

Advanced MPLS Technology

Overview

This chapter describes advanced concepts of MPLS technology and begins with an introduction to the concept of label switch paths (LSP) and a description of LSP diversion from the IGP shortest path (Traffic Engineering) and the potential way an LSP is broken (route summarization). The chapter also describes MPLS loop detection and prevention, both in packet-mode MPLS and cell-mode MPLS implementations. The chapter concludes with a description of the interaction between MPLS and exterior IP routing implemented with Border Gateway Protocol (BGP).

The chapter contains the following topics:

- Label Switch Paths in Unicast IP Routing
- Explicit Label Switch Paths (Traffic Engineering)
- Loop Detection in Packet Mode MPLS
- Loop Detection in Cell-Mode MPLS
- MPLS—BGP Interaction
- Summary

Objectives

Upon completion of this chapter, you will be able to perform the following tasks:

- Describe the concept of Label Switch Paths and the impact of route summarization on LSP
- Understand the basics of MPLS Traffic Engineering
- Understand the data-plane loop detection in MPLS and how it relates to IP TTL

- Explain the benefits and drawbacks of IP TTL propagation
- Understand the data-plane loop detection in the ATM environment and how it affects troubleshooting tools such as traceroute
- Explain the impacts of configuring MPLS in networks running BGP
- Design simplified BGP networks based on MPLS technology

Label Switch Paths in Unicast IP Routing

Objectives

Upon completion of this section, you will be able to perform the following tasks:

- Explain the concept of Label Switch Path
- Describe how the LSP is built in unicast IP routing
- Describe the impact of IP aggregation on Label Switch Paths

Label Switching Path

- **Label Switching Path (LSP)** is a sequence of LSRs that forward labeled packets of a certain forwarding equivalence class
- **MPLS unicast IP forwarding** builds LSPs based on the output of IP routing protocols
- **LDP/TDP** only advertises labels for individual segments in the LSP
- LSPs are **unidirectional**
- **Return traffic** uses a different LSP (usually the reverse path as most routing protocols provide symmetrical routing)
- **An LSP can take a different path** from the one chosen by an IP routing protocol (**MPLS Traffic Engineering**)

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-5

A **Labels Switching Path (LSP)** is a sequence of LSRs that forward labeled packets for a particular Forwarding Equivalence Class (FEC). Each LSR swaps the top label in a packet traversing the LSP. An LSP is similar to Frame Relay or ATM virtual circuits. In cell-mode MPLS, an LSP is a virtual circuit.

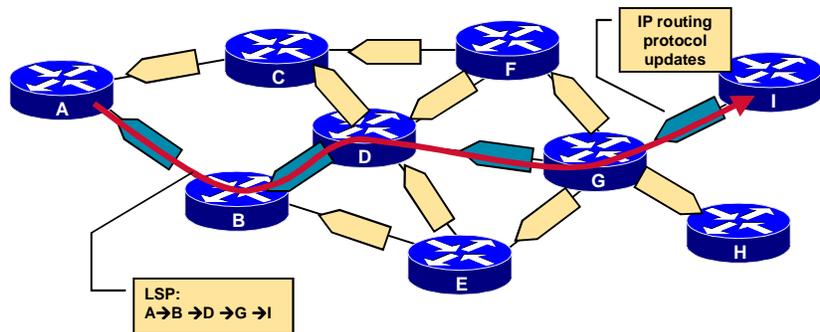
In **MPLS unicast IP forwarding** the **Forwarding Equivalence Classes** are determined by destination networks found in the main routing table. Therefore, an LSP is created for each entry found in the main routing table (BGP entries are the only exceptions and are covered later in this chapter).

An IGP is used to populate the routing tables in all routers in an MPLS domain. LDP or TDP is used to propagate labels for these networks and build LSPs.

LSPs are **unidirectional**. Each LSP is created over the shortest path, selected by the IGP, towards the destination network. Packets in the opposite direction use a different LSP. The return LSP is usually over the same LSRs except they form the LSP in the opposite order.

MPLS Traffic Engineering (MPLS/TE) can be used to change the default IGP shortest path selection.

LSP Building Example



- IP routing protocol determines the path
- LDP/TDP propagates labels to convert the path to a label switching path (LSP)

© 2002, Cisco Systems, Inc.

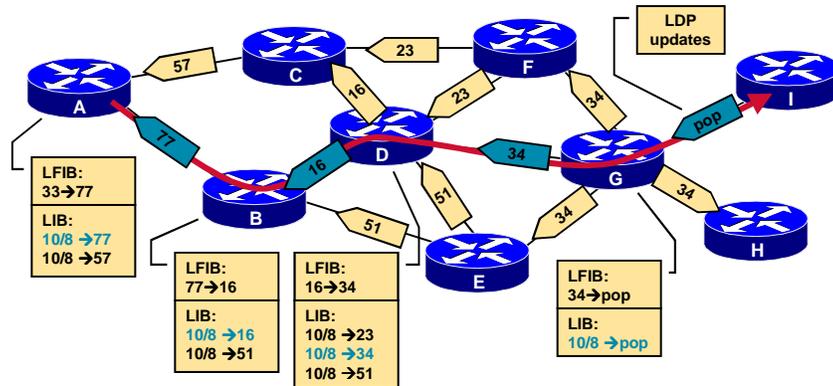
www.cisco.com

MPLS v2.1-6

The figure illustrates how an IGP such as OSPF, IS-IS, EIGRP, etc. propagates routing information to all routers in an MPLS domain. Each router determines its own shortest path. LDP or TDP that propagate labels for those networks and routers, add this information to the FIB and LFIB tables.

In the example in the figure, an LSP is created for a particular network. This LSP starts on router A and follows the shortest path, determined by the IGP.

LSP Building Example (Cont.)



- LDP/TDP propagates labels to convert the path into a label switching path (LSP)

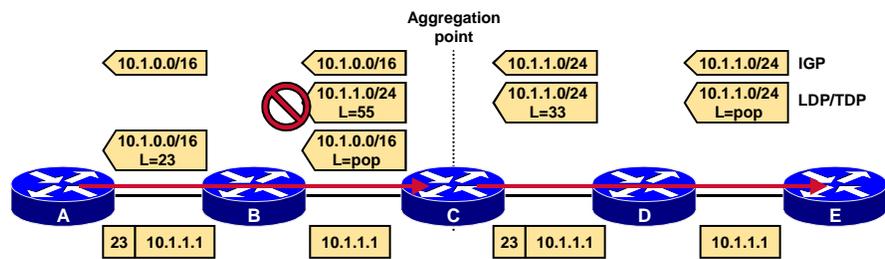
© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-7

The figure shows the contents of LFIB and LIB tables. Frame-mode MPLS uses liberal retention mode which is evident from the contents of the LIB tables. Only those labels that come from the next-hop router are inserted into the LFIB table.

Impacts of IP Aggregation on Label Switch Paths



- IP Aggregation breaks an LSP into two segments
- Router C is forwarding packets based on Layer-3 information

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-8

The figure illustrates a potential problem in an MPLS domain. An IGP propagates the routing information for network 10.1.1.0/24 from router E to other routers in the network. Router C uses a summarization mechanism to stop the proliferation of all subnets of network 10.1.0.0/16. Only the summary network 10.1.0.0/16 is sent to routers B and A.

LDP or TDP propagate labels concurrently with the IGP. The LSR that is the endpoint of an LSP always propagates the “pop” label (see “Penultimate Hop Popping” in the previous chapter).

Router C has both networks in the routing table:

- 10.1.1.0/24 (the original network)
- 10.1.0.0/16 (the summary)

Router C, therefore, sends a label, 55 in the example, for network 10.1.1.0/24 to router B. LDP also sends a “pop” label for the new summary network, because it originates on this router. Router B, however, can only use the “pop” label for the summary network 10.1.0.0/16 because it has no routing information about the more specific network 10.1.1.0/24, due to the fact that this information was suppressed on router C.

The summarization results in two LSPs for destination network 10.1.1.0/24. The first LSP ends on router C where a routing lookup is required to assign the packet to the second LSP.

Impacts of IP Aggregation on Label Switch Paths (Cont.)

- **ATM LSRs must not aggregate because they cannot forward IP packets**
- **Aggregation should not be used where end-to-end LSPs are required (MPLS VPN)**

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-9

When cell-mode MPLS is used, ATM switches are IP-aware, run an IP routing protocol, LDP or TDP and are generally seen as IP routers. In reality, however, ATM switches are only capable of forwarding cells, not IP packets.

Aggregation (or summarization) should not be used on ATM LSRs because it breaks LSPs in two, which means that ATM switches would have to perform layer-3 lookups.

Aggregation should also not be used where an end-to-end LSP is required. Typical examples of networks that require end-to-end LSPs are:

- A transit BGP autonomous system where core routers are not running BGP.
- An MPLS/VPN backbone.
- An MPLS-enabled ATM network.
- A network that uses MPLS Traffic Engineering.

Summary

A Label Switching Path (LSP) is a sequence of LSRs that forward labeled packets for a particular Forwarding Equivalence Class (FEC).

In MPLS unicast IP forwarding Forwarding Equivalence Classes are determined by destination networks found in the main routing table.

Summarization causes LSPs to break into two LSPs.

Lesson Review

1. What is an LSP?
2. Which mechanism determines the path?
3. What happens when IP aggregation (summarization) is used?

Explicit Label Switch Paths (Traffic Engineering)

Objectives

Upon completion of this section, you will be able to perform the following tasks:

- Explain the concept of explicit Label Switch Path
- Describe how an explicit LSP can be used for traffic engineering
- Describe the needs for running LDP/TDP across explicit LSP

Explicit LSP

- **LSPs are usually determined by IP routing protocols**
- **MPLS Traffic Engineering** can be used to diverge from the IGP-determined path
- **CR-LDP** or **RSVP** with extensions for Traffic Engineering is used to establish LSPs
- **LSPs can also be configured manually**

© 2002, Cisco Systems, Inc.

www.cisco.com

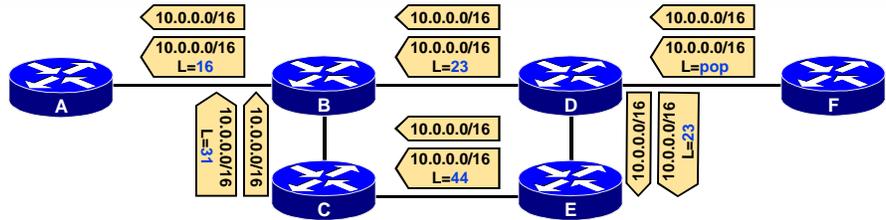
MPLS v2.1 -14

The default operation of MPLS is to construct LSPs that are equal to the shortest path selected by the IP routing protocol.

MPLS Traffic Engineering (MPLS/TE) is used to create LSPs that diverge from the shortest path. CR-LDP or RSVP with MPLS extensions are used to create those LSPs.

MPLS/TE supports automatic generation of LSPs where OSPF or IS-IS with MPLS/TE extensions must be used to propagate the information about the available resources and constraints in the network. An LSP can also be specified manually by listing LSRs in the LSP.

MPLS Traffic Engineering Example



- IGP and LDP/TDP create an LSP based on the shortest path determined by IGP

© 2002, Cisco Systems, Inc.

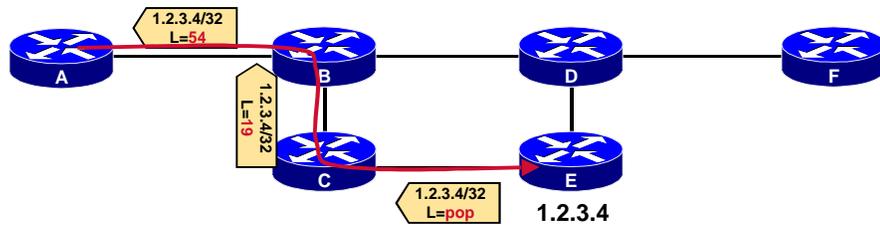
www.cisco.com

MPLS v2.1-15

The figure illustrates how an IGP and LDP propagate routing information and labels for network 10.0.0.0/16. If all inter-router links in the figure have the same IGP cost, the default LSP goes from router A through routers B and D to router F.

The next figure shows how a Traffic Engineering tunnel is established between routers A and E.

MPLS Traffic Engineering Example (Cont.)



- RSVP creates a Traffic Engineering tunnel between Routers A and E
- The new link can be included into IGP shortest path calculation
- RSVP uses downstream-on-demand label distribution
- The tunnel creation is initiated from Router A

© 2002, Cisco Systems, Inc.

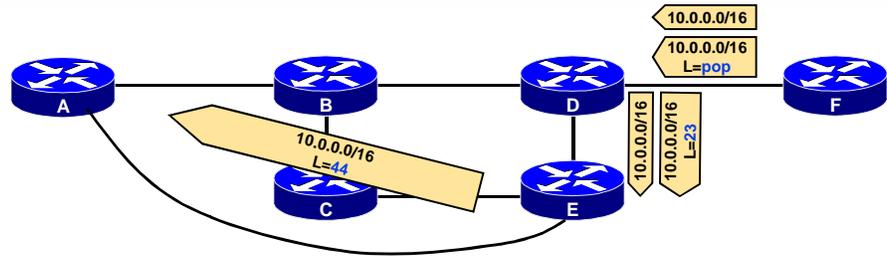
www.cisco.com

MPLS v2.1 -16

RSVP is used to create an additional LSP between routers A and E. This LSP appears as a leased line (point-to-point link) between these two routers.

The next page shows how the IGP now establishes a neighbor relationship across this link.

MPLS Traffic Engineering Example (Cont.)



- IGP and LDP/TDP create a new LSP based on the shortest path determined by IGP
- This LSP is going across the MPLS/TE LSP

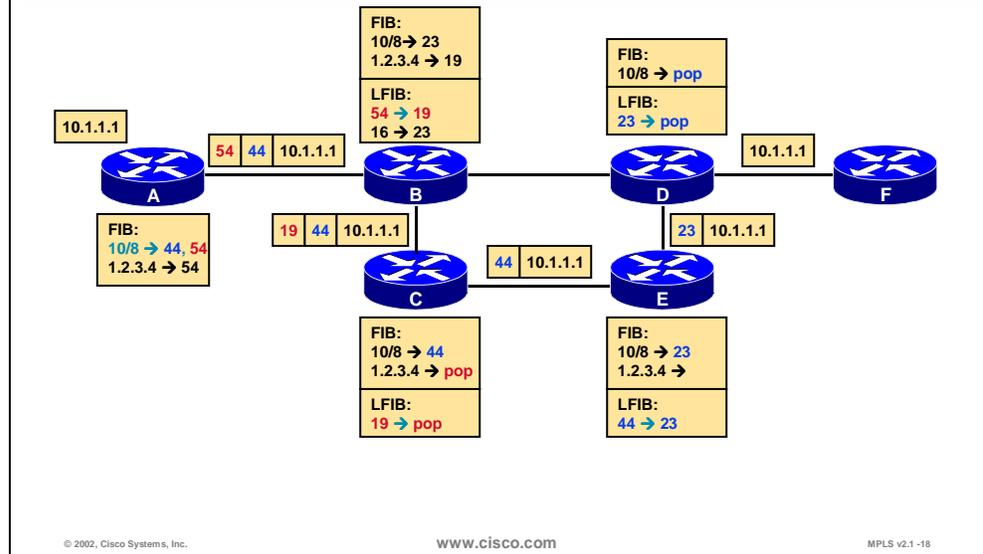
© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-17

After establishing an LDP neighbor relationship between router A and router E, router A receives another update for network 10.0.0.0/16 (label 44). A route for the destination can also be inserted into the IGP's topology database to consider this link in the SPF calculation. Router A can now choose between two available paths. Depending on the MPLS/TE configuration, router A may decide that 10.0.0.0/16 is closer through the MPLS/TE tunnel.

MPLS Traffic Engineering Example (Cont.)



This figure shows the contents of the FIB and the LFIB tables after the IGP, LDP or TDP and RSVP have propagated all the routing information and labels.

When router A forwards a packet to the destination network 10.0.0.0/16, it must put it into the LSP for that network. This LSP, however, goes across another LSP. Two labels must be used on that packet:

- The top label (54) is used for the LSP that was constructed by RSVP (MPLS/TE tunnel to address 1.2.3.4 on router E).
- The second label (44) was learned via LDP and represents the LSP for network 10.0.0.0/16.

Router B simply forwards the packet based on the top label (RSVP-derived label 19 replaces label 54).

Router C forwards the packet based on the top label that is also removed (label 19 is mapped to the *pop* action). The packet that is forwarded now has one single label.

Router E forwards the packet based on the remaining label (LDP-derived 44) and replaces it with the next-hop label 23.

Router D forwards the packet based on the label 23 and removes the label (penultimate hop popping).

Router F forwards the packet based on the destination address found in the IP header (traditional IP routing lookup).

Explicit LSPs

- **As seen in the previous example MPLS/TE can be used to implement load balancing across unequal paths**
- **Explicit paths are almost transparent to LDP/TDP**
- **LDP/TDP uses directed hello packets to find non-adjacent neighbors**

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-19

A network with redundant links may have some links that are under-utilized and some that are over-utilized. Based on a traffic analysis, MPLS/TE tunnels can be created to balance the load across unequal paths.

Explicit LSPs appear as unidirectional point-to-point links between non-adjacent routers. These LSPs are almost transparent. The only difference is that LDP and TDP use directed hello packets to establish LSR adjacency over traffic engineering tunnels.

Summary

MPLS Traffic Engineering can be used to create explicit LSPs that appear as point-to-point links between non-adjacent routers.

MPLS/TE tunnels can be used to provide load balancing across unequal paths for better link utilization.

MPLS/TE uses OSPF or IS-IS with MPLS/TE extensions to propagate the information about available resources and constraints in the network.

RSVP or CR-LDP is used to set up explicit LSPs and propagate labels.

Lesson Review

1. What is the purpose of using explicit LSPs?
2. Which technology makes use of explicit LSPs?
3. How does LDP/TDP find neighbors across an MPLS/TE tunnel?
4. Which protocols can be used to establish MPLS/TE tunnels?
5. What type of label propagation do these protocols use?

Loop Detection in Packet Mode MPLS

Objectives

Upon completion of this section, you will be able to perform the following tasks:

- Describe loop detection in packet-mode MPLS
- Explain the implications of IP TTL propagation into the TTL field of the label header
- Explain the interactions between IP TTL propagation and traceroute diagnostic tools

Loop Detection

- **LDP/TDP relies on loop-detection mechanisms built into IGPs that are used to determine the path**
- **If, however, a loop is generated (that is, misconfiguration with static routes), the TTL field in the label header is used to prevent indefinite looping of packets**
- **TTL functionality in the label header is equivalent to TTL in the IP headers**
- **TTL is usually copied from the IP headers to the label headers (TTL propagation)**

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1 -24

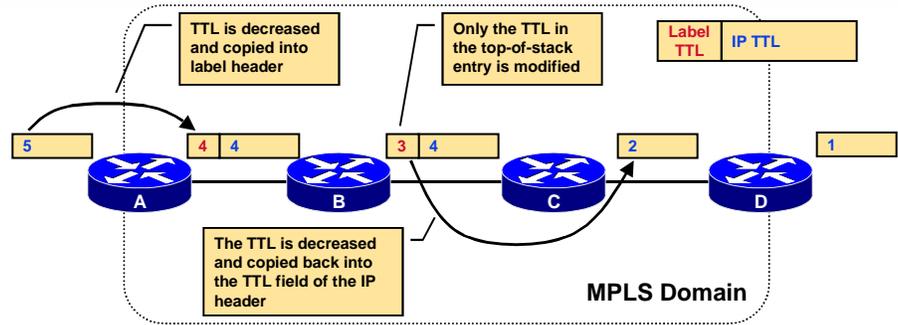
Loop detection in MPLS-enabled network relies on more than one mechanism.

Most routing loops are prevented by the IGP used in the network. MPLS for unicast IP forwarding simply uses the shortest paths determined by the IGP. These paths are typically loop-free.

If, however, a routing loop does occur (for example, due to misconfigured static routes) MPLS labels also contain a Time-to-live field (TTL) that prevents packets from looping indefinitely.

The TTL functionality in MPLS is equivalent to that of traditional IP forwarding. Furthermore, when an IP packet is labeled, the TTL value from the IP header is copied into the TTL field in the label. This is called **TTL propagation**.

Normal TTL Operation



- Cisco routers have TTL propagation enabled by default
- On ingress: TTL is copied from IP header to label header
- On egress: TTL is copied from label header to IP header

© 2002, Cisco Systems, Inc.

www.cisco.com

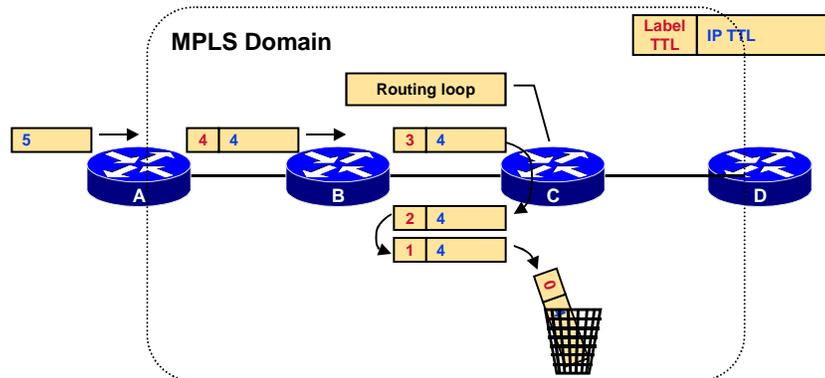
MPLS v2.1-25

The figure illustrates how the TTL value 5 in the IP header is decremented and copied into the label's TTL field when a packet enters an MPLS domain.

All other LSRs only decrement the TTL field in the label. The original TTL field is not changed until the last label is removed when the label TTL is copied back into the IP TTL.

TTL propagation provides a transparent extension of IP TTL functionality into an MPLS-enabled network.

Loop Detection



Labeled packets are dropped when the TTL is decremented to zero

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1 -26

The figure illustrates a routing loop between routers B and C. The packet looping between these two routers is eventually dropped because the value of its TTL field reaches zero.

Disabling TTL Propagation

- **TTL propagation can be disabled**
- **IP TTL value is not copied into the labels and label TTL is not copied back into IP TTL**
- **Instead, the value 255 is assigned to the label header TTL field on the ingress LSR**
- **Disabling TTL propagation hides core routers in the MPLS domain**
- **Traceroute across an MPLS domain does not show any core routers**

© 2002, Cisco Systems, Inc.

www.cisco.com

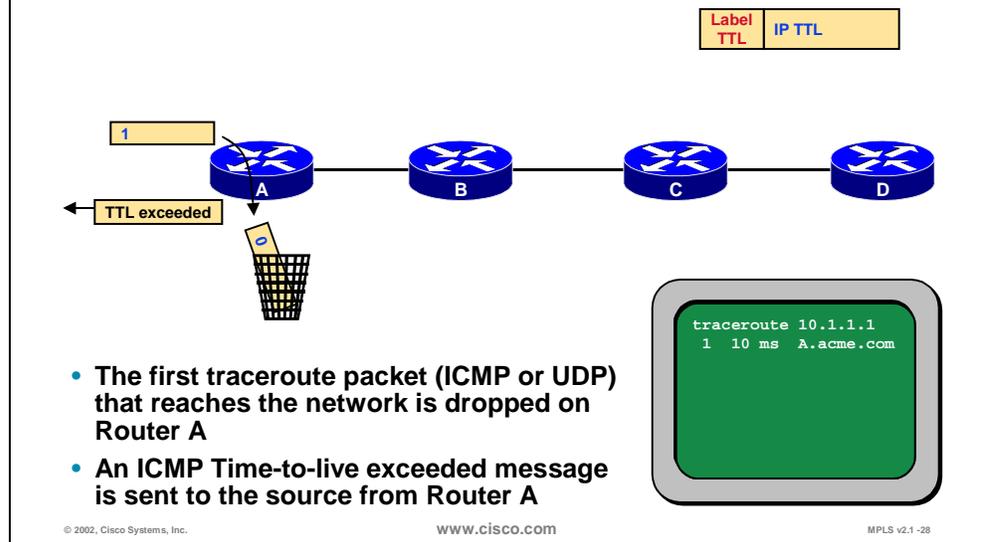
MPLS v2.1-27

TTL propagation can be disabled to hide the core routers from the end users. Disabling TTL propagation causes routers to set the value 255 into the label's TTL field when an IP packet is labeled.

The network is still protected against indefinite loops, but it is unlikely that the core routers will ever have to send an ICMP reply to user-originated traceroute packets.

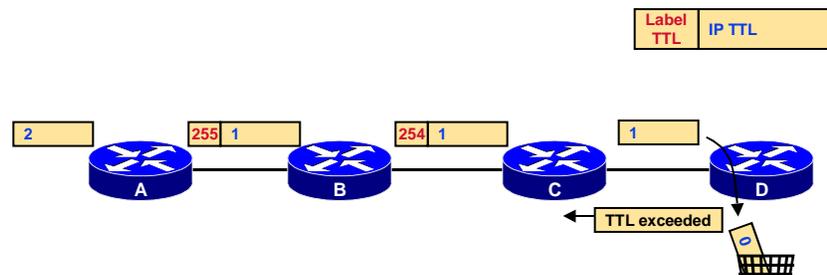
The following pages illustrate the result of a traceroute across an MPLS network that does not use TTL propagation.

Traceroute with Disabled TTL Propagation (1)



The first traceroute packet (ICMP or UDP) that reaches the MPLS network is dropped on the first router (A) and an ICMP reply is sent to the source. This results in an identification of router A by the traceroute application.

Traceroute with Disabled TTL Propagation (2)



- The second traceroute packet that reaches the network is dropped on Router D
- An ICMP Time-to-live exceeded message is sent to the source from Router D

```
traceroute 10.1.1.1
1 10 ms A.acme.com
2 10 ms D.acme.com
```

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-29

The traceroute application increases the initial TTL for every packet that it sends. The second packet, therefore, would be able to reach one hop further (router B in the example). However, the TTL value is not copied into the label's TTL field. Instead, router A sets the label's TTL field to 255. Router B decrements the label's TTL and router C removes the label without copying it back into the IP TTL. Router D then decrements the original (IP TTL), drops the packet because the TTL has reached zero, and sends an ICMP reply to the source.

The traceroute application has identified router D. The next packets would simply pass through the network.

The final result is that a traceroute application was able to identify the edge LSRs but not the core LSRs.

Impact of Disabling TTL Propagation

- **Traceroute across an MPLS domain does not show core routers**
- **TTL propagation has to be disabled on all label switch routers**
- **Mixed configurations (some LSRs with TTL propagation enabled and some with TTL propagation disabled) could result in faulty traceroute output**
- **TTL propagation can be enabled for forwarded traffic only—traceroute from LSRs does not use the initial TTL value of 255**

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-30

Cisco routers have TTL propagation enabled by default.

If TTL propagation is disabled it must be disabled on all routers in an MPLS domain to prevent unexpected behavior.

TTL can be optionally disabled for forwarded traffic only, which allows administrators to use traceroute from routers to troubleshoot problems in the network.

Summary

MPLS primarily relies on IP routing protocols to prevent routing loops. There are, however, additional loop prevention mechanisms built into MPLS architecture such as the TTL field in the MPLS label header.

MPLS uses the TTL field in the label header to prevent indefinite looping of forwarded packets. By default, the value of IP TTL field is copied into the TTL field in the label header (TTL propagation), resulting in total transparency to the end-user. If, however, the TTL propagation is disabled, the service provider is able to hide core routers from end-users.

Lesson Review

1. How are routing loops prevented in MPLS networks?
2. What is the purpose of the TTL field?
3. What is TTL propagation?
4. What is the result of disabling TTL propagation?
5. What can happen when some LSRs have TTL propagation disabled and some do not?

Loop Detection in Cell-Mode MPLS

Objectives

Upon completion of this section, you will be able to perform the following tasks:

- Explain the challenges of loop detection in cell-mode MPLS
- Describe how the label-distribution procedures enable loop detection in cell-mode MPLS
- List loop detection mechanisms available during TDP/LDP label distribution

Loop Detection in Cell-mode MPLS

- **VPI/VCI field in the ATM header is used for label switching**
- **ATM header does not contain a TTL field**
- **LDP/TDP still primarily relies on IGP to prevent routing loops**
- **There is an additional mechanism built into LDP/TDP to prevent loops**

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-35

Cell-mode MPLS uses the VPI/VCI fields in the ATM header to encode labels. These two fields do not include a TTL field. Therefore, the cell-mode MPLS must use other ways of preventing routing loops.

Again, most loops are prevented by the IGP, used in the network. However, if there is a loop, LDP can identify the LDP requests that were looped.

LDP Hop Count TLV

- **LDP uses an additional TLV to count the number of hops in an LSP**
- **The TTL field in the IP header or label header is decreased by the number of hops by the ingress ATM edge LSR before being forwarded through an LVC**
- **If the TTL field is zero or less the packet is discarded**
- **Maximum number of hops can also be specified for LDP**

© 2002, Cisco Systems, Inc.

www.cisco.com

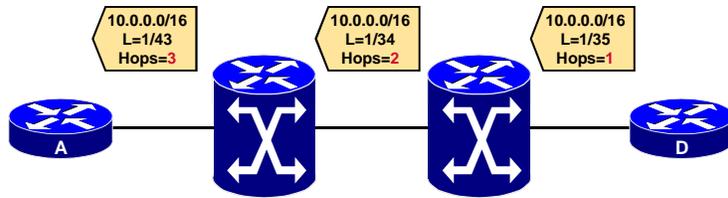
MPLS v2.1-36

LDP uses a hop-count TLV (type-length value or attribute) to count hops in the ATM part of the MPLS domain.

This hop-count can be used to provide correct TTL handling on ATM edge LSRs on behalf of ATM LSRs that cannot process IP packets.

A maximum limit in the number of hops can also be set.

LDP Hop Count Example



LSR A discovers the length of the LSP across the ATM domain to LSR D through LDP

© 2002, Cisco Systems, Inc.

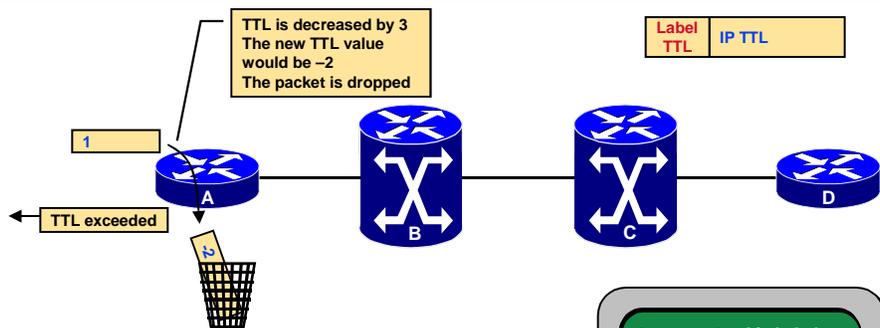
www.cisco.com

MPLS v2.1 -37

The figure illustrates how LDP, in addition to propagating the IP prefix-to-label mapping, counts hops across an MPLS-enabled ATM network.

The next page shows how traceroute is affected by this functionality.

Traceroute through ATM LSRs Example (1)



- The first traceroute packet that reaches the network is dropped on Router A
- An ICMP Time-to-live exceeded message is sent to the source from Router A

© 2002, Cisco Systems, Inc.

www.cisco.com

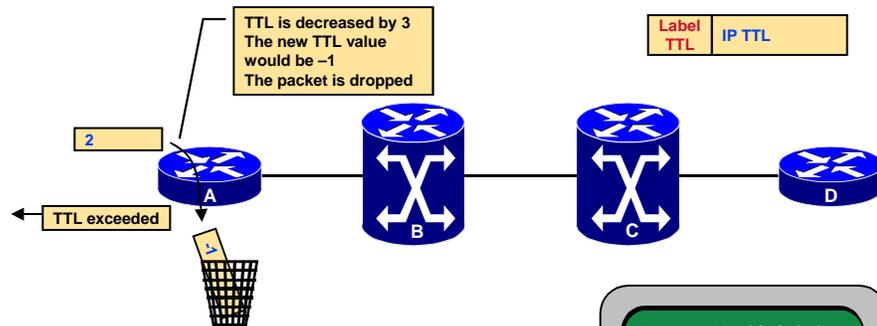
MPLS v2.1-38

The following pages illustrate how traceroute works across an IP-aware ATM network that is not capable of using the TTL field and generating ICMP replies.

The figure illustrates how an edge ATM LSR subtracts the hop-count value instead of simply decrementing the TTL value.

The first packet results in a TTL value -2 (less or equal to zero) and the packet is dropped. An ICMP reply is sent to the source.

Traceroute through ATM LSRs Example (2)



- The second traceroute packet that reaches the network is dropped on Router A
- An ICMP Time-to-live exceeded message is sent to the source from Router A

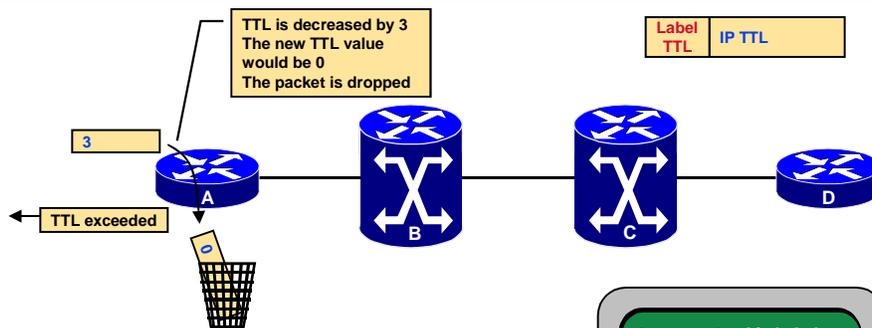
© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-39

The second packet is also dropped and another ICMP reply is sent from router A on behalf of the ATM switch B, which cannot identify the TTL field and send ICMP replies itself.

Traceroute through ATM LSRs Example (3)



- The third traceroute packet that reaches the network is dropped on Router A
- An ICMP Time-to-live exceeded message is sent to the source from Router A

```
traceroute 10.1.1.1
1 10 ms A.acme.com
2 10 ms A.acme.com
3 10 ms A.acme.com
```

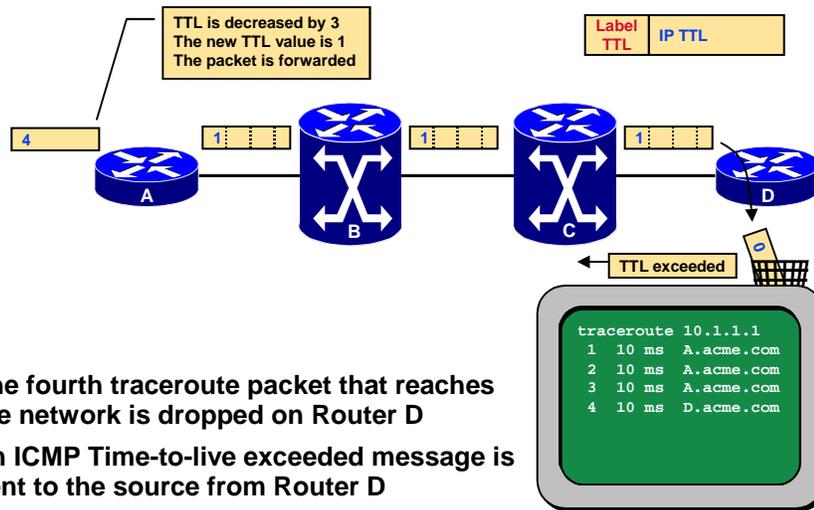
© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-40

The third packet is also dropped and the third ICMP reply is sent from router A on behalf of the ATM switch C.

Traceroute through ATM LSRs Example (4)



© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1 -41

The fourth packet can reach the other edge ATM LSR (a router), which is capable of identifying the TTL field and sending ICMP replies.

The traceroute application receives as many replies as there are hops in the network, even though there are two devices in the path that are not capable of identifying the TTL field.

LDP Path Vector TLV

- **Path Vector TLV is another safeguard that prevents loops in LDP**
- **This TLV is used to carry router IDs of all ATM LSRs in the path**
- **If an LSR receives an LDP update with its own router ID in the Path Vector TLV, the update is ignored**
- **Path Vector TLV is similar to BGP's AS-path or Cluster List attributes**
- **Path Vector TLV is not present in TDP**

© 2002, Cisco Systems, Inc.

www.cisco.com

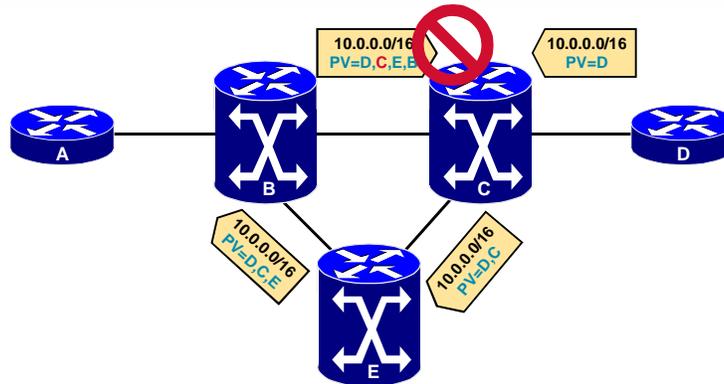
MPLS v2.1-42

The Path Vector TLV is another loop prevention mechanism that is used to prevent loops within LDP.

A Path Vector TLV is similar to BGP's AS path or Cluster List attributes. Each LSR adds its own router ID to the Path Vector TLV. If an LDP process receives an LDP label-mapping request (during the downstream-on-demand label allocation process) where its router ID can be found in the Path Vector TLV, the request is rejected.

Note Path Vector TLV is only supported by LDP. TDP relies only on Hop Count TLV to detect routing loops in the MPLS control plane.

Path Vector Example



- The LDP update is dropped because it contains the router ID of Router C in the Path Vector TLV

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1 -43

This figure illustrates how a label-mapping request looped back to the ATM LSR C that dropped it, because it found its own router id in the Path Vector TLV.

Summary

Loop Detection Summary

- **MPLS primarily relies on loop-detection mechanisms built into IGPs**
- **Hop Count TLV is used to simulate TTL functionality on ATM LSRs with the help of edge ATM LSRs**
- **Path Vector TLV is used to prevent loops in LDP updates**

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-44

Loop prevention in MPLS primarily relies on loop detection built into IP routing protocols. There are, however, several MPLS-specific loop detection mechanisms:

- Cell-mode MPLS with LDP uses a Path Vector TLV and Hop Count TLV to prevent loops in LDP.
- TTL field in the 32-bit label is used to prevent indefinite looping of packets if there is a loop in the network.

Lesson Review

1. Which mechanisms are used to prevent routing loops in MPLS-enabled networks using cell-mode MPLS?
2. Which TLVs in LDP are used to prevent loops?
3. Describe TTL operation in cell-mode MPLS.

MPLS—BGP Interaction

Objectives

Upon completion of this section, you will be able to perform the following tasks:

- Describe label allocation procedures for external IP routes
- Explain label sharing between external routes and BGP next hops
- Describe traditional BGP core design requirements
- Explain the relaxation of core design requirements made possible by MPLS
- List BGP design rules applicable in MPLS-based networks

Label Allocation in Unicast IP

- Labels are assigned to Forwarding Equivalence Classes
- **Forwarding Equivalence Class** in unicast IP routing is equal to a destination prefix found in an IP routing table
- This is true only for **IGP-derived** prefixes
- **BGP-derived** prefixes are assigned the label that is used for the BGP next-hop address
- **Result:** all prefixes learned from an external BGP neighbor use a single label

© 2002, Cisco Systems, Inc.

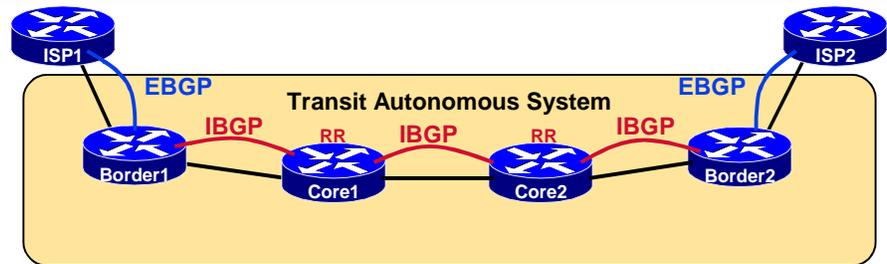
www.cisco.com

MPLS v2.1 -49

Unicast IP forwarding in MPLS networks assigns a unique label to every entry found in the main routing table. This simple rule causes a large number of labels in an ISP environment where a routing table may contain more than 100,000 networks.

To minimize the number of labels needed in such networks, an exception was made for BGP-derived routing information. All BGP-derived entries in the main routing table use the same label that is used to reach the BGP next-hop. This results in one single label being used for all networks learned from one BGP neighbor.

Traditional BGP Transit Autonomous System Design Requirements



- **All core routers** are required to run BGP
- **All core routers require full Internet routing information (more than 100.000 networks) to be able to forward IP packets between ISP1 and ISP2**

© 2002, Cisco Systems, Inc.

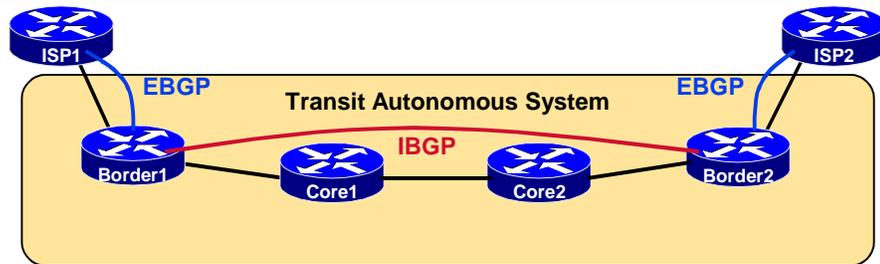
www.cisco.com

MPLS v2.1-50

One application of MPLS is in transit autonomous system where traditionally all routers had to run BGP to be able to forward packets to the correct border router.

The figure illustrates a transit autonomous system where all four routers are running BGP, which inserts more than 100.000 networks into the main routing table of each router.

Simplified BGP Network Design in MPLS-based Networks



- **Only border routers** are required to run BGP
- Core routers run an IGP to learn about BGP next-hop addresses
- Core routers run LDP/TDP to learn about labels for next-hop addresses

© 2002, Cisco Systems, Inc.

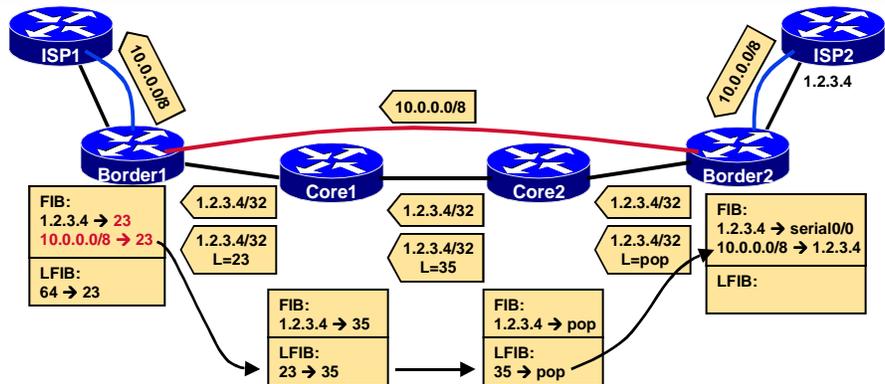
www.cisco.com

MPLS v2.1 -51

The figure shows how MPLS was used in the network to remove the need for the two core routers to run BGP. In the example, only border routers now have to run BGP.

Core routers are still capable of correctly forwarding labeled packets across the backbone even though they do not have the full routing information.

MPLS-based Transit AS Building FIB and LFIB



All routers are capable of forwarding packets to external destinations:

- Border (edge) routers label and forward IP packets
- Core routers forward labeled packets

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-52

The BGP propagation can be split into the following steps:

- ISP2 sends a BGP update about network 10.0.0.0/8.
- Router “Border2” inserts this network into the main routing table (and FIB table) and forwards it to router “Border1” over the IBGP session.
- Router “Border1” inserts this network into the main routing table (and FIB table) and forwards it to ISP1 over the EBGP session.

The relevant part of the IGP propagation can be split into the following steps:

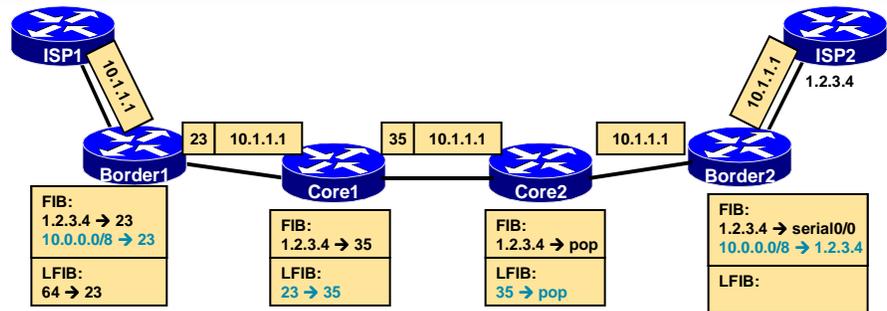
- Router “Border2” forwards the ISP2’s address (1.2.3.4) to router “Core2”.
- Router “Core2” forwards the ISP2’s address (1.2.3.4) to router “Core1”.
- Router “Core1” forwards the ISP2’s address (1.2.3.4) to router “Border1”.

The generation and propagation of labels can be split into the following steps:

- Router “Border2” advertises a “pop” label for ISP2’s address (1.2.3.4) to router “Core2”.
- Router “Core2” generates a local label “35” and advertises it to router “Core2”. A mapping from “35” to “pop” is inserted into the LFIB table.
- Router “Core1” generates a local label “23” and advertises it to router “Border1”. A mapping from “23” to “35” is inserted into the LFIB table.
- Router “Border1” inserts a mapping for IP address 1.2.3.4 to the next-hop label “23”. The BGP-derived network 10.0.0.0/8 is also mapped to the same label that is used for the BGP next-hop (10.0.0.0/8 is mapped to label “23”). These two mappings are inserted into the FIB table.

The figure on the next page illustrates how core routers are capable of forwarding labeled packet for destination network 10.0.0.0/8.

MPLS-based Transit AS Packet Propagation



© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-53

This figure illustrates how IP packets coming from ISP1 and going to 10.0.0.0/8, learned from ISP2, can be forwarded across the core routers even though they do not have the routing information for this network:

- Router “Border1” labels the packet with label “23” (the same label is used for networks 10.0.0.0/8 and 1.2.3.4 because 1.2.3.4 is the BGP next-hop for network 10.0.0.0/8).
- Router “Core1” has the mapping for label “23”. The label is swapped with the next-hop label “35”.
- Router “Core2” has the mapping for label “35”. The label is mapped to label “pop” which results in the label being removed.
- Router “Border2” performs a lookup in the FIB table where the destination 10.0.0.0/8 can be found because this router is running BGP.

Benefits of MPLS-based Transit AS

- **Simplified BGP topology (only AS edge routers are required to run BGP with full Internet routing)**
- **Core routers do not require a lot of memory (100,000 networks may require more than 50MB of memory for the BGP table, IP routing table and CEF's FIB table and distributed FIB tables)**
- **Changes in the Internet do not impact core routers**
- **Allows private addresses (RFC 1918) to be used in the core if TTL propagation is disabled (traceroute across the AS will not show any private addresses)**

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1 -54

As seen from the example, the BGP topology is simplified when combined with MPLS. Not all core routers are required to run BGP when MPLS is used. The decision where to enable BGP is primarily determined by the topology of the networks and the optimization requirements.

The main benefit of this is evident on the routers that would normally need to run BGP but no longer have to:

- Less memory is needed if BGP with full Internet routing information (more than 100,000 networks) is not used. It also reduces memory requirements on the distributed platforms if CEF is used.
- BGP flaps do not affect core routers.
- Private addresses can be used in the core if TTL propagation is disabled to hide the core routers.

Common Design and Configuration Errors

BGP next-hop addresses should not be summarized by the IGP used in the AS

- Summarization of next-hop addresses causes LSPs to break into two shorter LSPs
- The summarizing routers would have to run BGP to overcome the summarization problem

The recommendation is to have all BGP next-hops reachable as host routes or original subnets throughout the autonomous system (no summarization)

© 2002, Cisco Systems, Inc.

www.cisco.com

MPLS v2.1-55

The following precautions must be taken when combining BGP and MPLS to reduce the number of routers that run BGP:

- Do not summarize BGP next-hop addresses because summarization breaks LSPs into two LSPs.
- If summarization is configured then the router doing the summarization should also run BGP to be able to forward IP packets based on the destination IP address.

Summary

Labels that are assigned to BGP-derived networks are the same as those assigned to their next-hop addresses.

This approach allows a new way of designing BGP networks. Not all core routers are required to run BGP (depending on the topology of the network).

Lesson Review

1. What are the main benefits of using MPLS in transit autonomous systems?
2. What are the design requirements for MPLS-based transit AS?
3. What happens if BGP next-hop address is summarized somewhere in the AS?

Chapter Summary

After completing this chapter, you should be able to perform the following tasks:

- Describe the concept of Label Switch Paths and the impact of route summarization on LSP
- Explain the basics of MPLS Traffic Engineering
- Describe data-plane loop detection in MPLS and how it relates to IP TTL
- Explain the benefits and drawbacks of IP TTL propagation
- Describe data-plane loop detection in an ATM environment and how it affects troubleshooting tools such as traceroute
- Explain the impacts of configuring MPLS in networks running BGP
- Design simplified BGP networks based on MPLS technology

