Resilient MPLS Rings
draft-kompella-mpls-rmr-00

Abstract

This document describes the use of the MPLS control and data planes on ring topologies. It describes the special nature of rings, and proceeds to show how MPLS can be effectively used in such topologies. It describes how MPLS rings are configured, auto-discovered and signaled, as well as how the data plane works. Companion documents describe the details of discovery and signaling for specific protocols.

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].

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

Rings are a very common topology in transport networks. A ring is the simplest topology offering link and node resilience. Rings are nearly ubiquitous in access and aggregation networks. As MPLS increases its presence in such networks, and takes on a greater role in transport, it is imperative that MPLS handles rings well; this is not the case today.

This document describes the special nature of rings, and the special needs of MPLS on rings. It then shows how these needs can be met in several ways, some of which involve extensions to protocols such as IS-IS [RFC1195], OSPF[RFC2328], RSVP-TE [RFC3209] and LDP [RFC5036].

1.1. Definitions

A (directed) graph G = (V, E) consists of a set of vertices (or nodes) V and a set of edges (or links) E. An edge is an ordered pair of nodes (a, b), where a and b are in V. (In this document, the terms node and link will be used instead of vertex and edge.)

A ring is a subgraph of G. A ring consists of a subset of nodes {R_i, 1 <= i <= n} of V. For convenience, we define R_0 = R_n. The edges {(R_i, R_i+1) and (R_i+1, R_i), 0 <= i < n} must be a subset of E. We define the direction from node R_i to R_i+1 (0 <= i < n) (and hence from R_n = R_0 to R_1) as "downstream" (DS) and the reverse direction as "upstream" (US). As there may be several rings in a graph, we number each ring with a distinct "Ring ID" RID.

```
                  R1 . . . R2
                .             .
             R8                 R3
          |  .       Ring        .  |
 Upstream |  .                   .  | Downstream
          v  .     RID = 17      .  v
             R7                 R4
                .             .
                  R6 . . . R5
```

Figure 1: Ring with 8 nodes

The following terminology is used for ring LSPs:

RL_k: A Ring LSP anchored on node R_k is denoted RL_k.

DL_jk (UL_jk): A label allocated by R_j for RL_k in the DS (US) direction.

P_jk (Q_jk): A Path (Resv) message sent by R_j for RL_k.

2. Motivation

A ring is the simplest topology that offers resilience. This is perhaps the main reason to lay out fiber in a ring. Thus, effective mechanisms for fast failover on rings are needed. Furthermore, there are large numbers of rings. Thus, configuration of rings needs to be as simple as possible. Finally, bandwidth management on access rings is very important, as bandwidth is generally quite constrained here.

The goals of this document are to present mechanisms for improved MPLS-based resilience in ring networks (using ideas that are reminiscent of Bidirectional Line Switched Rings), automatic bring-up of LSPs, better bandwidth management and auto-hierarchy. These goals can be achieved using extensions to existing IGP and MPLS signaling protocols, using central provisioning, or in other ways.

3. Theory of Operation

Say a ring has n nodes R_i, 0 <= i <= n, where R_0 = R_n. Each node is the anchor of one or more Ring LSPs. A Ring LSP RL_i anchored on node R_i consists of two counter-rotating LSPs that start and end at R_i. A ring LSP is "multipoint"; any node R_j can use RL_i to send traffic to R_i in either direction. Bidirectional connectivity between nodes R_i and R_j is achieved by using ring LSPs RL_j (to reach R_j) and RL_i (to reach R_i); each can be used in either direction.

3.1. Configuration

An MPLS ring is configured by assigning RIDs to all the nodes in the ring. The links between adjacent ring nodes are ring links (unless told otherwise); this may also be configured, or it may be discovered, say by means of IGP hellos. Once ring nodes and ring links are identified, the ring has been defined.

Ring LSPs are not provisioned; they are created automatically when an MPLS ring is defined. Each node R_j allocates DS and US labels for each ring LSP RL_k and sends these to its ring neighbors. The signaling protocol used to send labels can be RSVP-TE or LDP; these extensions will be described later. When R_j receives DS and US labels for RL_k, it can install LFIB entries for RL_k.

3.2. Auto-discovery

A link-state IGP such as IS-IS or OSPF can be used to simplify the configuration of MPLS rings. Details will be given in a companion document.

3.3. Signaling

Both RSVP-TE and LDP, with appropriate extensions, can be used to signal ring LSPs. Details will be given in companion documents.

3.4. Installing Primary LFIB Entries

In setting up RL_k, a node R_j sends out two labels: DL_jk to R_j-1 and UL_jk to R_j+1. R_j also receives two labels: DL_j+1,k from R_j+1, and UL_j-1,k from R_j-1. R_j can now set up the forwarding entries for RL_k. In the DS direction, R_j swaps incoming label DL_jk with DL_j+1,k with next hop R_j+1. In the US direction, R_j swaps incoming label UL_jk with UL_j-1,k with next hop R_j-1. R_k does not install LFIB entries in this manner.

3.5. Installing FRR LFIB Entries

At the same time that R_j sets up its DS and US LFIB entries, it can also set up the protection forwarding entries for RL_k. In the DS direction, R_j sets up an FRR LFIB entry to swap incoming label DL_jk with UL_j-1,k with next hop R_j-1. In the US direction, R_j sets up an FRR LFIB entry to swap incoming label UL_jk with DL_j+1,k with next hop R_j+1. Again, R_k does not install FRR LFIB entries in this manner.

3.6. Protection

Note that in this scheme, there are no protection LSPs as such -- no node or link bypasses, nor detours, nor LFA-type protection. Protection is via the "other" direction around the ring, which is why ring LSPs are in counter-rotating pairs.

If a node R_j detects a failure DS from R_j+1, it switches traffic on all DS ring LSPs to the US direction using the FRR LFIB entries. This switchover can be very fast, as the FRR LFIB entries can be preprogrammed. If the detection is fast too, then traffic loss is minimal.

R_j then sends an indication to R_j-1 that the DS direction is not working, so that R_j-1 can similarly switch traffic to the US direction. These indications propagate US until each traffic source on the ring uses the US direction. Thus, within a short period, traffic will be flowing in the optimal path, given that there is a failure on the ring. This contrasts with (say) bypass protection, where until the ingress recomputes a new path, traffic will be suboptimal.

4. Security Considerations

It is not anticipated that either the notion of MPLS rings or the extensions to various protocols to support them will cause new security loopholes. As this document is updated, this section will also be updated.

5. IANA Considerations

There are no requests to IANA for this document.

6. References

6.1. Normative References

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

6.2. Informative References

[RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[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.

[RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", RFC 5036, October 2007.

Author's Address

Kireeti Kompella
Juniper Networks, Inc.
1194 N. Mathilda Avenue
Sunnyvale, CA 94089
USA

Email: kireeti.kompella@gmail.com