

# Tutorial: MPLS Fast Reroute

Joseph M. Soricelli (jms@juniper.net) NANOG 30, Miami, Florida





### **Caveats and Assumptions**

 The views presented here are those of the author and they do not necessarily represent the views of Juniper Networks

#### Basic understanding of RSVP and MPLS

- Labels (Push, Pop, Swap)
- Path, Resv, error messages
- TE extensions for the IGPs
- RSVP sessions

# You will ask a question when you don't understand!

### Why Use Fast Reroute?

#### Traffic protection

- For packets already passing through the LSP
- Continuous forwarding of labeled traffic

#### Application driven concerns

Real-time traffic

#### Non-protected transport network?

### Requirements

#### RSVP based solution only

- Needs an "outside" view of the network
- Traffic engineering capabilities
- Support for the RSVP-TE extensions and the Fast Reroute objects

# Agenda

#### Terminology

- "I want to be protected!"
- One-to-One backup
- Facility Backup
- RSVP packet dumps

# Agenda

#### → Terminology

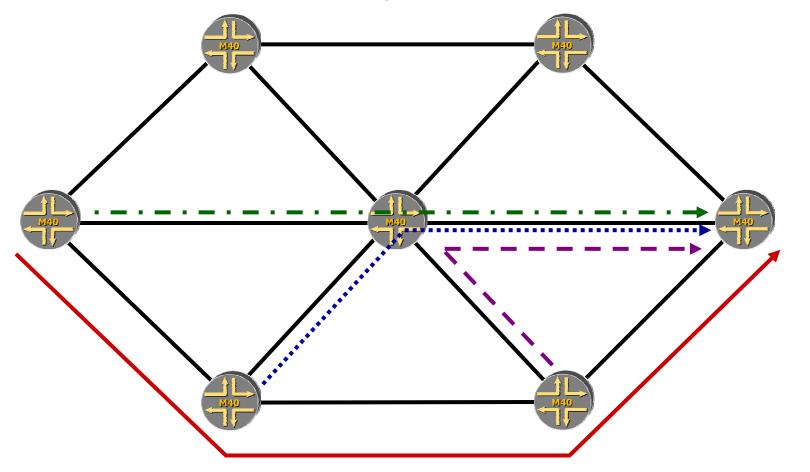
- "I want to be protected!"
- One-to-One backup
- Facility Backup
- RSVP packet dumps

### **One-to-One Backup**

- Each LSP gets it's own set of protected resources
- Allows for forwarding around the next downstream node and link
  - Except for egress node
- Paths established in the network to avoid the node
  - Always headed towards the egress router
  - Each LSP creates and uses its own Fast Reroute paths
- During a failure, the label is swapped and sent into the alternate path
  - No label stacking

### **One-to-One Node Protection**

 Each node along the LSP's path creates an alternate LSP around the downstream node and headed towards the egress router



# Facility Backup (Many-to-One)

Supports Node Protection

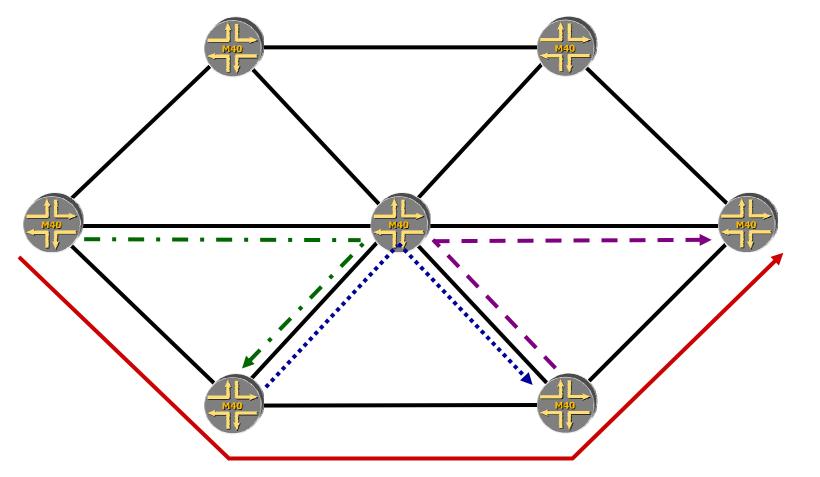
#### Supports Link Protection

- Allows for forwarding around the next downstream link
- Connects to the next downstream node
- Paths established in the network to avoid the link
- Each set of neighbors creates and uses its own Fast Reroute path

 During a failure, the label is swapped and a second label is pushed before being sent along the alternate path

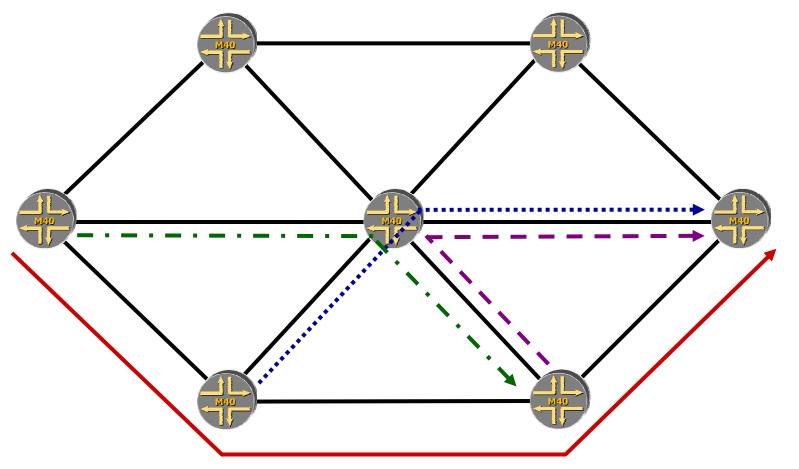
# **Facility Backup – Link Protection**

 Each node creates an alternate LSP around the downstream link



### **Facility Backup – Node Protection**

- Each node creates an alternate LSP around the downstream node and the interconnecting link
  - Penultimate node uses link protection



# **Point of Local Repair (PLR)**

- Node which notices the failure of:
  - Downstream link
  - Downstream node
- Begins forwarding traffic along the alternate path
- Notifies the ingress router that the main LSP has a problem
  - Sends PathErr upstream

# More Terms (1 of 2)

#### Protected LSP

- Has Fast Reroute enabled and alternate paths established in the network
- Both one-to-one and many-to-one provide protection for LSPs

#### Detour LSP

- Used in a one-to-one protection scheme
- Section 2.1 Sec

#### Next-Hop Bypass LSP

- Used in a facility backup link protection scheme
- LSP created between the two adjacent neighbors

# More Terms (2 of 2)

#### Next-Next-Hop Bypass LSP

- Used in a facility backup node protection scheme
- LSP created to avoid the downstream node
- Merge Point
  - Point where the alternate path rejoins the main LSP

#### Detour Merge Point

 Point where multiple detours join together along the protected path

# Agenda

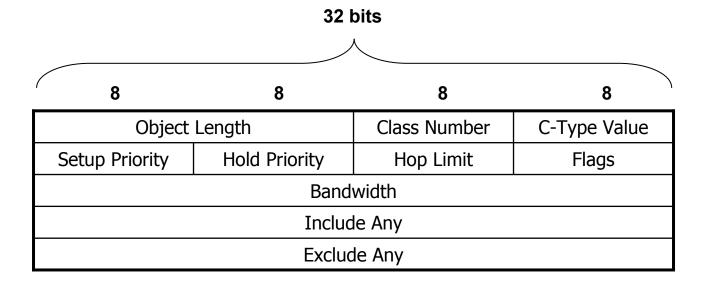
#### Terminology

- → "I want to be protected!"
- One-to-One backup
- Facility Backup
- RSVP packet dumps

### **Fast Reroute Object**

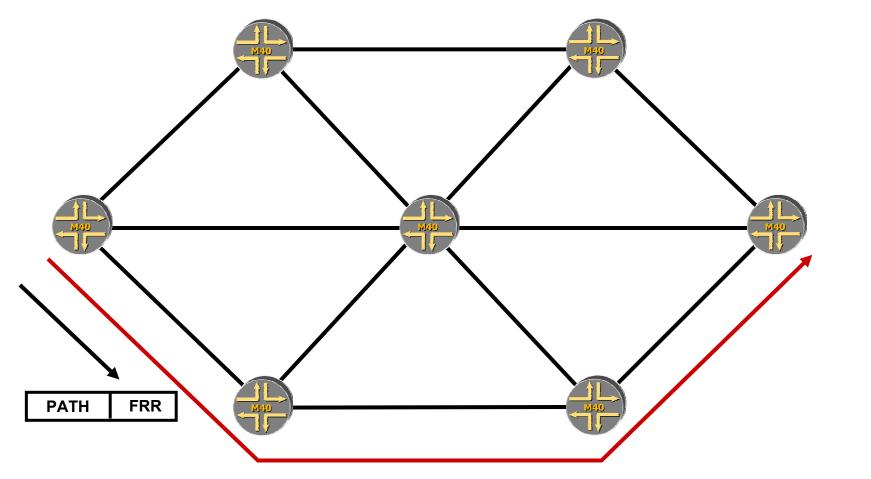
Included in Path messages for the protected LSP

- Signals the ingress router's desire to protect the LSP
- Contains information which each node uses to establish it's protected paths
- Flag values determine the type of protection requested
  - 0x01 for one-to-one backup
  - 0x02 for facility backup



### **LSP Path Message and FRR Object**

 The ingress router inserts the Fast Reroute Object within the Path message for the protected LSP



# Session Attribute Flags (1 of 2)

 Informs each router in the path about the characteristics of the LSP

#### 0x01 – Local protection desired

- Allows the LSP routers to establish detour paths in the network which violate the ERO of the LSP.
- Generic setting which allows the nodes to use either protection scheme
- 0x02 Label recording desired
  - Requests that the LSP routers include the label they assigned in the RRO

# **Session Attribute Flags (2 of 2)**

#### • 0x04 – SE style desired

- Bandwidth should not be double counted for the LSP and it's detours
- The ingress node for the protected LSP will reroute using "make before break"

#### 0x08 – Bandwidth protection desired

- The detour LSPs should reserve bandwidth in the network along their paths
- Inherit the BW reserved for the protected LSP or use the BW specified in the Fast Reroute Object

#### 0x10 – Node protection desired

 Explicitly requests that the LSP routers use node protection when establishing their detour paths

### **Constrained Shortest Path First**

- Each PLR ingress node consults the TED
  - Attempts to locate a path for the particular protection scheme requested
- If the FRR Object requests constraints, the protection path matches those requests
- A complete ERO for the protection path is created
- Protection path is signaled by RSVP

# Agenda

#### Terminology

"I want to be protected!"

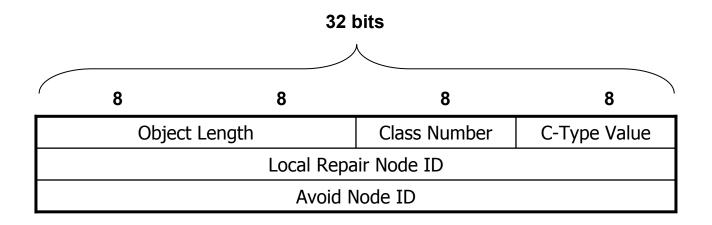
#### One-to-One backup

- Facility Backup
- RSVP packet dumps

### **Detour Object**

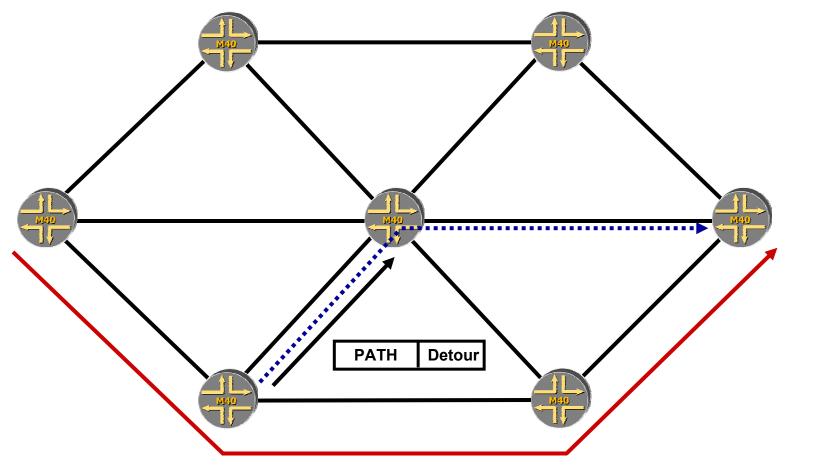
#### Included in Path messages for the detour LSP

- Allows routers along the detour path to associate multiple detours with the same RSVP session
- Includes information about the detour ingress
- Includes the node which the detour is avoiding



### **Detour Path Message and Object**

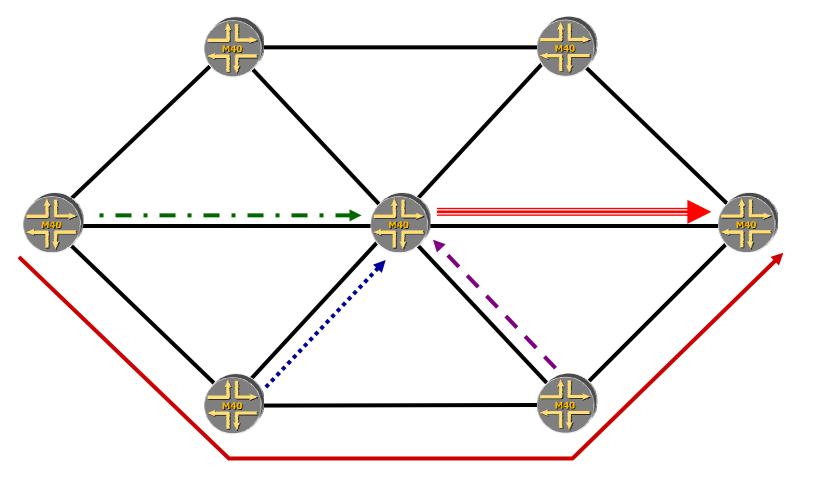
 After the routers along the Path assign resources to the LSP, each generates a Path message for the detour path and includes the Detour Object



# **Merging Detours**

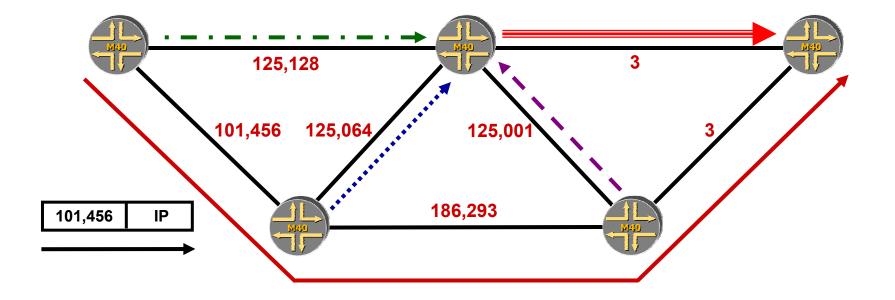
Detour merge points combine detours together

 Router in the middle of the network combines all detours into a single detour path

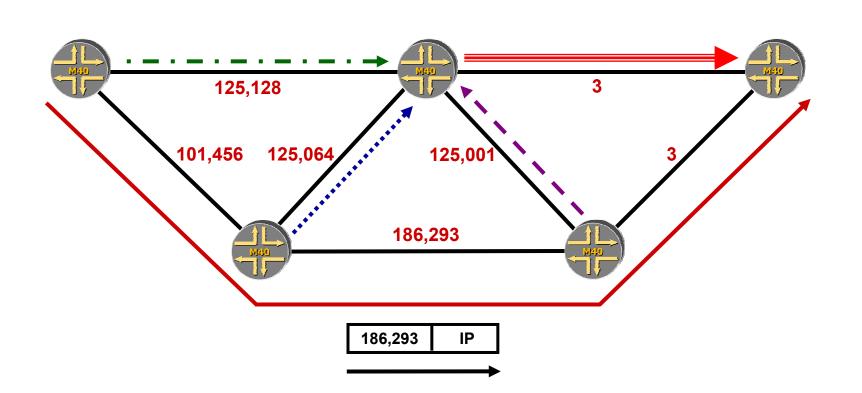


### LSP Label Operations (1 of 3)

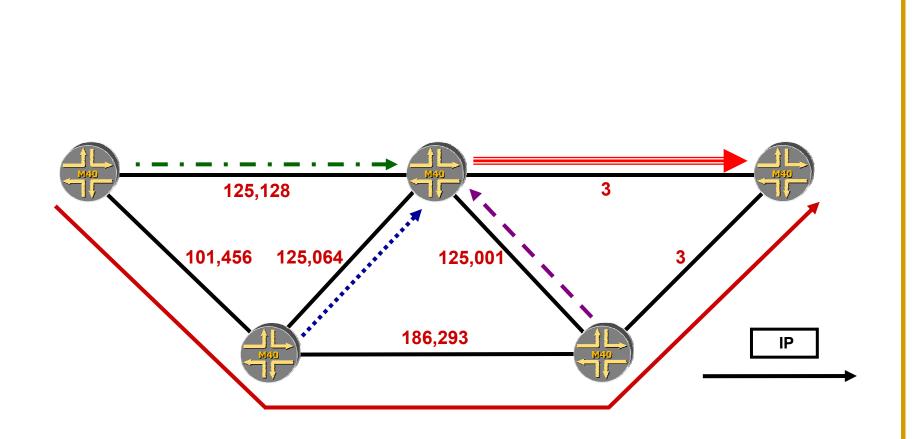
 In a normal operating environment, the routers perform label swaps as expected



### LSP Label Operations (2 of 3)



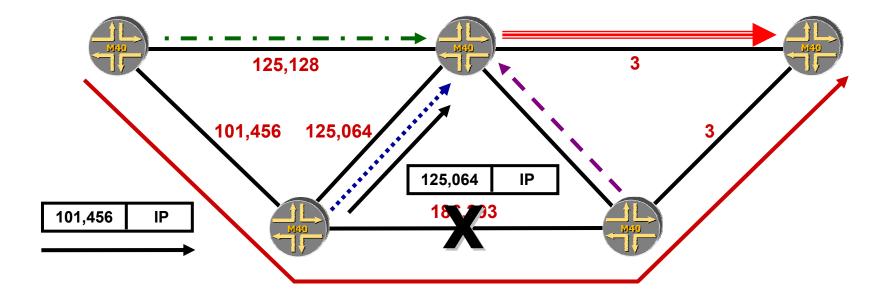
### LSP Label Operations (3 of 3)



# **Detour Label Operations (1 of 2)**

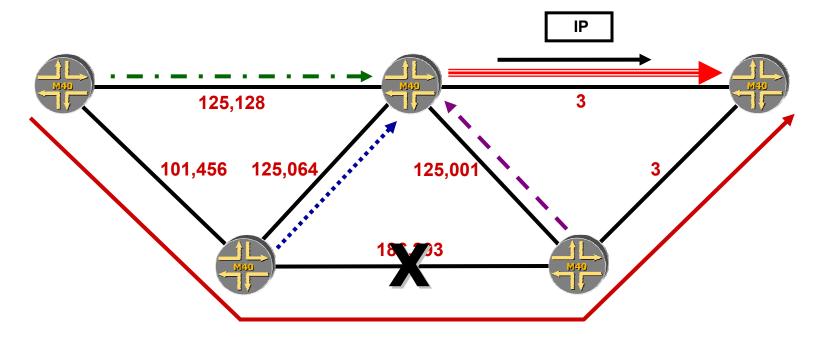
#### Point of local repair performs a label swap

 Incoming label exchanged for label advertised by the downstream router along the detour



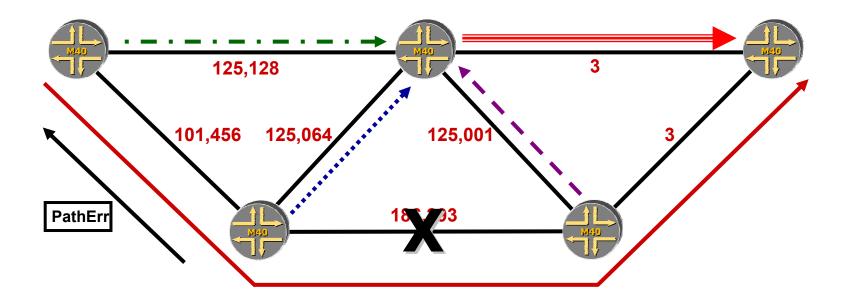
# **Detour Label Operations (2 of 2)**

- Merge point swaps label and forwards the packet along the detour path
  - In our example, the merge point is the penultimate hop so the label is popped



# **Notification of Local Repair (PathErr)**

- PLR sends a PathErr message to the ingress router of the protected LSP
  - Allows ingress to move to an alternate path for "permanent" recovery



# **Record Route Object Flags (1 of 2)**

 Allows each router in the path to determine the availability or use of a protection path

#### 0x01 – Local protection available

- Means that the downstream link from the router is protected by a protection mechanism
- Can be either node or link protection

#### 0x02 – Local protection in use

 Means that the PLR is actively using the protection path for the LSP

# **Record Route Object Flags (2 of 2)**

#### 0x04 – Bandwidth protection

Means that the router was able to successfully establish a backup path which meets the BW specified by the ingress router for the LSP

#### 0x08 – Node protection

- Means that the downstream node from the router is protected by node protection
- This is NOT set when only link protection is available

# Agenda

- Terminology
- "I want to be protected!"
- One-to-One backup
- Facility Backup
- RSVP packet dumps

### **Backup Paths Created**

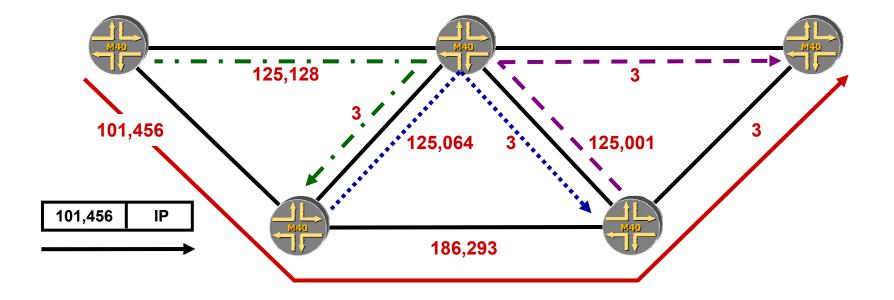
- Once the protected LSP is established, the routers along the path create the backup paths required
  - Could be a Next-Hop Bypass LSP
  - Could be a Next-Next-Hop Bypass LSP

#### Any existing Bypass LSPs are used

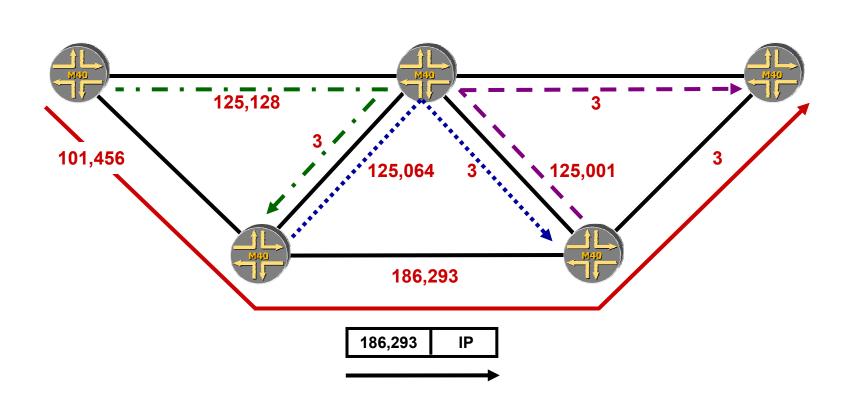
Key to the Many-to-One concept

### LSP Label Operations (1 of 3)

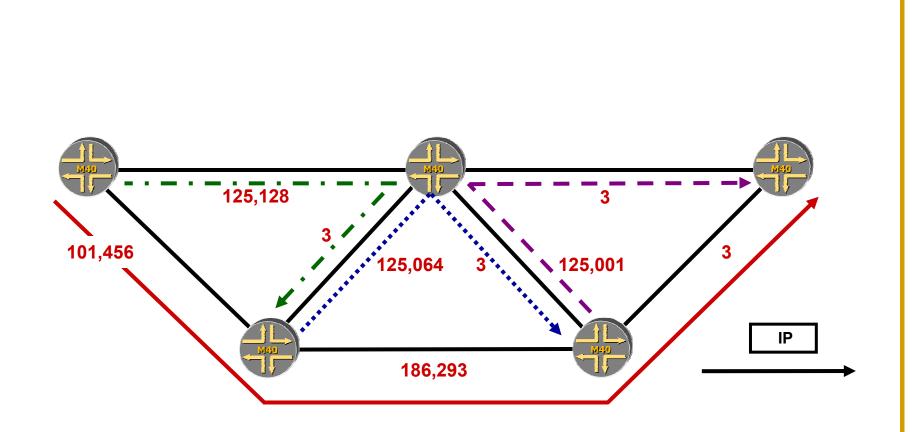
 In a normal operating environment, the routers perform label swaps as expected



### LSP Label Operations (2 of 3)



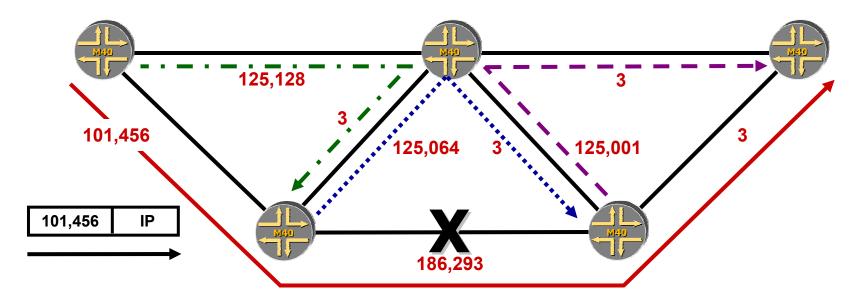
### LSP Label Operations (3 of 3)



# Link Facility Backup Operations (1 of 4)

#### Point of local repair performs a swap and a push

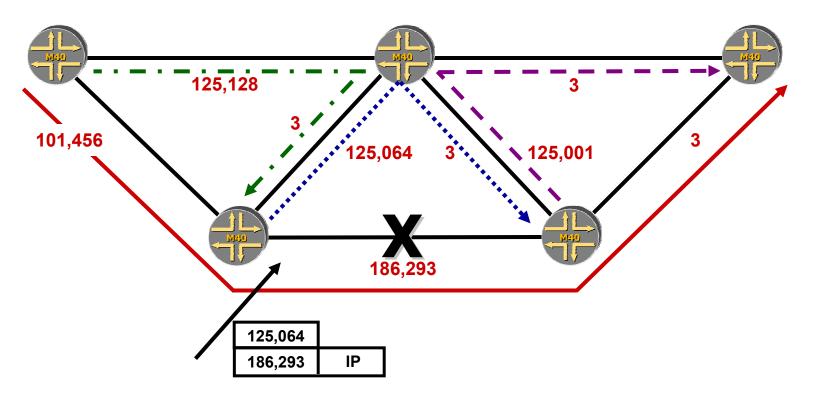
- Incoming label exchanged for label advertised by the downstream router along the protected LSP
- Adds the label representing the first hop along the Bypass LSP



# Link Facility Backup Operations (2 of 4)

#### Point of local repair performs a swap and a push

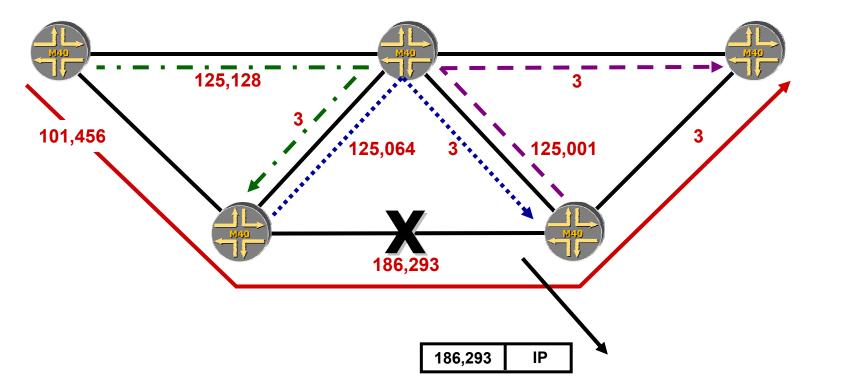
- Incoming label exchanged for label advertised by the downstream router along the protected LSP
- Adds the label representing the first hop along the Bypass LSP



# Link Facility Backup Operations (3 of 4)

#### Bypass LSP transit router performs a label pop

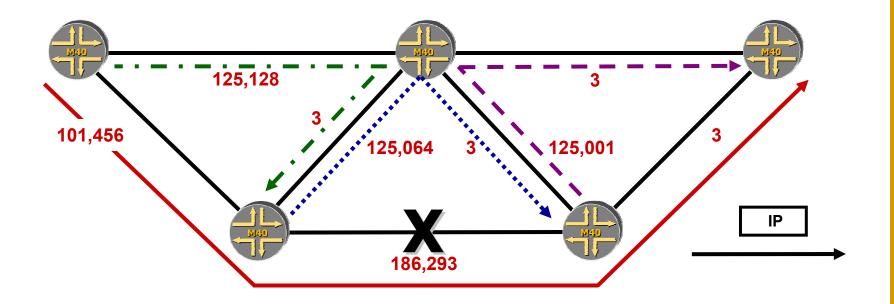
 Removes the incoming bypass label and forwards the remaining data to the merge point



# Link Facility Backup Operations (4 of 4)

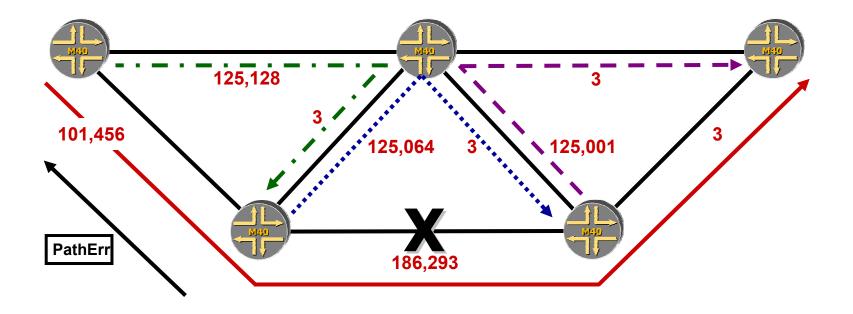
 Merge point pops the incoming label along the bypass LSP

- Merge point is the penultimate router in our case
- Label swaps are also possible



# **Notification of Local Repair (PathErr)**

- PLR sends a PathErr message to the ingress router of the protected LSP
  - Allows ingress to move to an alternate path for "permanent" recovery



# **Record Route Object Flags (1 of 2)**

 Allows each router in the path to determine the availability or use of a protection path

### 0x01 – Local protection available

- Means that the downstream link from the router is protected by a protection mechanism
- Can be either node or link protection

### 0x02 – Local protection in use

 Means that the PLR is actively using the protection path for the LSP

# **Record Route Object Flags (2 of 2)**

#### 0x04 – Bandwidth protection

Means that the router was able to successfully establish a backup path which meets the BW specified by the ingress router for the LSP

### 0x08 – Node protection

- Means that the downstream node from the router is protected by node protection
- This is NOT set when only link protection is available

# **Facility Backup – RSVP Operations**

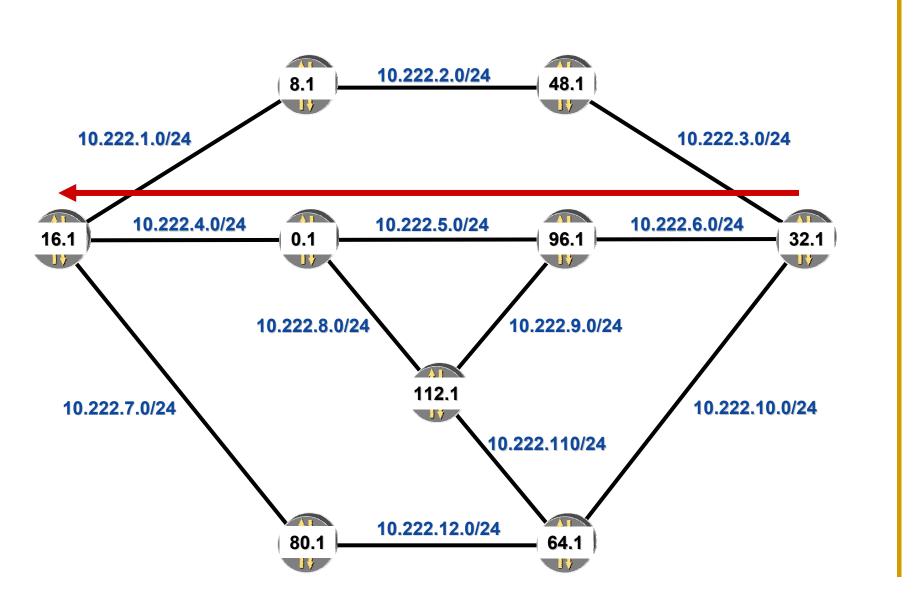
- In the event of a local repair RSVP messages are passed between neighbors across the bypass LSP
- Path, PathTear, and ResvConf messages use the bypass
- Resv, ResvTear, and PathErr messages are address to the Previous Next-Hop address for protected LSP
  - Uses best-effort routing
  - Extracts address from the RSVP-Hop object

# Agenda

### Terminology

- "I want to be protected!"
- One-to-One backup
- Facility Backup
- → RSVP packet dumps

### **Sample Network**



### **One-to-One Protection - Path**

### Initial Path message from the ingress is received on the 96.1 router

#### Contains Fast Reroute Object

RSVP recv Path 192.168.32.1->192.168.16.1 Len=236 so-0/1/2.0 Session7 Len 16 192.168.16.1 (port/tunnel ID 25124) Proto 0 Len 12 10.222.6.2/0x0857cd8c Hop Len 8 30000 ms Time SessionAttribute Len 16 Prio (7,0) flag 0x0 "FRR-Test" Sender7 Len 12 192.168.32.1 (port/lsp ID 1) Tspec Len 36 rate Obps size Obps peak Infbps m 20 M 1500 ADspec Len 48 SrcRoute Len 28 10.222.6.1 S 10.222.5.2 S 10.222.4.1 S LabelRequest Len 8 EtherType 0x800 Properties Len 12 Primary path RecRoute Len 12 10.222.6.2 FastReroute Len 20 Prio(7,0) Hop 6 BW Obps Include 0x0000000 Exclude 0x0000000

### **One-to-One Protection - Resv**

### Resv message for the protected LSP is transmitted back to the ingress router

```
RSVP send Resv 10.222.6.1->10.222.6.2 Len=136 so-0/1/2.0
Session7 Len 16 192.168.16.1(port/tunnel ID 25124) Proto 0
Hop Len 12 10.222.6.1/0x0857cd8c
Time Len 8 30000 ms
Style Len 8 FF
Flow Len 36 rate Obps size Obps peak Infbps m 20 M 1500
Filter7 Len 12 192.168.32.1(port/lsp ID 1)
Label Len 8 100144
RecRoute Len 28 10.222.6.1 10.222.5.2 10.222.4.1
```

### **One-to-One Protection - Detour**

#### PLR generates a Path message containing a Detour object

RSVP send Path 192.168.32.1->192.168.16.1 Len=244 so-0/1/2.0 Session7 Len 16 192.168.16.1 (port/tunnel ID 25124) Proto 0 Len 12 10.222.6.1/0x0857cd8c Hop Time Len 8 30000 ms SessionAttribute Len 16 Prio (7,0) flag 0x0 "FRR-Test" Sender7 Len 12 192.168.32.1 (port/lsp ID 1) Len 36 rate Obps size Obps peak Infbps m 20 M 1500 Tspec ADspec Len 48 SrcRoute Len 36 10.222.6.2 S 10.222.10.2 S 10.222.12.1 S 10.222.7.2 S LabelRequest Len 8 EtherType 0x800 Properties Len 12 Primary path RecRoute Len 20 10.222.6.1 10.222.6.2 Detour Len 12 Branch from 10.222.6.1 to avoid 192.168.0.1

### **One-to-One Protection - Detour**

#### PLR receives a Resv message which confirms the Detour LSP is established

RSVP recv Resv 10.222.6.2->10.222.6.1 Len=144 so-0/1/2.0 Session7 Len 16 192.168.16.1(port/tunnel ID 25124) Proto 0 Hop Len 12 10.222.6.2/0x0857cd8c Time Len 8 30000 ms Style Len 8 FF Flow Len 36 rate Obps size Obps peak Infbps m 20 M 1500 Filter7 Len 12 192.168.32.1(port/lsp ID 1) Label Len 8 100048 RecRoute Len 36 10.222.6.2 10.222.10.2 10.222.12.1 10.222.7.2

### **One-to-One Protection – Notify Ingress**

#### PLR sets flags in RRO stating that protection is available

```
RSVP send Resv 10.222.6.1->10.222.6.2 Len=136 so-0/1/2.0
Session7 Len 16 192.168.16.1(port/tunnel ID 25124) Proto 0
Hop Len 12 10.222.6.1/0x0857cd8c
Time Len 8 30000 ms
Style Len 8 FF
Flow Len 36 rate Obps size Obps peak Infbps m 20 M 1500
Filter7 Len 12 192.168.32.1(port/lsp ID 1)
Label Len 8 100144
RecRoute Len 28 10.222.6.1(flag=9) 10.222.5.2(flag=1) 10.222.4.1
```

### **Facility Link Protection - Path**

 Initial Path message from the ingress is received on the 96.1 router

 Session Attribute flags set to 0x7 for link (0x01), label recording (0x02), and SE reservation (0x04)

RSVP recv Path 192.168.32.1->192.168.16.1 Len=216 so-0/1/2.0 Session7 Len 16 192.168.16.1 (port/tunnel ID 25118) Proto 0 Hop Len 12 10.222.6.2/0x0857cd8c Time Len 8 30000 ms SessionAttribute Len 16 Prio (7,0) flag 0x7 "FRR-Test" Sender7 Len 12 192.168.32.1 (port/lsp ID 1) Tspec Len 36 rate Obps size Obps peak Infbps m 20 M 1500 ADspec Len 48 SrcRoute Len 28 10.222.6.1 S 10.222.5.2 S 10.222.4.1 S LabelRequest Len 8 EtherType 0x800 Properties Len 12 Primary path RecRoute Len 12 10.222.6.2

### **Facility Link Protection - Resv**

# Resv message for the protected LSP is transmitted back to the ingress router RRO contains labels assigned to the LSP

```
RSVP send Resv 10.222.6.1->10.222.6.2 Len=160 so-0/1/2.0
  Session7 Len 16 192.168.16.1 (port/tunnel ID 25118) Proto 0
          Len 12 10.222.6.1/0x0857cd8c
 Hop
  Time
          Len 8 30000 ms
  Style Len 8 SE
 Flow
          Len 36 rate Obps size Obps peak Infbps m 20 M 1500
 Filter7
          Len 12 192.168.32.1 (port/lsp ID
                                           1)
 Label
          Len 8 100112
 RecRoute Len 52 10.222.6.1(Label=100112) 10.222.5.2(Label=100096)
                    10.222.4.1(Label=3)
```

### **Facility Link Protection - Bypass**

#### PLR generates a Path message to establish the Next-Hop Bypass LSP

RSVP send Path 192.168.96.1->192.168.0.1 Len=208 so-0/1/1.0 Session7 Len 16 192.168.0.1 (port/tunnel ID 9546) Proto 0 Hop Len 12 10.222.9.1/0x0857ccc0 Time Len 8 30000 ms SessionAttribute Len 28 Prio (7,0) flag 0x4 "Bypass->10.222.5.2" Sender7 Len 12 192.168.96.1 (port/lsp ID 1) Tspec Len 36 rate Obps size Obps peak Infbps m 20 M 1500 ADspec Len 48 SrcRoute Len 20 10.222.9.2 S 10.222.8.1 S LabelRequest Len 8 EtherType 0x800 RecRoute Len 12 10.222.9.1

### **Facility Link Protection - Bypass**

### PLR receives a Resv message which confirms the Next-Hop Bypass LSP is established

RSVP recv Resv 10.222.9.2->10.222.9.1 Len=128 so-0/1/1.0
Session7 Len 16 192.168.0.1(port/tunnel ID 9546) Proto 0
Hop Len 12 10.222.9.2/0x0857ccc0
Time Len 8 30000 ms
Style Len 8 SE
Flow Len 36 rate Obps size Obps peak Infbps m 20 M 1500
Filter7 Len 12 192.168.96.1(port/lsp ID 1)
Label Len 8 100256
RecRoute Len 20 10.222.9.2 10.222.8.1

# **Facility Link Protection – Notify Ingress**

### PLR sets flags in RRO stating that protection is available

RSVP send H	Resv	10	.222.6.1->10.222.6.2 Len=160 so-0/1/2.0
Session7	Len	16	192.168.16.1(port/tunnel ID 25118) Proto 0
Нор	Len	12	10.222.6.1/0x0857cd8c
Time	Len	8	30000 ms
Style	Len	8	SE
Flow	Len	36	rate Obps size Obps peak Infbps m 20 M 1500
Filter7	Len	12	192.168.32.1(port/lsp ID 1)
Label	Len	8	100112
RecRoute	Len	52	10.222.6.1(flag=1 Label=100112) 10.222.5.2(flag=1 Label=100096)
			10.222.4.1(Label=3)

### **Facility Node Protection - Path**

#### Initial Path message from the ingress is received on the 96.1 router

 Session Attribute flags set to 0x17 for link (0x01), label recording (0x02), SE reservation (0x04), and node protection (0x10)

```
RSVP recv Path 192.168.32.1->192.168.16.1 Len=216 so-0/1/2.0
Session7 Len 16 192.168.16.1 (port/tunnel ID 25121) Proto 0
Hop Len 12 10.222.6.2/0x0857cd8c
Time Len 8 30000 ms
SessionAttribute Len 16 Prio (7,0) flag 0x17 "FRR-Test"
Sender7 Len 12 192.168.32.1 (port/lsp ID 1)
Tspec Len 36 rate Obps size Obps peak Infbps m 20 M 1500
ADspec Len 48
SrcRoute Len 28 10.222.6.1 S 10.222.5.2 S 10.222.4.1 S
LabelRequest Len 8 EtherType 0x800
Properties Len 12 Primary path
RecRoute Len 12 10.222.6.2
```

### **Facility Node Protection - Resv**

# Resv message for the protected LSP is transmitted back to the ingress router RRO contains labels assigned to the LSP

```
RSVP send Resv 10.222.6.1->10.222.6.2 Len=160 so-0/1/2.0
Session7 Len 16 192.168.16.1(port/tunnel ID 25121) Proto 0
Hop Len 12 10.222.6.1/0x0857cd8c
Time Len 8 30000 ms
Style Len 8 SE
Flow Len 36 rate Obps size Obps peak Infbps m 20 M 1500
Filter7 Len 12 192.168.32.1(port/lsp ID 1)
Label Len 8 100128
RecRoute Len 52 10.222.6.1(Label=100128) 10.222.5.2(Label=100112)
10.222.4.1(Label=3)
```

### **Facility Node Protection - Bypass**

### PLR generates a Path message to establish the Next-Next-Hop Bypass LSP

RSVP send Path 192.168.96.1->192.168.16.1 Len=236 so-0/1/1.0 Session7 Len 16 192.168.16.1(port/tunnel ID 9547) Proto 0 Hop Len 12 10.222.9.1/0x0857ccc0 Time Len 8 30000 ms SessionAttribute Len 40 Prio (7,0) flag 0x4 "Bypass->10.222.5.2->10.222.4.1" Sender7 Len 12 192.168.96.1(port/lsp ID 1) Tspec Len 36 rate Obps size Obps peak Infbps m 20 M 1500 ADspec Len 48 SrcRoute Len 36 10.222.9.2 S 10.222.11.2 S 10.222.12.1 S 10.222.7.2 S LabelRequest Len 8 EtherType 0x800 RecRoute Len 12 10.222.9.1

### **Facility Node Protection - Bypass**

### PLR receives a Resv message which confirms the Next-Next-Hop Bypass LSP is established

RSVP recv Resv 10.222.9.2->10.222.9.1 Len=144 so-0/1/1.0
Session7 Len 16 192.168.16.1(port/tunnel ID 9547) Proto 0
Hop Len 12 10.222.9.2/0x0857ccc0
Time Len 8 30000 ms
Style Len 8 SE
Flow Len 36 rate Obps size Obps peak Infbps m 20 M 1500
Filter7 Len 12 192.168.96.1(port/lsp ID 1)
Label Len 8 100272
RecRoute Len 36 10.222.9.2 10.222.11.2 10.222.12.1 10.222.7.2

# **Facility Node Protection – Notify Ingress**

### PLR sets flags in RRO stating that protection is available

RSVP send Resv 10.222.6.1->10.222.6.2 Len=160 so-0/1/2.0 Session7 Len 16 192.168.16.1(port/tunnel ID 25121) Proto 0 Hop Len 12 10.222.6.1/0x0857cd8c Time Len 8 30000 ms Style Len 8 SE Flow Len 36 rate Obps size Obps peak Infbps m 20 M 1500 Filter7 Len 12 192.168.32.1(port/lsp ID 1) Label Len 8 100128 RecRoute Len 52 10.222.6.1(flag=9 Label=100128) 10.222.5.2(flag=1 Label=100112) 10.222.4.1(Label=3)

### **Failure Mode Processing**

#### Protected LSP is using facility link protection

- Link failure between egress and penultimate hop generates PathErr messages upstream
- We see this from the viewpoint of the 96.1 router

RSVP recv PathErr 10.222.5.2->10.222.5.1 Len=84 so-0/1/0.0
Session7 Len 16 192.168.16.1(port/tunnel ID 25126) Proto 0
Error Len 12 code 25 value 3 flag 0 by 10.222.5.2
Sender7 Len 12 192.168.32.1(port/lsp ID 1)
Tspec Len 36 rate 0bps size 0bps peak Infbps m 20 M 1500

RSVP send PathErr 10.222.6.1->10.222.6.2 Len=84 so-0/1/2.0
Session7 Len 16 192.168.16.1(port/tunnel ID 25126) Proto 0
Error Len 12 code 25 value 3 flag 0 by 10.222.5.2
Sender7 Len 12 192.168.32.1(port/lsp ID 1)
Tspec Len 36 rate 0bps size 0bps peak Infbps m 20 M 1500

### **Failure Mode Processing**

### Next-Hop Bypass LSP is currently in use

 Flags in received RRO on the 96.1 router tell us that protection is in use

RSVP recv Resv 10.222.5.2->10.222.5.1 Len=144 so-0/1/0.0 Session7 Len 16 192.168.16.1(port/tunnel ID 25126) Proto 0 Hop Len 12 10.222.5.2/0x0857cbf4 Time Len 8 30000 ms Style Len 8 SE Flow Len 36 rate Obps size Obps peak Infbps m 20 M 1500 Filter7 Len 12 192.168.32.1(port/lsp ID 1) Label Len 8 100144 RecRoute Len 36 10.222.5.2(flag=3 Label=100144) 10.222.7.2(Label=3)

## **Questions and Comments**

### Feedback on this presentation is highly encouraged

\* jms@juniper.net

### Questions?





# Thank you!

http://www.juniper.net

