MPLS Traffic Engineering
Fast Reroute
Agenda

• Introduction

• Terminology of protection/restoration
  • Protection versus Restoration,
  • Protection: 1+1, M:N, 1:1,
  • Restoration,
  • Time to restore and Time to switch back,
  • Local versus Global repair,
  • Usage of protected Path,
  • Revertive versus Non revertive mode
Agenda

- **MPLS Traffic Engineering**
  - Global restoration: TE LSP rerouting
  - Global protection: Path protection
  - Local protection:
    - Link protection
    - Node protection
  - Backup path computation and provisioning

- **IETF Update**

- **Conclusion**
Agenda

• Introduction
• Terminology of protection/restoration
• MPLS Traffic Engineering Fast Reroute
• IETF Update
• Conclusion
Protection/Restoration in IP/MPLS networks

As IP/MPLS networks carry a very large amount of critical IP traffic (MPLS VPN, VoIP, …)

Protection/Restoration is a key component of the overall architecture just as Routing, QOS, …
Protection/restoration in IP/MPLS networks

- Many various protection/restoration schemes (co)exist today:
  - Optical protection
  - Sonet/SDH
  - IP
  - MPLS Traffic Engineering Fast Reroute

- The objective being to avoid double protection
Protection/Restoration in IP/MPLS networks

- IP routing protocol typically offers a convergence on the order of seconds (default=40s with OSPF, 30s with ISIS)
- IP restoration is **Robust** and protects against link **AND** node protection
- IP convergence may be dramatically improve and could easily offers a few seconds convergence (1, 2, 3 secs ?) using various enhancements:
  - fast fault detection,
  - fast SPF and LSA propagation triggering,
  - priority flooding,
  - Incremental Dijsktra,
  - Load Balancing
Protection/Restoration in IP/MPLS networks

- 1-3 secs may be sufficient for some traffic as others (ex: voice trunking) will require more aggressive target, typically 50 ms.

- Solutions?
  - Optical protection,
  - Sonet/SDH (GR 253)
  - MPLS protection/restoration
Protection/Restoration in IP/MPLS networks

MPLS Traffic Engineering Protection/Restoration

• Compared to lower layers mechanism, MPLS offers:
  • A protection against link AND node failures
  • A much better bandwidth usage
  • Finer granularity. Different level of protection may be applied to various classes of traffic.
    • Ex: an LSP carrying VoIP traffic will require a 50ms protection scheme as Internet traffic may rely on IP convergence
  • A much cost effective protection mechanism
Agenda

- Introduction
- **Terminology of protection/restoration**
- MPLS Traffic Engineering Fast Reroute
- IETF Update
- Conclusion
Protection/Restoration in GMPLS networks

Terminology

• **Protection**: a back-up path is pre-established to be used as soon as the failure has been detected.

• **Restoration**: set of mechanisms by which a new path is being dynamically calculated as soon as the failure has been detected and propagated.

• Protection is faster, requires more spare resources but provides stronger guarantees.

• Protection may be combined with Restoration.
Protection/Restoration in IP/MPLS networks

Protection

- **1+1** the traffic is being duplicated on the protected path (constantly bridged).

- The Path switch LSR performs the switching or replication of the traffic between the working and recovery path.

- The Path Merge LSR receives both the working and recovery path traffics and performs the selection.

- Switching is performed at the tail-end which does not require sophisticated signalling (also called single ended protocol)
Protection/Restoration in IP/MPLS networks

Example: 1+1 protection in Point to point DWDM system (similar protection scheme exists in Optical mesh network)

LoS
OR
LoModulation/Clock
OR
Signal degrade
=
Client Tx switch OFF

1+1 protection does not exist in IP/MPLS
Protection/Restoration in IP/MPLS networks

Protection (cont)

• **M:N**: M protected paths for N working paths are signalled but may be used for low priority traffic which makes a more efficient use of the spare resource.

• When a failure occurs, the protected path is requested and low priority traffic is preempted.

• Ex: **1:1**, 1 protected path being established for every working path
Protection/Restoration in IP/MPLS networks

Restoration

- Once the failure has been detected, propagated and signalled, a new path/route is dynamically calculated

- A well known example is IP

  The failure is detected (through the layer 2 protocol or IGP hellos)

  The failure is propagated (through the LSP flooding)

  A new route is dynamically calculated (SPF) and the routing table is updated
Protection/Restoration in IP/MPLS networks

Examples

• **Protection**
  - 1+1 Optical protection (single ended protocol)
  - Sonet/SDH BLSR and UPSR  \(\xrightarrow{\text{10-50 msecs}}\)
  - MPLS Fast Reroute (link and node protection)
  - MPLS TE Path protection, \(\Theta(100 \text{ (s) of msecs})\)

• **Restoration**
  - IP routing protocol \(\xleftarrow{\text{2 – 40 seconds}}\)
  - MPLS TE LSP reroute \(\Theta(s)\)
Protection/Restoration in IP/MPLS networks

TTR: Time to Restore (convergence time)

- TTR = time between the fault and traffic recovery

  **Fault detection** may differ from the lower layers

  **Hold-off timer.** Waiting time to let lower layers protection mechanisms (if any) to operate. May be 0

  **Fault localization**

  **Fault notification.** May be a non negligible factor, the propagation delay may be relatively high even compared to the path calculation in Restoration techniques.

  **Fault restoration.** Time once the fault has been detected, localized and notified for the LSR in charge of rerouting the traffic to actually reroute the traffic (also called switch over)
Protection/Restoration in IP/MPLS networks

TTS (Time to switch back)

- **TTS** = once the fault has been cleared, time to switch back to the previous working path.
  
  **Fault clearing time.** Time to detect the fault has been cleared. Highly depends on lowers layers.
  
  **Wait to Restore timer.** Allows not to switch back immediately to improve stability in case of flapping. A back-off mechanism may also be used there.
  
  **Traffic restoration time**
Protection/Restoration in IP/MPLS networks

Scope of recovery: **local repair** versus **global repair**

- **Local (link/node) repair**: the recovery is being performed by the node immediately upstream to the failure

  Protection (most of the time): the protected (back-up) path is pre-established and diversely routed from the working path

  Restoration: the back-up path is dynamically established around the failure network component (link or node)

- **Example**

  MPLS local repair FRR (link/node protection)
Protection/Restoration in IP/MPLS networks

- **Global repair**: the recovery is being performed by the head-end (where the LSP is initiated)
  
  Both restoration and protection may be used.
  
  The head-end needs a notification also called FIS (Fault indication signal).

- Then, the head-end may use restoration to reroute the traffic or protection to reroute the traffic onto a pre established protected path
Protection/Restoration in IP/MPLS networks

- Slower than local repair (propagation delay of the FIS may be a non negligible component)

- Examples of global repair mechanisms
  
  IP is a global repair mechanism using restoration. TTR is typically $\Theta(s)$
  
  MPLS TE Path protection is a global repair mechanism that may use
  
  Protection: the protected TE LSP is pre signalled
  
  Restoration: the protected TE LSP is dynamically established
Protection/Restoration in IP/MPLS networks

- **Path mapping**: refers to the method of mapping traffic from the faulty working path onto the protected path (1:1, M:N)

- **QOS of the protected path**: does the protected path offer an equivalent QOS as the working path during failure?

- **Recovery granularity**: from a portion of one working path to a bundle of working path.
Protection/Restoration in IP/MPLS networks

• **Usage of the protected path**

  Dedicated 1+1: the back-up LSP (protected) cannot be used for low priority traffic

  Dedicated 1:1 and shared M:N. The back-up path may be used for low priority traffic.
Protection/Restoration in IP/MPLS networks

Switch back operation

• Revertive versus non revertive

In **revertive mode**, once the failure is cleared the working path is being automatically re established (always preferred to the protected path)

In **non revertive mode**, when the faulty path is restored, it may become the recovery path.
Protection/Restoration in IP/MPLS networks

Performance

• The recovery class may or not be equivalent

• IP offer a lower class, MPLS TE may offer an equivalent or lower class
Protection/Restoration in IP/MPLS networks

A few comparison criterias

- Recovery time
- Setup vulnerability
- Back-up capacity
- Additive latency
- Protection QOS
- Re-ordering
- State overhead
- Loss

- Coverage (link/node, concurrent faults, % of coverage, number of recovery paths, number of protected paths, …)
Agenda

- Introduction
- Terminology of protection/restoration
- MPLS Traffic Engineering Fast Reroute
- IETF Update
- Conclusion
Terminology

- **Reroutable LSP**: TE LSP for which a local protection is desired
- **Protected LSP**: an LSP is being protected at a HOP H if and only if it does have a backup tunnel associated at hop H.
- **Primary LSP**: a protected LSP prior to any failure
- **PLR**: Point of local repair (head-end of the backup tunnel)
- **Backup tunnel/LSP**: TE LSP used to backup the protected LSP
Terminology

Terminology (cont)

- Merge point: Tail-end of the backup tunnel
- NHOP backup tunnel: a Backup Tunnel which bypasses a single link of the Primary Path.
- NNHOP backup tunnel: a Backup Tunnel which bypasses a single node of the Primary Path.
Terminology

- NNHOP Back-up LSP
- Protected LSP
- Reroutable LSP
- PLR
- NHOP backup LSP
- Merge Point
MPLS TE LSP rerouting
(Global restoration)
TE LSP rerouting (Global restoration)

- Controlled by the head-end of a trunk via the resilience attribute of the trunk
- Fallback to either (pre)configured or dynamically computed path. Pre-configured path may be either pre-established, or established “on demand”

```plaintext
interface Tunnel0
  ip unnumbered Loopback0
  no ip directed-broadcast
  tunnel destination 10.0.1.102
  tunnel mode mpls traffic-eng
  tunnel mpls traffic-eng autoroute announce
  tunnel mpls traffic-eng priority 3 3
  tunnel mpls traffic-eng bandwidth 10000
  tunnel mpls traffic-eng path-option 1 explicit name prim_path
  tunnel mpls traffic-eng path-option 2 dynamic
  ip explicit-path name prim_path enable
  next-address 10.0.1.123
  next-address 10.0.1.100
```
All the routers have standard MPLS TE configuration

```plaintext
interface Tunnel0
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 10.0.1.102
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng autoroute announce
 tunnel mpls traffic-eng priority 3 3
 tunnel mpls traffic-eng bandwidth 10000
 tunnel mpls traffic-eng path-option 1 dynamic
 tunnel mpls traffic-eng record-route

interface POS0/0
 ip address 10.1.33.5 255.255.255.252
 no ip directed-broadcast
 ip router isis
 mpls traffic-eng tunnels
crc 32
 clock source internal
 pos framing sdh
 pos scramble-atm
 pos flag s1s0 2
 ip rsvp bandwidth 155000 155000

router isis
 passive-interface Loopback0
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng level-2
 net 49.0001.0000.0000.0011.00
 is-type level-2-only
 metric-style transition
 log-adjacency-changes
```

R1  R2  R3  R4  R5
The FIS (failure indication signal)

* R1 receives a Path Error from R2 and a Resv Tear
* R1 will receive a new LSA/LSP indicating the R2-R4 is down and will conclude the LSP has failed (if R1 is in the same area as the failed network element)

Which one on those two events will happen first? It depends of the failure type and IGP tuning

An optimisation of the Path Error allows to remove the failed link from the TE database to prevent to retry the same failed link (if the ISIS LSP or the OSPF LSA has not been received yet).

`mpls traffic-eng topology holddown sigerr <seconds>`
MPLS TE rerouting

- Use RSVP pacing to limit the loss of RSVP message in case of rerouting of several TE LSP:
  
  ip rsvp msg-pacing [period msec [burst msgs [max_size qsize]]]

- ISIS scanner (controls the propagation of TE information from ISIS to the TE database) may be used to speed-up convergence:
  
  mpls traffic-eng scanner [interval <1-60>] [max-flash <0-200> ]
  
  Interval: 5 seconds
  
  Max-flash: 15 updates
MPLS TE rerouting

- R1 is now informed that the LSP has suffered a failure
- R1 clears the Path state with an RSVP Path Tear message
- R1 recalculates a new Path for the Tunnel and will signal the new tunnel. If no Path available, R1 will continuously retry to find a new path (local process)
- PATH Protection time = O(s).

Does it reach the target?
MPLS TE rerouting

MPLS Traffic Engineering TE LSP reroute

- **TTR** = time between the fault and restoration

  * **Fault detection** may differ from the lower layers. May be done by the IGP (hello’s), layer 2 triggers
  * **Hold-off timer**. 0
  * **Fault notification**. Fault Indication Signal may be
    * the IGP (LSA/LSP update)
    * RSVP Path Error/Resv Tear/RSVP notify message

  FIS should be reliably transmitted with high priority. RSVP notify message may also be used.
MPLS TE rerouting

MPLS TE reroute (cont)

Fault restoration.

Restoration: the head must recalculate a Path (CSPF), signal the LSP and reroute the traffic

TTR = O(seconds)
MPLS TE Path Protection (global protection)
MPLS TE Path Protection

• MPLS TE Path Protection is a global repair mechanism using protection switching

• The idea is to be able to set up a primary LSP AND a back-up LSP (pre-signalled) so once the failure has been detected and signalled (by the IGP or RSVP signalling) to the head-end the traffic can be switched onto the back-up LSP

• No path computation and signalling of the new LSP once the failure has been detected and propagated to the head-end (compared to LSP reroute)
MPLS TE Path Protection

- By configuration the TE back-up LSP attributes may or not be different as the primary TE LSP:
  - The bw of the back-up LSP may some % of the primary bw
  - RCA of the back-up LSP may or not be taken into account
- Diversely routed paths are calculated by the CSPF on the head-end (they may be link, node or SRLG diverse)
MPLS TE Path Protection

- Limitation of MPLS TE Path protection
  - The FIS propagation may be unacceptable especially for very sensitive traffic,
  - The number of states in the network is doubled!!
  - CSPF is likely to be highly inefficient in terms of bandwidth usage.

→ primary diversely routed paths may share backup bandwidth (under the assumption of single network element failure)
MPLS TE Path Protection

- Path protection may be an attractive solution if and only if:
  - Just a few LSPs require protection
  - A few hundreds of msecs convergence time is acceptable
Principles of MPLS TE Fast Reroute (local protection)
MPLS TE FRR – Local protection

MPLS Fast Reroute local repair

- **Link protection**: the backup tunnel tail-head (MP) is one hop away from the PLR.

- **Node protection + Enhancements**: the backup tunnel tail-end (MP) is two hops away from the PLR.

```
R1  R2  R3
R4  R5
```

```
R1  R2  R3  R4  R5

R6  R7  R8

R9
```

© 2001, Cisco Systems, Inc. All rights reserved.
MPLS TE FRR – Local protection

- MPLS Fast Reroute link and node protection is:
  - **LOCAL** (compared to IGP or Path protection which are global protection/restoration mechanisms) which allows to achieve the 50msecs convergence time
  - Uses **Protection** (to provide fast rerouting)
  - Non **Revertive** but the previous path may be reused if more optimal (via reoptimization)
  - Reoptimization with Make before break to find a more optimal path
MPLS TE FRR – Local protection

- A key principle of Local repair is to guaranty a very fast traffic recovery with or without QOS guaranty (bandwidth guaranty) during a transient phase while other mechanisms (reoptimization) are used over a longer time scale.
MPLS TE FRR Local repair

- Controlled by the PLR
  - local repair is configured on a per link basis
  - the resilience attribute of a trunk allows to control whether local repair should be applied to the trunk (\texttt{mpls tra fast-reroute}).

  \text{“Local Protection Desired”} bit of the SESSION\_ATTRIBUTE object flag is set.

- Just the reroutable LSPs will be backed-up (fine granularity)

- Uses nested LSPs (stack of labels)

  1:N protection is KEY for scalability. N protected LSP will be backed-up onto the SAME backup LSP
MPLS TE Fast Reroute
Link Protection
(local protection)
MPLS TE FRR – Link Protection

- Backup labels (NHOP Backup Tunnel)

![Diagram of MPLS TE FRR with R1, R2, R3, R4, R5, and R6 routers showing backup labels 10, 11, and 20. The label 'x' indicates the label for the protected LSP, and the label 'X' indicates the label for the bypass LSP.]

Label for the protected LSP

Label for the bypass LSP
MPLS TE FRR – **Link** Protection

- Backup labels (NHOP Backup Tunnel)

```
R1  R2  R3  R4  R5  R6
10  20  11  11  11
x   x
```

- Label for the protected LSP
- Label for the bypass LSP
MPLS TE FRR – Link Protection

• Backup labels (NHOP)

2 remarks:
* The path message for the old Path are still forwarded onto the Back-Up LSP
* Modifications have been made to the RSVP code so that
  - R2 could receive a Resv message from a different interface than the one used to send the Path message
  - R4 could receive a Path message from a different interface (R3-R4 in this case)
MPLS TE FRR – Link Protection

- The PLR SHOULD send a PathErr message with error code of "Notify" (Error code = 25) and an error value field of ss00 cccc cccc cccc where ss=00 and the sub-code = 3 ("Tunnel locally repaired").

→ This will trigger the head-end reoptimization

- Then the TE LSP will be rerouted over an alternate Path (may be identical) using Make Before Break.
MPLS TE FRR - Link Protection - Configuration

- **On R1**
  
  ```
  !
  interface Tunnel0
  ip unnumbered Loopback0
  no ip directed-broadcast
  tunnel destination 10.0.1.102
  tunnel mode mpls traffic-eng
  tunnel mpls traffic-eng autoroute announce
  tunnel mpls traffic-eng priority 3 3
  tunnel mpls traffic-eng bandwidth 10000
  tunnel mpls traffic-eng path-option 1 dynamic
  tunnel mpls traffic-eng record-route
  tunnel mpls traffic-eng fast-reroute
  ```

Tunnel 0 is configured as fast reroutable

“Local Desired Protection” flag set in the SESSION_ATTRIBUTE object
MPLS TE FRR - Link Protection - Configuration

A Back-Up Tunnel Tu99 explicitly routed is configured on R2

interface Tunnel99
  ip unnumbered Loopback0
  no ip directed-broadcast
tunnel destination 10.0.1.100
tunnel mode mpls traffic-eng
tunnel mpls traffic-eng priority 1 1
tunnel mpls traffic-eng bandwidth 10000
tunnel mpls traffic-eng path-option 1 explicit name secours
tunnel mpls traffic-eng record-route

Use also:
Router (cfg-ip-expl-path)#
exclude-address a.b.c.d
Where a.b.c.d is a link address or a router ID to exclude a node

No tunnel mpls traffic-eng autoroute announce!

ip explicit-path name secours enable
next-address 10.0.1.123
next-address 10.0.1.100
On R2

interface POS4/0
description Link to R4
ip address 10.1.13.2 255.255.255.252
no ip directed-broadcast
ip router isis
encapsulation ppp
mpls traffic-eng tunnels
mpls traffic-eng backup-path Tunnel99
tag-switching ip
no peer neighbor-route
crc 32
clock source internal
pos ais-shut
pos report lrdi
ip rsvp bandwidth 155000 155000
MPLS TE FRR - Link Protection
On R1
Show tag for 100.0.1.150 32 det
Local Outgoing Prefix Bytes tag Outgoing Next Hop
tag tag or VC or Tunnel Id switched interface
26 20 10.0.1.150/32 0 Tu0 point2point
MAC/Encaps=4/12, MTU=4466, Tag Stack{27 20}, via PO0/0
0F008847 0001B00000014000
Fast Reroute Protection via {UnknownIF, outgoing label 27}
Per-packet load-sharing, slots: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

View additional tagging info with the 'detail' option
On R2

sh mpls tra fast-reroute det

LFIB FRR Database Summary:

- Total Clusters: 1
- Total Groups: 1
- Total Items: 1

Link 10:: PO4/0 (Up, 1 group)

Group 16:: PO4/0->Tu99 (Up, 1 member)

- Transit Item 810D60 (complete) [FRR OutLabel: 22]
- Key {incoming label 27}

sh tag for

<table>
<thead>
<tr>
<th>Local</th>
<th>Outgoing</th>
<th>Prefix</th>
<th>Bytes tag</th>
<th>Outgoing</th>
<th>NextHop</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>22</td>
<td>10.0.1.127</td>
<td>0 [1] 16896</td>
<td>PO4/0</td>
<td>point2point</td>
</tr>
</tbody>
</table>
MPLS TE FRR - Link Protection

Traffic is running from R0’s loo to R6’s loo(10.0.1.150)

R0  R1  R2  R3  Tu99  Tu0  Tu50  R4  R5  R6

IP Packet

LDP

26  27  20  28  22  20

\[ t_0: \] R2-R4 link fails

\[ t_1: \]

**Data plane:** R2 will immediately swap 27 <-> 22 (as before) and Push 28 (This is of course done for all the protected LSPs crossing the R2-R4 link)

**Control Plane** registers for a link-down event. Once the RSVP process receives this event, it will send out an RSVP PERR msg (O(s))

\[ t_2: \] R3 will do PHP

\[ t_3: \] R4 receives an identical labeled packet as before (Global Label Allocation needed)
MPLS TE FRR - Link Protection

Traffic is running from R0's loo to R6's loo(10.0.1.150)

2 remarks:
* The path message for the old Path are still forwarded onto the Back-Up LSP
* Modifications have been made to the RSVP code so that
  - R2 could receive a Resv message from a different interface than the one used to send the Path message
  - R4 could receive a Path message from a different interface (R3-R4 in this case)
MPLS TE Fast Reroute
Node Protection
(local protection)
MPLS TE FRR – **Node Protection**

- Node protection allows to configure a back-up tunnel to the next-next-hop! This allows to protect against link AND node failure.

![Diagram of a network with routers R0 to R9 showing node protection against R6 failure.](image-url)
MPLS TE FRR – Node Protection

- Backup labels

```
R1  10  R2  11  R6  12  R7  12  R8
R9

x  Label for the protected LSP
```
MPLS TE FRR – Node Protection

- Backup labels

The PLR learns the label to use from the RRO object carried in the Resv message when the reroutable LSP is first established – With global label space allocation on the MP
MPLS TE FRR – Node Protection

For each fast reroutable LSP ("Local protection Desired" bit set in the SESSION_ATTRIBUTE in the Path message), the tailhead LSR must include an RRO object in its Resv message (with label sub-object)
MPLS TE FRR – Node Protection

Subobject 1: IPv4 address

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

RRO Object

```
+-----------------+-----------------+-----------------+-----------------+
<table>
<thead>
<tr>
<th>Type</th>
<th>Length</th>
<th>IPv4 address (4 bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv4 address (continued)</td>
<td>Prefix Length</td>
<td>Flags</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
<td>-----------------------------</td>
</tr>
</tbody>
</table>
```

Flags

0x01  Local protection available

Indicates that the link downstream of this node is protected via a local repair mechanism. This flag can only be set if the Local protection flag was set in the SESSION_ATTRIBUTE object of the corresponding Path message.

0x02  Local protection in use

Indicates that a local repair mechanism is in use to maintain this tunnel (usually in the face of an outage of the link it was previously routed over, or an outage of the neighboring node).
MPLS TE FRR – Node Protection

Subobject 1: IPv4 address

Flags (cont)

Bandwidth protection: 0x04

The PLR will set this when the protected LSP has a backup path which provides the desired bandwidth, which is that in the FAST_REROUTE object or the bandwidth of the protected LSP, if no FAST_REROUTE object was included. The PLR may set this whenever the desired bandwidth is guaranteed; the PLR MUST set this flag when the desired bandwidth is guaranteed and the "bandwidth protection desired" flag was set in the SESSION_ATTRIBUTE object.

Node protection: 0x08

When set, this indicates that the PLR has a backup path providing protection against link and node failure on the corresponding path section. In case the PLR could only setup a link-protection backup path, the "Local protection available" bit will be set but the "Node protection" bit will be cleared.
The PLR learns the label to use from the RRO object carried in the Resv message when the reroutable LSP is first established.

The “label recorded desired” bit must be set in the SESSION-ATTRIBUTE of the Path message.

<table>
<thead>
<tr>
<th>Subobject 0x03, Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</td>
</tr>
</tbody>
</table>

The Length contains the total length of the subobject in bytes, including the Type and Length fields.

Flags

- **0x01** = Global label
  - This flag indicates that the label will be understood if received on any interface.

C-Type

- The C-Type of the included Label Object. Copied from the Label Object.

Contents of Label Object

- The contents of the Label Object. Copied from the Label Object.
MPLS TE FRR – **Node Protection**

- **Backup labels**

![Diagram showing node protection](image)

- The PLR swaps 10 <-> 12, pushes 20 and forward the traffic onto the backup tunnel.
MPLS TE FRR – Node Protection

• Path states maintenances
  • As in the case of NHOP backup tunnel, the Path messages are sent onto the backup tunnel to refresh the downstream states
MPLS TE FRR – Node Protection

- When the failure occurs, the PLR also updates:
  - The ERO object,
  - The PHOP object,
  - The RRO object

- As with Link protection, the PLR should the Point of Local Repair SHOULD send a PathErr message with error code of "Notify" (Error code = 25) and an error value field of ss00 cccc cccc cccc cccc where ss=00 and the sub-code = 3 ("Tunnel locally repaired"). → This will trigger the head-end reoptimization
MPLS TE FRR – Local repair

• When the failed link or node comes UP again the new resources may be re used once reoptimization has been triggered on the head-ends.

• As a reminder, reoptimization is triggered:
  
  • Periodically: “mpls traffic-eng reoptimize timers frequency <0-604800>
  
  • When a link comes “UP” if “mpls traffic-eng reoptimize events link-up”
  
  • When explicitly triggered (exec mode): “mpls traffic-eng reoptimize <tunnel x>”

• Make before break prevents any traffic disruption
MPLS TE FRR – Node Protection

- **Backup tunnel**
- **Protected tunnel**

**Path Tear**
- **Path Error (Locally repaired)**

**Reoptimization triggered**
**Make before break**
MPLS TE FRR – Node Protection

- The number of back-up tunnels for an interface is no longer limited to one!

On R2

interface POS4/0
description Link to R4
ip address 10.1.13.2 255.255.255.252
no ip directed-broadcast
ip router isis
encapsulation ppp
mpls traffic-eng tunnels
mpls traffic-eng backup-path Tunnel10
mpls traffic-eng backup path Tunnel15
tag-switching ip
no peer neighbor-route
crc 32
clock source internal
pos ais-shut
pos report lrdi
ip rsrp bandwidth 155000 155000

- Which is mandatory for Node protection …
MPLS TE FRR – Node Protection

- Back-up tunnel selection for a given LSP

- Tu1 is chosen for LSP1
- Tu2 is chosen for LSP2
MPLS TE FRR – Node Protection

- One may combine tunnels terminating on the next hop and next-next-hop
- This allows to increase redundancy,
- In case of unavailability of a back-up tunnel the other one is used (order of preference is determined by the tunnel ID number)
- Load balancing between to back-up tunnels terminating on the same nhhop.
MPLS TE FRR – Node Protection

- Load balancing: Multiple back-up tunnels to the same destination may be created.
Backup tunnel path computation and provisioning

- **Packing algorithm**: refers to the method used to select the backup tunnel for each protected LSP.

- For each protected LSP at a given PLR:
  - Select the set of backup tunnel whose merge point crosses the primary path,
  - Find a backup tunnel whose remaining bandwidth is \( \geq \) of the protected LSP (if bandwidth protection is required)
  - Multiple backup tunnel selection policies are available
Per Class backup tunnel

• When using both regular and DS-TE tunnels, it may desirable to configure regular and DS-TE backup tunnels.
• Other combinations are also possible
• Packing algorithm enhancements
MPLS TE FRR Local repair

- Uses nested LSPs (stack of labels)

1:N protection is KEY for scalability. N protected LSP will be backed-up onto the SAME backup LSP
MPLS TE FRR Local repair

- Uses nested LSPs (stack of labels)

Convergence Time
- FRR LSPs: 50ms
- Non FRR LSPs: O(s)

Backup tunnel
- No states for the rerouted LSP

Fast reroutable LSPs
- NON Fast reroutable LSPs are rerouted using restoration
MPLS TE protection/restoration schemes

Link/Node Failure detection

- **Link** failure detection
  - On POS, link failure detection is handled by Sonet/SDH alarms
    - On Receive side: LOS/LOF/LAIS
    - On Transmit side: LRDI
  - Very fast.

- **Node** failure detection is a more difficult problem
  - Node hardware failure => Link failure
  - Software failure ... Need for a fast keepalive scheme (IGP, RSVP hellos)
RSVP Hellos

- RSVP Hellos extension is defined in RFC3209
- The RSVP hello extension enables an LSR to detect node failure detection
- Allows to detect:
  - Link failure when layer 2 does not provide failure detection mechanism,
  - Node failure when the layer 2 does not fail.
RSVP Hellos

- RSVP hello adjacency are brought up dynamically (if at least one protected LSP in READY state (with one backup tunnel operational))
- One RSVP hello adjacency per link per neighbor (not per protected LSP !!)

- An hello adjacency is removed when the last protected LSP in READY state is torn down
RSVP Hellos

- RSVP hello has been designed for Node failure detection. Fast link failure detection already exist on Sonet/SDH links.

- But can also be used as a fast link failure detection on GE links (point to point or behind a switch) → FRR over GE links
Using FRR without MPLS TE

- MPLS TE FRR is backing-up TE LSP. If MPLS TE is not used in the network, one may use Fast Reroute for fast convergence using unconstraint 2 hops protected tunnel.

- Ex:

```
R1 -- R2 -- R3 -- R4 -- R7 -- R8
     |                  |
     | 2 hops Protected |
     | LSP              |

R9

Protected LSP are defined w/o any constraint -> follow the shortest path !!

NNHOP backup tunnel
```
MPLS TE protection/restoration schemes

- MPLS TE FRR may be used in specific parts of the network where very fast convergence is required,
- Compared to other protection schemes (optical, Sonet/SDH) no backup bandwidth is wasted.
MPLS TE protection/restoration schemes

- Combining IP routing restoration and MPLS TE FRR fast protection

- No need for fast convergence
- O(s) convergence -> IGP tuning

- 50msecs convergence
MPLS TE protection/restoration schemes

- IP convergence versus MPLS TE FRR
  - IP convergence is O(s) and may be even speed-up around 1 second
  - For faster convergence (<50msec), MPLS TE Fast Reroute should be used.
Backup tunnel path computation and provisioning
MPLS TE protection/restoration schemes

- Backup tunnel path computation and provisioning is definitely an important topic,
- Complexity is driven by the parameters to take into account and the degree of optimality

- Back-up tunnel diversely routed from the protected section – no constraint
- Back-up tunnel diversely routed from the protected section – SRLG disjoint – Bandwidth protection – Bandwidth usage optimisation -…

Complexity (not linear)
MPLS TE protection/restoration schemes

• The level of complexity will also determine whether the backup tunnel complexity is done Off-line or On-line (distributed)

- Back-up tunnel diversely routed from the protected section – no constraint
- Back-up tunnel diversely routed from the protected section – SRLG disjoint
- Bandwidth protection – Bandwidth usage optimisation -...

Complexity (not linear)

Tunnel Builder Pro
Diversely routed paths

Optical plane

Fibber

Link to protect

NO!

Back-Up tunnel

FRR design

IP/MPLS view

Link to protect

Back-Up tunnel

FRR design
Diversely routed paths

- **SRLG are configured on each link so that:**
  - The back-up path is computed as SRLG disjoint from the protected LSP (Path protection),
  - The backup path is computed as SRLG disjoint from the protected section (Local repair)

- **SRLG are flooded by the IGP:**
  - New TLV for ISIS
  - Sub TLV of link TLV 18 (type 16)

- **More than one SRLG may be configured on a link**
Diversely routed paths

- An example with ISIS

<table>
<thead>
<tr>
<th>Sub-TLV</th>
<th>Type</th>
<th>Length</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4</td>
<td></td>
<td>Administrative group (color)</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td></td>
<td>Outgoing Interface Identifier</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td></td>
<td>Incoming Interface Identifier</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td></td>
<td>IPv4 interface address</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td></td>
<td>IPv4 neighbor address</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td></td>
<td>Maximum link bandwidth</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td></td>
<td>Reservable link bandwidth</td>
</tr>
<tr>
<td>11</td>
<td>32</td>
<td></td>
<td>Unreserved bandwidth</td>
</tr>
<tr>
<td>12</td>
<td>32</td>
<td></td>
<td>Maximum LSP Bandwidth</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td></td>
<td>TE Default metric</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td></td>
<td>Link Mux Capability</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td></td>
<td>Link Protection Type</td>
</tr>
<tr>
<td>250-254</td>
<td>-</td>
<td></td>
<td>Reserved for cisco specific extensions</td>
</tr>
<tr>
<td>255</td>
<td>-</td>
<td></td>
<td>Reserved for future expansion</td>
</tr>
</tbody>
</table>

+ 2 new TLVs.

<table>
<thead>
<tr>
<th>TLV Type</th>
<th>Length</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>136 (TBD)</td>
<td>variable</td>
<td>Link Descriptor</td>
</tr>
<tr>
<td>138 (TBD)</td>
<td>variable</td>
<td>Shared Risk Link Group</td>
</tr>
</tbody>
</table>
Backup tunnel path computation and provisioning

• Two possible approaches
  
  • Local repair without bandwidth protection
    
    • Once the link/node failure occurs, the protected LSP is rerouted within 50msecs but the rerouted LSP does not get any bandwidth guaranty. Note Diffserv should be used to protect sensitive traffic over the backup (potentially congested) path
  
  • Local repair with bandwidth protection
    
    • The protected LSP are rerouted onto a backup tunnel that provides bandwidth guaranty.
  
• This relates to the amount of bandwidth that the protected LSP will receive (before being reoptimized by the head-end (if possible)).
Backup tunnel path computation and provisioning

• Whether a protected LSP receives bandwidth protection or not depends on the backup tunnel constraints.

• Local repair without bandwidth protection
  • Does not require backup tunnel computation complexity.
  • Backup tunnel with 0 bandwidth
  • For each PLR, a NNHOP backup tunnel is configured to every NNHOP.
Backup tunnel path computation and provisioning

Local repair without bandwidth protection

- Backup tunnel path computation and provisioning is straightforward
Backup tunnel path computation and provisioning

Local repair with bandwidth protection

- **Problem definition**: find a set of backup tunnels between each PLR and its NNHop such that the protected LSPs could receive the appropriate amount of bandwidth when rerouted over the (those) backup LSPs.

- Note that between the PLR and the MP more than one backup tunnel may be used (Load balancing)
Backup tunnel path computation and provisioning

Local repair with bandwidth protection

• The problem of QOS guaranty can be reduced to a problem of bandwidth provisioning (provided the propagation delay is bounded)

• May also cover Propagation delay increase guaranty

• Requires much more complexity (complex backup tunnel path computation).
Backup tunnel path computation and provisioning

Local repair with bandwidth protection

- CSPF is likely to be highly inefficient.
- Other more sophisticated backup tunnel path computation methods may be required.
Backup tunnel path computation and provisioning

Local repair with bandwidth protection

NOT POSSIBLE TO PLACE A 10M backup tunnel
Backup tunnel path computation and provisioning

Local repair with bandwidth protection
Backup tunnel path computation and provisioning

Local repair with bandwidth protection

Tunnel Builder PRO
MPLS TE protection/restoration schemes

- Number of back-up LSPs required (impact on the number of states)
  - LSP reroute: 0
  - Path protection: $O(\# \text{ LSPs})$
  - FRR Link protection: $O(\# \text{ links})$
  - FRR Node protection: $O(\up to (\# \text{ Node})^2)$
Agenda

• Introduction
• Terminology of protection/restoration
• MPLS Traffic Engineering Fast Reroute
• IETF Update
• Conclusion
IETF Update

IETF

• WG IETF draft (adopted as a WG document at IETF 52, SLC):

draft-ietf-rsvp-lsp-fastreroute-00.txt

P Pan, DH. Gan: JUNIPER networks,

G. Swallow, JP Vasseur: CISCO SYSTEMS

D. Cooper: Global Crossing

A. Atlas, M. Jork: AVICI Systems
IETF Update

Common sections

Detour

Backup facility + enhancements

Common sections
MPLS TE protection/restoration schemes

- The Detour LSP solution overview

- The protected LSPs are signalled with a FAST_REROUTE object specifying the attributes of the detour:
  - priority, max hops, bandwidth.
- Detour LSP are set up at each PLR
- Detour path computed using CSPF (periodic …)
MPLS TE protection/restoration schemes

- The number of detour LSPs = nb of protected LSPs * (N – 1) w/o merging

  Where N : average number of hops per LSP

- With 5000 protected LSPs and an average diameter of 6 hops, this represents 25000 TE LSPs (w/o merging)

- So the main issue of this solution is the scalability.

- With bypass, a single backup tunnel is used to backup a set of protected LSPs.
Agenda

• Introduction
• Terminology of protection/restoration
• MPLS Traffic Engineering Fast Reroute
• IETF Update
• Conclusion
MPLS TE protection/restoration schemes

- In summary,
  - LSP reroute (Global Restoration) is the default TE rerouting mode (slow)
  - Path protection (Global protection) if just a few protected LSPs, no sub seconds TTR required,
  - FRR link protection (Global protection) provides 50 msecs (may replace SDH/Sonet protection), could be configured on a few specific links. Limit the number of extra states required (using M:N protection) – label stacking.
  - Node protection (Global protection) the most efficient protection scheme providing 50ms in case of link and node protection. Limit the number of extra states required (using M:N protection) – label stacking
Thank You !
Agenda

• Failure profiles
  ➔ Link failures,
  ➔ Hardware node failures,
  ➔ Software failures
    - Control plane Node failures (GRP frozen, …),
    - Forwarding plane node failures
  ➔ “Planned” Hw/Sw failures
    - Software upgrades,
    - Hardware upgrades (LC, GRP, Chassis, …)
Agenda

→ Determining the network failure profiles to key prior to determining the set of protection/restoration schemes to deploy
Agenda

- Improving network reliability
  - By network architecture (load balancing, elimination of central point of failure, ...),
  - Improving network element redundancy
    - Hardware redundancy (GRP, Chassis, ...),
    - Software redundancy (High Availability)
  - Protection network elements (links/nodes/SRLG) with Fast rerouting (IGP, FRR).

Phasing
Protection/restoration schemes

Improving network availability

Centralized

Protected link
unprotected link

Backup tunnels
PCS

MPLS TE FRR

Link/Node/SRLG Fast protection (50msecs) with BW guarantees
Agenda

**Failure profiles**

- **Link failure**
  - Link protection: Optical (1+1, 1:N, ...), SDH SONET, ... MPLS TE FRR

- **Hardware node failures**, ✓ IGP, MPLS TE Fast Reroute

- **Software failures**
  - Control plane Node failures (GRP frozen, ...), ✓ per box mechanism
  - Forwarding plane node failures

- **Planned Hw/Sw failure**
  - Software upgrades, ✓ Overload bit, MPLS TE Fast Reroute (RSVP hellos)
  - Hardware upgrades (LC, GRP, Chassis, ...)
    ✓ Overload bit, MPLS TE Fast Reroute (RSVP hellos)
Agenda

• Protection/Restoration scheme performances. Multidimensional problem …
  - convergence speed (50msecs – 2-3 secs – minutes). Controls packet loss.
  - QOS on the rerouted Path
    → Queueing algorithm (Forwarding plane)
    → Bandwidth guaranty (Control plane)
    → Bounded propagation delay (control plane)
Protection/restoration schemes

Improving network availability

- IGP Fast convergence
- Backup tunnels
- Distributed
- Centralized
- Link/Node/SRLG Fast protection (50msecs) with BW guaranties
- MPLS TE FRR

Protected link
unprotected link