Jean Philippe Vasseur

Zafar Ali

Siva Sivabalan
Cisco
Systems, Inc.

IETF Internet Draft
Expires: August, 2003

February, 2003

draft-vasseur-mpls-nodeid-subobject-00.txt

Definition of an RRO node-id subobject

Status of this Memo

Abstract

In the context of MPLS TE Fast Reroute ([FAST-REROUTE]), the Merge Point (MP) address is required at the Point of Local Repair (PLR) in order to select a backup tunnel intersecting a protected Traffic Engineering LSP on a downstream LSR. However, existing protocol mechanisms are not sufficient to find MP address multi-areas or multi-domain routing network. Hence, the current MPLS Fast Reroute mechanism cannot be used to protect inter-area or inter-AS TE LSPs from a failure of an ABR (Area Border Router) or ASBR (Autonomous System Border Router) respectively. This document specifies the use of existing RRO IPv4 and IPv6 subobjects (with a new flag defined) to define the node-id subobject in order to solve this issue.

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 [i].

1. Terminology

LSR - Label Switch Router

LSP - An MPLS Label Switched Path

PCS - Path Computation Server (may be any kind of LSR (ABR, ...)
or a centralized path computation server

PCC - Path Computation Client (any head-end LSR) requesting a
path
computation of the Path Computation Server.

Local Repair - Techniques used to repair LSP tunnels quickly
when a node or link along the LSPs path fails.

Protected LSP - An LSP is said to be protected at a given hop if
it has one or multiple associated backup tunnels
originating at that hop.

Bypass Tunnel - An LSP that is used to protect a set of LSPs
passing over a common facility.

Backup Tunnel - The LSP that is used to backup up one of the
many
LSPs in many-to-one backup.

PLR - Point of Local Repair. The head-end of a backup tunnel or
a detour LSP.

MP - Merge Point. The LSR where detour or backup tunnels meet

the protected LSP. In case of one-to-one backup, this is
where
multiple detours converge. A MP may also be a PLR.

NHOP Bypass Tunnel - Next-Hop Bypass Tunnel. A backup tunnel
which bypasses a single link of the protected LSP.

NNHOP Bypass Tunnel - Next-Next-Hop Bypass Tunnel. A backup
tunnel which bypasses a single node of the protected LSP.

Reroutable LSP - Any LSP for with the "Local protection desired" bit is set in the Flag field of the SESSION_ATTRIBUTE object of its Path messages.

CSPF - Constraint-based Shortest Path First.

Inter-AS MPLS TE: TE LSP whose Head-end LSR and Tail-end LSR do not reside within the same Autonomous System (AS) or both Head-end LSR and Tail-end LSR are in the same AS but the TE tunnel transiting path may be across different ASes

Interconnect or ASBR Routers: Routers used to connect to another AS of a different or the same Service Provider via one or more Inter-AS links.

2. Introduction

MPLS Fast Reroute (FRR) ([FAST-REROUTE]) is a fast recovery local protection technique used to protect Traffic Engineering LSPs from link/SRLG/node failure. One or more TE LSPs (called backup LSPs) are pre-established to protect against the failure of a link/node/ SRLG. In case of failure, every protected TE LSP traversing the failed resource is rerouted onto the appropriate backup tunnels in 10s of msecs.

There are a couple of requirements on the backup tunnel path. At least, a backup tunnel should not pass through the element it protects. Additionally, a primary tunnel and a backup tunnel should intersect at least at two points (nodes): Point of Local Repair (PLR) and Merge Point (MP). The former should be the head-end LSR of the backup tunnel, and the latter should be the tail-end LSR of the backup tunnel. The PLR is where FRR is triggered when link/node/SRLG failure happens. Furthermore, the MP address is also required to send RSVP Refresh messages of the rerouted TE LSPs when the FRR is active.

There are different methods for computing paths for backup tunnels. Specifically, a users can statically configures one or more backup tunnels at the PLR, with explicit path or the PLR can be configured to automatically compute a backup path or to send a path computation request to a PCS (which can be an LSR or an off-line tool). Consider the following scenario

Asumptions:
- a multi-area network made of three areas: 0, 1 and 2.
- a protected TE LSP T1 (TE LSP signaled with the "local Protection desired" bit set in the SESSION-ATTRIBUTE object or the FRR object) from R0 to R3
- a backup tunnel B1 from R1 to R2, not traversing ABR1, and following the R1-ABR3-R2 path. R1 reroutes any protected TE LSP traversing ABR1 onto the backup tunnel B1 in case of ABR1's failure.

```
           <--- area 1 --><---area 0---><---area 2--->
              R0-----R1-ABR1--R2------ABR2--------R3
                     \        /
                      \ ABR3 /
```

When T1 is first signaled, the PLR R1 needs to dynamically select an appropriate backup tunnel intersecting T1 on a downstream LSR. However, existing protocol mechanisms are not sufficient to unambiguously find MP address in a network with inter-area or inter-AS traffic engineering (although the example above was given in the context of multi-area networks, a similar reasoning applies to TE LSP spanning multiple ASes). This draft addresses these limitations.

R1 needs to ensure the following:

1. Backup tunnel intersects with the primary tunnel at the MP (and thus has a valid MP address), e.g., in Figure 1, R1 needs to determine that T1 and B1 share the same MP node R2,

2. Backup tunnel satisfies the primary LSP's request with respect to the bandwidth protection request (i.e., bandwidth guaranteed for the primary tunnel during failure), and the type of protection (preferably, protecting against a node failure versus a link failure), as specified in [FAST-REROUTE].

A PLR can make sure that condition (1) is met by examining the Record Route Object (RRO) of the primary tunnel to see if any of the addresses specified in the RRO is attached to the tail-end of the backup tunnel. As per [RSVP-TE], the addresses specified in the RRO IPv4 subobjects can be node-ids and/or interface addresses, with specific recommendation to use the interface address of the outgoing Path messages. Hence, in Figure 1, router R2 is more likely to specify interface addresses in the RROs for T1 and B1. Note that these interface addresses are different in this example.

The problem of finding the MP using the interface addresses or node-ids can be easily solved in a single area TE. Specifically, in the case of single area TE, the PLR has the knowledge of all the interfaces attached to the tail-end of the backup tunnel. This information is available in PLR's IGP topology database. Thus, the PLR can determine whether a backup tunnel intersecting a protected TE LSP on a downstream node exists and can also find the MP address regardless of how the addresses contained in the RRO IPv4 or IPv6 subobjects are specified (i.e., whether using the interface addresses or the node IDs). However, such routing information is not available in a multi-area and inter-AS traffic-engineering environments. Hence, unambiguously making sure that condition (1) above is met with inter-area TE and inter-AS traffic-engineering TE LSPs is not possible with existing mechanisms.

In this draft, we define extensions to and describe the use of RSVP [RSVP, RSVP-TE] to solve the above-mentioned problem.

3. Signaling node-ids in RROs

As mentioned above, the limitation that we need to address is the generality of the contents of the RRO IPv4 and IPv6 subobjects, as defined in [RSVP-TE].

The IPv4 and IPv6 RRO subobjects are currently defined in [RSVP-TE] and have the following flags defined:

Local protection available: 0x01

Indicates that the link downstream of this node is protected via a local repair mechanism, which can be either one-to-one or facility backup.

Local protection in use: 0x02

Indicates that a local repair mechanism is in use to maintain this tunnel (usually in the face of an outage of the link it was previously routed over, or an outage of the neighboring node).

Bandwidth protection: 0x04

The PLR will set this when the protected LSP has a
backup path
which is guaranteed to provide the desired bandwidth
specified
in the FAST_REROUTE object or the bandwidth of the
protected
LSP, if no FAST_REROUTE object was included. The PLR
may set
this whenever the desired bandwidth is guaranteed; the
PLR
MUST set this flag when the desired bandwidth is
guaranteed
and the "bandwidth protection desired" flag was set in
the
SESSION_ATTRIBUTE object. If the requested bandwidth is
not
guaranteed, the PLR MUST NOT set this flag.

Node protection: 0x08

The PLR will set this when the protected LSP has a
backup path
which provides protection against a failure of the next
LSR
along the protected LSP. The PLR may set this whenever
node
protection is provided by the protected LSP's backup
path; the

PLR MUST set this flag when the node protection is
provided
and the "node protection desired" flag was set in the
SESSION_ATTRIBUTE object. If node protection is not
provided,
the PLR MUST NOT set this flag. Thus, if a PLR could
only
setup a link-protection backup path, the "Local
protection
available" bit will be set but the "Node protection" bit
will
be cleared.

An additional flag is specified:

Node-id: 0x10

When set, this indicates that the address specified in the RRO IPv4 or IPv6 subobject is a node-id address, which refers to the "Router Address" as defined in [OSPF-TE], or "Traffic Engineering Router ID" as defined in [ISIS-TE]. A node MUST use the same address consistently.

An IPv4 or IPv6 RRO subobject with the node-is flag set is also called a node-id subobject.

The problem of finding MP address in a network with inter-area or inter-AS traffic engineering is solved by adding a node-id subobject (an RRO "IPv4" and "IPv6" sub-object with the 0x10 flag set).

Any Head-end LSR of a protected TE LSP MUST include an RRO object, as defined in [FAST-REROUTE]. In addition, any LSR compliant with this draft must systematically include a node-id IPv4 or IPv6 subobject in the RRO object for each protected TE LSP (in addition to the sub-objects required by MPLS TE Fast Reroute as defined in [FAST-REROUTE]).

A node MAY decide to include its node-id subobject in the RRO object only for the TE LSP whose IPv4 or IPv6 address source address (specified in the SENDER-TEMPLATE object of the RSVP Path message) does not belong to its local area/AS.

4. Processing RRO with node-id subobjects

The node-id subobject is added into the RECORD_ROUTE object after the Label Record subobject. A node MUST not push a node-id subobject without also pushing an IPv4 or IPv6 subobjects, as defined in [FAST-REROUTE]. A node may push both IPv4 node-id and IPv6 node-id sub-objects, but that MUST be done on consistent basis.

4.1. Finding Merge Point

A PLR can find the MP and suitable backup tunnel by simply comparing the node-id of the backup tunnel's tail-end with Node IDs included in the RRO of the primary tunnel. When both IPv4 node-id and IPv6 node-id

sub-objects are present, a PLR may use any or both of them in finding the MP address.

4.2. Processing at the border nodes

In a network with inter-AS traffic engineering, there may be some concerns about leaking the RRO information, including node-id subobjects, outside the autonomous system (see [INTER-AS-TE-REQS]). In such cases, before forwarding the RRO object outside of an AS, the ASBR may filter some/all node-id subobjects pertaining to the downstream nodes in the AS. The RRO node-id subobjects filtration can be based on a local policy configured on the ASBR.

How an ASBR handles/filters the contents of the RRO objects is outside of the scope of this draft.

5. Backward Compatibility Note

To remain compatible with the nodes that do not support the RRO IPv4 or

IPv6 node-id subobjects, a node can safely ignore these objects. The implication of this limitation will be that these nodes could not be MP in a network with inter-area or inter-AS traffic engineering.

6. Security Considerations

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

7. Intellectual Property Considerations

Cisco Systems may have intellectual property rights claimed in regard to some of the specification contained in this document

8. Acknowledgments

We would like to acknowledge input and helpful comments from Carol Iturralde, Anca Zamfir, and Reshad Rahman.

References

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

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

[FAST-REROUTE] Ping Pan, et al, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", draft-ietf-mpls-rsvp-lsp-fastreroute-01.txt, May

2003.

[OSPF-TE] Katz, D., Yeung, D., Kompella, K., "Traffic Engineering Extensions to OSPF Version 2", draft-katz-yeung-ospf-traffic-09.txt(work in progress).

[ISIS-TE] Li, T., Smit, H., "IS-IS extensions for Traffic Engineering", draft-ietf-isis-traffic-04.txt (work in progress)

[MULTI-AREA-TE] Kompella at all,"Multi-area MPLS Traffic Engineering", draft-kompella-mpls-multiarea-te-03.txt, June 2002.

[LOOSE-PATH-REOPT] Vasseur and Ikejiri, "Reoptimization of an explicit loosely routed MPLS TE paths", draft-vasseur-mpls-loose-path-reopt-00.txt, February 2003, Work in Progress.

[ABR-ASBR-FRR] Vasseur, Ali and Sivabalan, "Support of MPLS TE Fast Reroute protection of ABR/ASBR LSRs", draft-vasseur-mpls-abr-asbr-frr-00.txt, February 2003, Work in progress.

[INTER-AS-TE-REQS] Zhang et al, "MPLS Inter-AS Traffic Engineering requirements", draft-tewg-interas-te-req-01.txt (work in progress).

[INTER-AS-TE] Vasseur and Zhang, "MPLS Inter-AS Traffic Engineering", draft-vasseur-mpls-inter-as-te-00.txt, February 2003, work in progress.

Authors' Address:

Jean Philippe Vasseur
Cisco Systems, Inc.
300 Apollo Drive
Chelmsford, MA 01824
USA
Email: jpv@cisco.com

Zafar Ali
Cisco Systems, Inc.
100 South Main St. #200
Ann Arbor, MI 48104
USA
zali@cisco.com

Siva Sivabalan
Cisco Systems, Inc.
2000 Innovation Drive
Kanata, Ontario, K2K 3E8
Canada
msiva@cisco.com