CCAMP
Internet-Draft
Updates: 3209 (if approved)
Intended status: Informational
Expires: May 23, 2013

S. Balls, Ed.
J. Hardwick
Metaswitch
C. Margaria
Nokia Siemens Networks

November 19, 2012

Relaxing RSVP Loop Checking
draft-balls-ccamp-relax-loop-check-02

Abstract

This specification relaxes the rules governing loop checking within RSVP. These were originally defined in RFC3209 and are too strict for the requirements of today's data planes.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 23, 2013.

Copyright Notice

Table of Contents

1. Introduction

Generalized MPLS (GMPLS) Traffic Engineering (TE) Label Switched Paths (LSPs) are prohibited from passing through a single node more than once. Today's data planes are such that allowing spiral paths through a control plane node should be allowed in order to set up LSPs.

1.1. Requirements Language

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].

2. General Overview

With today's data planes it is acceptable for a single data flow (LSP) to pass through a single control plane node on more than one occasion on the path from source to destination. Currently control plane protocols will prevent such a path being managed in the control plane as they explicitly detect this as a loop. However, this may not necessarily be a loop in the data plane and it is desirable for such LSPs to be able to be managed in the same way as non-looping LSPs. This document refers to such LSPs as spiralling LSPs.

2.1. Example in WDM networks

In WDM networks it can be necessary to route the data via an additional box in order to fulfil regeneration or wavelength conversion requirements. For example, consider the following simple example.

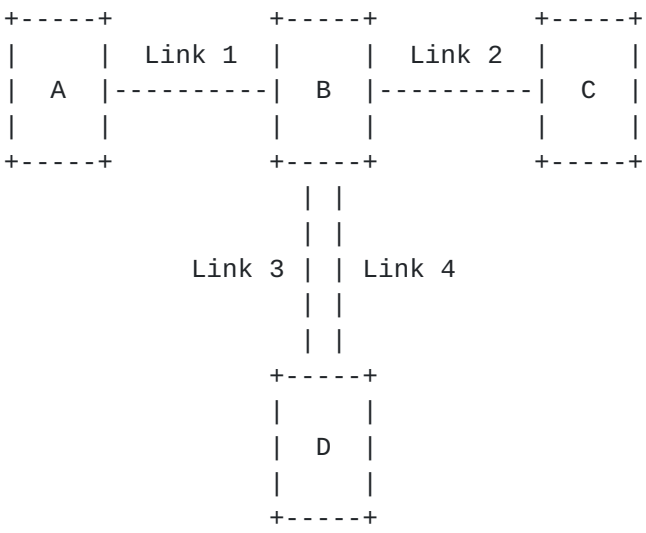

Figure 1

If node B cannot perform wavelength conversion but Link 1 and Link 2 do not have a common free wavelength then the only way to set up a path from node A to node C will be via node D. This requires two passes through node B which to RSVP looks like a loop, but is a spiral.

2.2. Example using Connectivity Matrices

In any type of network a specific node may have connectivity restrictions that limit the output ports available given the input ports. Connectivity Matrices are described in [RFC6163] For example, given the above network, where node B has the following connectivity restrictions.

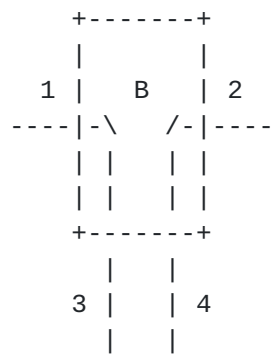

Figure 2

As in the above example, the only way to set up a path from node A to node C will be via node D. This requires two passes through node B which to RSVP looks like a loop, but is a sprial.

2.3. Example with additional label restrictions

Connections between ports on a node may be restricted based on labels. Consider the following network.

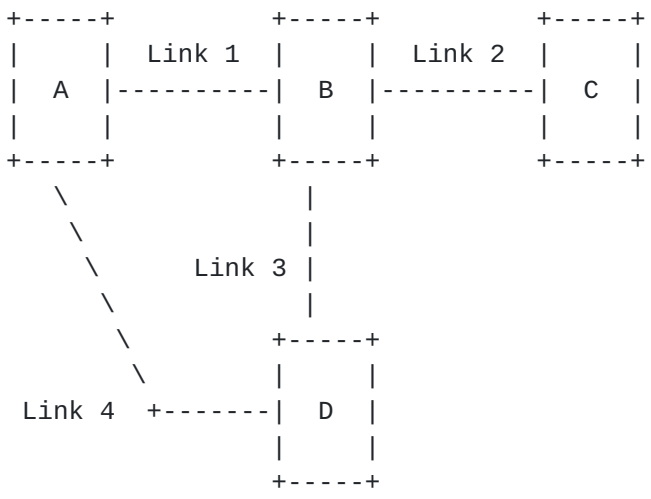

Figure 3

This network has the following properties.

- Node A is electro-optical outputting Lambda 1 and can switch Lambda 2.

o Node D can convert between Lambda 1 and Lambda 2.

o Link 1 and Link 3 have Lambda 1 available.

o All links have Lambda 2 available.

To setup a path from A to C in this network, the LSP must pass through Link 1 twice: once using Lambda 1 and once using Lambda 2. This results in the path A-B-D-A-B-C being taken which requires two passes through node A. This looks like a loop, but due to the different lambdas used on each pass is a spiral.

2.4. Example In Distributed Networks

In networks where the control plane and data plane are physically distinct, it is possible that a single control plane element will be controlling multiple data plane elements. This is the case now in some ASON networks, and will increasingly be the case with the move towards SDN networks. Consider the following network.

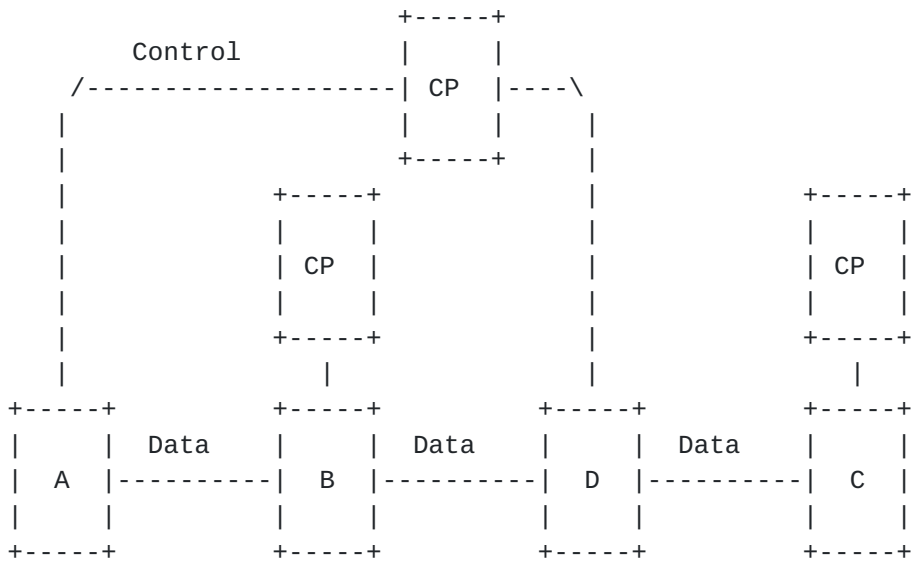

Figure 4

CP are the control planes instances, with A, B, D and C the data plane. Since data nodes A and D are managed by one control plane, an LSP from A to C would appear as a loop, where it is clear that this is not the case in the data plane. As different interfaces are being used the control plane could treat such an LSP as a spiral.

2.5. Example with Ingress Protection

If performing ingress protection with an off-forwarding-path backup node, as described by [I-D.torvi-mpls-rsvp-ingress-protection], then the ingress node will see a Path message for the same session twice. Preventing a data plane loop, but allowing a spiral is also required in this case.

3. Existing workaround

In current networks it is possible to support such paths either through management configuration at each node, or splitting the path into two or more signalling sessions. In the above examples this can be achieved with one session from A to D, and a second session from D to C. It would also require management on node D to join the data paths together. It is desirable that a single signalling session can be used to set up such paths, thus only requiring management input at the ingress.

4. Solution

4.1. Overview

To support such networks, the rules governing RSVP loop checking are relaxed to allow spirals, but still prevent loops. No changes to protocol messages are made.

4.2. Assumptions and limitations

These changes are only applicable to GMPLS out of band signalling when using point to point data links.

4.3. General Rules

The following rules govern the changes in behaviour that allow RSVP loop checking to be relaxed while still setting up non-looping data paths in RSVP.

- For each pass through the control plane node, the pair of inbound and outbound data interfaces and labels must be different.

4.4. RRO handling

Section 4.4.4 of [RFC3209] states that RSVP must reject a Path message if the receiving router is already in the RRO. This is now relaxed to allow such a condition provided a different interface-

label pair is used in each case. If the router has existing session state for a received Path message, and it MUST verify that the newly requested data path (input and output interface and label) is different from the existing data path(s) for that session, and the existing data path(s) is (are) present earlier in the RRO. If this is not the case, the router MUST return a "Routing problem" PathErr message with the error value "loop detected".

In order to carry out this checking correctly, specific interfaces and labels SHOULD be recorded in the RRO. If this is not the case, each node can only verify the path is acceptable against local state and should not reject the RRO if the local state is valid.

It is allowable for local policy to exist to limit the number of different paths through a router in a single LSP instance. If this limit is exceeded the router SHOULD return a "Routing problem" PathErr message with the error value "loop detected". This local policy is not intended to be advertised in routing. It is present as a backstop to protect against malicious Path messages consuming all resources on the router.

4.5. ERO handling

Sections 4.3.4.1 and 4.3.5 of [RFC3209] also state that RSVP must detect and avoid loops. This checking is also relaxed to allow spirals in the cases stated above. Again, local policy can limit the number of different paths through a router in a single LSP instance. A router may "look ahead" in the ERO to determine such local policy will be exceeded in advance of it happening and SHOULD return a "Routing problem" PathErr message with the error value "loop detected" in such a case.

When calculating or expanding an ERO a router may include multiple entries through a single router. If the ERO contains loose hops that form a loop, and a node determines a non-looping route is available, it MAY remove the loop from the ERO.

4.6. Interface handling

As stated in the general rules, an implementation supporting multiple passes through a node must ensure that for each pass the input and output interfaces and labels are different.

Internally, this means that if a Path message is received using a different input interface this may no longer mean the LSP has been rerouted upstream. Implementations must check the RRO to determine the correct behaviour when processing such a Path message. Care must be taken to handle valid cases where the incoming label can change.

4.7. Signalling

For the avoidance of doubt, no new signalling is being defined in this draft.

The behaviour of refresh or error messages is unchanged and should therefore be sent along the looped path (if present). Nodes SHOULD NOT shortcut the loop.

4.8. Error Handling

How to behave when receiving a PathErr with error value "loop detected" is out of scope of this draft and is a local implementation decision. For example, it may choose to try and recalculate the path mandating that the error node is avoided, or does not support looping.

5. Acknowledgements

With thanks to Jonathan Sadler and Yimin Shen for their input when discussing this draft.

6. IANA Considerations

This memo includes no request to IANA.

7. Security Considerations

In principle these changes to RSVP pose no security exposures over and above [RFC3209]. However, by allowing loops a single LSP can now consume multiple resources. As suggested local policy can limit the number of paths and thus the resource a single LSP can consume.

8. References

8.1. Normative References

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

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

8.2. Informative References

[I-D.torvi-mpls-rsvp-ingress-protection]
Atlas, A., Torvi, R., and M. Jork, "Ingress Protection for RSVP-TE p2p and p2mp LSPs", draft-torvi-mpls-rsvp-ingress-protection-00 (work in progress), July 2012.

[RFC6163] Lee, Y., Bernstein, G., and W. Imajuku, "Framework for GMPLS and Path Computation Element (PCE) Control of Wavelength Switched Optical Networks (WSONs)", RFC 6163, April 2011.

Authors' Addresses

Steve Balls (editor)
Metaswitch
100 Church Street
Enfield, EN2 6BQ
UK

Phone: +44 208 366 1177
Email: steve.balls@metaswitch.com

Jonathan Hardwick
Metaswitch
100 Church Street
Enfield, EN2 6BQ
UK

Phone: +44 208 366 1177
Email: jonathan.hardwick@metaswitch.com

Cyril Margaria
Nokia Siemens Networks
St Martin Strasse 76
Munich, 81541
Germany

Phone: +49 89 5159 16934
Email: cyril.margaria@nsn.com