Internet Engineering Task Force
Internet Draft
Expires: May 2001

Dimitry Haskin
Ram Krishnan
Axiowave Networks

November 2000

A Method for Setting an Alternative Label Switched Paths to Handle Fast Reroute

draft-haskin-mpls-fast-reroute-05.txt

Status

Abstract

This document describes a method for setting up an alternative label switched path to handle fast reroute of traffic upon a failure in a primary label switched path in Multi-protocol Label Switching (MPLS) network.

Internet Draft draft-haskin-mpls-fast-reroute-05.txt November 2000

Table of Contents

1. Introduction

The ability to quickly reroute traffic around a failure or congestion in a label switched path (LSP) can be important in mission critical MPLS networks. When an established label switched path becomes unusable (e.g. due to a physical link or switch failure) data may need to be re-routed over an alternative path. Such an alternative path can be established after a primary path failure is detected or, alternatively, it can be established beforehand in order to reduce the path switchover time.

Pre-established alternative paths are essential where packet loss due to an LSP failure is undesirable. Since it may take a significant time for a device on a label switched path to detect a distant link failure, it may continue sending packets along the primary path. As soon as such packets reach a switch that is aware of the failure, packets must be immediately rerouted by the switch to an alternative path away from the failure if loss of data is to

be avoided. Since it is impossible to predict where failure may occur along an LSP tunnel, it might involve complex computations and extensive signaling to establish alternative paths to protect the entire tunnel. In the extreme, to fully protect an LSP tunnel, alternative paths might be established at each intermediate switch along the primary LSP.

This document defines a method for setting alternative label switched paths with the objective to provide a single failure protection in such a manner that facilitates quick restoration comparable to 50 milliseconds provided in SONET self-healing rings and at the same time minimizes alternative path computation complexity and signaling requirements. It also can provide in-band means for quick detection of link and switch failures or congestion along a primary path without resorting to an out of band signaling mechanism. Both one-to-one (1:1) protection and many-to-one (1:N) protection can be achieved with the proposed approach as described in this document.

In order for the presented method to work, it is important that network topology and policy allow the establishment of a backup LSP between the endpoint switches of the protected LSP tunnel such that, with the exception of the tunnel endpoint switches, the backup LSP does not share any resources with the path that it intends to protect.

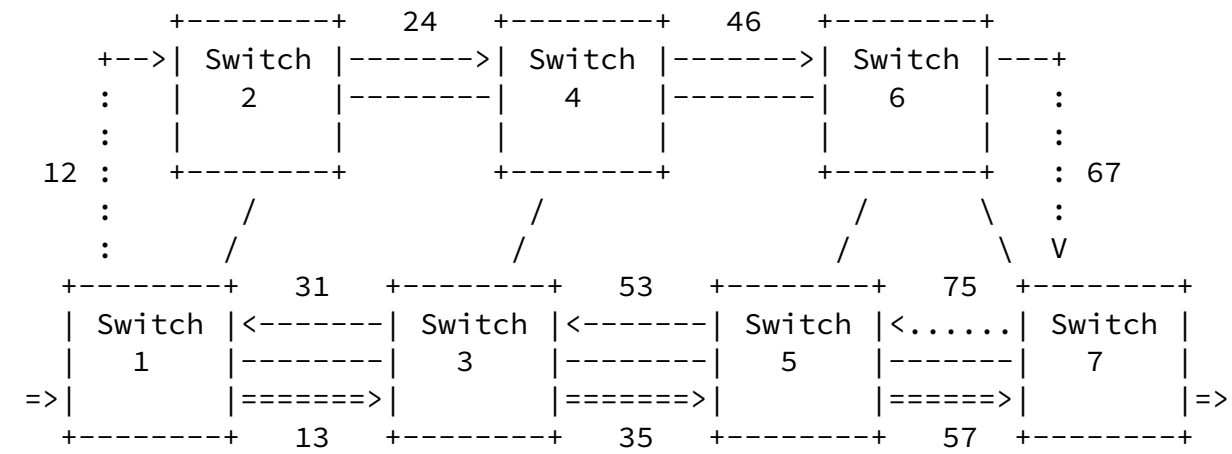
The fast reroute support can be facilitated with additional extensions incorporated in the MPLS signaling protocols such as RSVP or CR-LDP. These extensions are not defined in this document.

2. Alternative Path Arrangement

The main idea behind the presented method is to reverse traffic at the point of failure of the protected LSP back to the source switch of the protected LSP such that the traffic flow can be then redirected via a parallel LSP between source and destination switches of the protected LSP tunnel.

Referring to Figure 1, there is an MPLS network consisting of 7 interconnected switches.

Figure 1:



The following terminology is used for purpose of describing the method:

A portion of a label switched path that is to be protected by an alternative path is referred as 'protected path segment'. Only failures within the protected path segment, which may include the entire primary path, are subject to fast reroute to the alternative path. A primary LSP between switches 1 and 7 is shown by a double-dashed links labeled 13, 35, and 57. Arrows indicate direction of the data traffic.

The switch at the ingress endpoint of the protected path segment is referred as 'the source switch'. Switch 1 in Figure 1 is the source switch in our example of a protected path.

The switch at the egress endpoint of the protected path segment is referred as 'the destination switch'. Switch 7 in Figure 1 is the destination switch in our example of a protected path.

The switches between the source switch and the destination switch along the protected path are referred as protected switches.

The switch immediately preceding the destination switch along the protected path segment is referred as the last hop switch. Switch 5 in Figure 1 is the last hop switch for the protected path.

The essence of the presented method is that an alternative

unidirectional label switched path is established in the following way:

The initial segment of the alternative LSP runs between the last hop switch and the source switch in the reverse direction of the protected path traversing through every protected switch between the last hop switch and the source switch. The dashed line between switches 5 and 1 illustrates such a segment of the alternative path. Alternatively, the initial LSP segment can be set from the destination switch to the source switch in the reverse direction of the protected path traversing through every protected switch between the destination switch and the source switch. The dashed line between switches 7 and 1 illustrates the initial path segment that is set in this way.

The second and final segment of the alternative path is set between the source switch and the destination switch along a transmission path that does not utilize any protected switches. It is not an intention of this document to specify procedures for calculating such a path. The dashed line between Switches 1 and 7 through Switches 2, 4, and 6 illustrates the final segment of the alternative path.

The initial and final segments of the alternative path are linked to form an entire alternative path from the last hop switch to the destination switch. In Figure 1 the entire alternative path consists of the LSP links labeled 53, 31, 12, 24, 46, and 67 if the alternative path originates at the last hop switch. Alternatively, the entire alternative path consists of the LSP links labeled 75, 53, 31, 12, 24, 46, and 67 if the alternative path originates at the destination switch of the primary path.

As soon as a failure along the protected path is detected, an operational switch at the ingress of the failed link reroutes incoming traffic around the failure or congestion by redirecting this traffic into the alternative LSP traversing the switch in the reverse direction of the primary LSP according to the procedures described in the following sections of the document.

The presented method of setting the alternative label switched path has the following benefits:

- Path computation complexity is greatly reduced. Only a single additional path between the source and destination switches of the protected path segment needs to be calculated. Moreover, both primary and alternative path computations can be localized at a single switch avoiding problems that can arise when computations are distributed among multiple switches.

- The amount of LSP setup signaling is minimized. With small extensions to RSVP or LDP (described in separated documents), a single switch at ingress of the protected path can initiate label allocations for both primary and alternative paths.

- Optionally, presence of traffic on the alternative path segment that runs in the reverse direction of the primary path can be used as an indication of a failure or congestion of a downstream link along the primary path. As soon as the source switch detects the reverse traffic flow, it may stop sending traffic downstream of the primary path and start sending data traffic directly along the final alternative path segment.

  It is fair to note that this technique increases the likelihood of data packet reordering during the path rerouting process. Therefore benefits of the reducing the alternative path latency should be weighed against possible problems associated with short term packet reordering. On a positive side, if multiple microflows are aggregated in a single protected LSP tunnel, only a very limited number of microflows may be affected by such packet reordering. Additionally, the impact of reordering on any single microflow may be minimal.

The described in-band signaling of an LSP failure to the source switch does not exclude other methods of propagating an error condition back to the source.

It also can be noted that if the alternative label switched path is originated at the destination switch of the primary path, it forms a 'loop-back' LSP that originates and terminates at this switch. Therefore in this case it is possible to verify integrity of the entire alternative path by simply sending a probe packet from the destination switch along the alternative path and asserting that the packet arrives back to the destination switch. When this technique is used to assert the path integrity, the care must be taken that the limited diagnostic traffic is not interpreted as an indication of a primary path failure that triggers data rerouting at the source switch.

3. 1:1 protection

If the 1:1 path protection is desired, an individual backup LSP is set for each LSP that needs to be protected as described in section 2. When a switch detects that a downstream link has failed, it simply splices the traffic onto the alternative LSP. Referring to Figure 1, if the link between the Switch 3 and Switch 5 fails, Switch 3 accomplishes the fast reroute by swapping the incoming MPLS label 13 of the primary path with the outgoing MPLS label 31 of the alternative path. In this example the primary and alternative paths are linked at Switch 3 forming the following label switched path for the traffic flow: 13->31->12->24->46->67.

4. 1:N protection

In the case of the 1:N protection a single alternative path can be used for protection of more than one LSP between the same source and destination switches. The difference in rerouting LSPs the 1:N protection case is that, rather than splicing protected traffic into the alternative LSP, it may be necessary to use the MPLS label stacking to tunnel protected traffic via the backup LSP to the destination switch as described below.

A switch detecting failure of a downstream link, first swaps the incoming MPLS label of each protected LSP with the respective incoming label that identifies that LSP at the destination switch and then pushes the outgoing label of the backup LSP to the top of the forwarded MPLS packets. In essence, the protected MPLS packets are encapsulated inside of the backup LSP and emerge at the backup tunnel tail at the egress switch with their respective labels known to that switch.

Referring to Figure 1 and assuming that global label space is used at the destination switch, if the link between the Switch 3 and Switch 5 fails, Switch 3 swaps incoming MPLS label 13 of the protected LSP with label 57 (incoming label at Switch 7) and then encapsulates the resulting packet into the backup tunnel by pushing label 31 to the top of the forwarded MPLS packets.

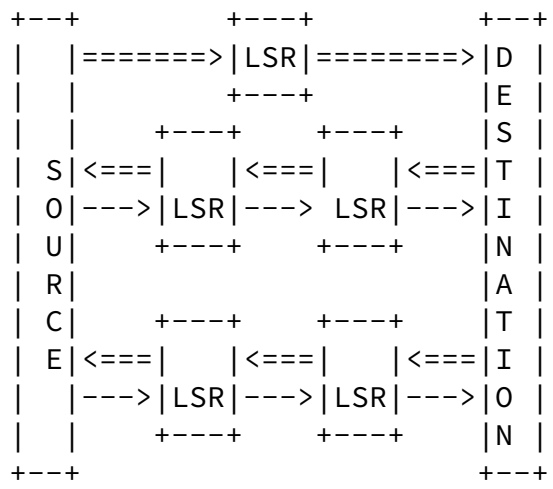
Needless to say in order for this scheme to work, each router in the protected path must be aware what labels are used at the egress LSR for each protected LSP. Such knowledge can be propagated with the appropriate extensions incorporated into signaling protocols such as

RSVP or CR-LDP.

A single segment of a tunnel between source and destination switches can be used to protect multiple LSP segments that originate and terminate on these switches as long as this segment of the backup tunnel is completely disjoint from each protected LSP segment except for the source and destination switches. In such a case the reverse segments of backup path merge into the disjoint segment of the backup path at the source switch of the protected LSPs as

illustrated in Figure 2. In Figure 2, dashed lines represent protected LSPs and double-dashed lines represent backup LSP tunnels.

Figure 2:



5. Restoration Shortcuts

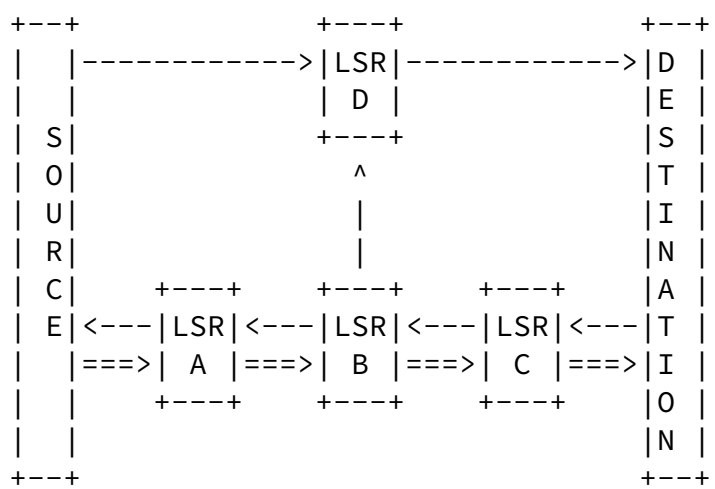
Some types of applications require bounded end-to-end transmission delays to deliver useful services. A notable example is the Voice over IP (VoIP) service which requires end-to-end delays that do not exceed 400 ms for an acceptable level of service. VoIP is also a prime candidate for the fast reroute services. Since most of the voice codecs in use today operate in the range of 20-50 ms latency, the network component is left with around 300 ms of the end-to-end delay limit.

Given the above considerations, it is important that, when restoration provisions are made for a delay sensitive service,

transmission delays over an alternative path would not exceed an acceptable limit. Since a number of the current network providers are capable to guarantee network transport delay that do not exceed 80 ms on their backbone, it appears that in some cases it will be possible to use the proposed restoration technique with a single alternative path. It allows for at most 200 ms round trip delay over a reverse path segment plus at most 100 ms delay over a disjoint backup path segment. However in other cases it may be necessary to introduce restoration shortcuts as described below to satisfy the VoIP latency requirement during restoration.

Restoration shortcuts are achieved by allowing selected transit routers in the primary LSP to establish one or more 'shortcut' alternative LSPs to the egress router as illustrated in Figure 3. In this illustration, primary link failures that may occur downstream of LSR B are rerouted over the shortcut LSP from LSR B to the destination of LSP being backed up. In illustrated example the shortcut LSP merges into the backup LSP at LSR D.

Figure 3:



6. Elementary link level protection scheme

If only link-level protection is desired, an alternative path between link endpoints can be set up to protect each link. Such a scheme can be viewed as a degenerate case of this proposal in which the link endpoints constitute the source and destination endpoints

in the described approach.

7. Bandwidth Reservation Considerations

Generally there is no need to exclusively allocate bandwidth resources to the alternate LSP. The holding priority of the primary LSP can be used as traffic-triggered resource preemption priority for the alternate LSP in case the primary LSP fails and traffic is switched to the alternate LSP as described in this document. What we call here the traffic-triggered priority is the preemption priority assigned to an LSP that is utilized only when there is traffic present on that LSP. When there is no traffic, other LSPs sharing the interface should get full access to bandwidth and other system resources. Consequently, if the traffic-triggered priority of the alternative LSP is greater than the holding priorities of the other LSPs using an interface in the alternate path, the alternate LSP can preempt bandwidth and other system resources as soon as traffic gets rerouted via the alternate LSP. This enables high-priority LSPs, which are being rerouted, to preempt resources from lower priority LSPs without explicit bandwidth reservation for the alternate path. Of course, if bandwidth efficiency is not an issue, bandwidth resources can be explicitly reserved for the alternate LSP also.

An extension to existing signaling protocols such as RSVP and LDP may be needed to indicate that traffic-triggered resource preemption is requested for a particular LSP as opposed to the setup priority preemption.

---

8. Intellectual Property Considerations

IETF has been informed of possible intellectual property protection for some or all of the technologies disclosed in this document.

9. Acknowledgments

This document has benefited from discussions with Jim Boyle, Robert Boyd, and Alan Hannan. We also thank Ken Schroder, Jeff Parker and Yanhe Fan for their comments on the document.

10. References

[1] Rosen, E. et al., "Multiprotocol Label Switching Architecture", Internet Draft, draft-ietf-mpls-arch-07.txt, July 2000.

[2] Awduche, D. et al., "Requirements for Traffic Engineering over MPLS", RFC-2702.

11. Authors' Addresses

Dimitry Haskin
Axiowave Networks, Inc.
100 Nickerson Road
Marlborough, MA 01752
E-mail: dhaskin@axiowave.com

Ram Krishnan
Axiowave Networks, Inc.
100 Nickerson Road
Marlborough, MA 01752
E-mail: ram@axiowave.com