Implementing MPLS Label Distribution Protocol

Implementing MPLS Label Distribution Protocol

The Multiprotocol Label Switching (MPLS) is a standards-based solution driven by the Internet Engineering Task Force (IETF) that was devised to convert the Internet and IP backbones from best-effort networks into business-class transport mediums.

MPLS, with its label switching capabilities, eliminates the need for an IP route look-up and creates a virtual circuit (VC) switching function, allowing enterprises the same performance on their IP-based network services as with those delivered over traditional networks such as Frame Relay or ATM.

Label Distribution Protocol (LDP) performs label distribution in MPLS environments. LDP provides the following capabilities:

- LDP performs hop-by-hop or dynamic path setup; it does not provide end-to-end switching services.
- LDP assigns labels to routes using the underlying Interior Gateway Protocols (IGP) routing protocols.
- LDP provides constraint-based routing using LDP extensions for traffic engineering.

Finally, LDP is deployed in the core of the network and is one of the key protocols used in MPLS-based Layer 2 and Layer 3 virtual private networks (VPNs).

Prerequisites for Implementing MPLS Label Distribution Protocol

The following are the prerequisites to implement MPLS LDP:

- You must be in a user group associated with a task group that includes the proper task IDs. The command reference guides include the task IDs required for each command. If you suspect user group assignment is preventing you from using a command, contact your AAA administrator for assistance.
- You must be running Cisco IOS XR software.
- You must install a composite mini-image and the MPLS package.

  Note This point is not appplicable for a Cisco NCS 540 Series Router.

- You must activate IGP.
- We recommend to use a lower session holdtime bandwidth such as neighbors so that a session down occurs before an adjacency-down on a neighbor. Therefore, the following default values for the hello times are listed:
  - Holdtime is 15 seconds.
  - Interval is 5 seconds.

  For example, the LDP session holdtime can be configured as 30 seconds by using the **holdtime** command.

Overview of Label Distribution Protocol

In IP forwarding, when a packet arrives at a router the router looks at the destination address in the IP header, performs a route lookup, and then forwards the packet to the next hop. MPLS is a forwarding mechanism in which packets are forwarded based on labels. Label Distribution Protocols assign, distribute, and install the labels in an MPLS environment. It is the set of procedures and messages by which Label Switched Routers (LSRs) establish LSPs through a network by mapping network-layer routing information directly to data-link layer switched paths. These LSPs may have an endpoint at a directly attached neighbor (comparable to IP hop-by-hop forwarding), or may have an endpoint at a network egress node, enabling switching via all intermediary nodes.

LSPs can be created statically, by RSVP traffic engineering (TE), or by LDP. LSPs created by LDP perform hop-by-hop path setup instead of an end-to-end path. LDP enables LSRs to discover their potential peer routers and to establish LDP sessions with those peers to exchange label binding information. Once label bindings are learned, the LDP is ready to setup the MPLS forwarding plane.

For more information about setting up LSPs, see MPLS Label Distribution Protocol : Details, on page 13.

Label Distribution Protocol Discovery Parameters

Discovery parameter specifies the time periods between transmitted and not received hello messages.

Configuration Example

A discovery parameter specifies time of the discovered neighbor (15 seconds) which is kept without receipt of any subsequent hello messages. After the specified time period, there is an interval of 5 seconds between the transmission of consecutive hello messages.

Configuration of Label Distribution Protocol Discovery Parameters

```
Router(config)#mpls ldp
Router(config-ldp)#router-id 192.168.70.1
Router(config-ldp)#discovery hello holdtime 15
Router(config-ldp)#discovery targeted-hello holdtime 5
Router(config-ldp)#commit
```

Verification

Displays all the current MPLS LDP parameters.

```
RP/0/RP0/CPU0:router# show mpls ldp parameters
LDP Parameters:
Role: Active
Protocol Version: 1
Router ID: 192.168.70.1

Discovery:
Link Hellos:     Holdtime:15 sec, Interval:5 sec
Targeted Hellos: Holdtime:5 sec, Interval:10 sec
Quick-start: Enabled (by default)
Transport address:       IPv4: 192.168.70.1
```

Enable Label Distribution Protocol Over an Interface

This section explains how to enable LDP over an interface. LDP should be enabled on all interfaces that connects the router to potential LDP peer routers.

Configuration Example

This example shows how to enable LDP over a tengigabit ethernet interface.

```
Router(config)# mpls ldp
Router(config-ldp)# router-id 192.168.70.1
Router(config-ldp)# interface TenGigE 0/0/0/5
Router(config-ldp-if)# commit
```

Verification

```
RP/0/RP0/CPU0:router# show mpls ldp parameters
Discovery:
Link Hellos:     Holdtime:15 sec, Interval:5 sec
Targeted Hellos: Holdtime:5 sec, Interval:10 sec
Accepting:
IPv4: all;
```

Label Distribution Protocol Discovery for Targeted Hellos

This topic explains LDP configuration for non-directly connected routers. Targeted hellos are unicast.

Configuration Example

The origin router (router 1) actively sends an hello message to the destination router (router 2). The destination router accepts the incoming hello message to respond with hello toward its discovered neighbor.

Router 1

```
Router(config)# mpls ldp
Router(config-ldp)# router-id 192.168.70.1
Router(config-ldp)# neighbor 12.12.12.12 targeted
Router(config-ldp)# interface TenGigE 0/0/0/5
Router(config-ldp-if)# commit
```

Router 2

```
Router(config)# mpls ldp
Router(config-ldp)# router-id 12.12.12.12
Router(config-ldp)# neighbor 192.168.70.1 targeted
Router(config-ldp)# address-family ipv4
Router(config-ldp-af)# discovery targeted-hello accept
Router(config-ldp)# commit
```

Note **discovery targeted-hello accept** directs the system to accept targeted hello messages from any source and activates passive mode on the LSR for targeted hello acceptance. This command is executed on the receiver node.

Label Distribution Protocol Graceful Restart

Label Distribution Protocol (LDP) graceful restart is a way to enable non-stop forwarding for label distribution protocol. The correct way to set up non-stop forwarding using label distribution protocol graceful restart is to bring up label distribution protocol neighbors (link or targeted) with additional configuration related to graceful restart.

Setting up Label Distribution Protocol Graceful Restart

Configuration Example

This example shows how to configure LDP graceful restart. In this example, the amount of time that a neighboring router maintains the forwarding state about the gracefully restarting router is specified as 180 seconds. Also, the amount of time the LDP neighbor should wait for a reconnection from the gracefully restarting router in the event of a LDP session failure is specified as 169 seconds.

```
Router(config)#mpls ldp
Router(config-ldp)#interface TenGigE 0/0/0/5
Router(config-ldp-if)#exit
Router(config-ldp)#graceful-restart
Router(config-ldp)#graceful-restart forwarding-state-holdtime 180
Router(config-ldp)#graceful-restart reconnect-timeout 169
Router(config-ldp)#commit
```

Verification

```
RP/0/RP0/CPU0:router#show mpls ldp graceful-restart
Forwarding State Hold timer : Not Running
GR Neighbors : 1

Neighbor ID      Up  Connect Count  Liveness Timer      Recovery Timer
---------------  --  -------------  ------------------  ------------------
8.8.8.8          Y        1               -                   -
```

```
RP/0/RP0/CPU0:router#show mpls ldp parameters
Graceful Restart:Enabled
Reconnect Timeout:169 sec, Forwarding State Holdtime:180 sec
NSR: Disabled, Not Sync-ed
```

Label Advertisement Control

You can control the exchange of label binding information by using label advertisement control (outbound filtering) or label acceptance control (inbound filtering)

Label Advertisement Control (Outbound Filtering)

Label Distribution Protocol advertises labels for all the prefixes to all its neighbors. When this is not desirable (for scalability and security reasons), you can configure LDP to perform outbound filtering for local label advertisement for one or more prefixes to one more peers. This feature is known as *LDP outbound label filtering, or local label advertisement control*.

Label Acceptance Control (Inbound Filtering)

LDP accepts labels (as remote bindings) for all prefixes from all peers. LDP operates in liberal label retention mode, which instructs LDP to keep remote bindings from all peers for a given prefix. For security reasons, or to conserve memory, you can override this behavior by configuring label binding acceptance for set of prefixes from a given peer. The ability to filter remote bindings for a defined set of prefixes is also referred to as *LDP inbound label filtering*.

Configure Label Advertisement Control

Label switched router (LSR) advertises all incoming label prefixes to each neighboring router. The exchange of label binding information can be controlled using the **mpls ldp label advertise**. Using the optional keywords, you can advertise selective prefixes to all neighbors, advertise selective prefixes to defined neighbors, or disable label advertisement to all peers for all prefixes.

Configuration Example

Configure Label Advertisement Control (Outbound Filtering)

Outbound label advertisement control can be configured by specifying one of the following options:

- **interface**

  Specifies an interface for label advertisement.

- **to** *ldp-id*

- **for** *prefix-acl*

  Specifies neighbors to advertise and receive label advertisements.

```
Router(config)#mpls ldp
Router(config-ldp)#address-family ipv4
Router(config-ldp-af)#label local advertise to 1.1.1.1:0 for pfx_ac11
Router(config-ldp-af)#label local advertise interface TenGigE 0/0/0/5
Router(config-ldp-af)#commit
```

Configure Label Advertisement Control (Inbound Filtering)

Inbound label acceptance control configures all prefixes specified by prefix-acl from neighbor (as specified by its LDP ID).

```
Router(config)#mpls ldp
Router(config-ldp)#address-family ipv4
Router(config-ldp-af)#label remote accept from 192.168.1.1:0 for acl_1
Router(config-ldp-af)#label remote accept from 192.168.2.2:0 for acl_2
Router(config-ldp-af)#label remote advertise to 1.1.1.1:0 for acl_1
Router(config-ldp-af)#commit
```

Local Label Allocation Control

Label Distribution Protocol allocates local labels for all prefixes that are not Border Gateway Protocol (BGP) prefixes[1]. This is acceptable when LDP is used for applications other than Layer 3 virtual private networks (L3VPN) core transport. When LDP is used to set up transport LSPs for L3VPN traffic in the core, it is not efficient or even necessary to allocate and advertise local labels for, potentially, thousands of IGP prefixes. In such a case, LDP is typically required to allocate and advertise local label for loopback /32 addresses for PE routers. This is accomplished using LDP local label allocation control, where an access list can be used to limit allocation of local labels to a set of prefixes. Limiting local label allocation provides several benefits, including reduced memory usage requirements, fewer local forwarding updates, and fewer network and peer updates.

[1] For L3VPN Inter-AS option C, LDP may also be required to assign local labels for some BGP prefixes.

Configure Local Label Allocation Control

Configuration Example

Label allocation can be configured using an IP access list to specify a set of prefixes that local labels can allocate and advertise. By default, local label allocation control is disabled and all non-BGP prefixes are assigned local labels.

The size of the local label can be configured by using the local label range table. The size of the local label space can be minimum 16000 to maximum 1048575.

```
Router(config)#mpls ldp
Router(config)#mpls label range ?
<16000-1048575>  Minimum label value
  table          Specify label table
Router(config)# mpls label range 100000 150000
Router(config)#commit
```

```
Router#conf t
Sat Dec 19 05:20:14.174 UTC
Router(config-ldp)#address-family ipv4
Router(config-ldp-af)# label local allocate for pfx_acl_1
```

Verification

```
RP/0/RP0/CPU0:router#show mpls ldp summary
Mon Dec  7 03:44:05.332 UTC

AFIs      : IPv4
Routes    : 12 prefixes
Bindings  : 18128 prefixes
Local     : 12
Remote    : 18127
Neighbors : 1
Hello Adj : 1
Addresses : 4
Interfaces: 5 LDP configured
```

Displays the dynamic range of local labels that are available on packet interface

```
RP/0/RP0/CPU0:ios#show mpls label range
Sat Dec 19 05:21:19.610 UTC
Range for dynamic labels: Min/Max: 100000/150000
```

Session Protection

After the link comes up IP converges earlier and much faster than MPLS LDP and may result in MPLS traffic loss until MPLS convergence. If a link flaps, the LDP session also flaps due to loss of link discovery. LDP session protection minimizes traffic loss, provides faster convergence, and protects existing LDP (link) sessions by means of “parallel” source of targeted discovery hello.

You can configure LDP session protection to automatically protect sessions with all or a given set of peers (as specified by peer-acl). When configured, LDP initiates backup targeted hellos automatically for neighbors

for which primary link adjacencies already exist. These backup targeted hellos maintain LDP sessions when primary link adjacencies go down.

Configure Session Protection

Configuration Example

As per the configuration, LDP session protection for peers specified by peer-acl is configured to maximum duration of 60 seconds.

```
Router(config)#mpls ldp
Router(config-ldp)#session protection for peer_acl_1 duration 60
Router(config-ldp)#commit
```

Label Distribution Protocol Interior Gateway Protocol Synchronization

Lack of synchronization between LDP and Interior Gateway Protocol (IGP) can cause MPLS traffic loss. Upon link up, for example, IGP can advertise and use a link before LDP convergence has occurred or, a link may continue to be used in IGP after an LDP session goes down.

LDP IGP synchronization coordinates LDP and IGP so that IGP advertises links with regular metrics only when MPLS LDP is converged on that link. LDP considers a link converged when at least one LDP session is up and running on the link for which LDP has sent its applicable label bindings and received at least one label binding from the peer. LDP communicates this information to IGP upon link up or session down events and IGP acts accordingly, depending on sync state.

In the event, an LDP graceful restart session disconnect, a session is treated as converged as long as the graceful restart neighbor is timed out. Additionally, upon local LDP restart, a check-point recovered LDP graceful restart session is used and treated as converged and is given an opportunity to connect and resynchronize.

Under certain circumstances, it might be required to delay declaration of re-synchronization to a configurable interval. LDP provides a configuration option to delay declaring synchronization up for up to 60 seconds. LDP communicates this information to IGP upon linkup or session down events.

Configure Interior Gateway Protocol (IGP) Synchronization

The configuration for LDP IGP synchronization resides in respective IGPs (OSPF and IS-IS) and there is no LDP-specific configuration for enabling this feature. However, there is a specific LDP configuration for IGP sync delay timer.

Configure LDP IGP Synchronization: Open Shortest Path First (OSPF) Example

Perform the following steps to synchronize LDP IGP for OSPF:

1. LDP IGP synchronization identifies the OSPF routing process.
2. Enters the OSPF configuration mode.

3. Enables LDP IGP synchronization on an interface.

As per the configuration, the sync delay is 30 seconds.

```
Router(config)#router ospf 100
Router(config-ospf)#mpls ldp sync
Router(config-ospf)#mpls ldp igp sync delay 30
Router(config-ospf)#commit
```

Configure LDP IGP Synchronization: Intermediate System to Intermediate System (IS-IS) Example

Perform the following steps to synchronize LDP IGP for IS-IS:

- LDP IGP synchronization enables the Intermediate System-to-Intermediate System (IS-IS) routing protocol and defines an IS-IS instance.
- Configures the IS-IS protocol on an interface and enters ISIS interface configuration mode.
- Enters address family configuration mode for configuring IS-IS routing for a standard IP version 4 (IPv4) or IP version 6 (IPV6) address prefix.
- Enables LDP IGP synchronization.

As per the configuration, the sync delay is 30 seconds.

```
Router(config)#router isis 100
Router(config-isis)#interface TenGigE 0/0/0/5
Router(config-isis-if)#address-family ipv4 unicast
Router(config-isis-if-af)#mpls ldp sync
Router(config-isis-if-af)#commit
```

Label Distribution Protocol Interior Gateway Protocol Auto-configuration

Interior Gateway Protocol (IGP) auto-configuration allows you to automatically configure LDP on all interfaces associated with a specified IGP interface; for example, when LDP is used for transport in the core network. However, there needs to be one IGP set up to enable LDP auto-configuration.

Typically, LDP assigns and advertises labels for IGP routes and must often be enabled on all active interfaces by an IGP. Without IGP auto-configuration, you must define the set of interfaces under LDP, a procedure that is time-intensive and error-prone.

Enable/Disable Interior Gateway Protocol(IGP) Auto-Configuration

Enable LDP Auto-Configuration global or in an Area for a specified OSPF Instance

LDP auto-configuration is supported for IPv4 unicast family in the default VRF. The IGP is responsible for verifying and applying the configuration.

Configuration Example

Perform the following steps to enable IGP auto-configuration globally for a specified OSPF process name.

- Enter an uniquely identifiable OSPF routing process. The process name is any alphanumeric string no longer than 40 characters without spaces.
- Enable LDP auto-configuration.
- To enable LDP auto-configuration in an area use the **area id** (either a decimal value or an IP address)
- Enables LDP auto-configuration globally or in an area for specified OSPF instance.

Note Enabling Auto-Configuration in an Area for a specified OSPF Instance feature is supported for IPv4 unicast family in default VRF only.

```
Router(config)#router ospf 190
Router(config-ospf)#mpls ldp auto-config
Router(config-ospf)#area 8
Router(config-ospf-ar)#interface TenGigE 0/0/0/5
Router(config-ospf-ar-if)#commit
```

Disable LDP Auto-Configuration for a specified OSPF Instance

You can also disable auto-configuration on a per-interface basis. This permits LDP to enable all IGP interfaces except those that are explicitly disabled and prevents LDP from enabling an interface when LDP auto-configuration is configured under IGP.

Configuration Example

Enter the interface configuration mode to configure an interface, specify an interface to disable the auto-configuration.

```
Router(config)#mpls ldp
Router(config-ldp)#interface TenGigE 0/0/0/5
Router(config-ldp-if)#igp auto-config disable
Router(config-ldp-if)#commit
```

LDP Nonstop Routing

LDP nonstop routing (NSR) functionality makes failures, such as Route Processor (RP) or Distributed Route Processor (DRP) fail over, invisible to routing peers with minimal to no disruption of convergence performance. By default, NSR is globally enabled on all LDP sessions except AToM.

A disruption in service may include any of these events:

- Route processor (RP) or distributed route processor (DRP) failover
- LDP process restart
- Minimum disruption restart (MDR)

Note Unlike graceful restart functionality, LDP NSR does not require protocol extensions and does not force software upgrades on other routers in the network, nor does LDP NSR require peer routers to support NSR. L2VPN configuration is not supported on NSR. Process failures of active LDP results in session loss and, as a result, NSR cannot be provided unless RP switchover is configured as a recovery action.

Configure Label Distribution Protocol Non-Stop Routing

Configuration Example

Perform this task to configure LDP Non-Stop Routing.

```
Router(config)#mpls ldp
Router(config-ldp)#nsr
Router(config-ldp)#commit
```

Verification

```
RP/0/RP0/CPU0:router#show mpls ldp nsr summary
Mon Dec  7 04:02:16.259 UTC
Sessions:
Total: 1, NSR-eligible: 1, Sync-ed: 0
(1 Ready)
```

Downstream on Demand

The Downstream on demand feature adds support for downstream-on-demand mode, where the label is not advertised to a peer, unless the peer explicitly requests it. At the same time, since the peer does not automatically advertise labels, the label request is sent whenever the next-hop points out to a peer that no remote label has been assigned.

To enable downstream-on-demand mode, this configuration must be applied at mpls ldp configuration mode:

mpls ldp downstream-on-demand with *ACL*

The ACL contains a list of peer IDs that are configured for downstream-on-demand mode. When the ACL is changed or configured, the list of established neighbors is traversed. If a session's downstream-on-demand configuration has changed, the session is reset in order that the new down-stream-on-demand mode can be configured. The reason for resetting the session is to ensure that the labels are properly advertised between the peers. When a new session is established, the ACL is verified to determine whether the session should negotiate for downstream-on-demand mode. If the ACL does not exist or is empty, downstream-on-demand mode is not configured for any neighbor.

For it to be enabled, the Downstream on demand feature has to be configured on both peers of the session. If only one peer in the session has downstream-on-demand feature configured, then the session does not use downstream-on-demand mode.

If, after, a label request is sent, and no remote label is received from the peer, the router will periodically resend the label request. After the peer advertises a label after receiving the label request, it will automatically readvertise the label if any label attribute changes subsequently.

Configure Downstream on Demand

Configuration Example

Perform this task to configure LDP Downstream on Demand.

```
Router(config)#mpls ldp
Router(config-ldp)#session downstream-on-demand with ABC
Router(config-ldp)#commit
```

Explicit-Null and Implicit-Null Labels

Cisco MPLS LDP uses null label, implicit or explicit, as local label for routes or prefixes that terminate on the given LSR. These routes include all local, connected, and attached networks. By default, the null label is **implicit-null** that allows LDP control plane to implement penultimate hop popping (PHP) mechanism. When this is not desirable, you can configure **explicit-null** that allows LDP control plane to implement ultimate hop popping (UHP) mechanism. You can configure this explicit-null feature on the ultimate hop LSR. This configuration knob includes an access-list to specify the IP prefixes for which PHP is desired.

This new enhancement allows you to configure implicit-null local label for **non-egress (ultimate hop LSR)** prefixes by using the **implicit-null-override** command. This enforces implicit-null local label for a specific prefix even if the prefix requires a non-null label to be allocated by default. For example, by default, an LSR allocates and advertises a non-null label for an IGP route. If you wish to terminate LSP for this route on penultimate hop of the LSR, you can enforce implicit-null label allocation and advertisement for this prefix using **implicit-null-override** feature.

Note If a given prefix is permitted in both explicit-null and implicit-null-override feature, then implicit-null-override supercedes and an implicit-null label is allocated and advertised for the prefix.

This feature works with any prefix including static, IGP, and BGP, when specified in the ACL.

Configure Explicit Null and Implicit Null Labels

Perform this task to configure Explicit Null and Implicit Null labels.

Configuration Example for Explicit Null

```
Router(config)#mpls ldp
Router(config-ldp)#address-family ipv4
Router(config-ldp-af)#label
Router(config-ldp-af-lbl)#local
Router(config-ldp-af-lbl-lcl)#advertise
Router(config-ldp-af-lbl-lcl-advt)#explicit-null
```

Configuration Example for Implicit Null

```
Router(config)# mpls ldp
Router(config-ldp)# address-family ipv4
Router(config-ldp-af)#label local implicit-null-override for acl1
Router(config-ldp-af)#commit
```

MPLS Label Distribution Protocol : Details

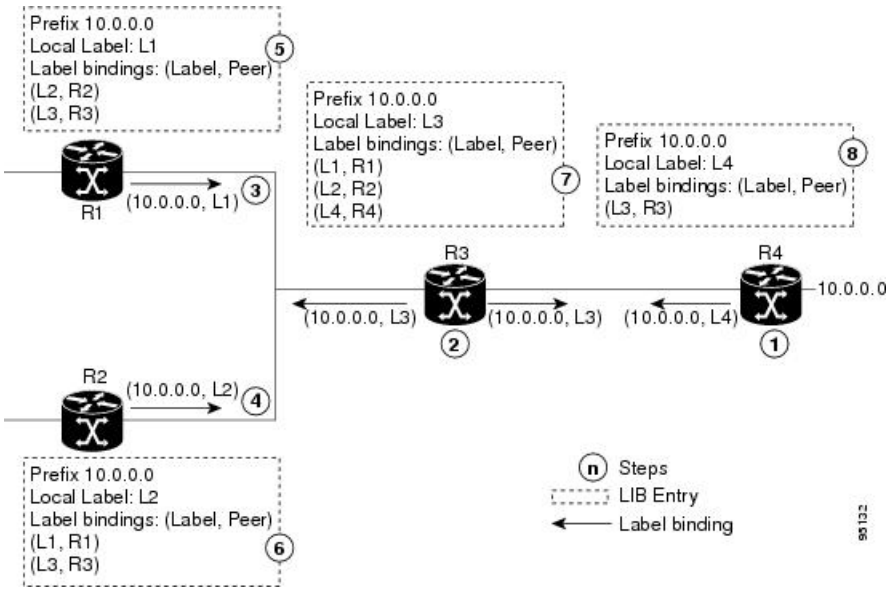
This section provides detailed conceptual information about setting up LSPs, LDP graceful restart, and LDP session protection.

Setting Up Label Switched Paths

MPLS packets are forwarded between the nodes on the MPLS network using Label Switched Paths(LSPs). LSPs can be created statically or by using a label distribution protocol like LDP. Label Switched Paths created by LDP performs hop-by-hop path setup instead of an end-to-end path. LDP enables label switched routers (LSRs) to discover their potential peer routers and to establish LDP sessions with those peers to exchange label binding information.

The following figure illustrates the process of label binding exchange for setting up LSPs.

Figure 1: Setting Up Label Switched Paths



For a given network (10.0.0.0), hop-by-hop LSPs are set up between each of the adjacent routers (or, nodes) and each node allocates a local label and passes it to its neighbor as a binding:

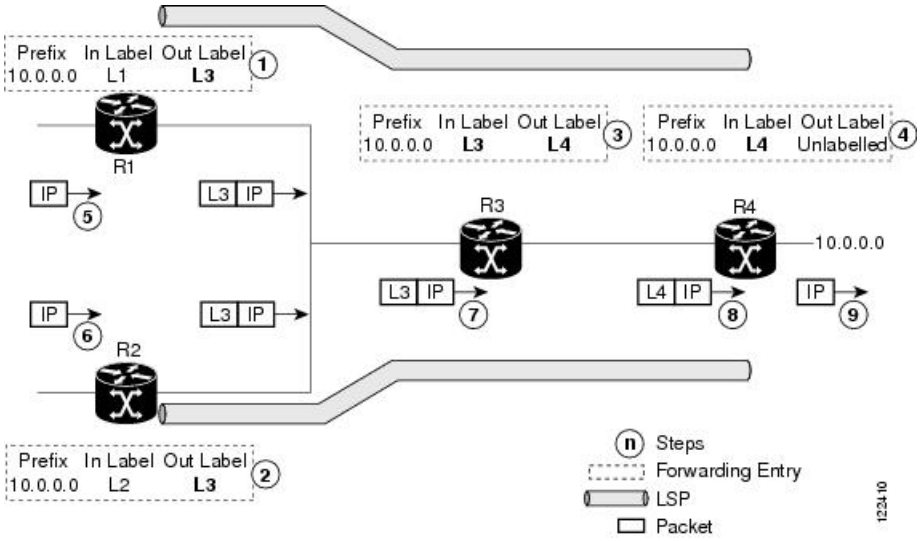
1. R4 allocates local label L4 for prefix 10.0.0.0 and advertises it to its neighbors (R3).
2. R3 allocates local label L3 for prefix 10.0.0.0 and advertises it to its neighbors (R1, R2, R4).
3. R1 allocates local label L1 for prefix 10.0.0.0 and advertises it to its neighbors (R2, R3).

4. R2 allocates local label L2 for prefix 10.0.0.0 and advertises it to its neighbors (R1, R3).
5. R1's label information base (LIB) keeps local and remote labels bindings from its neighbors.
6. R2's LIB keeps local and remote labels bindings from its neighbors.
7. R3's LIB keeps local and remote labels bindings from its neighbors.
8. R4's LIB keeps local and remote labels bindings from its neighbors.

MPLS Forwarding

Once the label bindings are learned, MPLS forwarding plane is setup and packets are forwarded as shown in the following figure.

Figure 2: MPLS Forwarding



1. Because R3 is next hop for 10.0.0.0 as notified by the FIB, R1 selects label binding from R3 and installs forwarding entry (Layer 1, Layer 3).
2. Because R3 is next hop for 10.0.0.0 (as notified by FIB), R2 selects label binding from R3 and installs forwarding entry (Layer 2, Layer 3).
3. Because R4 is next hop for 10.0.0.0 (as notified by FIB), R3 selects label binding from R4 and installs forwarding entry (Layer 3, Layer 4).
4. Because next hop for 10.0.0.0 (as notified by FIB) is beyond R4, R4 uses NO-LABEL as the outbound and installs the forwarding entry (Layer 4); the outbound packet is forwarded IP-only.
5. Incoming IP traffic on ingress LSR R1 gets label-imposed and is forwarded as an MPLS packet with label L3.
6. Incoming IP traffic on ingress LSR R2 gets label-imposed and is forwarded as an MPLS packet with label L3.
7. R3 receives an MPLS packet with label L3, looks up in the MPLS label forwarding table and switches this packet as an MPLS packet with label L4.

8. R4 receives an MPLS packet with label L4, looks up in the MPLS label forwarding table and finds that it should be Unlabeled, pops the top label, and passes it to the IP forwarding plane.
9. IP forwarding takes over and forwards the packet onward.

Details of Label Distribution Protocol Graceful Restart

LDP (Label Distribution Protocol) graceful restart provides a control plane mechanism to ensure high availability and allows detection and recovery from failure conditions while preserving Nonstop Forwarding (NSF) services. Graceful restart is a way to recover from signaling and control plane failures without impacting forwarding.

Without LDP graceful restart, when an established session fails, the corresponding forwarding states are cleaned immediately from the restarting and peer nodes. In this case LDP forwarding restarts from the beginning, causing a potential loss of data and connectivity.

The LDP graceful restart capability is negotiated between two peers during session initialization time, in FT SESSION TLV. In this typed length value (TLV), each peer advertises the following information to its peers:

Reconnect time

Advertises the maximum time that other peer will wait for this LSR to reconnect after control channel failure.

Recovery time

Advertises the maximum time that the other peer has on its side to reinstate or refresh its states with this LSR. This time is used only during session reestablishment after earlier session failure.

FT flag

Specifies whether a restart could restore the preserved (local) node state for this flag.

Once the graceful restart session parameters are conveyed and the session is up and running, graceful restart procedures are activated.

When configuring the LDP graceful restart process in a network with multiple links, targeted LDP hello adjacencies with the same neighbor, or both, make sure that graceful restart is activated on the session before any hello adjacency times out in case of neighbor control plane failures. One way of achieving this is by configuring a lower session hold time between neighbors such that session timeout occurs before hello adjacency timeout. It is recommended to set LDP session hold time using the following formula:

```
Session Holdtime <= (Hello holdtime - Hello interval) * 3
```

This means that for default values of 15 seconds and 5 seconds for link Hello holdtime and interval respectively, session hold time should be set to 30 seconds at most.

Phases in Graceful Restart

The graceful restart mechanism is divided into different phases:

Control communication failure detection

Control communication failure is detected when the system detects either:

- Missed LDP hello discovery messages

- Missed LDP keepalive protocol messages
- Detection of Transmission Control Protocol (TCP) disconnection a with a peer

Forwarding state maintenance during failure

Persistent forwarding states at each LSR are achieved through persistent storage (checkpoint) by the LDP control plane. While the control plane is in the process of recovering, the forwarding plane keeps the forwarding states, but marks them as stale. Similarly, the peer control plane also keeps (and marks as stale) the installed forwarding rewrites associated with the node that is restarting. The combination of local node forwarding and remote node forwarding plane states ensures NSF and no disruption in the traffic.

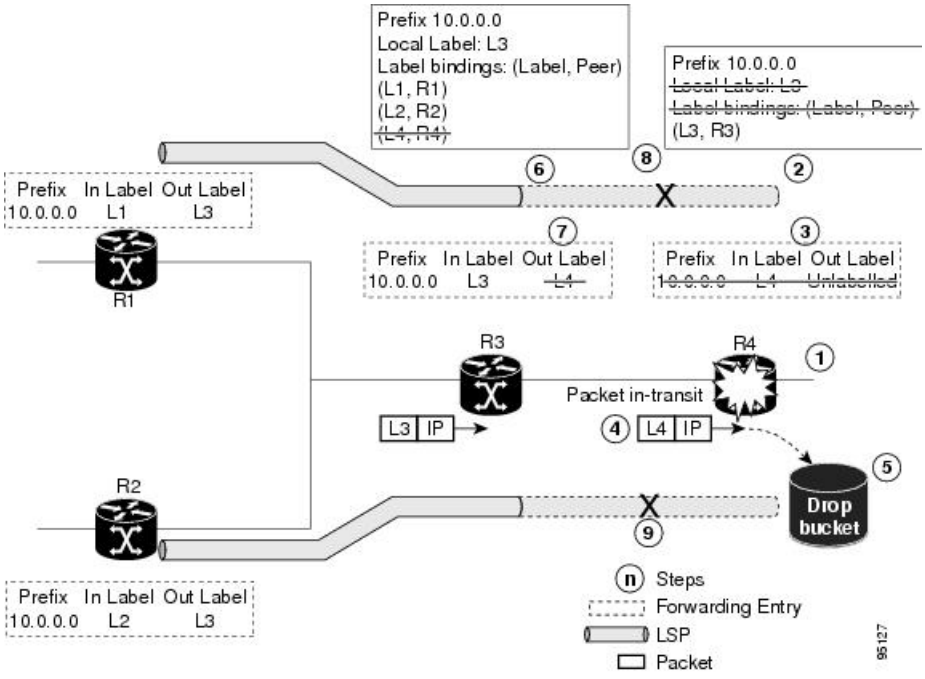
Control state recovery

Recovery occurs when the session is reestablished and label bindings are exchanged again. This process allows the peer nodes to synchronize and to refresh stale forwarding states.

Control Plane Failure

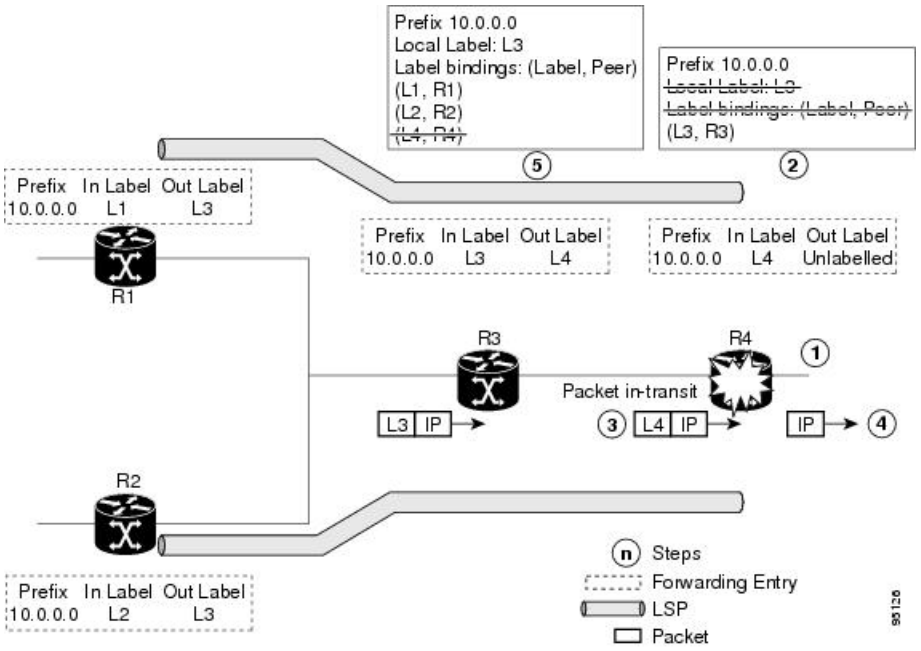
When a control plane failure occurs, connectivity can be affected. The forwarding states installed by the router control planes are lost, and the in-transit packets could be dropped, thus breaking NSF. The following figure illustrates control plane failure and recovery with graceful restart and shows the process and results of a control plane failure leading to loss of connectivity and recovery using graceful restart.

Figure 3: Control Plane Failure



Recovery with Graceful Restart

Figure 4: Recovering with Graceful Restart



1. The R4 LSR control plane restarts.
2. LIB is lost when the control plane restarts.
3. The forwarding states installed by the R4 LDP control plane are immediately deleted.
4. Any in-transit packets flowing from R3 to R4 (still labeled with L4) arrive at R4.
5. The MPLS forwarding plane at R4 performs a lookup on local label L4 which fails. Because of this failure, the packet is dropped and NSF is not met.
6. The R3 LDP peer detects the failure of the control plane channel and deletes its label bindings from R4.
7. The R3 control plane stops using outgoing labels from R4 and deletes the corresponding forwarding state (rewrites), which in turn causes forwarding disruption.
8. The established LSPs connected to R4 are terminated at R3, resulting in broken end-to-end LSPs from R1 to R4.
9. The established LSPs connected to R4 are terminated at R3, resulting in broken LSPs end-to-end from R2 to R4.

When the LDP control plane recovers, the restarting LSR starts its forwarding state hold timer and restores its forwarding state from the checkpointed data. This action reinstates the forwarding state and entries and marks them as old.

The restarting LSR reconnects to its peer, indicated in the FT Session TLV, that it either was or was not able to restore its state successfully. If it was able to restore the state, the bindings are resynchronized.

The peer LSR stops the neighbor reconnect timer (started by the restarting LSR), when the restarting peer connects and starts the neighbor recovery timer. The peer LSR checks the FT Session TLV if the restarting

peer was able to restore its state successfully. It reinstates the corresponding forwarding state entries and receives binding from the restarting peer. When the recovery timer expires, any forwarding state that is still marked as stale is deleted.

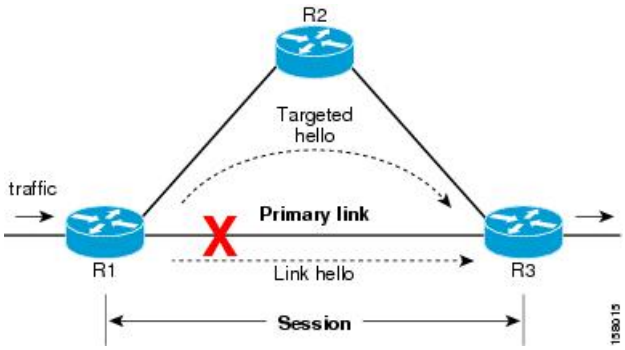
If the restarting LSR fails to recover (restart), the restarting LSR forwarding state and entries will eventually timeout and is deleted, while neighbor-related forwarding states or entries are removed by the Peer LSR on expiration of the reconnect or recovery timers.

Details of Session Protection

LDP session protection lets you configure LDP to automatically protect sessions with all or a given set of peers (as specified by peer-acl). When configured, LDP initiates backup targeted hellos automatically for neighbors for which primary link adjacencies already exist. These backup targeted hellos maintain LDP sessions when primary link adjacencies go down.

The Session Protection figure illustrates LDP session protection between neighbors R1 and R3. The primary link adjacency between R1 and R3 is directly connected link and the backup; targeted adjacency is maintained between R1 and R3. If the direct link fails, LDP link adjacency is destroyed, but the session is kept up and running using targeted hello adjacency (through R2). When the direct link comes back up, there is no change in the LDP session state and LDP can converge quickly and begin forwarding MPLS traffic.

Figure 5: Session Protection



Note When LDP session protection is activated (upon link failure), protection is maintained for an unlimited period time.