MPLS HIGH AVAILABILITY: COMPONENTS AND MANAGEMENT

MARCH 2004

Agenda

Cisco.com

- MPLS HA Components
  - MPLS HA – Border Gateway Protocol VPNv4 NSF with SSO
  - MPLS HA – Traffic Engineering NSF with SSO
  - MPLS HA – Fast Reroute (FRR) NSF with SSO
  - MPLS HA – Any Transport over MPLS (AToM) NSF with SSO
  - MPLS HA – LC-ATM NSF with SSO
- MPLS HA – Management
- Summary

MPLS HA – BGP VPNV4
CISCO NSF WITH SSO

MPLS BGP VPNv4 Graceful Restart Benefits

- **Minimizes the negative effects on MPLS forwarding caused by the LSRs control plane restart**
- **Preserves the MPLS forwarding state (VPN routes), across BGP restart (RP switchover- control plane)**
- **The graceful restart mechanism for BGP VPNv4 routes is identical to the mechanism used by BGP for IPv4 routes, as both preserve IP forwarding state**

  draft-ietf-mpls-bgp-mpls-restart-02.txt

  draft-ietf-idr-restart-08.txt

MPLS Forwarding State

MPLS forwarding state could include:

- incoming label → (outgoing label, next hop)
- address prefix → (outgoing label, next hop)

MPLS BGP VPNv4 Graceful Restart Mechanisms

Cisco.com

- Graceful Restart Mechanism for BGP

  draft-ietf-idr-restart-08.txt

- Graceful Restart Mechanism for BGP with MPLS

  draft-ietf-mpls-bgp-mpls-restart-02.txt

- New Graceful Restart Capability is defined and carried in BGP Open message

  BGP Update message with no reachable NLRI and empty withdrawn NLRI is specified as an End-of-RIB marker

  (Applicable) AFI/SAFI pairs in the Graceful Restart Cap is used by an LSR for indicating its ability to preserve MPLS forwarding state across BGP restart

  SAFI in the advertise capability indicates that NLRI field carries addressing information as well as labels

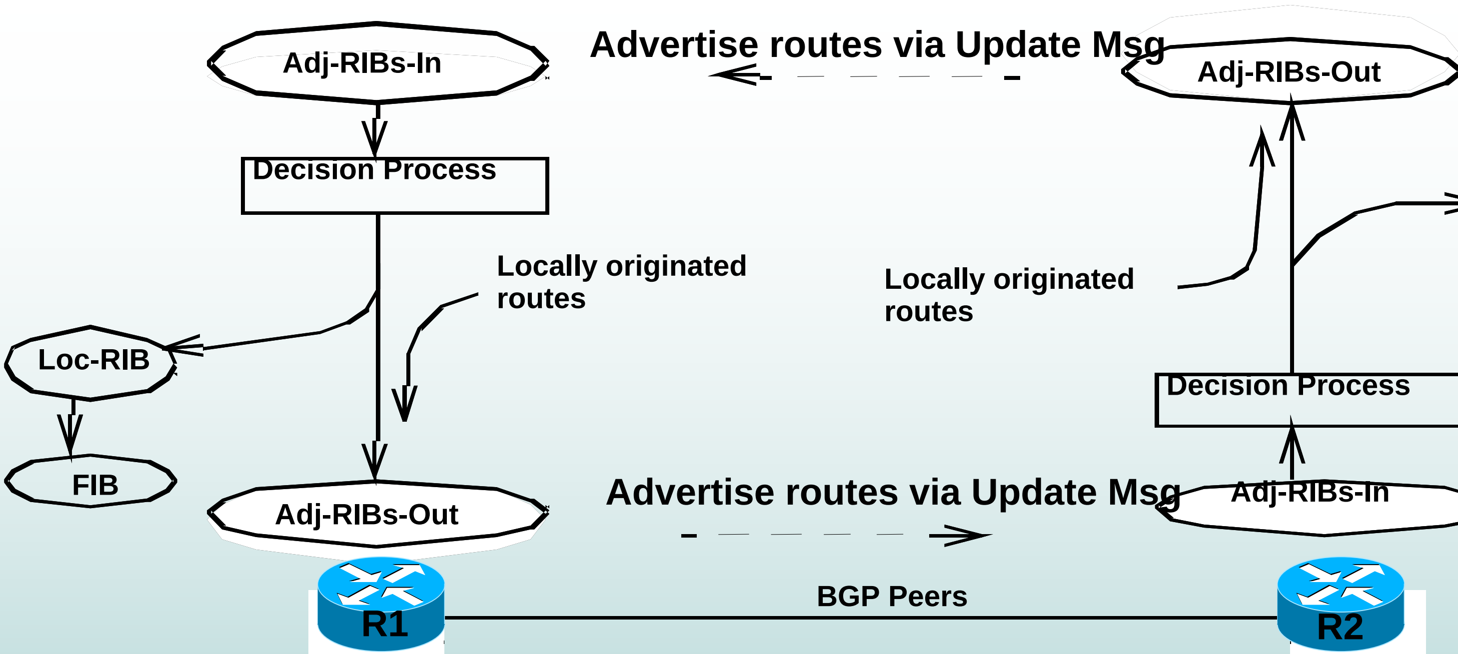
- Mechanism does not require any of the BGP-related state to be preserved across the restart

  This should be handled by BGP GR mechanism

MPLS BGP VPNv4 Graceful Restart Mechanism Key points

- Defines how to preserve forwarding state across BGP restart
- Allows a router to create MPLS forwarding entries for VPNv4 prefixes in NSF mode
    - Forwarding entries are preserved during a restart
- Saves prefix and corresponding label information and recovers the information after a restart
- Assumes that only the actual MPLS forwarding state must be preserved
- Does not require any of the BGP-related state to be preserved across the restart
- If label binding on an LSR is created/maintained by multiple protocols (LDP, RSVP-TE) individually, LSR must preserve the information about which protocol has assigned which labels across the restart

BGP Functional Overview



- **Adj-RIBs-In: Routing info learned from BGP peers via Update message.**
- **Decision Process: Takes Adj_RIBs-In as input and selects the best routes**
- **Loc-RIB: Local BGP routing info selected by the decision process**
- **FIB: Info used for actual forwarding**
- **Adj-RIBs-Out: Routing info selected for advertisement to BGP peers.**

MPLS BGP VPNv4 Graceful Restart Process

- LSR Restarts
- BGP GR executed and best route is selected
- Restarting LSR performs one the additional BGP VPNv4 GR Procedures based on following conditions:
  1. The best route selected by the restarting LSR was received with a non-Implicit Null label, and the LSR advertises this route with itself as the Next Hop
  2. Best route selected by the restarting LSR was originated by LSR or received either without a label or with an Implicit NULL label, the LSR generates a (non Implicit NULL) label for the route and the LSR advertises this route with itself as the Next Hop
  3. Restarting LSR does not set BGP NH to itself
  4. Alternate method is not supported because it requires pre-allocating range of labels

MPLS BGP VPNv4 Graceful Restart Process Condition

- **If the best route selected by the restarting LSR was received with a non-Implicit Null label, and the LSR advertises this route with itself as the Next Hop:**

  Restarting LSR searches its MPLS forwarding state (what was preserved across the restart) for an entry with <outgoing label, NH> equal to the one in the received route

  If found, the LSR removes stale state

  If the entry is <incoming label, (outgoing label, NH)> rather than <prefix, (outgoing label, NH)>, the LSR uses the incoming label from the entry when advertising the route to its neighbors

  If the found entry has no incoming label, the LSR uses any unused label when advertising the route to its neighbors

MPLS BGP VPNv4 Graceful Restart Process Condition (Cont.)

- If the best route selected by the restarting LSR was originated by LSR or received either without a label or with an Implicit NULL label, the LSR generates a (non Implicit NULL) label for the route and the LSR advertises this route with itself as the Next Hop
- LSR searches its MPLS forwarding state for an entry that indicates that the LSR has to perform label pop, and the NH equal to the NH of the route in consideration

    If this entry is found, the LSR uses the incoming label from the entry when advertising the route to its neighbors

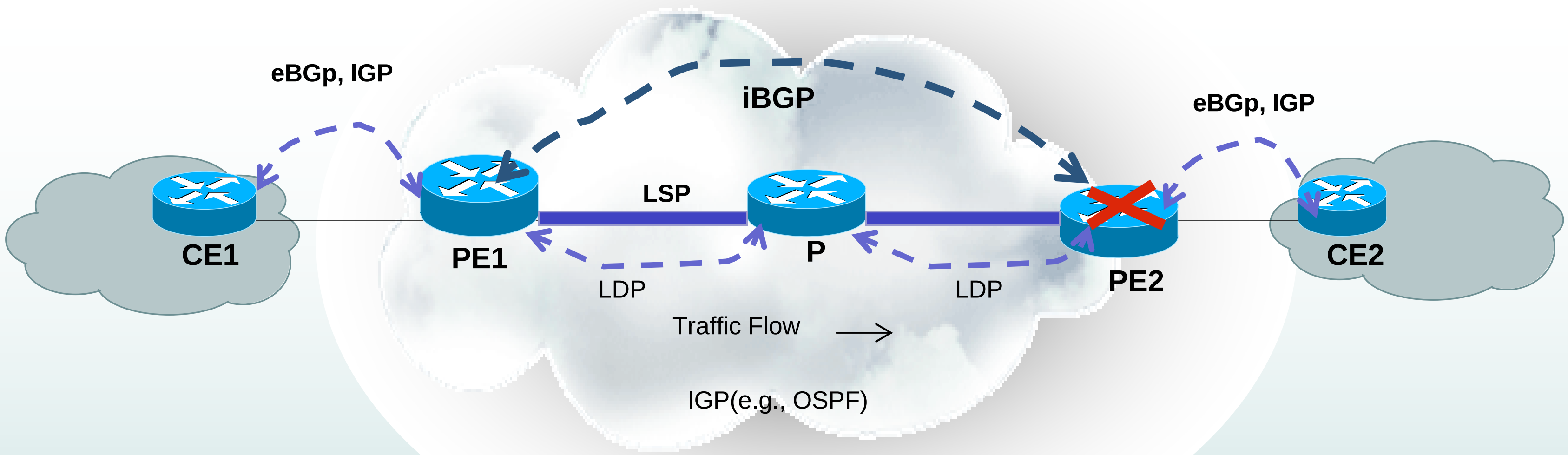
    Otherwise, LSR takes any unused label when advertising the route to its neighbors

MPLS BGP VPNv4 Graceful Restart Process Condition (Cont.)

Cisco.com

- **If restarting LSR does not set BGP NH to itself, the restarting LSR will use the label received with the route or advertise the route without a label**

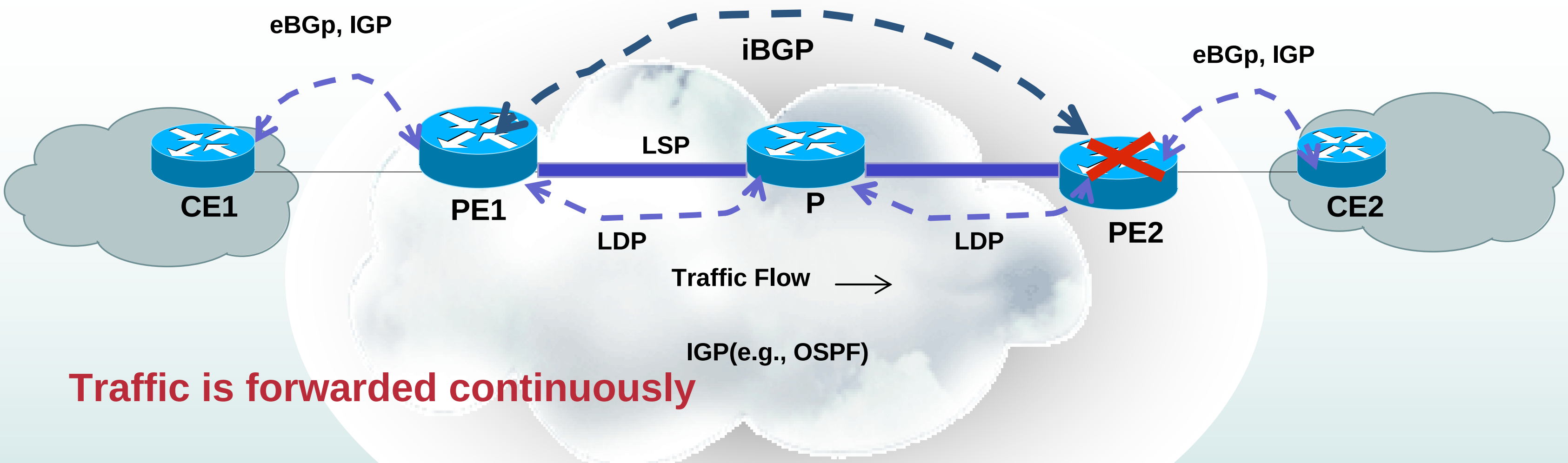
MPLS VPN – BGP Graceful Restart Procedure Overview



1. BGP routing information (select routes, advertise label, etc.) and BGP Restart Capability (restart timer, AFI/SAFI, etc.) exchanged between customer edge (CE) and provider edge (PE) routers
2. Link-state packets (LSPs) exchanged (via LDP) from PE1 to PE2
3. BGP/MPLS in PE2 restarts (RP failure)
4. PE1 detects restart, marks MPLS forwarding information from PE2 as stale, continues forwarding traffic using stale information
5. PE1 sends BGP updates to PE2, & EoRib marker
6. PE2 reconverges and sends current MPLS forwarding information to PE1
7. PE1 reconverges and flushes stale information

Traffic is continuously forwarded

MPLS VPN – BGP Graceful Restart Procedure



Traffic is forwarded continuously

1. PE1 and PE2 are IBGP neighbors and exchange VPNv4 routes
2. Since PE1 and PE2 are configured for BGP GR, they also exchange the GR capability in the OPEN messages they send to each other during BGP session initialization
3. PE2 is the router which will be restarted (active RP fails, switches over over to backup RP)
4. PE2 syncs the local label to prefix mapping in its BGP VPN table to the standby RP

MPLS VPN – BGP Graceful Restart Procedure (Cont.)

5. CEF table and the Label Forwarding Database (LFD) are also synced to the standby RP
6. Label switching database (LSD) which is responsible for label allocation to the LDM (Label Distribution modules) also syncs over blocks of allocated labels to the standby RP. It does this so that after switchover, the new active RP does not allocate the already allocated labels to another LDM or another prefix.
7. Now a switchover happens on PE2
8. The BGP session between PE1 and PE2 goes down.
9. PE1 marks the entries in its BGP table which it learnt from PE2 as stale but does not delete the routes from its VRF tables. Hence it continues to forward traffic to PE2.
10. PE2 also maintains its forwarding capability by maintaining its CEF and LFD on the linecards. Hence it is capable of forwarding traffic arriving from CE and going towards PE1 as well as traffic coming from PE1 and going towards CE.
11. The BGP session between PE1 and PE2 comes back up. PE1 needs to see the session come back up within the restart time (default 180s). If not, it is going to delete all the stale routes from the BGP table and hence the routing table.

MPLS VPN – BGP Graceful Restart Procedure (Cont.)

12. Once the BGP session comes back up, PE1 advertises all the routes in its Adj-RIB-out to PE2 along with the label mapping.
13. PE2 receives these updates which contain the prefix to outgoing label mapping. BGP has synced over the prefix to incoming label mapping to the RP prior to switchover.
14. BGP on PE2 will wait for all of its restarting peers to complete resending their updates to PE2. When all the updates are received, BGP starts its route selection process.
15. BGP on PE2 now lets IPRM (IP Resource Manager) know that it has learnt a new outgoing label for the prefix. IPRM is the module which sits in the middle of the LDM (LDP, TE, VPN etc) and MFI and handles the interaction between them.
16. IPRM now installs the rewrite (incoming and outgoing labels for the prefix) into the LSD which then distributes it to the LFD on the LC.
17. LFD on the LC installs the rewrite into the CEF on the LC.
18. On PE2, after BGP has run its route selection process, it populates its Adj-RIB-out which it advertises to PE1.
19. Once PE1 receives the updates from PE2, it removes the stale marking from the BGP prefixes. If PE1 does not receive these updates within the stalepath time (360s by default), it deletes all its stale entries from its BGP table and hence the routing table.

BGP VPNv4 Cisco NSF with SSO Support Requirements

- Elements that enable VPN NSF
- BGP GR ensures prefix information is preserved during a restart
- BGP VPNv4 Checkpoint
- What happens if a router does not have VPN NSF enabled?

Elements that enable BGP VPNv4 NSF

- **BGP GR mechanisms defined in IETF draft and in the Cisco Nonstop Forwarding feature module**
- **NSF support for the label distribution protocol in the core network**
    - **Either MPLS LDP, TE, or static labels**
- **NSF support for the Internal Gateway Protocol (IGP) used in the core (Open Shortest Path First (OSPF) or Intermediate System-to-Intermediate System (ISIS)**
- **NSF support for the routing protocols between the PE and CE routers**

BGP GR Ensures Prefix Information Is Preserved During a Restart

BGP GR as explained in earlier slides.

BGP VPNv4 Checkpoint

Cisco.com

- **VPN Prefix Information Is Checkpointed from the active Route Processor to the backup Route Processor**
  - **Checkpointing is a function that copies state information from the active RP to the backup RP to ensure the backup RP has an identical copy of the latest information**
  - **Checkpointing begins when the active RP does a bulk synchronization, which copies all the local label bindings to the backup RP**
  - **Active RP dynamically checkpoints individual prefix label bindings when a label is allocated or freed**
- **Allows forwarding of labeled packets to continue before BGP re-converges**
- **When BGP allocates local labels for prefixes, it checkpoints the local label binding {prefix, label} in the backup RP**

What Happens If A Router Does Not Have VPN NSF Enabled?

- **If a router is not configured for VPN NSF and it attempts to establish a BGP session with a router that is configured with VPN NSF, the two routers create a normal BGP session but do not have the ability to perform VPN NSF**

MPLS L3 VPN Features

- MPLS VPN BGP GR for VPNv4
- MPLS VPN BGP Checkpointing
- MPLS VPN SSO/NSF Support for VRF
- MPLS VPN SSO/NSF Support for I-AS
- MPLS VPN SSO/NSF Support for CSC
- Refer to Cisco Feature Navigator for the latest information:

    www.cisco.com/go/fn/

How to enable BGP VPNv4 NSF with SSO

Cisco.com

- **Enabling SSO**

  www.cisco.com/univercd/cc/td/doc/product/software/ios122s/122snwft/release/122s18/sso18s.htm

- **Enabling LDP Graceful Restart**

  www.cisco.com/univercd/cc/td/doc/product//product/software/ios122s/122snwft/release/122srls4

- **Enabling Nonstop Forwarding for Routing Protocols**

  www.cisco.com/univercd/cc/td/doc/product/software/ios122s/122snwft/release/122s18/nsf18s.htm

- **Enabling NSF Support for Basic VPNs**
- **Configuring NSF support for VPN interfaces that use BGP for label distribution**
- **Verifying VPN NSF (Optional)**

BGP VPNv4 GR Configuration

- **Enable MPLS nonstop forwarding on an interface that uses BGP as the label distribution protocol**

```
mpls forwarding bgp
```

BGP VPNv4 GR Restrictions

- LDP sessions are supported, but Tag Distribution Protocol (TDP) sessions are not supported
- VPN NSF cannot be configured on label-controlled ATM (LC-ATM) interfaces
- VPN NSF requires that neighbor networking devices be NSF-aware

  Peer routers must support the graceful restart of the protocol used to communicate with the VPN NSF-capable router

BGP VPNv4 Graceful Restart Troubleshooting

Show ip bgp labels	Displays information about MPLS labels from eBGP route table
Show ip bgp vpnv4 all labels	Displays info on the active and standby RPs when they are configured for MPLS VPN NSF
Debug ip bgp vpnv4 checkpoint	Displays the events for the VRF checkpointing system between the active and standby RP
Debug ip bgp vpnv4 nsf	Displays the NSF events for the VRF tableid synch subsystem between the active and standby RPs
Debug mpls checkpoint label-binding	Displays the events for the checkpoint label bindings of MPLS applications running on the router

Show ip bgp labels

Displays EBGP labels associated with an ASBR

```
Router#show ip bgp labels
Network            Next Hop           In Label/Out Label
3.3.0.0/16         0.0.0.0            imp-null/exp-null
15.15.15.15/32     15.15.15.15        18/exp-null
16.16.16.16/32     0.0.0.0            imp-null/exp-null
17.17.17.17/32     34.0.0.1           20/exp-null
18.18.18.18/32     43.0.0.1           24/31
18.18.18.18/32     38.0.0.1           24/33
19.19.19.19/32     43.0.0.1           25/32
19.19.19.19/32     38.0.0.1           25/34
20.20.20.20/32     43.0.0.1           21/30
20.20.20.20/32     38.0.0.1           21/32
33.0.0.0           15.15.15.15        19/exp-null
34.0.0.0           0.0.0.0            imp-null/exp-null
35.0.0.0           43.0.0.1           22/29
```

Show ip bgp vpnv4 all labels

Cisco.com

- NH and network prefixes are from BGP table
- From the same LSR, Info on Standby RP
- InLabels are local labels, which can be verified by comparing the log to SHOW IP BGP LABELS. OutLabels are assigned by BGP NH router. They should be the same, because BGP is used as a label distribution protocol
- From the same LSR, Info on Active RP

```
Router#show ip bgp vpnv4 all labels
Network          Next Hop    In label/Out label
Route Distinguisher: 100:1 (vpn1)
12.12.12.12/32   0.0.0.0     16/aggregate(vpn1)
135.0.0.0/8      0.0.0.0     17/aggregate(vpn1)
Route Distinguisher: 609:1 (vpn0)
13.13.13.13/32   0.0.0.0     18/aggregate(vpn0)
```

- From the same LSR, Info on Standby RP

```
Router#show ip bgp vpnv4 all labels
Network        Masklen    In label
Route Distinguisher: (dec) 0001000001
12.12.12.12    /32         16
135.0.0.0      /8          17
Route Distinguisher: (dec) 002970001
13.13.13.13    /32         18
```

Debug ip bgp vpnv4 checkpoint

Displays the events for the VRF checkpointing system between the active and standby RP

```
Router#debug ip bgp vpnv4 checkpoint
3d18h: %HA-5-SYNC_NOTICE: Config sync started.
3d18h: vrf-nsf: vrf vpn2 tableid 1 send OK
3d18h: vrf-nsf: vrf tableid bulk sync
complete msg send OK
3d18h: vrf-nsf: CF send ok
3d18h: vrf-nsf: CF send ok
3d18h: %HA-5-SYNC_NOTICE: Config sync completed.
3d18h: %HA-5-SYNC_NOTICE: Standby has restarted.
3d18h: %HA-5-
MODE: Operating mode is sso,config mode is sso
```

Debug ip bgp vpnv4 nsf on the Active RP

Router#debug ip bgp vpnv4 nsf

```
MPLS VPN NSF Processing debugging is on
Router(config)#ip vrf vpn3

3d18h: vrf-nsf: vrf vpn3 tableid 2 send rpc OK

Router(config-vrf)#no ip vrf vpn3

%IP addresses from all interfaces in VRF vpn3 have been rem
oved

3d18h: vrf-nsf:
rx vrf tableid delete complete msg, tid = 2, name = vpn3
```

Debug ip bgp vpnv4 nsf on the Standby RP

Router#debug ip bgp vpnv4 nsf

```
MPLS VPN NSF Processing debugging is on
00:05:21: vrf-
nsf: rx vrf tableid rpc msg, tid = 2, name = vpn3
%IP addresses from all interfaces in VRF vpn3 have been rem
oved 00:06:22: vrf-

nsf: vrf vpn3 tableid 2 , delete complete, send OK
```

Debug mpls checkpoint label-binding

Cisco.com

Output log when the command is issued on the standby route processor

Router#debug mpls checkpoint label-binding

```
MPLS Label Binding Checkpoint debugging is on
3d17h: mpls_lbl_bind_chkpt: client ID 13 up, total client 1
3d17h: mpls_lbl_bind_chkpt: msg rx for 1D, vers 0, type 1
action 56, len 0, state 4, peer 13
3d17h: mpls_lbl_bind_chkpt: post msg type 1
3d17h: mpls_lbl_bind_chkpt: msg rx for 1D, vers 0, type 1
action 56, len 0, state 4, peer 13
3d17h: mpls_lbl_bind_chkpt: post msg type 1
3d17h: mpls_lbl_bind_chkpt: msg rx for 1D, vers 0, type 1
action 56, len 0, state 4, peer 13
3d17h: mpls_lbl_bind_chkpt: post msg type 1
3d17h: mpls_lbl_bind_chkpt: appl_id 13, KEY 000C800018888200
3d17h: mpls_chkpt_db: AVL insert successful, Key 000C800018888200
action Add, label 19
3d17h: mpls_lbl_bind_chkpt: appl_id 13, KEY 000C800013200080
3d17h: mpls_chkpt_db: AVL insert successful, Key 000C800013200080
action Add, label 20
3d17h: mpls_lbl_bind_chkpt: appl_id 13, KEY 000C80001383838200
3d17h: mpls_chkpt_db: AVL insert successful, Key 000C80001383838200
action Add, label 21 3d17h: Stby RP OR CF peer not ready, don't send msg
3d17h: mpls_lbl_bind_chkpt: client ID 13 down, total client 0
3d17h: mpls_lbl_bind_chkpt: msg rx for 1D, vers 0, type 1
action 56, len 2, state 4, peer 13
3d17h: mpls_lbl_bind_chkpt: post msg type 1
3d17h: mpls_lbl_bind_chkpt: appl_id 13, KEY  action NSF unconfig, appl id 13
```

Requirements to support VPNv4 GR

- **All peering LSRs need to support Restart to take full advantage OF VPNv4 GR**
- **Default to nonGR BGP if one of the neighbors doesn’t support it, it will default back to original BGP**
- **If label binding on an LSR is created/maintained by multiple protocols(LDP, RSVP-TE) individually, LSR needs to preserve across the restart the info about which protocol has assigned which labels**

MPLS HA – AToM
CISCO NSF WITH SSO

MPLS HA – AToM

- **GR is exactly the same because it is directed, AToM only checks for local labels**

MPLS HA: RSVP-TE
CISCO NSF WITH SSO

Why MPLS TE NSF with SSO?

- **No forwarding impact to existing tunnels**
    - **Subject to hardware restrictions**
- **Fast recovery of signaling and control state for existing tunnels**
- **Basic RSVP signaling must be recovered from neighbors**
    - **Re-establish conversations with neighbors**
    - **Rebuild control state for existing LSPs**
    - **Reclaim existing local labels and bindings**
- **System ready for new requests immediately**

MPLS/TE Checkpoint Requirements

- Endpoint information is rebuilt from configuration
  - No checkpointing required
- Topology DB and PCALC only used for new tunnel setup
  - Users can wait for IGP resynchronization
- There is a priority inversion problem in RRR/LM
  - Solved by checkpointing RRR/LM bandwidth information
- Fast reroute tunnel selection is a local hop decision
  - Needs to be checkpointed for actively rerouted tunnels
- Good to avoid local label reallocation
  - Solved by checkpointing local labels
  - Alternative: RSVP protocol extension

MPLS HA–TE–RSVP Mechanism

- Defined in section 9 of GMPLS-RSVP(TE) RFC 3473
- Restart_Cap object is carried in RSVP-TE Hello messages
- Presence of Restart_Cap object indicates support for RSVP-TE HA

0 7 15	23 31
Length	Class-Num(131) C-Type 1
Restart Time (in milliseconds)	
Recovery Time (in milliseconds)	

MPLS HA–TE–RSVP Mechanism (Cont.)

Restart_Cap includes two timers

- Restart time

  Should be set to the sum of the time it takes the sender of the object to restart its RSVP-TE component and the communication channel (used for RSVP communication)

  Sender uses a value of 0xffffffff to indicate GR restart

- Recovery time

  Period of time that the sender desires for the recipient to re-synchronize RSVP and MPLS forwarding state with the sender after the re-establishment of Hello synchronization

  Value of zero (0) indicates that MPLS forwarding state was not preserved across a particular reboot

MPLS HA–TE–RSVP Key points

Restart_Cap object supports two types of control communication faults recovery

- **Control Channel Faults**

  Control communication is lost between two nodes

  Node refreshes all the states shared with the neighbors within the Recovery time advertised to restart and recover by the neighbor

- **Nodal Faults**

  Node maintains data forwarding state even though the control state is lost (ie: after a restart)

  Recovery_Label object is used during the nodal fault recovery process (discussed later in Nodal Fault Recovery slides)

MPLS HA–TE–RSVP
Modifications to Hello Processing

- **Invoking Restart Procedures**

  Node will wait the Restart Time indicated by the neighbor before invoking procedures related to communication loss

- **Wait time**

  During the Restart Time wait, all Hello messages are sent with Dst_Instance value to 0 and original Src_Instance value and current states are preserved

  Refreshing of RESV and Path state are suppressed during the Restart Time(The established LSPs, node will continue to receive periodic RSVP refresh messages from the neighbor)

- **Informing the neighbor**

  Communication loss(or recovered state) can be informed to upstream neighbor via "Control Channel Degraded State" indication, during and after Restart Time

- **After wait time**

  A new Hello message is received from a neighbor and a type of fault is determined based on Src_Instance value. If the value is different, the neighbor is concluded to have restarted

MPLS HA–TE–RSVP: Procedures for the Restarting Node (Nodal fault)

- **Node restarts its control plane**
- **Node checks if it was able to preserve its MPLS forwarding state (set recovery time to 0 if failed to preserve**
- **Node initiates the state recovery process**
- **State that is not resynchronized during the recovery period is removed at the end of Recovery Timer expiration**

MPLS HA–TE–RSVP: Procedures for a neighbor of a Restarting Node

- **Outgoing Path messages include a Recovery_Label object containing a label value corresponding to the label value received in the most recently received corresponding Resv message**
- **A node refreshes all Path state shared with the restarting neighbor within 1/2 of the Recovery time advertised by the restarted neighbor**
- **All Resv state shared with the restarting node are refreshed only after a corresponding Path message is received**
- **Normal state processing is re-enabled right after the corresponding path message is received**

MPLS HA – FRR
CISCO NSF WITH SSO

MPLS HA – FRR

- **FRR is same as TE**

MPLS HA – LC-ATM
CISCO NSF WITH SSO

MPLS HA – LC-ATM

- **LC-ATM is DoD LDP GR**

ATM Edge LSR

- ATM edge LSRs are routers with ATM MPLS sub interfaces
- Segmentation and reassembly is done at the ATM edge LSRs
- In the ATM MPLS, all ATM MPLS LSPs are initiated and terminated at the ATM edge LSR
- In the ATM edge LSR, traffic is processed in the packet level
- ATM edge LSR uses the Downstream On Demand LDP protocol

ATM-LSR

- ATM-LSR is a core ATM MPLS device
- There is no segmentation or reassembly at the ATM-LSR
- MPLS LSPs should not be initiated or terminated at the ATM LSRs
- LSPs are only established between the ATM edge LSRs
- Traffic is only handled in the cell level. Packet level processing is not possible
- ATM-LSR uses the Downstream On Demand LDP protocol

Label Switch Controller

- **In the Label Switch Controller (LSC), the data forwarding is done by the ATM switch and control protocols are run at the routers**

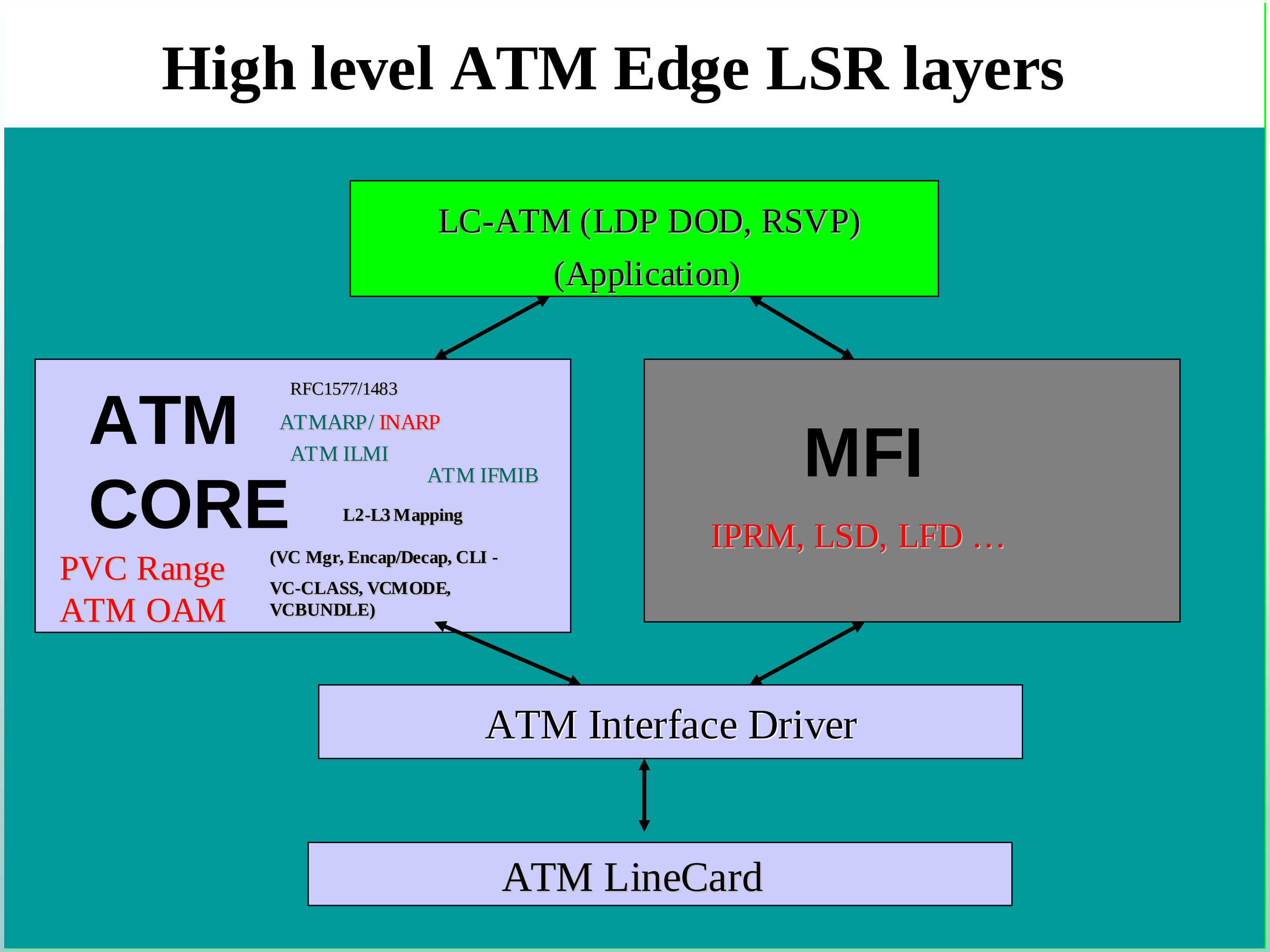
  Data and control planes are separated by two different hardware

- **Inherently these switches are capable of NSF**

  It can forward traffic, even in the absence of the controller

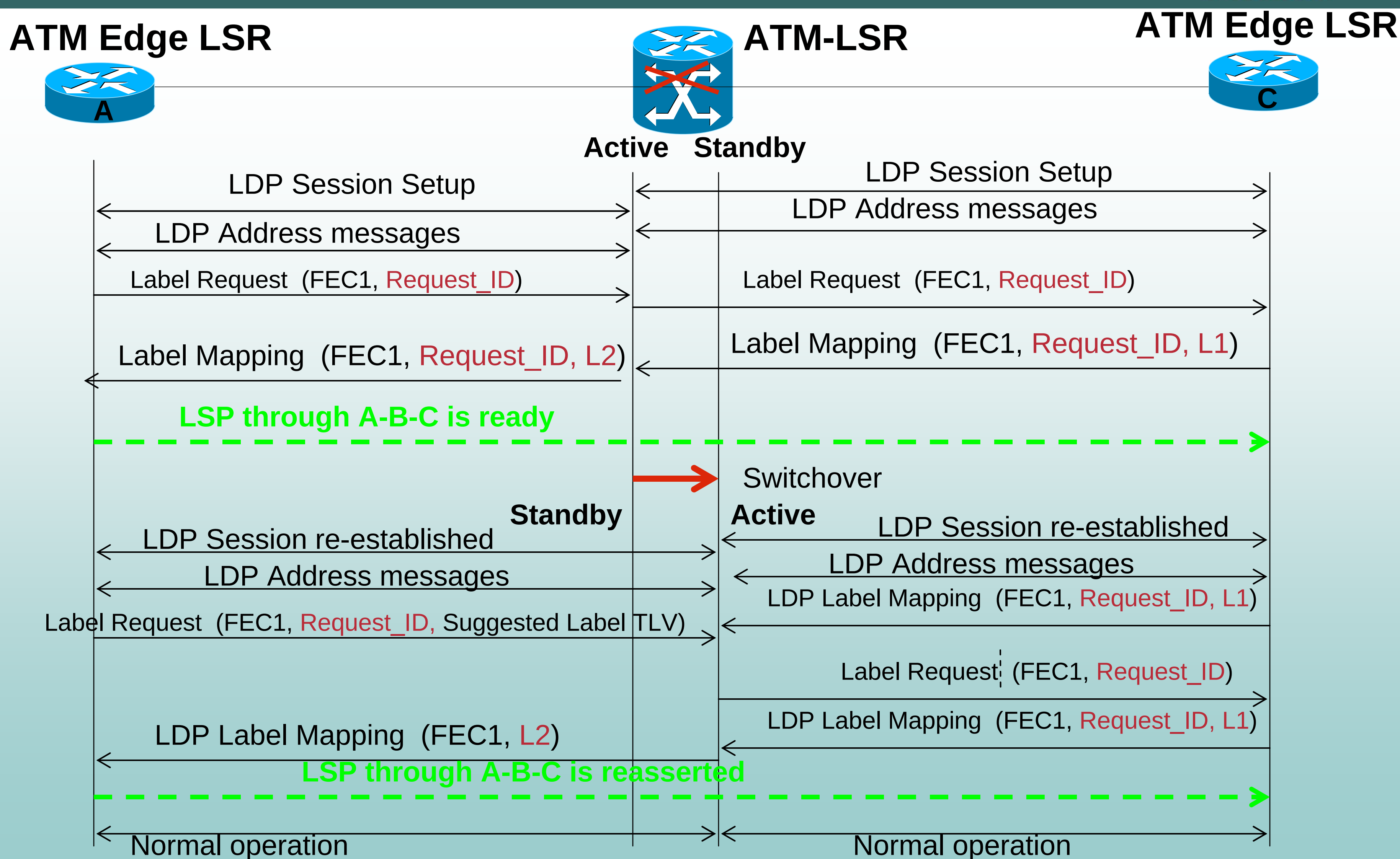
- **Switch naturally has the capabilities for recovery**

ATM Edge LSR Control Flow



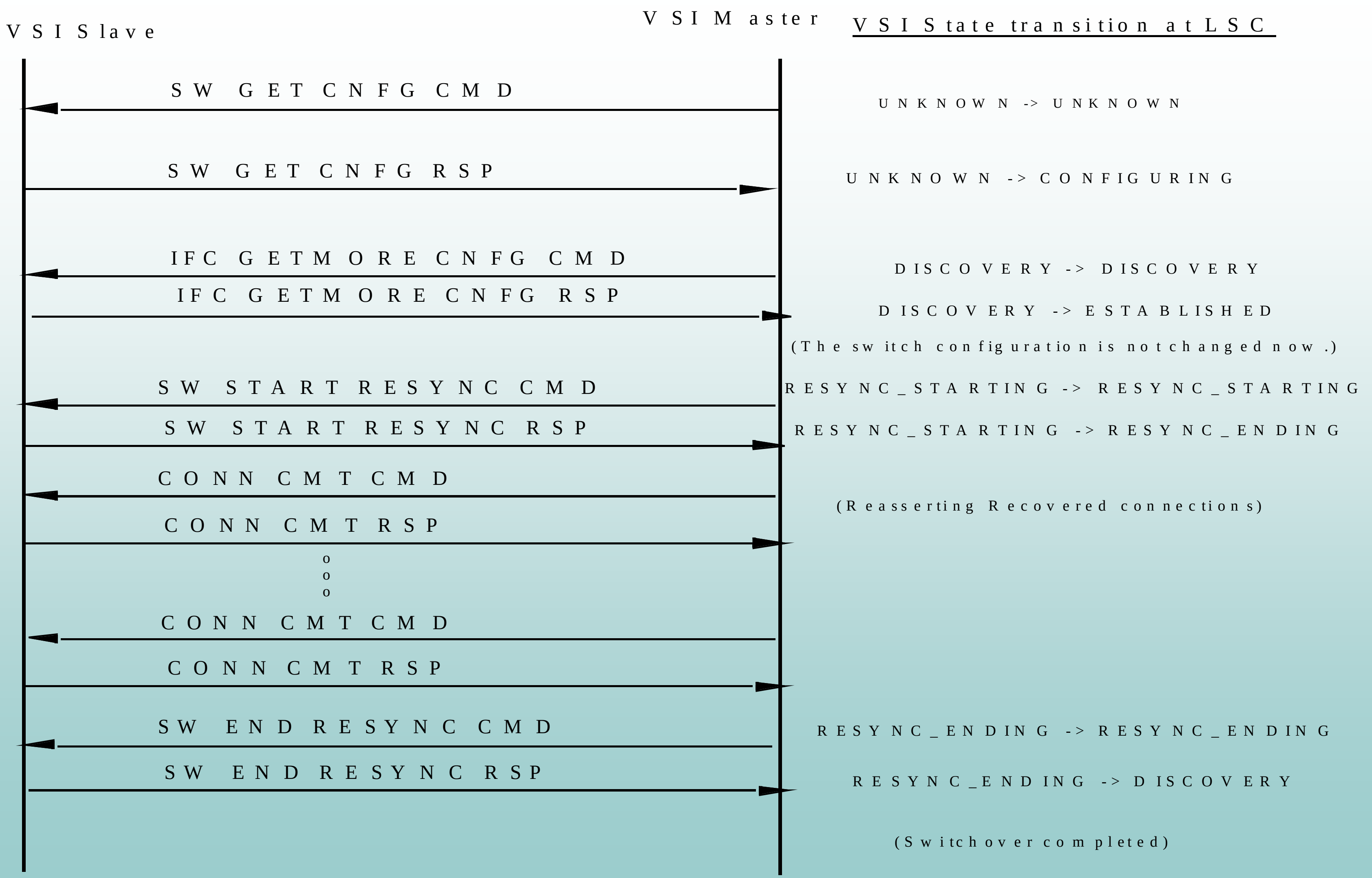
LDP DOD Graceful Restart

Cisco.com



VSI Protocol Changes for ATM-LSR SSO

Cisco.com

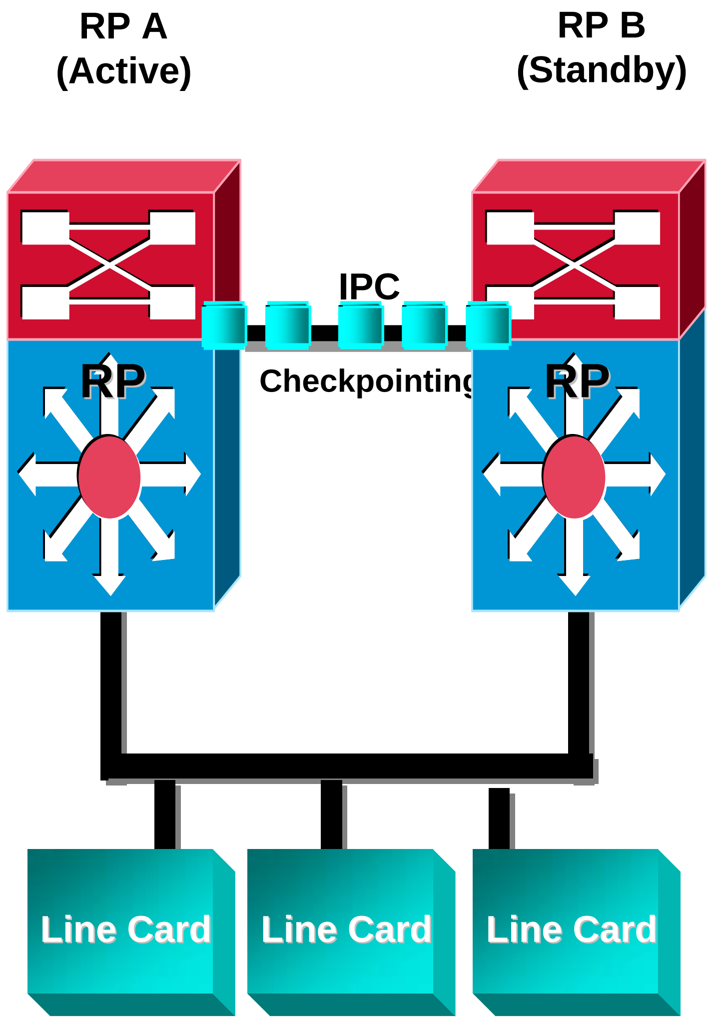


ATM Edge LSR HA

- Cell-based MPLS needs to recover ATM label (VPI/VCI), VCD and local packet label
- Layer 2 ATM SSO already saves and recover the VPI/VCI and associated VCD for the PVCs
- LVCs are similar to the UBR PVCs. Therefore, the new type PVC SSO can be used to recover the LVCs Layer 2 states
- PVC-based packet MPLS needs to recover PVCs Layer 2 states along with the MPLS SSO. Layer 2 PVC SSO is done by ATM SSO

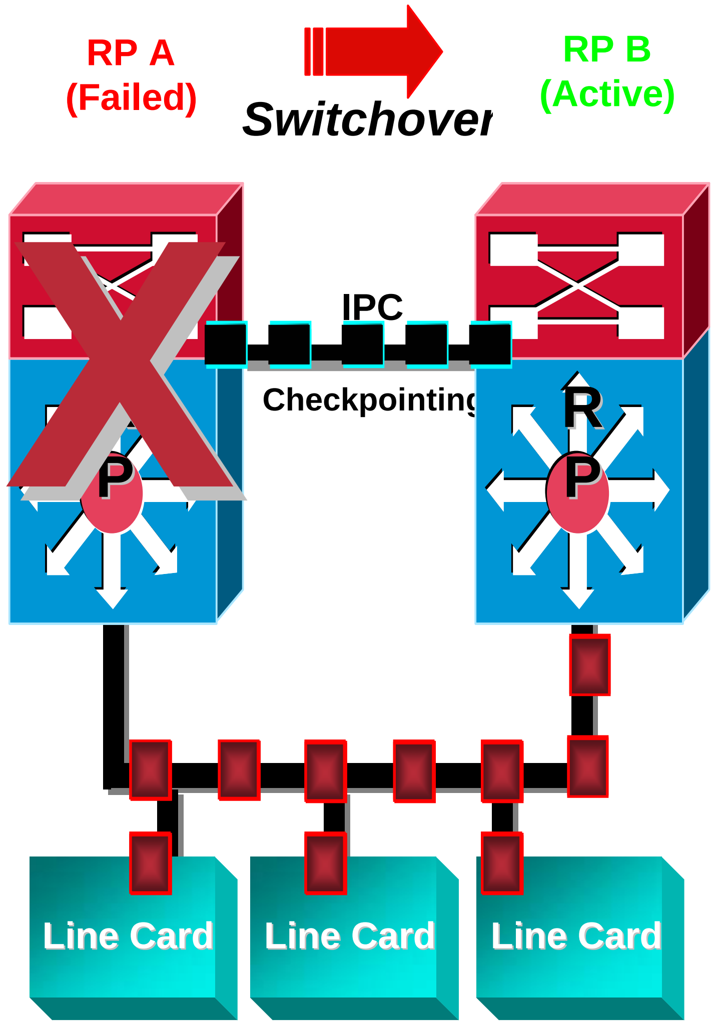
ATM Edge LSR SSO Before switchover

- MPLS LDP DOD starts in the Active RP and LVCs are established normally
- LSR can be a ingress LSR for some Head end LSPs and it can be a egress LSR for other Tail end LSPs
- ATM SSO checkpoints all the information needed for recovering the PVCs and LVCs on the standby RP
- Active RP LDP checkpoints the packet local label on the standby RP
- MFI also checkpoints the packet local label for the forwarding entries



ATM Edge LSR SSO
RP A Failed

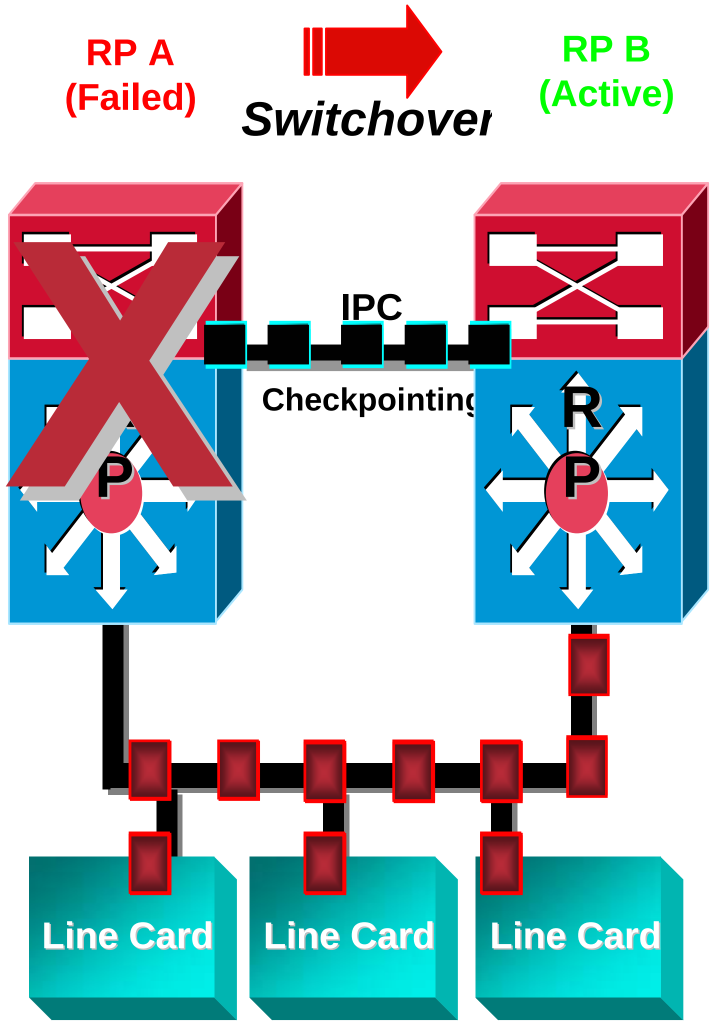
- Active RP fails and Standby RP becomes Active RP. MPLS Forwarding continues on LCs
- MFI marks the HA application’s MPLS forwarding entries as stale



ATM Edge LSR SSO
After switchover

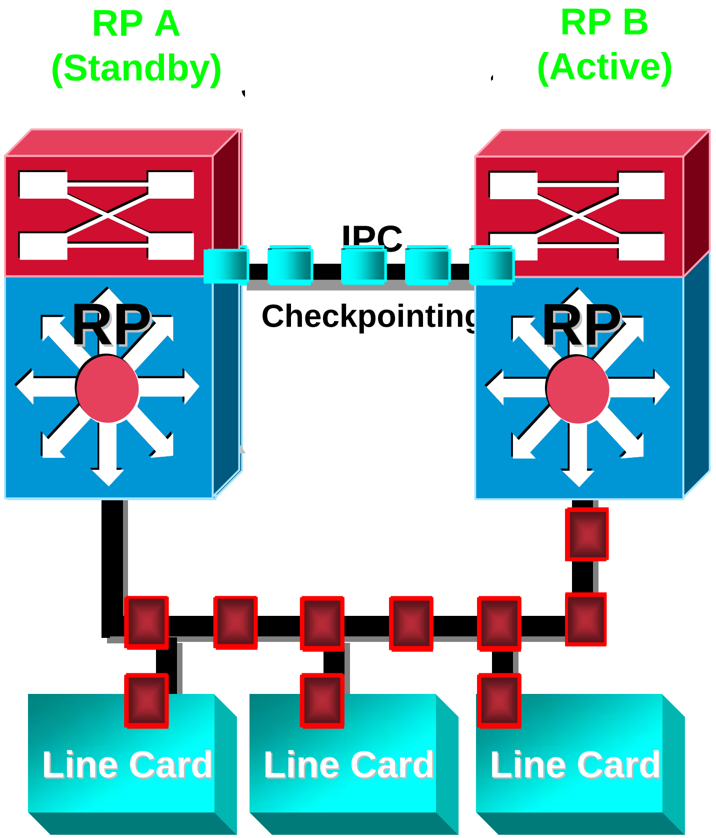
Cisco.com

- Active RP fails and Standby RP becomes Active RP
  - MPLS Forwarding continues on LCs
- Head end LVC Recovery at RP B
- After switchover, standby RP becomes active RP. Tail end needs
  - Outgoing ATM label
  - Local packet label
- LSR sends Label Requests for the eligible FECs
- LSR receives the Label Mapping from neighbors and recovers the outgoing labels from the neighbor LSR

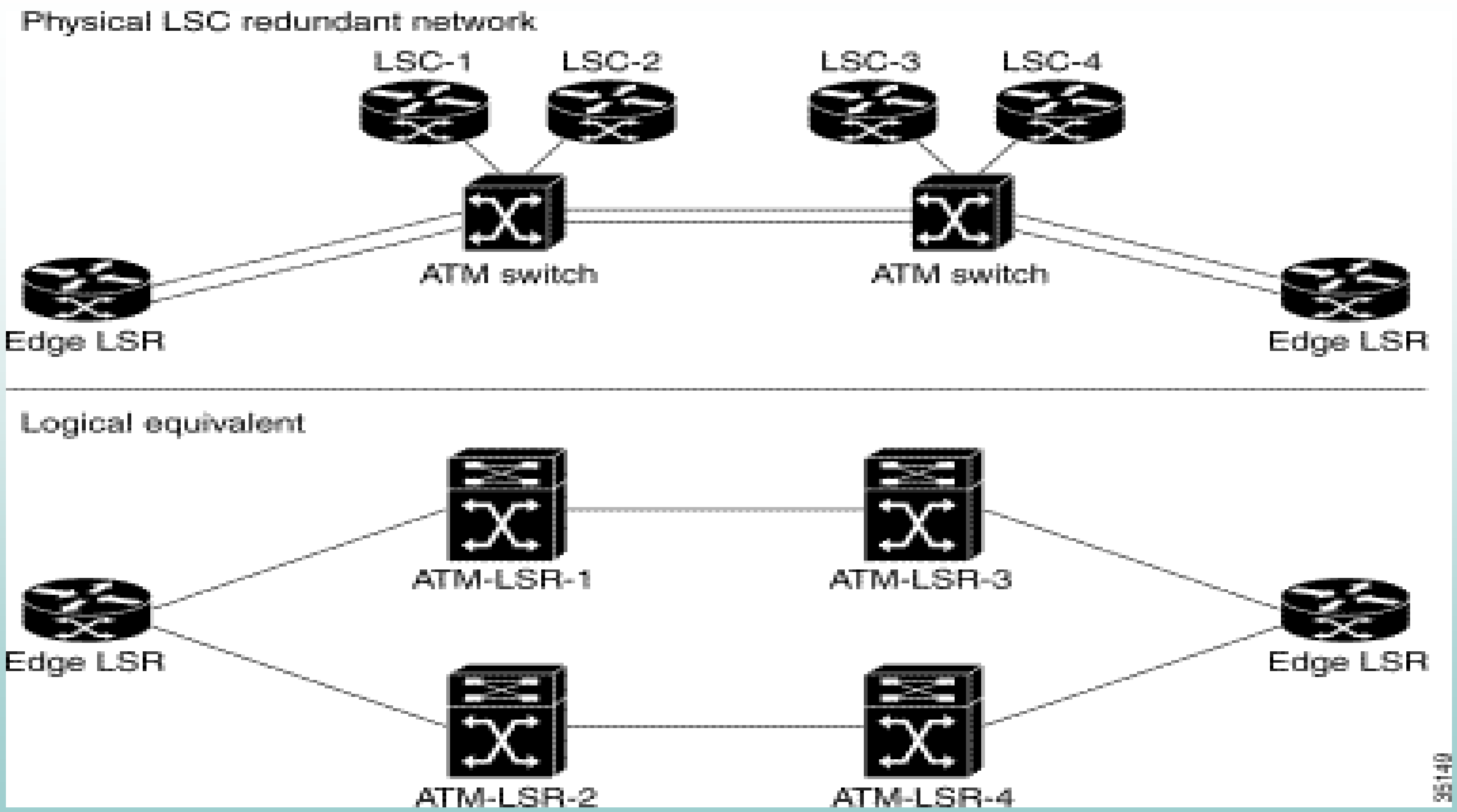


ATM Edge LSR SSO: Switchover completed

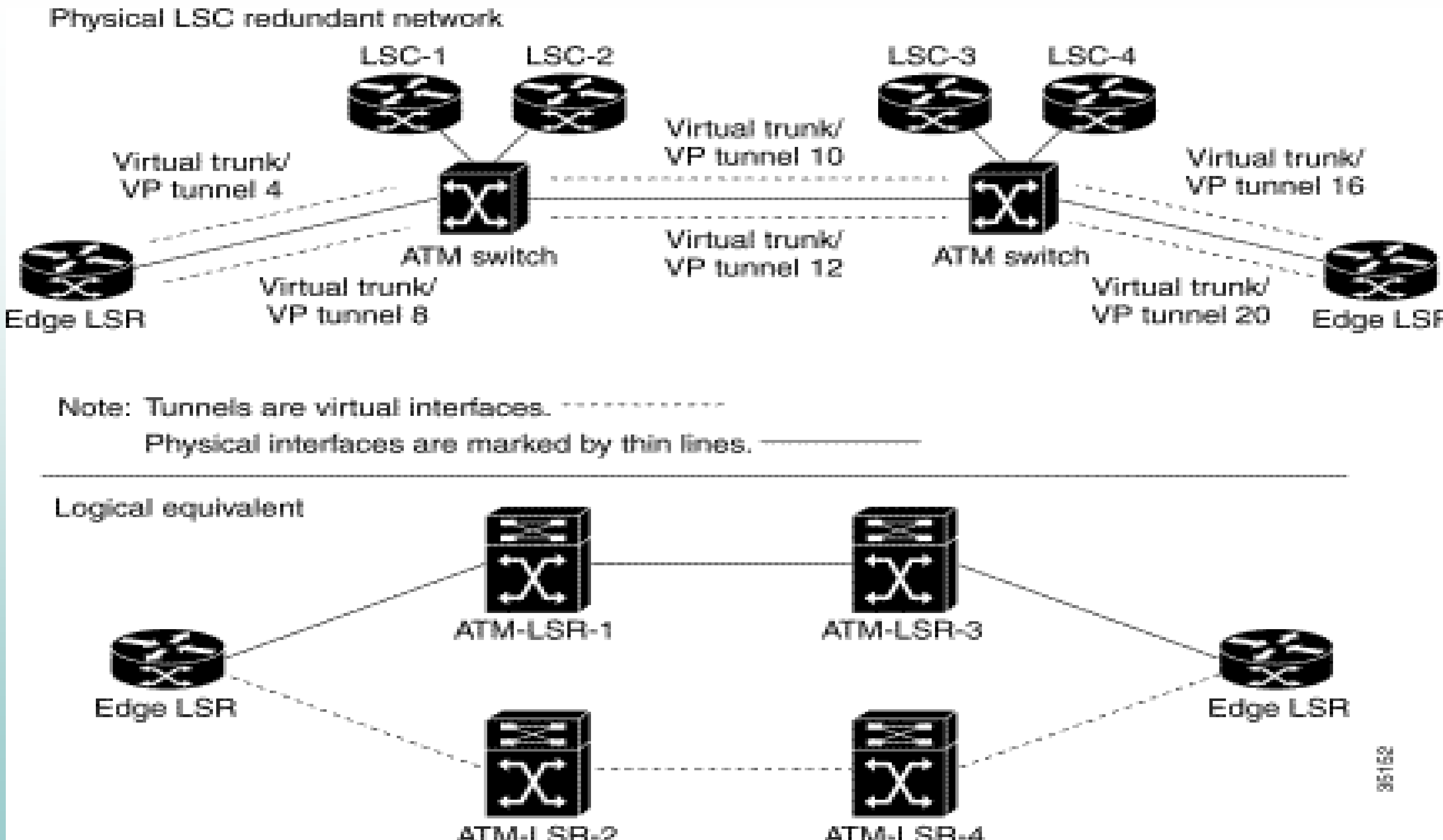
- Tail end LVC Recovery at RP B:
- Information needs to be recovered for the Tail end LVC
    - Incoming ATM label
    - Local packet label
- Upstream LSR sends Suggestive Label Request
- Restarting LSR receives the Suggestive Label Request for all eligible FECs
- Restarting LSR sends the Label Mapping reply with the same Message ID
- Local packet label information is recovered using the LDP
- MFI Forwarding entries are reasserted



SSO Neighbor based (ATM-LSR) LSC Hot Redundancy



SSO Neighbor Based (ATM-LSR) LSC Warm Redundancy



Agenda

Cisco.com

- MPLS HA Components

    MPLS HA – Border Gateway Protocol VPNv4 NSF with SSO

    MPLS HA – Traffic Engineering NSF with SSO

    MPLS HA – Fast Reroute (FRR) NSF with SSO

    MPLS HA – Any Transport over MPLS (AToM) NSF with SSO

    MPLS HA – LC-ATM NSF with SSO

- MPLS HA – Management
- Summary

MPLS HA – Management MIBs

- MPLS VPN: Cisco NSF with SSO aware VPN MIB Traps
- MPLS TE: Cisco NSF with SSO aware TE MIB Traps
- MPLS: Cisco NSF with SSO aware LDP MIB Traps

Agenda

- **MPLS HA Components**
  - MPLS HA – Border Gateway Protocol VPNv4 NSF with SSO
  - MPLS HA – Traffic Engineering NSF with SSO
  - MPLS HA – Fast Reroute (FRR) NSF with SSO
  - MPLS HA – Any Transport over MPLS (AToM) NSF with SSO
  - MPLS HA – LC-ATM NSF with SSO
- **MPLS HA – Management**
- **Summary**

Summary

Cisco.com

- Cisco is enhancing its portfolio to add features for improved full HA solution
- MPLS HA features provide stateful switchover and NSF capability for VPN, LDP, TE, etc
- MPLS VPN HA requires MFI HA, LDP HA, BGP HA
- Need IP HA enabled to support MPLS HA
  - GR must be enabled on all participating RPs (OSPF, BGP, IS-IS) on P, PE, and CE routers
- HA Capable system is enabled with full Cisco NSF with SSO
  - Peers only need to support Graceful Restart

Summary

- ATоM Cisco NSF with SSO is exactly the same as directed LDP
- AToM application will only check for local labels
- TE Cisco NSF with SSO is defined in GMPLS-RSVP(TE) RFC 3473
- FRR Cisco NSF with SSO is same as NSF/SSO for TE
- LC-ATM Cisco NSF with SSO is DoD LDP GR

References

- www.cisco.com/go/mpls

Cisco Systems

