Network Working Group   Anca Zamfir
Internet Draft   Zafar Ali
Expires: January 6, 2009   Cisco Systems
                                             Dimitri Papadimitriou
                                                     Alcatel-Lucent
                                                       July 7, 2008

Component Link Recording and Resource Control for TE Link Bundles

draft-ietf-mpls-explicit-resource-control-bundle-04.txt

Status of this Memo

Copyright Notice

Abstract

   Record Route is a useful administrative tool that has been used
   extensively by the service providers. However, when TE links are
   bundled, identification of label resource in Record Route Object
   (RRO) is not enough for the administrative purpose. Network service

   providers would like to know the component link within a TE link that
   is being used by a given LSP. In other words, when link bundling is
   used, resource recording requires mechanisms to specify the component
   link identifier, along with the TE link identifier and Label. As it
   is not possible to record component link in the RRO, this draft
   defines the extensions to RSVP-TE [RFC3209] and [RFC3473] to specify
   component link identifiers for resource recording purposes.

   This draft also defines the Explicit Route Object (ERO) counterpart
   of the RRO extension. The ERO extensions are needed to perform
   explicit label/ resource control over bundled TE link. Hence, this
   document defines the extensions to RSVP-TE [RFC3209] and [RFC3473] to
   specify component link identifiers for explicit resource control and
   recording over TE link bundles.

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 RFC 2119 [RFC2119].

Table of Contents

[1](#). Terminology

TE Link: Unless specified otherwise, it refers to a bundled Traffic Engineering link as defined in [RFC4201]. Furthermore, the terms TE Link and bundled TE Link are used interchangeably in this draft.

Component (interface) link: refers (locally) to a component link as part of a bundled TE link. A component link is numbered/ unnumbered in its own right. For unnumbered component links, the component link ID is assumed to be unique on an advertising node. For numbered component links, the component link ID is assumed to be unique within a domain.

Component Interface Identifier: Refers to an ID used to uniquely identify a Component Interface. On a bundled link a combination of <component link identifier, label> is sufficient to unambiguously identify the appropriate resources used by an LSP [RFC4201].

[2](#). Resource Control and Recording

In GMPLS networks [RFC3945] that deals with unbundled (being either PSC, L2SC, TDM or LSC) TE Links, one of the types of resources that an LSP originator can control and would like to record are the TE Link interfaces used by the LSP. The resource control and recording is done by the use of an explicit route, i.e., Explicit Route (ERO) Object and record Route, i.e., Record Route Object (RRO) object, respectively.

Link Bundling, introduced in [RFC4201], is used to improve routing scalability by reducing the amount of TE related information that needs to be flooded and handled by IGP in a TE network. This is accomplished by aggregating and abstracting the TE Link resource. In some cases the complete resource identification is left as a local decision. However, as described above there are cases when it is desirable for a non-local (e.g., LSP head-end) node to identify completely or partially the LSP resources. In either case, and for administrative reasons, it is required to know which component link within a bundled TE link has been used for a given LSP.

When link bundling is used to aggregate multiple component links into a TE link, label is not the only resource that needs to be identified and recorded. In other words, the TE Link and the Label specified in the ERO/ RRO objects are not enough to completely identify the resource. For the bundled TE link case, in order to fully specify the resources on a link for a given LSP, the component link needs to be specified along with the label. In the case of bi-directional LSPs both upstream and downstream information may be specified. Therefore, explicit resource control and recording over a bundled TE link also

requires ability to specify a component link within the TE link.

This draft defines extensions to and describes the use of RSVP-TE
[RFC3209], [RFC3471], [RFC3473] to specify the component link
identifier for resource recording and explicit resource control over
TE link bundles. Specifically, in this document, component interface
identifier RRO and ERO subobjects are defined to complement their
Label RRO and ERO counterparts. Furthermore, procedures for
processing component interface identifier RRO and ERO subobjects and
how they can co-exist with the Label RRO and ERO subobjects are
specified.

3. LSP Resource Recording

LSP Resource Recording refers to the ability to record the resources
used by an LSP.

The procedure for unbundled numbered TE links is described in
[RFC3209] and for unbundled unnumbered TE links in [RFC3477]. For the
purpose of recording LSP resources used over bundled TE Links, the
Component Interface Identifier RRO sub-object is introduced.

3.1 Component Interface Identifier RRO subobject

A new subobject of the Record Route Object (RRO) is used to record
component interface identifier of a (bundled) TE Link. This subobject
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    |U| Reserved  (must be zero)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//   IPv4, IPv6 or unnumbered Component Interface Identifier   //
|                             . . .                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   L: 1 bit

       This bit must be set to 0.

   Type

       Type 10 (TBD): Component Interface identifier IPv4
       Type 11 (TBD): Component Interface identifier IPv6
       Type 12 (TBD): Component Interface identifier Unnumbered

   Length

The Length contains the total length of the subobject in
bytes, including the Type and Length fields. The Length is
8 bytes for the Component Interface identifier IPv4 and
Component Interface identifier Unnumbered types. For
Component Interface identifier IPv6 type of sub-object, the
length field is 20 bytes.

U: 1 bit

This bit indicates the direction of the component
interface. It is 0 for the downstream interface. It is
set to 1 for the upstream interface and is only used for
bi-directional LSPs.

3.2 Processing of Component Interface identifier RRO Subobject

If a node desires component link recording, the "Component Link
Recording desired" flag (value TBD) should be set in the
LSP_ATTRIBUTES object, object that is defined in [RFC4420].

Setting of "Component Link Recording desired" flag is independent of
the Label Recording flag in SESSION_ATTRIBUTE object as specified in
[RFC3209]. Nevertheless, the following combinations are valid:

1) If both Label and Component Link flags are clear, then neither
   Labels nor Component Links are recorded.

2) If Label Recording flag is set and Component Link flag is
   clear, then only Label Recording is performed as defined in
   [RFC3209].

3) If Label Recording flag is clear and Component Link flag is
   set, then Component Link Recording is performed as defined in this
   proposal.

4) If both Label Recording and Component Link flags are set, then
   Label Recording is performed as defined in [RFC3209] and also
   Component Link recording is performed as defined in this proposal.

In most cases, a node initiates recording for a given LSP by adding
the RRO to the Path message. If the node desires Component Link
recording and if the outgoing TE link is bundled, then the initial
RRO contains the Component Link identifier (numbered or unnumbered)
as selected by the sender. As well, the Component Link Recording
desired flag is set in the LSP_ATTRIBUTE object. If the node also
desires label recording, it sets the Label_Recording flag in the
SESSION_ATTRIBUTE object.

When a Path message with the "Component Link Recording desired" flag
set is received by an intermediate node, if a new Path message is to
be sent for a downstream bundled TE link, the node adds a new
Component Link subobject to the RECORD_ROUTE object (RRO) and appends
the resulting RRO to the Path message before transmission.

Note also that, unlike Labels, Component Link identifiers are always
known on receipt of the Path message.

When the destination node of an RSVP session receives a Path message
with an RRO and the "Component Link Recording desired" flag set, this
indicates that the sender node needs TE route as well as component
link recording. The destination node initiates the RRO process by
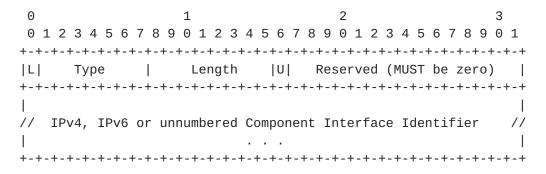adding an RRO to Resv messages. The processing mirrors that of the
Path messages

The Component Interface Record subobject is pushed onto the
RECORD_ROUTE object (RRO) prior to pushing on the node's IP address.
A node MUST NOT push on a Component Interface Record subobject
without also pushing on the IP address or unnumbered Interface Id
subobject that identifies the TE Link.

When component interfaces are recorded for bi-directional LSPs,
component interface RRO subobjects for both downstream and upstream
interfaces MUST be included.

[4]. Signaling Component Interface Identifier in ERO

A new OPTIONAL subobject of the EXPLICIT_ROUTE Object (ERO) is used
to specify component interface identifier of a bundled TE Link.

This Component Interface Identifier subobject 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |    Length     |U|   Reserved (MUST be zero)   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
//  IPv4, IPv6 or unnumbered Component Interface Identifier    //
|                         . . .                                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   L: 1 bit

      This bit must be set to 0.

   Type

Type 10 (TBD): Component Interface identifier IPv4
Type 11 (TBD): Component Interface identifier IPv6
Type 12 (TBD): Component Interface identifier Unnumbered

Length

The Length contains the total length of the subobject in
bytes, including the Type and Length fields. The Length is
8 bytes for the Component Interface identifier types: IPv4
and Component Interface identifier Unnumbered. For Component
Interface identifier IPv6 type of sub-object, the length field
is 20 bytes.

U: 1 bit

This bit indicates the direction of the component interface.
It is 0 for the downstream interface. It is set to 1 for the
upstream interface and is only used for bi-directional LSPs.

4.1 Processing of Component Interface Identifier ERO Subobject

The Component Interface Identifier ERO subobject follows a subobject
containing the IP address, or the link identifier [RFC3477],
associated with the TE link on which it is to be used. It is used to
identify the component of a bundled TE Link.

The following SHOULD result in "Bad EXPLICIT_ROUTE object" error
being sent upstream by a node processing an ERO that contains the
Component Interface ID sub-object:

o) The first component interface identifier subobject is not
   preceded by a sub-object containing an IPv4 or IPv6 address, or
   an interface identifier [RFC3477], associated with a TE link.

o) The Component Interface Identifier ERO subobject follows a
   subobject that has the L-bit set.

o) On unidirectional LSP setup, there is a Component Interface
   Identifier ERO subobject with the U-bit set.

o) Two Component Interface Identifier ERO subobjects with the same
   U-bit values exist.

If a node implements the component interface identifier subobject, it
MUST check if it represents a component interface in the bundled TE
Link specified in the preceding subobject that contains the IPv4/IPv6
address or interface identifier of the TE Link. If the content of the
component interface identifier subobject does not match a component

interface in the TE link, a "Bad EXPLICIT_ROUTE object" error SHOULD
be reported as "Routing Problem" (error code 24).

If U-bit of the subobject being examined is cleared (0) and the
upstream interface specified in this subobject is acceptable, then
the value of the upstream component interface is translated locally
in the TLV of the IF_ID RSVP HOP object [RFC3471]. The local
decision normally used to select the upstream component link is
bypassed except for local translation into the outgoing interface
identifier from the received incoming remote interface identifier. If
this interface is not acceptable, a "Bad EXPLICIT_ROUTE object" error
SHOULD be reported as "Routing Problem" (error code 24).

If the U-bit of the subobject being examined is set (1), then the
value represents the component interface to be used for upstream
traffic associated with the bidirectional LSP. Again, if this
interface is not acceptable or if the request is not one for a
bidirectional LSP, then a "Bad EXPLICIT_ROUTE object" error SHOULD be
reported as "Routing Problem" (error code 24). Otherwise, the
component interface IP address/ identifier is copied into a TLV sub-
object as part of the IF_ID RSVP_HOP object. The local decision
normally used to select the upstream component link is bypassed
except for local translation into the outgoing interface identifier
from the received incoming remote interface identifier.

The IF_ID RSVP_HOP object constructed as above MUST be included in
the corresponding outgoing Path message.

Note that, associated with a TE Link sub-object in the ERO, either
the (remote) upstream component interface or the (remote) downstream
component interface or both may be specified. As specified in
[RFC4201] there is no relationship between the TE Link type (numbered
or unnumbered) and the Link type of any one of its components.

The Component Interface Identifier ERO subobject is optional.
Similarly, presence of the Label ERO sub-objects is not mandatory
[RFC3471], [RFC3473]. Furthermore, component interface identifier
ERO subobject and Label ERO subobject may be included in the ERO
independently of each other. One of the following alternatives
applies:

o) When both sub-objects are absent, a node may select any
   appropriate component link within the TE link and any label on the
   selected component link.

o) When the Label subobject is only present for a bundled link, then
   the selection of the component link within the bundle is a local
   decision and the node may select any appropriate component link,

which can assume the label specified in the Label ERO.

o) When only the component interface identifier ERO subobject is
   present, a node MUST select the component interface specified in
   the ERO and may select any appropriate label value at the
   specified component link.

o) When both component interface identifier ERO subobject and Label
   ERO subobject are present, the node MUST select the locally
   corresponding component link and the specified label value on that
   component link. When present, both subobjects may appear in any
   relative order to each other but they MUST appear after the TE
   Link subobject that they refer to.

After processing, the component interface identifier subobjects are
removed from the ERO.

Inferred from above, the interface subobject should never be the
first subobject in a newly received message. If the component
interface subobject is the first subobject in a received ERO, then it
SHOULD be treated as a "Bad strict node" error.

Note: Information to construct the Component Interface ERO subobject
may come from the same mean used to populate the label ERO subobject.
Procedures by which an LSR at the head-end of an LSP obtains the
information needed to construct the Component Interface subobject are
outside the scope of this document.

5. Forward Compatibility Note

The extensions specified in this draft do not affect the processing
of the RRO, ERO at nodes that do not support them. A node that does
not support the Component Interface RRO subobject but that does
support Label subobject SHOULD only insert the Label subobject in the
RRO as per [RFC3471] and [RFC3473].

A node that receives an ERO that contains a Component Link ID
subobject SHOULD send "Bad EXPLICIT_ROUTE object" if it does not
implement this subobject.

Per [RFC3209], Section 4.4.5, a non-compliant node that receives an
RRO that contains Component Interface Identifier sub-objects should
ignore and pass them on. This limits the full applicability of if
nodes traversed by the LSP are compliant with the proposed
extensions.

6. Security Considerations

This document does not introduce new security issues. The security
considerations pertaining to the original RSVP protocol [RFC2205]
remain relevant.

7. IANA Considerations

This document introduces the following RSVP protocol elements:

o) Component Interface Identifier RRO subobject of the Record Route
   Object (RRO). The following Types are defined:

       Type 10 (TBD): Component Interface identifier IPv4
       Type 11 (TBD): Component Interface identifier IPv6
       Type 12 (TBD): Component Interface identifier Unnumbered

o) Component Interface Identifier subobject of the Explicit Route
   Object (ERO). The following Types are defined:

       Type 10 (TBD): Component Interface identifier IPv4
       Type 11 (TBD): Component Interface identifier IPv6
       Type 12 (TBD): Component Interface identifier Unnumbered

o) A new "Component Link Recording desired" flag (value TBD)
   of the LSP_ATTRIBUTES object [RFC4420]

8. References

8.1 Normative Reference

[RFC2205] R. Braden, et al., "Resource ReSerVation Protocol (RSVP)
           - Version 1, Functional Specification", RFC 2205,
           September 1997.

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

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

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

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

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

[RFC4201]   K. Kompella, et al., "Link Bundling in MPLS Traffic
            Engineering", RFC 4201, January 2003.

[RFC4420]   A. Farrel, et al., "Encoding of Attributes for
            Multiprotocol Label Switching (MPLS) Label Switched Path
            (LSP) Establishment Using Resource ReserVation Protocol-
            Traffic Engineering (RSVP-TE)", RFC 4420, February 2006.

8.2 Informative Reference

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

9. Author's Addresses

Anca Zamfir
Cisco Systems Inc.
2000 Innovation Dr.,
Kanata, Ontario, K2K 3E8
Canada.
Phone: (613)-254-3484
Email: ancaz@cisco.com

Zafar Ali
Cisco Systems Inc.
2000 Innovation Dr.,
Kanata, Ontario, K2K 3E8
Canada.
Phone: (613) 889-6158
Email: zali@cisco.com

Dimitri Papadimitriou
Alcatel-Lucent
Copernicuslaan 50
B-2018 Antwerpen
Belgium
Phone: +32 3 240-8491
Email: dimitri.papadimitriou@alcatel-lucent.be

10. Full Copyright Statement

11. Intellectual Property

Acknowledgement