MPLS Part II - Recovery
Outline

• Introduction
• MPLS Recovery Framework
• MPLS Mechanism for Protection/Restoration
  – Shared Backup LSP Restoration
  – Fast reroute
  – RSVP-TE Recovery
• A Heuristic Restoration Approach
  – Analysis
  – Simulations
  – Comparison
Current Backbone Networks Protection

• **Link layer protection** *(SONET/SDH)*
  – capable of service restoration within few tens of milliseconds
  – The scope of the protection is limited, has no visibility into higher layer operations

• **Layer 3 protection**
  – Routing protocol provide much greater flexibility
  – restoration time in the order of seconds to minutes
Motivation: MPLS-Based Recovery

- MPLS is the lowest layer with the knowledge of the entire network topology
- MPLS provide necessary traffic engineering capabilities
- MPLS has desirable attributes when applied to the purpose of recovery for connectionless networks
- MPLS provide restoration times significantly shorter than the convergence times of IP routing protocols
1+1 (Dedicated) Protection

Before

splitter

1x2 optical switch

After
Link (vs Path) Protection

• Use an alternate detour if a link failed, or
• use a disjoint backup path (path protection)
Mesh Network Protection
Non-Dedicated Protection

- 1:1 Protection
- 1:N Protection
Path Mapping

• 1-to-1 Protection
  recovery path that is only to be used to recover that specific working path
  the recovery path can carry pre-emptable/low-priority traffic

• 1-to-n Protection
  n working paths are protected using only one recovery path
  As to backup bandwidth sharing, as long as the n working paths are disjoint (and hence at most one can fail at a time), their protection paths can share backup bandwidth on common links.
LSP1: E-F-G
LSP1 Backup: E-C-D-G
LSP2: A-B
LSP2 Backup: A-C-D-B
Restoration vs Protection

• Restoration
  
  the recovery path or path segments are created dynamically after the detection of a fault on the working path. (the recovery path is not pre-computed)

• Protection Switching
  
  In contrast to restoration, the recovery path is pre-computed (but may not be pre-established).
Advantages & Disadvantages of Protection

• Simple and Quick: especially if it uses 1+1
• Do not require much extra process time, except to signal (set up) the switches along the pre-determined backup path (1:1 or 1:N)
• Usually can only recover from single link fault (what if the pre-computed path fails?)
• Inefficient usage of resource
Restoration

- Path Restoration
  - Route can be computed after failure

- Link Restoration
  - Path is discovered at the end nodes of the failed link
  - More practical than path restoration

- Advantages & Disadvantages of Restoration
  - Usually can recover from multiplex element faults
  - More efficient usage of resource
  - Complex
  - Slow: require extra process time to setup path and reserve resource
Comparison between Protection & Restoration

• Characteristic: Protection -- the resource are reserved before the failure, they may be not used; Restoration -- the resource are reserved and used after the failure

• Route: Protection -- predetermined; Restoration -- can be dynamically computed

• Resource Efficiency: Protection -- Low; Restoration -- High
Comparison between Protection & Restoration (Cont’)

• Time used: Protection -- Short; Restoration -- Long

• Reliability: Protection -- mainly for single fault; Restoration -- can survive under multiplex faults

• Implementation: Protection -- Simple; Restoration -- Complex
Definitions and Terminology

- **Path Switch LSR (PSL)**
  - The PSL is responsible for switching or replicating the traffic between the working path and the recovery path.
  - Normally chosen as the Ingress LSR or the nearest upstream LSR to the failure (link or node).

- **Path Merge LSR (PML)**
  - The PML is responsible for receiving the recovery path traffic, and either merges the traffic back onto the working path, or, if it is itself the destination, passes the traffic on to the higher layer protocols.
  - Normally chosen as the Egress LSR.
Definitions and Terminology

• **Fault Indication Signal (FIS)**
  A signal that indicates that a fault along a path has occurred. It is relayed by each intermediate LSR to its upstream or downstream neighbor, until it reaches the PSL.

• **Fault Recovery Signal (FIS)**
  A signal that indicates a fault along a working path has been repaired. Like the FIS, it is relayed by each intermediate LSR to its upstream or downstream neighbor, until it reaches the PSL.
Fault Detection

- **Link/Path Failure**: detected by a link probing mechanism (hello liveness message) between neighbor LSRs. Path failures can also be detected by Src/Dest (time-out, NAK etc) or by the Dest if 1+1 path protection is used.

- **Path Degraded**: path has connectivity, but that the quality of the connection is unacceptable (e.g., high error bit rate, label mismatch or due to TTL errors). Need performance monitoring mechanisms.

- **Link Degraded**: the link over which the working path is carried is performing below an acceptable level.
Scope of Recovery

• Local Repair
  – protect against a link or neighbor node fault and to minimize the amount of time required for failure propagation
  – Fast but not optimized

• Global Repair
  – the PSL is usually distant from the failure and needs to be notified by a FIS
  – the recovery path can be made completely link and node disjoint with its working path
  – slower than local repair
Post Recovery Operation

- When traffic is flowing on the recovery path, and the failure on the working path is repaired, one can consider the recovery path as a new working path (Non-Revertive Mode).
- Switch to the old working path (Revertive Mode).
- Switch to a more preferred working path.

(make before break ----- RSVP TE Recovery)
Outline

• Introduction
• MPLS Recovery Framework
• MPLS Mechanism for Protection/Restoration
  – Shared Backup LSP Restoration
  – Fast reroute
  – RSVP-TE Recovery
• A Heuristic Restoration Approach
  – Analysis
  – Simulations
  – Comparison
MPLS-based Recovery Principles

- **Configuration of Recovery**
  - Default-recovery (No MPLS-based recovery enabled)
  - Recoverable (MPLS-based recovery enabled)

- **Initiation of Path Setup**
  - Pre-established
  - Pre Qualified
  - Established-on-Demand

- **Initiation of Resource Allocation**
  - Pre-reserved
  - Reserved-on-Demand
MPLS Recovery Cycle Model

- Network Impairment
- Fault Detected
  - Start of Notification
  - Start of Recovery Operation
- Recovery Operation Complete
  - Path Traffic Restored

T1  Fault Detection Time  T2  Hold-off Time
T3  Notification Time    T4  Recovery Operation Time
T5  Traffic Restoration Time
Main Comparison Criteria

- **Recovery Time**
  the time required for a recovery path to be activated (and traffic flowing) after a fault

- **Loss**
  Recovery schemes may introduce a certain amount of packet loss during switchover to a recovery path.

- **Backup Capacity**: corresponding to (Quality of Protection)
  The capacity will be dependent on the traffic characteristics of the network, the particular protection plan selection algorithms as well as the signaling and re-routing methods.

- **Re-ordering**
  the action of putting traffic back on preferred paths might cause packet re-ordering

- **State Overhead**
  As the number of recovery paths in a protection plan grows, the state required to maintain them also grows
Outline

• Introduction

• MPLS Recovery Framework

• MPLS Mechanism for Protection/Restoration
  – Shared Backup LSP Restoration
  – Fast reroute
  – RSVP-TE Recovery

• A Heuristic Restoration Approach
  – Analysis
  – Simulations
  – Comparison
Shared Backup LSP Restoration

- links on the backup path can be shared between different active paths
- Single link failure restoration is guaranteed
- using only aggregate network usage information without requiring per-LSP routing information for all current active LSPs
  - Aggregate information is obtainable by adding a few new information elements to the link state advertisement of a link state routing protocol like OSPF or ISIS
Model

LSP1: E-F-G
LSP1 Backup: E-C-D-G

LSP2: A-B
LSP2 Backup: A-C-D-B
Algorithm With Aggregated Partial Information

For link (i,j)

• the cumulative bandwidth allocated for active paths is $A(i,j)$
• the cumulative bandwidth allocated for backup paths is $B(i,j)$
• the residual bandwidth free for allocation is $R(i,j)$
Estimate the Additional Amt of Backup Bandwidth Needed (or link cost)

- For a request of bandwidth $b$ the active path is calculated as the shortest path on the topology of links that have $R(i,j) > b$
- Let $M$ be the max of the $A$ values along the active path
- The backup path is calculated as follows. The cost of a link $(u,v)$ is now taken as:
  - $0$ if $M+b < B(u,v)$
  - $b$ if $B(u,v) \leq M$ and $b \leq R(u,v)$
  - $M+b - B(u,v)$ if $M \leq B(u,v)$ and $M+b \leq B(u,v)+R(u,v)$
  - infinity in all other cases
- The backup path is calculated as the shortest path on the topology with the cost of links calculated as above
Fast Reroute

• Main Idea:
  reverse traffic at the point of failure of the protected LSP back to the source switch, where the traffic flow can be redirected

• Objective:
  – provide a single failure protection with quick restoration comparable to the order of milliseconds
  – Minimize the alternative path computation complexity and signaling requirements
  – 1:1 protection and 1:N protection can be achieved
Fast Reroute Model

1:1 protection

From LSR 1 to LSR 7       LSR1: PSL ,       LSR7 : PML
Fast Reroute Model

1:N protection  (Using Label Stack)
Restoration Shortcuts

- LSR B setup a shortcut alternative LSP
- Applied for Voice over IP service
Pros. and Cons.

• Path computation complexity is greatly reduced
  – both primary and alternative path computations can be
    localized at a single switch
• The amount of LSP setup signaling is minimized
• presence of traffic on the alternative segment path
  can be used as an FIS of the downstream primary
  path
• Data packets need reordering during the path
  rerouting process
• Less resource efficient (total length protection)
RSVP Detour

- To achieve timely detour path setup, using pre-computed and pre-established detour path is essential for data traffic where packet loss is undesirable.
- Detour decision must be made as close to the failure point as possible.
- Ideal detour mechanism is to protect the entire LSP by establishing detour paths throughout the LSP.
- To minimize the path computation overhead, it is desirable for the detour paths to merge back to the main LSP as soon as possible.
- Only protect unidirectional LSP.
RSVP Detour Model

Activate RSVP Tunnel (LSP)
RSVP Detour (LSP)
detour (Ingress, LSR2)

Ingress  LSR 1  LSR 2  LSR 3  Egress
RSVP Extension

Two new objects are defined to support LSP fast-reroute

- **FAST_REROUTE** object
  - Setup (Holding) Priority: The priority of the detour with respect to taking(holding) resources
  - Hop – limit: The maximum number of extra hops the detour is allowed to take
  - Bandwidth

- **DETOUR** object
  - Source ID: IPv4 address identifying the beginning point of detour
  - Downstream Node ID: IP address identifying the downstream node that source is trying to avoid
Make before break

• In general, it is highly desirable not to disrupt traffic, or adversely impact network operations while TE tunnel rerouting is in progress.
• This adaptive and smooth rerouting requirement necessitates establishing a new LSP tunnel and transferring traffic from the old LSP tunnel onto it before tearing down the old LSP tunnel.
• The principle (implemented by RSVP Tunnel) applies not only in the case of failure but also when better routes are available than the existing ones.
Outline

• Introduction
• MPLS Recovery Framework
• MPLS Mechanism for Protection/Restoration
  – Shared Backup LSP Restoration
  – Fast reroute
  – RSVP-TE Recovery
• A Heuristic Restoration Approach
  – Analysis
  – Simulations
  – Comparison
Analysis

Basic Idea

• Providing fault tolerance in MPLS networks based on the concept of “domain protection”

• In that domain, protection paths for all working paths that terminate in an egress router are calculated simultaneously

• The algorithm attempts to locate two trees in the domain such that no single link failure would disconnect a node from the root of the tree (the egress node)

• Permit the decoupling the protection path placement from the working path (much greater flexibility)
Analysis

Conditions

• Assume MPLS domain represented by graph $G(N, L)$
  – $N$ is the set of $n$ nodes
  – $L$ is the set of $l$ links between the nodes
• Assume that graph $G$ is two-edge redundant and therefore can be protected against any single link failure
Analysis

• Input : The MPLS domain D and the egress router e
• Output : two collections of protection path connecting ingress routers to egress router e
• Initialization:
  – Find a spanning tree of graph G rooted in the egress router e.
  – Let \( P \) be the set of nodes for which the protection paths have been established
  – Initially it contains the egress router: \( p = \{ e \} \)
Heuristic Algorithm

Repeat until all nodes are protected (P=N):

1. Select one of the branches of the spanning tree attached to the egress node and mark all its nodes except for the egress node

2. Scan all marked nodes to find node i that has a link to an unmarked node j

3. Find a ring path consisting of the links of the spanning tree leading from e to i, the link between i and j, and the links of the spanning tree between j and e (note that if j = e, this segment of the ring is empty)
4. Place two ring paths along the ring:
   1) one in clockwise, the other in counterclockwise direction
   2) Merge the created protection paths with protection paths established in the previous iterations of the algorithm for the protected nodes that are now a part of the egress node
   3) All nodes on the ring are added to $p$

5. Consider a new graph constructed by treating all nodes in $P$ as a single node that will act as the egress node

6. Begin the subsequent iteration
1. Find a branch
2. Find node C has a link with an unmarked node D
3. construct the ring
4. All nodes on the ring are now a part of the egress node
5. Next iteration : Node F assume both node C and D as egress node, and setup the protection path respectively
Quality of a protection Scheme

Finding optimal protection path

• The length of the protection path
  – Indicate the delay of the traffic after a link failure
  – Reflect the amount of the resources required to protect the domain

• The number of protection paths pre link
  – Indicate the amount of resources (eg. Label table size, signalling overhead need to maintain the protection)
  – Indicate that how well the protection paths are distributed throughout the network
Heuristic Decisions

• Spanning tree selection
  – The protection paths are routed mostly along the branched of the spanning tree
  – Try to find up a spanning tree with many short branches
  – Using the Dijkstra’s algorithm to calculate the shortest path spanning tree

• Finding the smallest possible ring
  – Reduce the maximum and the average protection path lengths
Heuristic Decisions Example

- Find the short branch A-B-C (step 1)
- Comparison the length of the ring (step 2 & 3)
  - A-B-C-F selected
Simulations

Maximum length of protection paths for meshes

Average length of protection paths for meshes
Simulations

Number of protection paths per link for meshes

Avg. number of protection paths per link for meshes
## Comparison of MPLS Protection Schemes

<table>
<thead>
<tr>
<th>Options Schemes</th>
<th>Fast Reroute</th>
<th>Backup Tunnels</th>
<th>Heuristic Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complexity</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Path Placement</td>
<td>Restoration</td>
<td>Restoration</td>
<td>Flexible</td>
</tr>
<tr>
<td>Restoration Time</td>
<td>Fast</td>
<td>Slow</td>
<td>Fast</td>
</tr>
<tr>
<td>Packet Loss</td>
<td>Min.</td>
<td>Yes</td>
<td>Min</td>
</tr>
<tr>
<td>Packet Reordering</td>
<td>Yes</td>
<td>Min</td>
<td>Min</td>
</tr>
<tr>
<td>Restoration Requirement</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
</tbody>
</table>
Conclusion

- The scheme considers protection of all paths leading from ingress routers to a common egress router.
- Protection using two paths allow for greater flexibility in protection path placement.
- It provides the protection is better than Fast Reroute scheme.
- The algorithmic complexity is less than that of RSVP Backup Tunnel while providing comparably good protection.
- Unlike the FR and BT schemes, it guarantees independence of the working and protection path placement.
MPLS Recovery Goals

- Using traffic engineering to optimal use of resources
- Aim to facilitate restoration times that are sufficiently fast for the end user application
- Aim to maximize network reliability and availability
- Aim to be applicable for protection of traffic at various granularities
  - for a portion of the traffic on an individual path
  - for all traffic on an individual path
  - for all traffic on a group of paths
- Be applicable for segments or an entire end-to-end path
References

- Extensions to RSVP-TE for MPLS Path Protection Ken Owens, Vishal Sharma <Draft-chang-mpls-path-protection-02>, July 2001
- “A heuristic approach to service restoration in MPLS networks,” in proc. Of the 2001 IEEE International Conference on Communications (ICC), Helsinki, Finland, June 2001