Network Working Group
Internet Draft
<draft-iwata-mpls-crankback-rsvp-te-00.txt>
Expiration Date: June 2001

Atsushi Iwata
Norihito Fujita
NEC Corporation

Gerald R. Ash
AT&T

Adrian Farrel
Data Connection Ltd.

December 2000

Crankback Routing Extensions for MPLS Signaling with RSVP-TE

<draft-iwata-mpls-crankback-rsvp-te-00.txt>

Status of this Memo

Abstract

This draft proposes crankback routing extensions for RSVP-TE signaling and in a companion document for CRLDP signaling. Recently, several routing protocol extensions for advertising resource information in addition to topology information have been proposed for use in distributed constraint-based routing. In such a distributed routing environment, however, the information used to compute a constraint-based path may be out of date. This means that LSP setup requests may be blocked by links or nodes without sufficient resources. This draft specifies crankback routing extensions for CR-LDP and RSVP-TE so that the label request can be retried on an alternate path that detours around the blocked link or node upon a setup failure. Furthermore, the crankback routing schemes can also be applied to LSP restoration by indicating the location of the failure link or node. This would significantly improve the successful recovery ratio for failed LSPs, especially in situations where a large number of setup requests are triggered at the same time.

Table of Contents

1. Introduction

CR-LDP (Constraint-based Routing Label Distribution Protocol) [CR-LDP] and RSVP-TE (RSVP Extension for LSP Tunnel) [RSVP-TE] can be used for establishing explicitly routed LSPs (CR-LSPs) in an MPLS network. Using CR-LDP and RSVP-TE, resources can also be reserved along a path to guarantee or control QoS for traffic carried on the LSP. To designate an explicit path that satisfies QoS constraints, it is necessary to discern the resources available to each link or node in the network. For the collection of such resource information, routing protocols, such as OSPF [OSPF] and IS-IS [ISIS], can be extended to distribute additional state information [AWDUCHE1]. Explicit paths can be designated based on the distributed information at the LSR initiating a LSP and, if necessary, intermediate area border LSRs.

In a distributed routing environment, however, the resource information used to compute a constraint-based path may be out of date. This means that a setup request may be blocked, for example, because a link or node along the selected path has insufficient resources. In the current CR-LDP specification, when an LSP setup failure occurs, a Notification message is returned to the setup initiator (ingress LSR), which terminates this message and gives up the LSP establishment. In RSVP-TE, a blocked LSP setup may result in a PathErr message sent to the initiator or a ResvErr sent to the terminator (egress LSR). These messages may result in the LSP setup being abandoned. In Generalized MPLS [GMPLS] the Notify message can be used in RSVP-TE networks to expedite notification of LSP failures to ingress and egress LSRs, or to a specific "repair point".

If the ingress or intermediate area border LSR knows the location of the blocked link or node, the LSR can designate an alternate path and then reissue the setup request, which can be achieved by the mechanism known as crankback routing [PNNI, ASH1, ASH2]. We propose the use of crankback routing in RSVP-TE and, in a companion document, in CRLDP [CRNKBK-CRLDP]. Crankback routing requires notifying an upstream LSR of the location of the blocked link or node.

On the other hand, various restoration schemes for link or node failures have been proposed in [MAKAM, SHARMA] and others. Fast restoration by pre-establishing a backup LSP is useful for failures on a primary LSP. If both the primary and backup paths fail, however, it is necessary to reestablish the LSP on an end-to-end basis. End-to-end restoration for alternate routing requires the location of the failed link or node. The crankback routing schemes could also be used to notify upstream LSRs of the location of the failure, and this notification would likely occur more quickly than for the IGP to detect the failure and reconverge to new routing around the failure.

Furthermore, in situations where many link or node failures occur at the same time, the difference between the distributed routing information and the real-time network state becomes much greater than

in normal LSP setups. The LSP restoration must therefore be performed with inaccurate information, which is likely to cause setup blocking. Crankback routing would also improve failure recovery in these situations.

Recently, Multi-Protocol Lambda Switching has also been discussed [AWDUCHE2]. In a network without wavelength converters, setup requests are likely to be blocked more often than in a conventional MPLS environment because the same wavelength must be allocated at each OXC on an end-to-end explicit path. Furthermore, end-to-end restoration is the only way to recover LSP failures [CHAUDHURI]. This implies that crankback routing would also be useful in an MPLambdaS network, and again, the use of crankback could result in selecting an alternate path more quickly in a failure situation. This draft proposes a crankback routing system that is an extension of RSVP-TE, and of CRLDP in [CRNKBK-CRLDP], and discusses the identification of blocked links or nodes in an MPLS network.

See [CRNKBK-CRLDP] for a general discussion of crankback functionality, including

o explicit versus implicit rerouting,
o crankback routing behaviors due to LSP setup blockage,
o new status codes for rerouting, location identifiers of blocked links or nodes, and
o LSP restoration considerations.

In the following sections we identify the specific protocol extensions for crankback functionality within RSVP-TE.

2. RSVP-TE considerations

Nothing precludes the use of crankback routing mechanism and LSP restoration mechanism with appropriate TLV structures in the RSVP-TE messages.

We illustrate a simple case of setting up an LSP on an explicit route from router-A to router-B, in which certain Tspec and Flowspec requirements must be met on each link in the LSP. The steps are as follows [RSVP][RSVP-TE]:

(1) router-A (ingress LSR) sends an RSVP Path message to router-B (egress LSR) with

a) a TSPEC object that specifies the traffic characteristics of the data flow that the route-A will generate on the LSP

b) an EXPLICIT_ROUTE object (ERO) that specifies the explicit route of the LSP

(2) along the way accumulate Adspec to describe the network resources available

(3) at the router-B (the egress LSR) map these to an Rspec to build a FLOWSPEC object which includes a service class and the TSPEC object and the RSPEC object.

(4) router-B sends the FLOWSPEC object back in a Resv message along the specified explicit route, hop-by-hop in reverse order and reserves the resources link by link

Resource allocation failure may occur at two points.

(a) On the reverse path, as the Resv is processed, resources are reserved and, if they are not available on the required link or at a specific node, a ResvErr message is returned to the egress node (router-B) indicating "Admission Control failure" [RSVP]. Router-B is allowed to change the Rspec and try again, but in the event that this is not practical or not supported, router-B may choose to take any one of the following actions.

- Ignore the situation and allow recovery to happen through Path refresh and refresh timeout [RSVP].

- Send a PathErr towards router-A indicating "Admission Control failure".

- Send ResvTear towards router-A to abort the LSP setup.

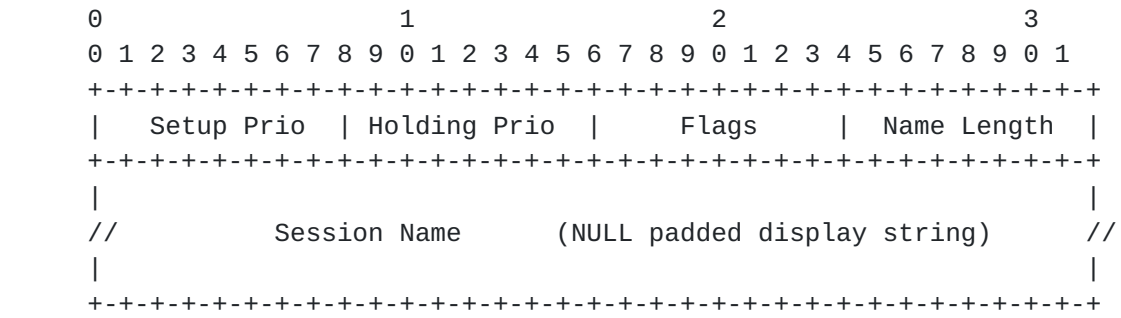
(b) It is also possible to make resource reservations on the forward path as the Path message is processed. This choice is made in many RSVP-TE implementations and is compatible with LSP setup in optical networks [GMPLS]. In this case if resources are not available, a PathErr message is returned to router-A indicating "Admission Control failure".

2.1 Indication of Rerouting Behavior

The behavior specified by the RtgFlg described in Section 6 of [CRNKBK-CRLDP] is achieved in RSVP-TE by a change to the LSP_TUNNEL Session Attribute Object. A new RtgFlg field is added in what was previously a reserved area in the object.

The object is now specified as follows.

class = SESSION_ATTRIBUTE, LSP_TUNNEL C-Type = 7



Setup Priority
The priority of the session with respect to taking resources, in the range of 0 to 7. The value 0 is the highest priority. The Setup Priority is used in deciding whether this session can preempt another session.

Holding Priority
The priority of the session with respect to holding resources, in the range of 0 to 7. The value 0 is the highest priority. Holding Priority is used in deciding whether this session can be preempted by another session.

Flags
0x01 Local protection desired

This flag permits transit routers to use a local repair mechanism which may result in violation of the explicit route object. When a fault is detected on an adjacent downstream link or node, a transit router can reroute traffic for fast service restoration.

0x02 Label recording desired

This flag indicates that label information should be included when doing a route record.

0x04 SE Style desired

This flag indicates that the tunnel ingress node may choose to reroute this tunnel without tearing it down. A tunnel egress node SHOULD use the SE Style when responding with a Resv message.

0x08 End-to-end rerouting desired

This flag indicates the end-to-end rerouting behavior for an LSP under establishment. This can also be used for specifying the behavior of end-to-end LSP restoration for established LSPs.

0x10 Segment-based rerouting (hierarchical rerouting) desired.

This flag indicates the segment-based rerouting (hierarchical rerouting) behavior for an LSP under establishment. This can also be used for specifying the segment-based (hierarchical) LSP restoration for established LSPs.

Name Length

The length of the display string before padding, in bytes.

Session Name

A null padded string of characters.

2.2 Rerouting Object

We define a new object, the "rerouting Object", analagous to the "Rerouting TLV" defined in Section 7.1 of [CRNKBK-CRLDP], to explicitly indicate that crankback rerouting is allowed at router-A. The new object is added to the definition of the PathErr message specifying it as optional.

When a PathErr message carrying the Rerouting object is received at router-A, router-A MAY attempt crankback rerouting. It can choose from the following actions.

- Send PathTear to give up

- Do nothing to give up (soft state times out)

- Pick a new route and send a new Path

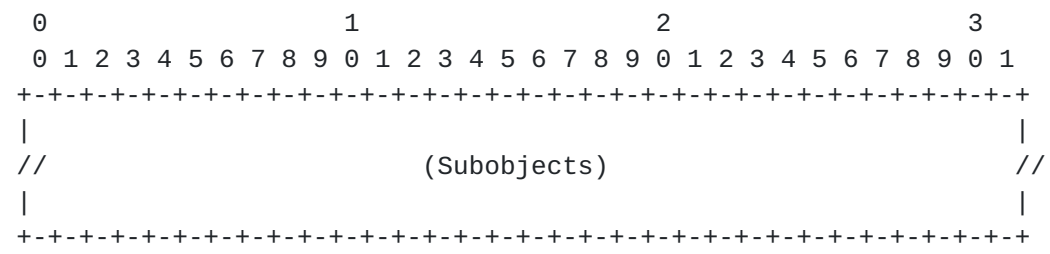
- Change the resource requirements and send a new Path message

The PathErr message is defined as follows:

```
<PathErr message> ::= <Common Header> [ <INTEGRITY> ]
                      <SESSION> <ERROR_SPEC>
                      [ <REROUTING> ]
                      [ <POLICY_DATA> ...]
                      [ <sender descriptor> ]
```

The Rerouting Object is defined as follows:

IPv4 REROUTING object: Class = TBD, C-Type = 1

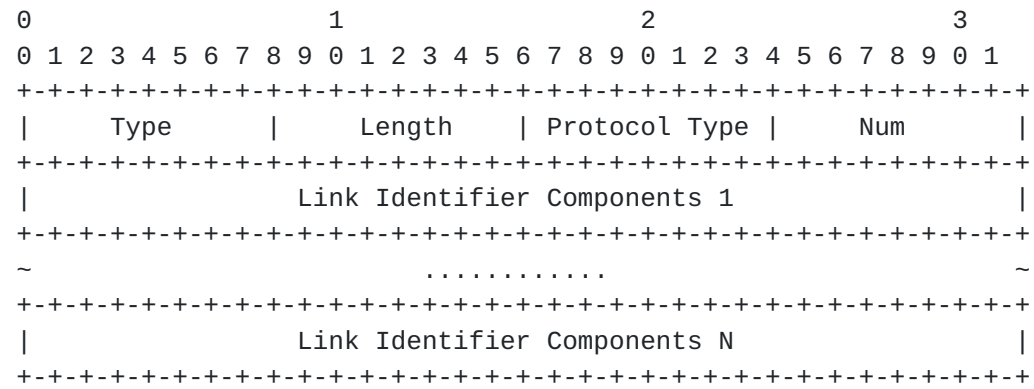


The Rerouting object may contain one or more subobjects of the following types:

Link Rerouting Subobject: See below for the format.
Node Rerouting Subobject: See below for the format.

2.3 Link-Rerouting Subobject

This subobject is analogous to the Link TLV in Sec.7.2 of [CRNKBK-CRLDP]. It may be carried in the Rerouting object in a PathErr message.



Type
The Type indicates the type of contents of the subobject.
1 Link-Rerouting subobject

Length
The length in bytes of the subobject

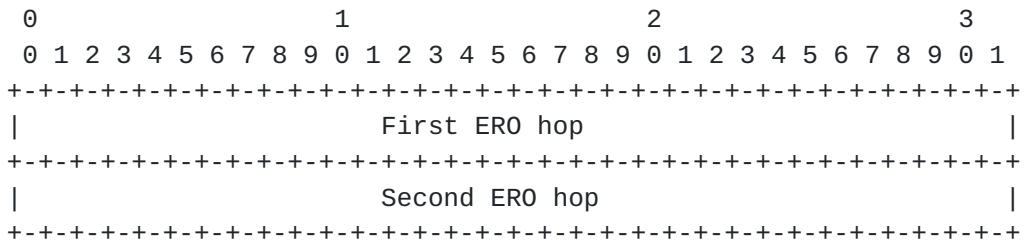
Protocol Type
A one-octet unsigned integer containing the unique identifier of a protocol used for QoS routing. The following type numbers are defined. If a constraint-based route is to be calculated with no routing protocols, type 1 should be used. A location is identified using the ERO, independent of protocols. If a type other than 1 is used, or the routing-protocol-specific link identifiers are used, the Flag of the SESSION_ATTRIBUTE in the LSP_TUNNEL_IPv4 Session Object carried in the Path message must be set 0x10, which implies that segment-based rerouting (hierarchical rerouting) is used.

1: RSVP-TE for IPv4 (EXPLICIT ROUTE Object)
2: OSPFv2 for IPv4 [OSPF]
3: Integrated IS-IS (or IS-IS) for IPv4 [ISIS]
4: OSPF for IPv6 [OSPF6]
5: Integrated IS-IS (or IS-IS) for IPv6 [ISIS6]
128-255: Reserved for private and experimental use.

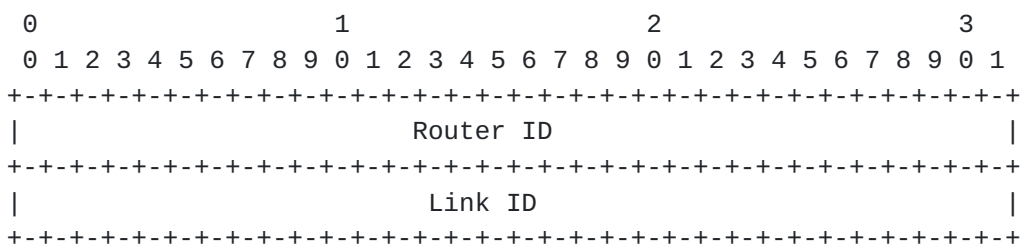
Num: Number of Link Identifier Components
A one-octet unsigned integer specifying the number of Link Identifier Components included in the object.

One or more Link Identifier Components
A variable length field containing link identifiers for relevant protocol types.

When the type is RSVP-TE for IPv4 (1), the following format is used. The first and second ERO hop upon setup blockage are included.



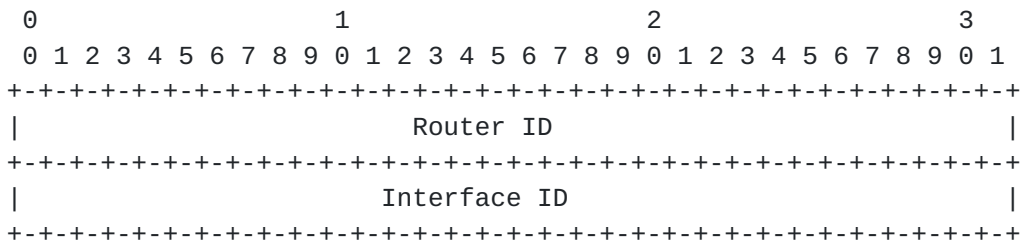
When the protocol type is OSPFv2 for IPv4 (2), the following 8-octet format is used.



Router ID
The same value as the Router ID used in OSPFv2 for IPv4.

Link ID
The same value as the Link ID used in OSPFv2 for IPv4.

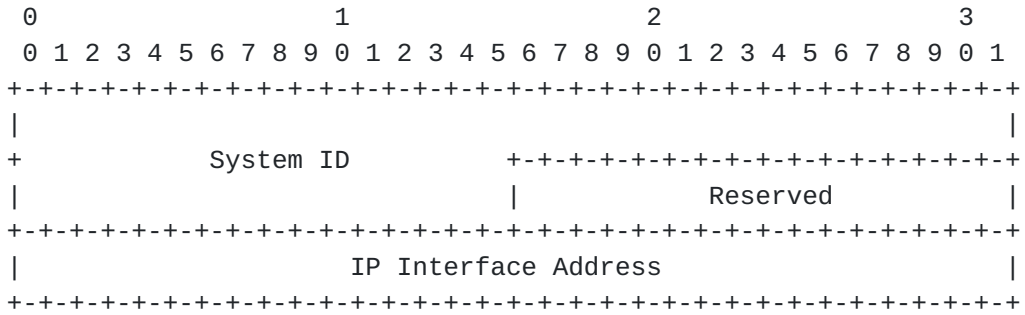
When the protocol type is OSPF for IPv6 (4), the following 8-octet format is used.



Router ID
The same value as the Router ID used in OSPF for IPv6.

Interface ID
The same value as the Interface ID used in OSPF for IPv6.

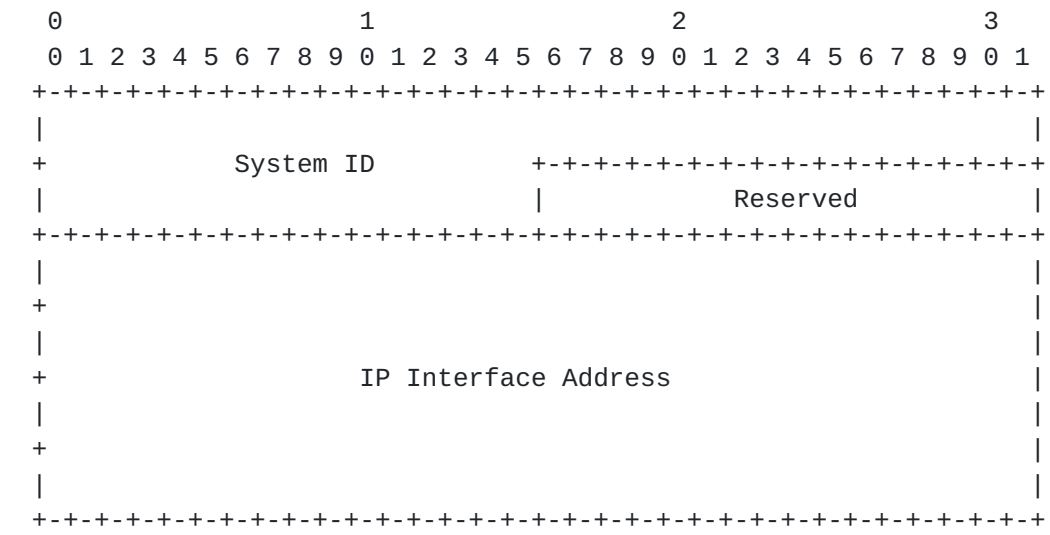
When the protocol type is IS-IS for IPv4 (3), the following 12-octet format is used.



System ID
The same value as System ID used in IS-IS for IPv4 (3).

IP Interface Address
The IP address assigned to the interface advertised in IS-IS for IPv4 (3).

When the protocol type is IS-IS for IPv6 (5), the following 24-octet format is used.

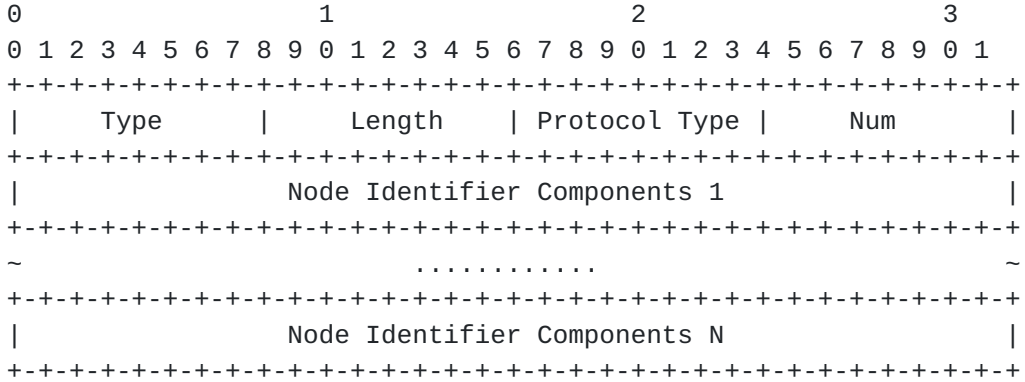


System ID
The same value as System ID used in IS-IS for IPv6 (5).

IP Interface Address
The IP address assigned to the interface advertised in IS-IS for IPv6 (5).

2.4 Node-Rerouting Subobject

This subobject is analogous to the Node TLV in Sec.7.3 of [CRNKBK-CRLDP]. It may be carried in the Rerouting object in a PathErr message.



Type
The Type indicates the type of contents of the subobject.
2 Node-Rerouting subobject

Length
The length in bytes of the subobject

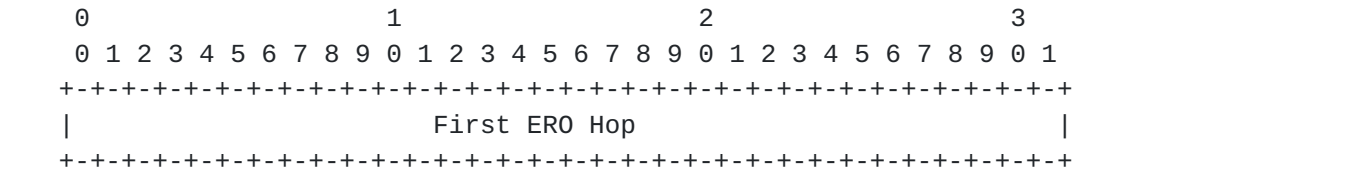
Protocol Type

A one-octet unsigned integer containing the unique identifier of the protocol used for QoS routing. The same value as defined in the Link-Rerouting subbject is used.

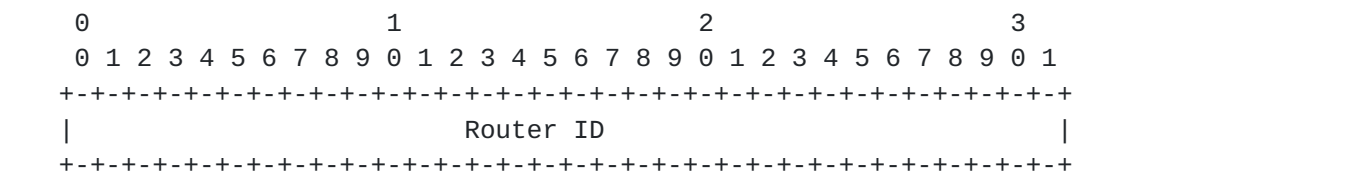
Num: Number of Node Identifier Components
A one-octet unsigned integer specifying the number of Node Identifier Components included in the object.

One or mode Node Identifier Components
A variable length field containing the node identifier for relevant protocol types.

When the type is RSVP-TE for IPv4 (1), the following format is used. The first ERO Hop upon setup blockage is included.

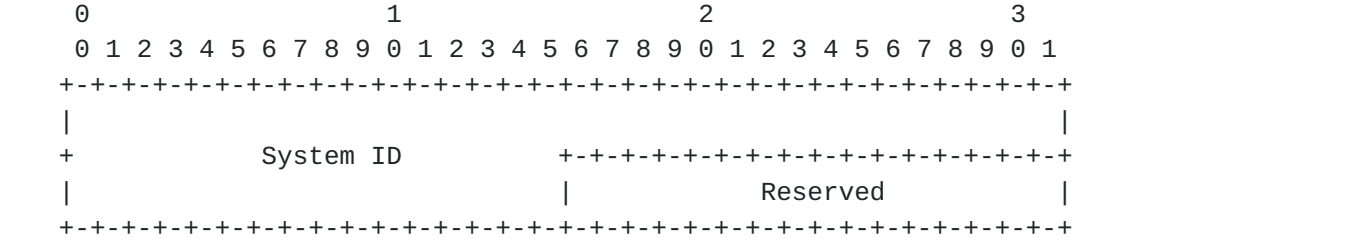


When the protocol type is OSPFv2 for IPv4 (2) or OSPF for IPv6 (4), the following 4-octet format is used.



Router ID
The same value as the Router ID used in OSPFv2 for IPv4 or in OSPF for IPv6.

When the protocol type is IS-IS for IPv4 (3) or IS-IS for IPv6 (5), the following 8-octet format is used.



System ID
The same value as the System ID used in IS-IS for IPv4 (3) or in

IS-IS for IPv6 (5).

2.5 Error Values

Error values for the Error Code "Admission Control Failure" are defined in [RSVP]. It may be appropriate to define new additional subcodes to provide additional action information.

This area is for future study.

2.6 ResvErr Usage

As described above, the resource allocation failure for RSVP-TE may occur when the reverse path when the Resv message is being processed. In this case, it is still useful to collect the crankback information and return it to the ingress LSR.

This can be achieved by specifying additions to the ResvErr message as follows.

```
<ResvErr Message> ::= <Common Header> [ <INTEGRITY> ]
                      <SESSION>  <RSVP_HOP>
                      <ERROR_SPEC>
                      [ <SCOPE> ]
                      [ <REROUTING> ]
                      [ <POLICY_DATA> ...]
                      <STYLE> [ <error flow descriptor> ]
```

The Rerouting object carried on the ResvErr is exactly as described above.

When a ResvErr carrying a Rerouting object is received at an egress LSR, the egress LSR MAY ignore this object and perform the same actions as for any other ResvErr. However, if the egress LSR supports the crankback extensions defined in this draft, it SHOULD generate a PathErr message and send it to the ingress LSR.

Such a PathErr should contain

- an Error_Spec object that shows the egress LSR as the Error Node

- the Rerouting object unchanged from the ResvErr.

2.7 Notify Message Usage

[GMPLS] defines a new message, the Notify message, to supplement error reporting in RSVP-TE networks. The Notify message is sent to

addresses requested on the Path and Resv messages. These addresses could (but need not) identify the ingress and egress LSRs respectively.

When a network error occurs, such as the failure of link hardware, the LSRs that detect the error MAY send Notify messages to the requested addresses. The type of error that causes a Notify message to be sent is an implementation detail.

The Notify message is not intended to replace the PathErr and ResvErr messages.

In the event that an LSR that supports [GMPLS] and the crankback extensions defined in this draft, it MAY choose to send a Notify message in the event of resource allocation failure. This would ensure a speedier report of the error to the ingress/egress LSRs and might make LSP restoration faster (see Section 10 of [CRNKBK-CRLDP]).

To facilitate crankback after a Notify message, the Notify message is extended to optionally carry the Rerouting object as follows.

```
<Notify message> ::= <Common Header> [<INTEGRITY>] <MESSAGE_ID>
                     <ERROR_SPEC>
                     [<REROUTING>]
                     <notify session list>

<notify session list> :== [<notify session list>] <notify session>

<notify session> :== <SESSION> [<POLICY_DATA>...]
                     <sender descriptor>
```

2.8 Backward Compatibility

It is recognized that not all nodes in an RSVP-TE network will necessarily support the extensions defined in this draft. It is important that an LSR that does not support these extensions can continue to process a PathErr, ResvErr or Notify message even if it carries the new Rerouting object.

This is achieved by using the rules set out in Section 3.10 of [RSVP] to define the value of the Class of the new Rerouting object. The value shall be selected from the set of values as follows.

- Class-Num = 11bbbbbb

  The node should ignore the object but forward it, unexamined and unmodified, in all messages resulting from this message.

3. Security Considerations

Both the crankback routing extensions and LSP restoration extensions for RSVP-TE inherit the same security mechanisms described in [RSVP] to protect against spoofing attacks of a session, the privacy of signaling messages, and the denial of service (DoS) attacks.

4. Acknowledgments

We would like to thank Juha Heinanen and Srinivas Makam for their review and comments.

5. References

[ASH1] G. Ash, "Routing Guidelines for Efficient Routing Methods," work in progress <draft-ash-itu-sg2-routing-guidelines-00.txt>, October 1999.

[ASH2] G. Ash, "Traffic Engineering & QoS methods for IP-, ATM-, & TDM-Based Multiservice Networks," work in progress <draft-ash-te-qos-routing-01.txt>, July 2000.

[ASH3] G. Ash, et al., "LSP Modification Using CR-LDP", work in progress <draft-ietf-mpls-crlsp-modify-02.txt>, October 2000.

[ASHW] P. Ashwood-Smith, et al., "Improving Topology Data Base Accuracy with LSP Feedback," work in progress <draft-ietf-mpls-te-feed-01.txt>, July 2000.

[AWDUCHE1] D. Awduche, et al., "Requirements for Traffic Engineering Over MPLS," RFC2702, September 1999.

[AWDUCHE2] D. Awduche, et al., "Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control With Optical Crossconnects", work in progress <draft-awduche-mpls-te-optical-02.txt>, July 2000.

[CHAUDHURI] S. Chaudhuri, et al., "Control of Lightpaths in an Optical Network," work in progress <draft-chaudhuri-ip-olxc-control-00.txt>, February 2000.

[CR-LDP] B. Jamoussi, et al., "Constraint-Based LSP Setup using LDP," work in progress <draft-ietf-mpls-cr-ldp-04.txt>, July 2000.

[GMPLS] P. Ashwood-Smith and L. Berger, et al., "Generalized MPLS - Signaling Functional Description," work in progress, <draft-ietf-mpls- generalized-signaling-00.txt>, October 2000.

[INTER-AREA-MPLS1] S. Venkatachalam, "OSPF, IS-IS, RSVP, CR-LDP extensions to support inter-area traffic engineering using MPLS TE," work in progress, <draft-dharanikota-interarea-mpls-te-ext-01.txt>, December 2000.

[INTER-AREA-MPLS2] S. Venkatachalam, "A Framework for the LSP Setup Across IGP Areas for MPLS Traffic Engineering," work in progress, <draft-venkatachalam-interarea-mpls-te-01.txt>, December 2000.

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

[ISIS6] C. E. Hopps, "Routing IPv6 with IS-IS," work in progress <draft-ietf-isis-ipv6-01.txt>, July 2000.

[LDP] L. Andersson, et al., "LDP Specification," work in progress <draft-ietf-mpls-ldp-11.txt>, August 2000.

[MAKAM] S. Makam, et al., "Framework for MPLS-based Recovery," work in progress <draft-makam-mpls-recovery-frmwrk-01.txt>, July 2000.

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

[OSPF6] R. Coltun, et al., "OSPF for IPv6," RFC2740, December 1999.

[PNNI] ATM Forum, "Private Network-Network Interface Specification Version 1.0 (PNNI 1.0)," <af-pnni-0055.000>, May 1996.

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

[RSVP-TE] D. Awduche, et al., "Extensions to RSVP for LSP Tunnels," work in progress <draft-ietf-mpls-rsvp-lsp-tunnel-07.txt>, March 2000.

[SHARMA] V. Sharma, et al., "An Assessment of QoS and Protection in MPLS," MPLS'99 Conference, June 1999.

[CRNKBK-CRLDP] A. Iwata, et. al., ôCrankback Routing Extensions for MPLS Signaling with CRLDP,ö work in progress <draft-ietf-mpls-crankback-crldp-00.txt>, December 2000.

6. Authors' Addresses

Atsushi Iwata
NEC Corporation
Computer & Communication Media Research
1-1, Miyazaki, 4-Chome, Miyamae-ku,
Kawasaki, Kanagawa, 216-8555, JAPAN

Phone: +81-(44)-856-2123
Fax: +81-(44)-856-2230
Email: a-iwata@ah.jp.nec.com

Norihito Fujita
NEC Corporation
Computer & Communication Media Research
1-1, Miyazaki, 4-Chome, Miyamae-ku,
Kawasaki, Kanagawa, 216-8555, JAPAN

Phone: +81-(44)-856-2123
Fax: +81-(44)-856-2230
Email: n-fujita@bk.jp.nec.com

Gerald R. Ash
AT&T
Room MT D5-2A01
200 Laurel Avenue
Middletown, NJ 07748, USA

Phone: +1-(732)-420-4578
Fax: +1-(732)-368-8659
Email: gash@att.com

Adrian Farrel
Network Convergence Group
Data Connection Ltd.
Windsor House
Pepper Street
Chester, UK

Phone: +44-(0)-1244-313440
Fax: +44-(0)-1244-312422
Email: af@dataconnection.com