered links in the control plane MAY be supported. The control channels are advertised by the routing protocol as normal links, which allows the routing of RSVP-TE and other control messages between the LSRs over the control plane network.

A physical interface address or a physical interface identifier is assigned to each physical interface connected to the data plane. An interface address or an interface identifier is logically assigned to each TE-link end associated with the physical data channel in the GMPLS domain. A TE link may be installed as a logical interface.

A numbered link is identified by a network unique identifier (e.g., an IP address) and an unnumbered link is identified by the combination of TE Router ID and a node-unique Interface ID. The existence of both numbered and unnumbered links in the data plane SHOULD be accepted. The recommended addressing for the numbered and unnumbered links is also suggested in this document.

4.1. Interior Gateway Protocols

We address in this section unnumbered addressing using two Interior Gateway Protocols (IGPs) that have extensions defined for GMPLS: OSPF-TE and IS-IS/TE [RFC3784].

4.1.1. Router Address

The Router Address is advertised in OSPF-TE using the Router Address TLV structure [RFC3630].

It is referred to as the Addressing Router that is advertised in IS-IS [RFC1195].

The IGP protocols use this as a means to advertise the TE Router ID. The TE Router ID is used in constrained-based path computation.

4.1.2. Link ID sub-TLV

The Link ID sub-TLV [RFC3630] advertises the Router ID of the remote end of the TE link. For point-to-point links, this is the Router ID of the neighbor. Multi-access links are left for further study.

Note that there is no correspondence in IS-IS to the Link ID sub-TLV in OSPF-TE.

4.1.3. Local Interface IP Address

The Local Interface IP Address is advertised in:
- the Local Interface IP Address sub-TLV in OSPF-TE
- the IPv4 Interface Address sub-TLV in IS-IS/TE

This is the ID of the local end of the numbered TE link. It MUST be a network unique number.

4.1.4. Remote Interface IP Address

The Remote Interface IP Address is advertised in:
- the Remote Interface IP Address sub-TLV in OSPF-TE
- the IPv4 Neighbor Address sub-TLV in IS-IS/TE

This is the ID of the remote end of the numbered TE link. It MUST be a network unique number.

4.2. Use of Addresses in RSVP-TE

4.2.1. IP Tunnel End Point Address in Session Object

The IP tunnel end point address of the Session Object is either an IPv4 or IPv6 address.

It is RECOMMENDED that the IP tunnel endpoint address in the Session Object [RFC3209] be set to the TE Router ID of the egress since the TE Router ID is a unique routable ID per node.

Alternatively, the tunnel end point address MAY also be set to the destination data plane address if the ingress knows that address or the TE Router ID.

4.2.2. IP Tunnel Sender Address in Sender Template Object

The IP tunnel sender address of the Sender Template Object is also either an IPv4 or IPv6 address.

It is RECOMMENDED that the IP tunnel sender address in the Sender Template Object [RFC3209] specifies the TE Router ID of the ingress since the TE Router ID is a unique routable ID per node.

Alternatively, the tunnel sender address MAY also be set to the sender data plane address or the TE Router ID.

4.2.3. IF_ID RSVP_HOP Object for Numbered Interfaces

1. IPv4/IPv6 Next/Previous Hop Address: the IPv4/IPv6 Next/Previous Hop Address [RFC3473] in the Path and Resv messages specifies the IP reachable address of the control plane interface used to send those messages, or the TE Router ID of the node that is sending those messages.

2. IPv4/IPv6 address in Value Field of the Interface_ID TLV: In both the Path and Resv messages, the IPv4/IPv6 address in the value field of TLV [RFC3471] specifies the associated data plane local interface address of the downstream data channel belonging to the node sending the Path message and receiving the Resv message. In the case of bi-directional LSPs, if the interface upstream is different from that downstream, then another TLV including the local interface address of the upstream data channel belonging to the node sending the Path message and receiving the Resv message MAY be also added to the TLV for downstream. Note that identifiers of both downstream and upstream data channels are specified in each TLV from the viewpoint of the sender of the Path message, in both the sent Path and received Resv messages.

4.2.4. Explicit Route Object (ERO)

The IPv4/IPv6 address in the ERO provides a data-plane identifier of an abstract node, TE node or TE link to be part of the signaled LSP.

We describe in section 6 the choice of addresses in the ERO.

4.2.5. Record Route Object (RRO)

The IPv4/IPv6 address in the RRO provides a data-plane identifier of either a TE node or TE link that is part of the established LSP.

We describe in section 6 the choice of addresses in the RRO.

4.3. IP Packet Destination Address

The IP destination address of the packet that carries the RSVP-TE message is a control-plane reachable address of the next hop or previous hop node along the LSP. It is RECOMMENDED that a stable control plane IP address of the next/previous hop node be used as an IP destination address in RSVP-TE message.

A Path message is sent to the next hop node. It is RECOMMENDED that the TE router ID of the next hop node be used as an IP destination address in the packet that carries the RSVP-TE message.

A Resv message is sent to the previous hop node. It is RECOMMENDED that the IP destination of a Resv message be any of the following:
- The IP source address of the received IP packet containing the Path message,
- A stable IP address of the previous node (found out by consulting the TED and referencing the upstream data plane interface,
- The value in the received phop field.

4.4. IP Packet Source Address

The IP source address of the packet that carries the RSVP-TE message MUST be a reachable address of the node sending the RSVP-TE message. It is RECOMMENDED that a stable IP address of the node be used as an IP source address of the IP packet. In case the IP source address of the received IP packet containing the Path message is used as the IP destination address of the Resv message, setting a stable IP address in the Path message is beneficial for reliable control-plane transmission. This allows for robustness when one of control-plane interfaces is down.

5. Unnumbered Addressing

In this section, we describe unnumbered addressing used in GMPLS to refer to different objects and their significance. Unnumbered addressing is supported for a data plane.

5.1. IGP

Since GMPLS caters to PSC and non-PSC links, a few enhancements have been made to the existing OSPF-TE and ISIS-TE protocols. The routing enhancements for GMPLS are defined in [GMPLS-RTG], [RFC3784] and [GMPLS-OSPF].

5.1.1. Link Local/Remote Identifiers in OSPF-TE

Link Local/Remote Identifiers advertise IDs of an unnumbered TE link's local and remote ends respectively. Those are numbers unique within the scopes of the advertising LSR and the LSR managing the remote end of the link respectively. Note that these numbers are not network unique and therefore could not be used as TE link end addresses on their own. An unnumbered TE link end address is comprised of a Router ID associated with the link local end, followed by the link local identifier [GMPLS-OSPF].

5.1.2. Link Local/Remote Identifiers in IS-IS/TE

Link local/remote identifiers in IS-IS/TE are exchanged in the Extended Local Circuit ID field of the "Point-to-Point Three-Way Adjacency" [RFC3373] IS-IS Option type. The same discussion in 5.1.1 applies here.

5.2. Use of Addresses in RSVP-TE

An unnumbered address used for data plane identification consists of the TE router ID and the associated interface ID.

5.2.1. IF_ID RSVP_HOP Object for Unnumbered Interfaces

The interface ID field in the IF_INDEX TLV specifies the interface of the data channel for that unnumbered interface.

In both the Path message and the Resv message, the value of the interface ID in the IF_INDEX TLV specifies the associated local interface ID of the downstream data channel belonging to the node sending the Path message and receiving the Resv message. In case of bi-directional LSPs, if the interface upstream is different from that downstream, then another IF_INDEX TLV including the local interface ID of the upstream data channel belonging to the node sending the Path message and receiving the Resv message MAY be also added to the IF_INDEX TLV for downstream. Note that identifiers of both downstream and upstream data channels are specified in each IF_INDEX TLV from the viewpoint of the sender of the Path message, in both sent Path and received Resv messages.

5.2.2. Explicit Route Object (ERO)

For unnumbered interfaces in the ERO, the interface ID is either the incoming or outgoing interface of the TE-link w/r to the GMPLS-capable LSR.

We describe in section 6 the choice of addresses in the ERO.

5.3. Record Route Object (RRO)

For unnumbered interfaces in the RRO, the interface ID is either the incoming or outgoing interface of the TE-link w/r to the GMPLS-capable LSR.

We describe in section 6 the choice of addresses in the RRO.

5.4. LSP_Tunnel Interface ID Object

The LSP_TUNNEL_INTERFACE_ID Object includes the LSR's Router ID and Interface ID as described in [RFC3477] to specify an unnumbered Forward Adjacency Interface ID. The Router ID of the GMPLS-capable LSR MUST be set to the TE router ID.

6. RSVP-TE Message Content

6.1. ERO and RRO Addresses

6.1.1. Strict Subobject in ERO

Implementations making limited assumptions about the content of an ERO when processing a received Path message may cause interoperability issues.

A subobject in the Explicit Route Object (ERO) includes an address specifying an abstract node (i.e., a group of nodes), a simple abstract node (i.e., a specific node), or a specific interface of a TE-link in the data plane, depending on the level of control required [RFC3209].

In case one subobject is strict, any of the following options are valid:

1. Address or AS number specifying a group of nodes
2. TE Router ID
3. Incoming TE link ID
4. Outgoing TE link ID optionally followed by one or two Label sub-objects

5. Incoming TE link ID and Outgoing TE link ID optionally followed by one or two Label sub-objects
6. TE Router ID and Outgoing TE link ID optionally followed by one or two Label sub-objects
7. Incoming TE link ID, TE Router ID and Outgoing TE link ID optionally followed by one or two Label sub-objects

The label value that identifies a single unidirectional resource between two nodes may be different from the perspective of upstream and downstream nodes. This is typical in the case of fiber switching, because the label value is a number indicating the port/fiber. This is also possible in case of lambda switching, because the label value is a number indicating the lambda, but may not be the value directly indicating the wavelength value (e.g., 1550 nm).

The value of a label in RSVP-TE object MUST indicate the value from the perspective of the sender of the object/TLV [RFC3471]. This rule MUST be applied to all types of object/TLV.

Therefore, the label field in the Label ERO subobject MUST include the value of the label for the upstream nodeÆs identification of the resource.

6.1.2. Loose Subobject in ERO

There are two differences between Loose and Strict subobject.

A subobject marked as a loose hop in an ERO MUST NOT be followed by a subobject indicating a label value [RFC3473].

A subobject marked as a loose hop in an ERO SHOULD never include an identifier (i.e., address or ID) of outgoing interface.

There is no way to specify in the ERO whether a subobject is associated with the incoming or outgoing TE link. This is unfortunate because the address specified in a loose subobject is used as a target for the path computation, and there is a big difference for the path selection process whether the intention is to reach the target node over the specified link (the case of incoming TE link) or to reach the node over some other link, so that it would be possible to continue the path to its final destination over the specified link (the case of outgoing TE link).

In the case where a subobject in an ERO is marked as a loose hop and identifies an interface, the subobject SHOULD include the address of the Incoming interface specifying the TE-link in the data plane.

6.1.3. RRO

When a node adds one or more subobjects to an RRO and sends the Path or the Resv message including the RRO for the purpose of recording the node's addresses used for an LSP, the subobjects (i.e. number, each type, and each content) added by the node SHOULD be the same in the Path ERO and Resv ERO. The intention is that they report the path of the LSP, and that the operator can use the results consistently. At any transit node, it SHOULD be possible to construct the path of the LSP by joining together the RRO from the Path and the Resv messages.

It is also important that a whole RRO on a received Resv message can be used as input to an ERO on a Path message.

Therefore, in case that a node adds one or more subobjects to an RRO, any of the following options are valid:

1. TE Router ID
2. Incoming TE link ID
3. Outgoing TE link ID optionally followed by one or two Label sub-objects
4. Incoming TE link ID and Outgoing TE link ID optionally followed by one or two Label sub-objects
5. TE Router ID and Outgoing TE link ID optionally followed by one or two Label sub-objects
6. Incoming TE link ID, TE Router ID and Outgoing TE link ID optionally followed by one or two Label sub-objects

Option (4) is RECOMMENDED considering the importance of the record route information to the operator.

6.2. Forwarding Destination of Path Message with ERO

The final destination of the Path message is the Egress node that is specified by the tunnel end point address in the Session object. The Egress node MUST NOT forward the corresponding Path message downstream, even if the ERO includes the outgoing interface ID of the Egress node for the Egress control [RFC4003].

7. GMPLS Control Plane

7.1. Control Channel Separation

In GMPLS, a control channel can be separated from the data channel and there is not necessarily a one-to-one association of a control channel to a data channel. Two adjacent nodes may have multiple IP hops between them in the control plane.

There are two broad types of separated control planes: native and tunneled. These differ primarily in the nature of encapsulation used for signaling messages, which also results in slightly different address handling with respect to the control plane address.

7.2. Native and Tunneled Control Plane

It is RECOMMENDED that all implementations support a native control plane.

If the control plane interface is unnumbered, the RECOMMENDED value for the control plane address is the TE Router-ID.

For the case where two adjacent nodes have multiple IP hops between them in the control plane, then as stated in section 9 of [RFC3945], implementations SHOULD use the mechanisms of section 8.1.1 of [MPLS-HIER] whether they use LSP Hierarchy or not. Note that section 8.1.1 of [MPLS-HIER] applies for "FA-LSP" as stated in that section but also to "TE-LINK" for the case where a normal TE link is used. Note also that a hop MUST NOT decrement the TTL of the received RSVP-TE message.

For a tunneled control plane, the inner RSVP and IP messages traverse exactly one IP hop. The IP TTL of the outermost IP header SHOULD be the same as for any network message sent on that network. Implementations receiving RSVP-TE messages on the tunnel interface MUST NOT compare the RSVP TTL to either of the IP TTLs received. Implementations MAY set the RSVP TTL to be the same as the IP TTL in the innermost IP header.

7.3. Separation of Control and Data Plane Traffic

Data traffic MUST NOT be transmitted through the control plane.

8. Security Considerations

In the interoperability testing we conducted, the major issue we found was the use of control channels for forwarding data. This was due to the setting of both control and data plane addresses to the same value in PSC (Packet-Switching Capable) equipment. This led to major problems that could be avoided simply by keeping the addresses for the control and data plane separate.

9. Full Copyright Statement

10. Intellectual Property

11. Acknowledgement

The authors would like to thank Adrian Farrel for the helpful discussions and the feedback he gave on the document. We'd also like to thank Jonathan Sadler and Hidetsugu Sugiyama for their feedback.

12. Authors' Addresses

Kohei Shiomoto
NTT Network Service Systems Laboratories
3-9-11 Midori
Musashino, Tokyo 180-8585
Japan
Email: shiomoto.kohei@lab.ntt.co.jp

Rajiv Papneja
Isocore Corporation
12359 Sunrise Valley Drive, Suite 100
Reston, Virginia 20191
United States of America
Phone: +1-703-860-9273
Email: rpapneja@isocore.com

Richard Rabbat
Fujitsu Labs of America, Inc.
1240 East Arques Ave, MS 345
Sunnyvale, CA 94085
United States of America
Phone: +1-408-530-4537
Email: richard.rabbat@us.fujitsu.com

13. Contributors

Yumiko Kawashima
NTT Network Service Systems Lab
Email: kawashima.yumiko@lab.ntt.co.jp

Ashok Narayanan
Cisco Systems
Email: ashokn@cisco.com

Eiji Oki
NTT Network Service Systems Laboratories
Midori 3-9-11
Musashino, Tokyo 180-8585, Japan
Email: oki.eiji@lab.ntt.co.jp

Lyndon Ong
Ciena Corporation
Email: lyong@ciena.com

Vijay Pandian
Sycamore Networks

Email: Vijay.Pandian@sycamorenet.com

Hari Rakotoranto
Cisco Systems
Email: hrakotor@cisco.com

14. References

14.1. Normative References

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

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

[RFC3209] Awduche, D., et al, "RSVP-TE: Extensions to RSVP for LSP Tunnels," RFC 3209, December 2001.

[RFC3471] Berger, L., ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description," RFC 3471, January 2003.

[RFC3473] Berger, L., ed. "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions," RFC 3473, January 2003.

[RFC3477] Kompella, K., Rekhter, Y., "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)," RFC 3477, January 2003.

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

[RFC3667] Bradner, S., "IETF Rights in Contributions", BCP 78, IETF RFC 3667, February 2004.

[RFC3668] Bradner, S., "Intellectual Property Rights in IETF Technology", BCP 79, IETF RFC 3668, February 2004.

[RFC3945] Mannie, E., ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," RFC 3945, October 2004.

[RFC4003] Berger, L., "GMPLS Signaling Procedure for Egress Control," RFC 4003, February 2005.

[GMPLS-OSPF] K. Kompella, Y. Rekhter (Eds.), "OSPF Extensions in Support of Generalized Multi-protocol Label Switching," work in progress, draft-ietf-ccamp-gmpls-ospf-gmpls-extensions-12.txt, October 2003.

[GMPLS-RTG] K. Kompella, Y. Rekhter (Eds.), "Routing Extensions in Support of Generalized Multi-protocol Label Switching," draft-ietf-ccamp-gmpls-routing-09.txt, October 2003.

[MPLS-HIER] Kompella, K. and Rekhter, Y., "LSP Hierarchy with Generalized MPLS TE," work in progress, draft-ietf-mpls-lsp-hierarchy-08.txt, March 2002.

14.2. Informative References

[RFC1195] Callon, R., "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments," RFC 1195, December 1990.

[RFC3373] Katz, D. and R. Saluja, "Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies," RFC 3373, September 2002.

[RFC3784] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)," RFC 3784, June 2004.