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

draft-haskin-mpls-fast-reroute-01.txt

Abstract

This document describes a method for setting an alternative label switched path to handle fast data packet reroute upon a failure in a primary label switched path in Multi-protocol Label Switching (MPLS) network.
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 rerouted 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 in such a matter that 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.

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 links or switches with the path that it intends to protect.

2. Alternative Path Arrangement

The main idea behind the presented method is to reverse traffic at the point 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:

```
+--------+   24   +--------+   46   +--------+
| Switch |------->| Switch |------->| Switch |
:  2     |--------|  4     |--------|  6     |
:  12 :  ++---++  ++---++  ++---++ :  67
:     /       /       / \  
:    /       /       /   V
+--------+   31   +--------+   53   +--------+
| Switch |<------| Switch |<------| Switch |<.....| Switch |
|  1     |------->|  3     |------->|  5     |------->|  7     |
| ==><| | ==><| | ==><| | ==><|
+--------+   13   +--------+   35   +--------+   57   +--------+
```

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 at its extreme 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 link failure or congestion along the protected path is detected an operational switch at ingress of failed link reroutes incoming traffic around of the failure or congestion by linking upstream portion of the primary path to the downstream portion of the alternative path. Thus if the link between Switches 3 and 5 fails, 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.

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.

- 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 a switch along the primary path sees the reverse traffic flow, it may stop sending traffic downstream
of the primary path by initiating an immediate rerouting of data traffic to the alternative path. As the result of this "crank back" process the source switch may start sending data traffic directly along the final alternative path segment. It is fair to note that the crank-back 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 tend to be minimal.

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 has to be taken that the limited diagnostic traffic is not interpreted as an indication of a primary path failure that triggers data rerouting at the intermediate switches.

3. 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.

4. Bandwidth Reservation Considerations

Generally there is no need to specifically 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. 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.

5. Intellectual Property Considerations

Nexabit Networks may seek patent or other intellectual property protection for some or all of the technologies disclosed in this document. In the event that Nexabit Networks obtains such patent rights, Nexabit Networks intends to license them on reasonable and non-discriminatory terms in accordance with the intellectual property rights procedures of the IETF standards process.

6. Acknowledgments

This document has benefited from discussions with Jim Boyle, Robert Boyd, and Alan Hannan.

7. References


7. Authors’ Addresses

Dimitry Haskin
Nexabit Networks, Inc.
200 Nickerson Road
Marlborough, MA 01752
E-mail: dhaskin@nexabit.com

Ram Krishnan
Nexabit Networks, Inc.
200 Nickerson Road
Marlborough, MA 01752
E-mail: ram@nexabit.com