Using OSPF in an MPLS VPN Environment

Overview

This module introduces the interaction between multi-protocol Border Gateway Protocol (MP-BGP) running between Provider Edge routers (PE-routers) and Open Shortest Path First (OSPF) protocol running inside a Virtual Private Network (VPN) implemented with MPLS VPN technology.

Objectives

Upon completion of this module, the learner will be able to perform the following tasks:

■   Describe the OSPF operation inside a VPN

■   Describe enhanced OSPF hierarchical model

■   Describe the interactions between OSPF and MP-BGP

■   Use OSPF as the PE-CE routing protocol in complex MPLS VPN environment

# Using OSPF as the PE-CE Protocol in an MPLS VPN Environment

Objectives

Upon completion of this lesson, the learner will be able to perform the following tasks:

- Describe the enhanced OSPF hierarchical model

- Describe the propagation of OSPF customer routes across the MPLS VPN backbone

- Explain why the OSPF routes propagated through MP-BGP are not reinserted into OSPF as external (LSA type-5) routes

- Describe the route selection process in PE routers

- Explain the loop prevention mechanisms

- Describe the need for and the mechanisms used by sham links

Traditional OSPF Routing Model

OSPF Area 0 (backbone area)

Area Border Router   Area Border Router

Area   Area   Area
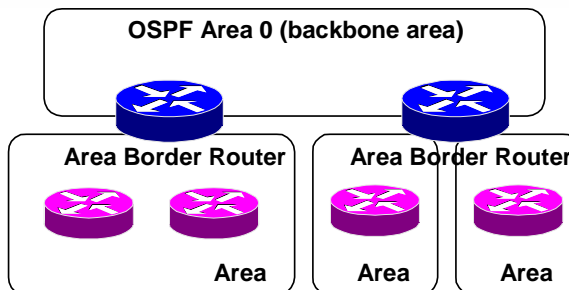
- **OSPF divides a network into areas, all of them linked through the backbone (area 0)**
- **Areas could correspond to individual sites from MPLS VPN perspective**

www.cisco.com   MPLS v2.1 -5

The Open Shortest Path First (OSPF) routing protocol was designed to support hierarchical networks with a central backbone. The network running OSPF is divided into *areas*. All areas have to be directly connected to the *backbone area* (area 0). The whole OSPF network (backbone area and any other areas connected to it) is called the *OSPF domain*.

The OSPF areas in the customer network could correspond to individual sites, but there are also other often-encountered options:

■ A single area could span multiple sites (for example, the customer decides to use an area per region, but the region contains multiple sites)

■ The backbone area could be extended into individual sites

Note   Please refer to the **Building Scalable Cisco Networks (BSCN)** course or OSPF curriculum for background information on OSPF.

MPLS VPN Routing Model

BGP backbone

PE-router   PE-router

CE-router   Site IGP   Site IGP   Site IGP
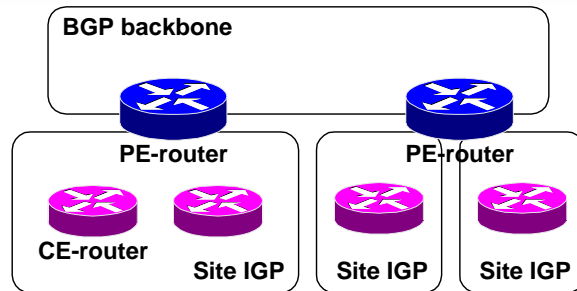
- **From the customer perspective, a MPLS VPN-based network has BGP backbone with IGP running at customer sites**
- **Redistribution between IGP and BGP is performed to propagate customer routes across MPLS VPN backbone**
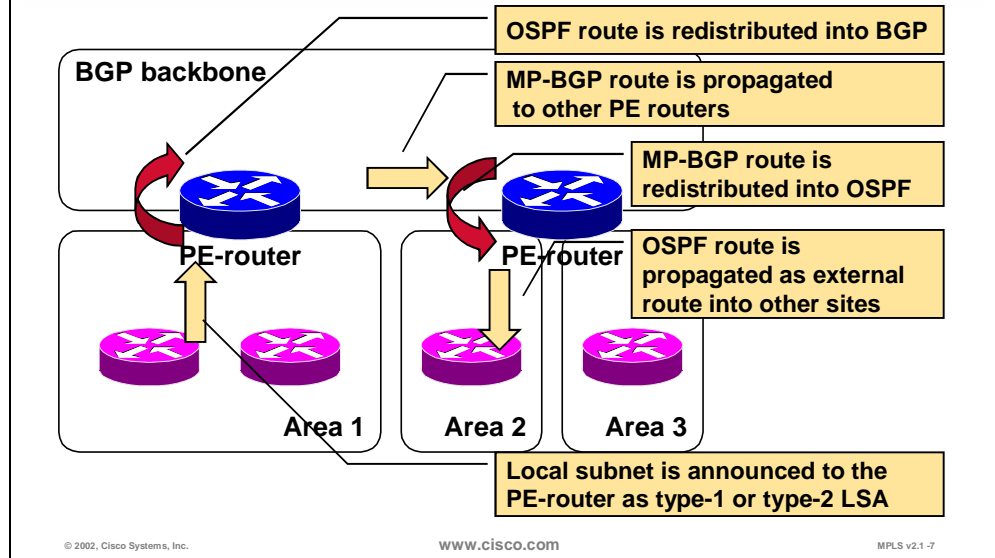
www.cisco.com   MPLS v2.1 -6

The MPLS VPN routing model introduces a BGP backbone into the customer network. Isolated copies of IGP run at every site and the multi-protocol BGP is used to propagate routes between sites. Redistribution between customer IGP, running between PE-routers and CE-routers and the backbone MP-BGP, is performed at every PE-router.

OSPF-BGP Redistribution Issue

BGP backbone

OSPF route is redistributed into BGP

MP-BGP route is propagated to other PE routers

MP-BGP route is redistributed into OSPF

PE-router

OSPF route is propagated as external route into other sites

PE-router

Area 1   Area 2   Area 3

Local subnet is announced to the PE-router as type-1 or type-2 LSA

www.cisco.com

MPLS v2.1 -7

The IGP - BGP redistribution introduced by the MPLS VPN routing model does not fit well into the customer networks running OSPF. Whenever a route is redistributed into OSPF from another routing protocol, it's redistributed as an *external* OSPF route, and this is what would happen when the customer is migrated to MPLS VPN service. The OSPF routes received by one PE-router would be propagated across the MPLS backbone and redistributed back into OSPF at another site as external OSPF routes.
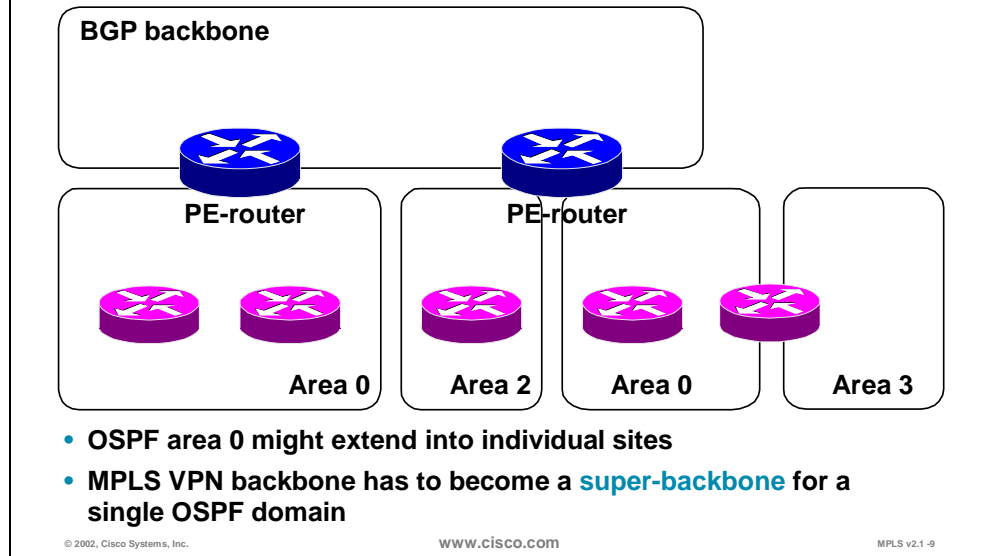
With the traditional OSPF to BGP redistribution, the OSPF route type (internal or external route) is not preserved when the OSPF route is redistributed into BGP. When that same route is redistributed back in OSPF, it's always redistributed as an external OSPF route.

There are a number of caveats associated with external OSPF routes:

■   External routes cannot be summarized

■   External routes are flooded across all OSPF areas

■   External routes could use a different metric type that is not comparable to OSPF cost

■   External routes are not inserted in *stub* areas or *not-so-stubby* (NSSA) areas

■   Internal routes are always preferred over external routes, regardless of their cost.

Because of all these caveats, migrating an OSPF customer toward MPLS VPN service might severely impact a customer's routing. The MPLS VPN architecture must therefore extend the classic OSPF - BGP routing model to support transparent customer migration.

OSPF-BGP Hierarchy Issue

BGP backbone

PE-router   PE-router

Area 0   Area 2   Area 0   Area 3

- OSPF area 0 might extend into individual sites
- MPLS VPN backbone has to become a **super-backbone** for a single OSPF domain

© 2002, Cisco Systems, Inc.   www.cisco.com   MPLS v2.1 -9

The MPLS VPN architecture extends the OSPF architecture by introducing another backbone above OSPF area 0 (*superbackbone*). The OSPF super-backbone is implemented with MP-BGP between the PE-routers, but is otherwise completely transparent to the OSPF routers. The architecture even allows disjointed OSPF backbone areas (area 0) at MPLS VPN customer sites.

With the superbackbone, the OSPF hierarchical model now has three levels. The superbackbone is always on top. Area 0 (backbone) connects directly to the superbackbone and ordinary areas (non-backbone) connect to area 0. However, when the superbackbone is present, the old backbone (area 0) is not required. As a design option, ordinary areas are allowed to connect directly to the superbackbone.

In the figure, area 3 is connected to area 0 which is connected to the superbackbone. There is also another site where area 0 is used. This discontinuous area 0 is not a problem because the superbackbone connects them together. Without the concept of the superbackbone, a discontinuous area 0 would in fact turn a traditional OSPF network into two separate OSPF networks. But the MPLS/VPN superbackbone extensions keep it as a single unified network.

The superbackbone implementation is using Multi Protocol BGP. This means that the OSPF information that is propagated between different OSPF sites is actually encoded into BGP. This is totally transparent to the customer routers. This fact is only apparent to the PE routers.

- **OSPF between sites in the same OSPF domain shall not use normal OSPF-BGP redistribution**
- **OSPF continuity must be provided across MPLS VPN backbone:**
  - **Internal OSPF routes should remain internal OSPF routes**
  - **External routes should remain external routes**
  - **OSPF metrics should be preserved**
- **CE routers run standard OSPF software**

The goals that have to be met by the OSPF super-backbone are as follows:

- The super-backbone shall not use standard OSPF - BGP redistribution

- OSPF continuity must be provided between OSPF sites:

  - Internal OSPF routes must remain internal OSPF routes

  - External OSPF routes must remain external OSPF routes

  - Non-OSPF routes redistributed into OSPF must appear as external OSPF routes in OSPF

  - OSPF metrics and metric types (External 1 or External 2) have to be preserved

- The OSPF super-backbone shall be transparent to the CE-routers that run standard OSPF software.

The concept of the superbackbone is only applicable when the OSPF sites belong to the same OSPF domain. OSPF sites that belong to different OSPF Domains should still use traditional OSPF-BGP redistribution.

MPLS VPN Backbone as OSPF Super-Backbone

MPLS VPN backbone = OSPF super-backbone

ABR   ABR

Area 0   Area 2   Area 0   Area 3
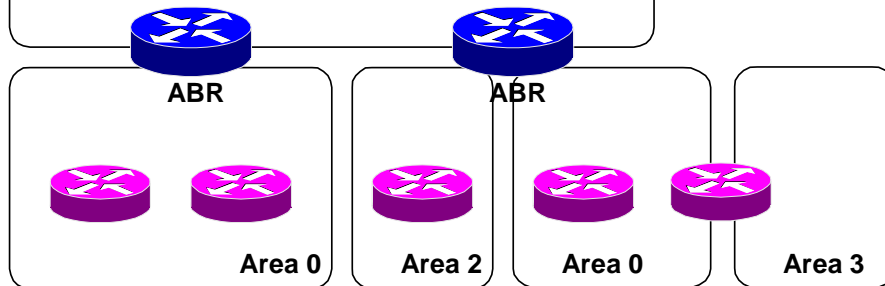
- **The MPLS VPN backbone appears as a backbone above OSPF area 0**
- **PE routers act as OSPF Area Border Routers (ABR)**
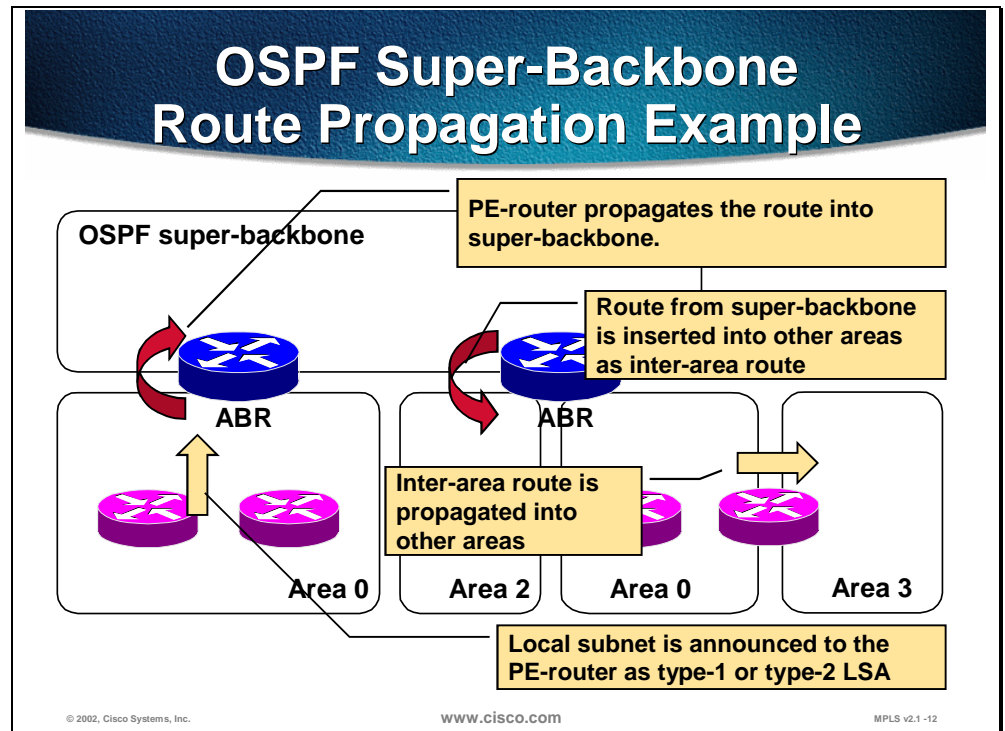
The MPLS VPN super-backbone appears as another layer of hierarchy in the OSPF architecture. The PE-routers that connect regular OSPF areas to the super-backbone therefore appear as OSPF Area Border Routers (ABR) in the OSPF areas to which they are attached. In Cisco IOS implementation, they also appear as AS Boundary Routers (ASBR) in non-stub areas.

From the perspective of a standard OSPF-speaking CE-router, the PE-routers insert inter-area routes from other areas into the area in which the CE-router is present. The CE-routers are not aware of the super-backbone or of other OSPF areas present beyond the MPLS VPN super-backbone.

The CE router views the PE router as an ABR. If the CE router is in area 0 (backbone), the CE router will believe that there is an ordinary area on the other side of the PE router. If the CE router is in a non-backbone area, the CE router will believe that the backbone is on the other side of the PE router. In reality, the superbackbone is behind the PE router. And the MP-BGP updates, with OSPF extensions, that the PE router has received from the MPLS/VPN backbone is converted into OSPF and fed into the area. The CE router can never reveal the PE router's disguise.

OSPF Super-Backbone
Route Propagation Example

PE-router propagates the route into super-backbone.

OSPF super-backbone

Route from super-backbone is inserted into other areas as inter-area route

ABR   ABR

Inter-area route is propagated into other areas

Area 0   Area 2   Area 0   Area 3

Local subnet is announced to the PE-router as type-1 or type-2 LSA

© 2002, Cisco Systems, Inc.   www.cisco.com   MPLS v2.1 -12

With the OSPF super-backbone architecture, the continuity of OSPF routing is preserved:

■   The OSPF intra-area route (described in OSPF router LSA, type 1, or network LSA, type 2) is inserted into the OSPF super-backbone by redistributing the OSPF route into MP-BGP.

| Note | If route summarization is performed at the ingress PE router, the OSPF continuity is dropped. |

■   The MP-BGP route is propagated to other PE-routers and inserted as an OSPF route into other OSPF areas. Since the super-backbone appears as another area behind the PE-router (acting as ABR), the MP-BGP route derived from intra-area route is always inserted as an inter-area route (described in OSPF summary LSA, type 3). The inter-area route could then be propagated into other OSPF areas by ABRs within the customer site.

OSPF Super-Backbone Rules

OSPF super-backbone behaves exactly like area 0 in regular OSPF:

- **PE-routers are advertised as Area Border Routers**
- **Routes redistributed from BGP into OSPF appear as inter-area summary routes or as external routes (based on their original LSA type) in other areas**
- **Routes from area 0 at one site appear as inter-area routes in area 0 at another site**

www.cisco.com

MPLS v2.1 -13

The OSPF super-backbone rules could be summarized as follows:

- PE-routers are advertising themselves as Area Border Routers. The super-backbone appears as another area to the CE-routers

- Routes redistributed into MP-BGP from OSPF will appear as inter-area routes in other OSPF sites if the original route was an intra-area or inter-area route and as external routes if the original route was an external route.

As a consequence to the second rule, routes from the backbone area at one site appear as inter-area routes (not as backbone routes) in backbone areas at other sites.

OSPF Domain

- **OSPF route types should be preserved if the two customer sites are in the same OSPF domain**
- **A customer that is migrating from an overlay VPN was probably running a single OSPF process. This process was the OSPF Domain**
- **But when redistribution occurs between two different OSPF processes, then routes should be external**
- **Conclusion: There must be a concept of OSPF Domain**

www.cisco.com MPLS v2.1 -14

The concept of the superbackbone is only applicable when two OSPF sites belong to the same OSPF domain. The OSPF route types should only be preserved if the sites belong to the same domain.

A customer that is migrating from an overlay VPN, where it was using Frame-Relay PVCs between the CE routers, was probably using one single OSPF process for the entire customer network. This OSPF process constitutes the OSPF Domain. In such a case, when the customer starts using the MPLS/VPN services, the superbackbone functions should be used.

But there is another cases where the customer was doing redistribution between different OSPF processes before the migration to MPLS/VPN started. In that case, the customer was already using several OSPF domains. Redistribution between them resulted in external routes being propagated. In this case, the superbackbone functions should not be used.

There must be a mechanism to differentiate between these two alternative behaviors. The solution is to introduce an OSPF DOMAIN ID. This identifier is carried with the MP-BGP update and indicates the source OSPF Domain. When the routes are being redistributed into the same OSPF Domain, the concepts of superbackbone apply and the route type is preserved. When redistribution occurs into another OSPF Domain, the traditional redistribution rules apply and the route type will be external.

A new BGP extended community, called OSPF DOMAIN ID, indicates the source domain of the route. In the Cisco implementation, the IOS process number of the OSPF process is used to identify the OSPF domain.

Note	The OSPF protocol itself does not communicate the process number. In the superbackbone implementation, the process number is carried by MP-BGP as an extended community attribute.

If the extended community value matches the receiving OSPF Domain, then the OSPF super-backbone is implemented with the help of several BGP attributes:

■ A new BGP extended community was defined to carry OSPF route type and OSPF area across the BGP backbone. The format of this community is defined in the following table:

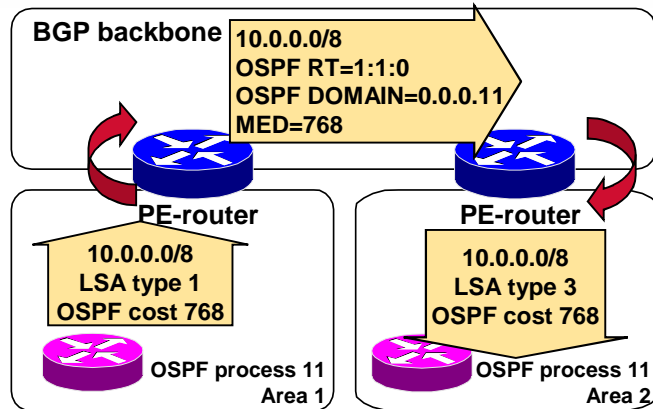| Field	Number of bytes	Comments
Community type	2	The community type is 0x8000
OSPF area	4	This field carries the OSPF area from which the route was redistributed into MP-BGP
LSA type	1	This field carries the OSPF LSA type from which the route was redistributed into MP-BGP
Option	1	This field is used for external metric type. Low-order bit is set for External Type 2 routes.

Note	The Option field in OSPF route type extended community is not equivalent to the Option field in the OSPF Link State Advertisement (LSA).

- As in the standard OSPF - BGP redistribution, the OSPF cost is carried in the MED attribute.

OSPF Super-Backbone Implementation (Cont.)

BGP backbone

10.0.0.0/8
OSPF RT=1:1:0
OSPF DOMAIN=0.0.0.11
MED=768

PE-router

10.0.0.0/8
LSA type 1
OSPF cost 768

OSPF process 11
Area 1

PE-router

10.0.0.0/8
LSA type 3
OSPF cost 768

OSPF process 11
Area 2

- OSPF route type is copied into extended BGP community on redistribution into BGP
- OSPF process number is copied into extended BGP community OSPF DOMAIN
- Egress PE-router performs inter-area transformation

www.cisco.com

MPLS v2.1 -16

This figure illustrates the propagation of internal OSPF routes across the MPLS VPN super-backbone. The sending PE-router redistributes the OSPF route into MP-BGP, copies OSPF cost into MED attribute, and sets the BGP extended community to indicate the LSA type from which the route was derived and the process number of the source OSPF process.

The receiving PE-router redistributes the MP-BGP route back into OSPF, verifies the domain ID and uses the original LSA type and the MED attribute to generate an inter-area summary LSA. An inter-area summary LSA is always generated, because the receiving PE-router acts as an ABR between the super-backbone and the OSPF area(s).

OSPF Super-Backbone Propagation of External Routes

BGP backbone

10.0.0.0/8
OSPF RT=1:5:1
OSPF DOMAIN=0.0.0.11
MED=20

PE-router

10.0.0.0/8
LSA type 5
E2 metric 20

OSPF process 11
Area 1

PE-router

10.0.0.0/8
LSA type 5
E2 metric 20

OSPF process 11
Area 2
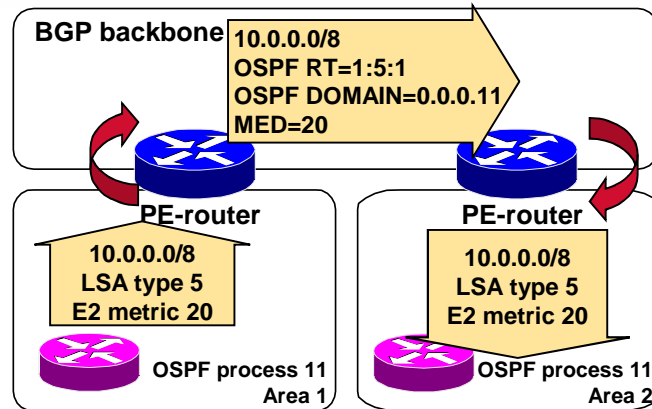
- **External OSPF routes are propagated in the same way as internal OSPF routes across super-backbone**
- **External metric and route type are preserved**

www.cisco.com

The external OSPF routes are redistributed into the MP-BGP in exactly the same way as the internal OSPF routes. The process changes slightly on the receiving PE-router:

■ For external routes (LSA type 5), the LSA is re-originated with the receiving PE-router being the ASBR. The external metric type is copied from the BGP extended community and the external cost is copied from the MED.

■ For NSSA external routes (LSA type 7), the route is announced to the other OSPF sites as LSA type-5 external route, since the route has already crossed the area boundary.

Different OSPF Domains

BGP backbone
10.0.0.0/8
OSPF RT=1:1:0
OSPF DOMAIN=0.0.0.11
MED=768

PE-router

PE-router

10.0.0.0/8
LSA type 1
OSPF cost 768

10.0.0.0/8
LSA type 5
OSPF cost 20

OSPF process 11
Area 1

OSPF process 12
Area 2
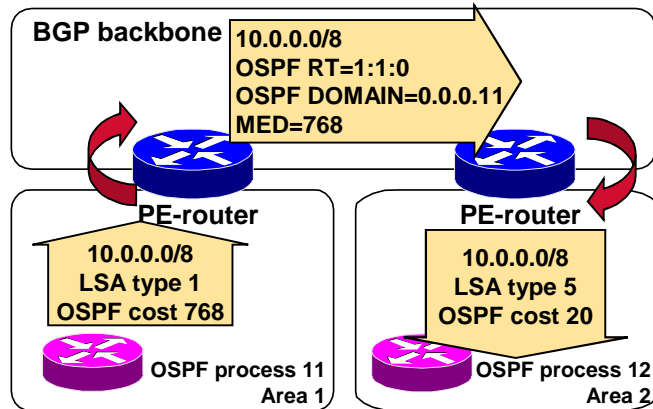
- **Routes from MP-BGP backbone that did not originate in the same OSPF Domain are subject to standard redistribution behavior when inserted into OSPF**

www.cisco.com   MPLS v2.1 -18

In the case where the OSPF Domain does not match, the MPLS VPN super-backbone still retains the traditional BGP - OSPF route redistribution. These routes are inserted into the OSPF topology database as type-5 external routes (or type-7 external routes for NSSA areas), with the default OSPF metric (not the value of MED).

This is the same behavior as if the BGP route did not carry the OSPF related extended communities at all.

OSPF Super-Backbone
Mixing Routing Protocols

BGP backbone

10.0.0.0/8
MED=3

PE-router   **PE-router**

10.0.0.0/8
Hop count 3

10.0.0.0/8
LSA type 5
E2 metric 20

RIP   **Area 2**
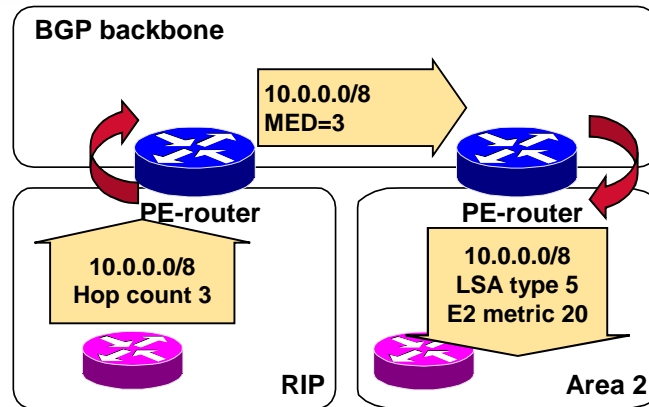
- **Routes from MP-BGP backbone that did not originate in OSPF are still subject to standard redistribution behavior when inserted into OSPF**

The MPLS VPN super-backbone still retains the traditional BGP - OSPF route redistribution behavior for routes that did not originate in OSPF at other sites (and therefore do not carry the OSPF related extended BGP communities). These routes are inserted into the OSPF topology database as type-5 external routes (or type-7 external routes for NSSA areas), with the default OSPF metric (not the value of MED).

OSPF-BGP Routing Loops

BGP backbone

The OSPF route is received by a PE-router, redistributed into MP-BGP and propagated across the MPLS VPN backbone

The route from super-backbone is inserted as inter-area route

PE-router **PE-router** **PE-router**

The OSPF route is propagated across the area

The other PE router would redistribute the route back into BGP

Area 1 **Area 2**

Local subnet is announced to the PE-router

www.cisco.com MPLS v2.1 -20

OSPF developers took many precautions to avoid routing loops between OSPF areas —for example, intra-area routes are always preferred over inter-area routes. These rules don't work after the super-backbone is introduced. Consider, for example, a network in the figure above, where the receiving OSPF area has two PE-routers attached to it.

Step 1 The sending PE-router receives an intra-area OSPF route.

Step 2 The intra-area OSPF route is redistributed into MP-BGP. OSPF community is attached to the route to indicate it was an OSPF route before being redistributed.

Step 3 Receiving PE-router redistributes the MP-BGP route into OSPF as an internal inter-area summary route.

Step 4 The summary route is propagated across OSPF area and received by the other PE-router attached to the same area.

Step 5 The administrative distance of the OSPF route is better than the administrative distance of the MP-IBGP route; therefore the PE-router selects the OSPF route and redistributes the route back into the MP-BGP process, potentially resulting in a routing loop.

- **An additional bit (Down bit) has been introduced in the Options field of the OSPF LSA header**

- **PE-routers set the Down bit when redistributing routes from MP-BGP into OSPF**

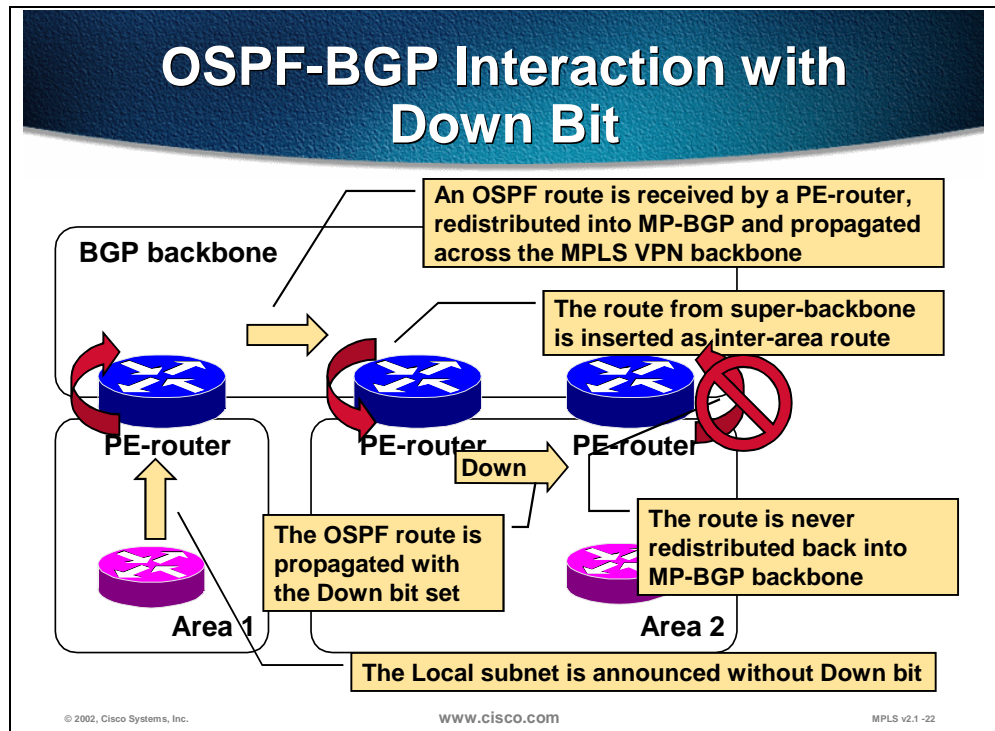- **PE-routers never redistribute OSPF routes with Down bit set into MP-BGP**

www.cisco.com
MPLS v2.1 -21

One of the options bit in the LSA header has been allocated to be the *down* bit.

The *down* bit is used between the PE-routers to indicate which routes were inserted into the OSPF topology database from the MPLS VPN super-backbone and thus shall not be redistributed back in the MPLS VPN super-backbone. The PE-router that redistributes the MP-BGP route as OSPF route into the OSPF topology database sets the *down* bit. Other PE-routers use the *down* bit to prevent this route from being redistributed back into MP-BGP.

OSPF-BGP Interaction with Down Bit

An OSPF route is received by a PE-router, redistributed into MP-BGP and propagated across the MPLS VPN backbone

The route from super-backbone is inserted as inter-area route

The route is never redistributed back into MP-BGP backbone

The OSPF route is propagated with the Down bit set

The Local subnet is announced without Down bit

BGP backbone

PE-router   PE-router   PE-router

Down

Area 1   Area 2

www.cisco.com
MPLS v2.1 -22

The typical usage of the *down* bit is shown in the figure:

Step 1 PE-router receives an OSPF route

Step 2 PE-router redistributes OSPF route into MP-BGP. The MP-BGP route is propagated to other PE-routers

Step 3 The MP-BGP route is inserted as inter-area route into an OSPF area by the receiving PE-router. The receiving PE-router sets the *down* bit in the summary (type-3) LSA.

Step 4 When the other PE-routers receive the summary LSA with the *down* bit set, they do not redistribute the route back into MP-BGP.

Routing Loops Across Multiple OSPF Domains

BGP backbone

A non-OSPF route is redistributed as an external OSPF route into the OSPF domain by a PE router

The OSPF route is propagated with the Down bit set

PE-router

PE-router

Down

The route is redistributed back into MP-BGP

The route is propagated without the Down bit

OSPF Domain 1 - Area 0

OSPF Domain 2 - Area 0

External OSPF route is redistributed into another OSPF domain. Down bit is cleared

© 2002, Cisco Systems, Inc.   www.cisco.com   MPLS v2.1 -23

The *down* bit stops the routing loops between MP-BGP and OSPF. It cannot, however, stop the routing loops when redistribution between multiple OSPF domains is involved, as is the case in the network in the figure above.

The routing loop in the network above occurs in these steps:

Step 1   The PE-router redistributes a non-OSPF route into an OSPF domain as an external route. The *down* bit is set because the route should not be redistributed back into MP-BGP.

Step 2   A CE-router redistributes the OSPF route into another OSPF domain. The *down* bit is lost if the CE-router does not understand this OSPF extension.

Step 3   The OSPF route is propagated through the other OSPF domain with the *down* bit cleared.

Step 4   A PE-router receives the OSPF route, *down* bit is not set, so the route is redistributed back into the MP-BGP backbone, resulting in a routing loop.

OSPF Tag Field

- **The Tag field in external OSPF routes is used to detect cross-domain routing loops**
- **PE routers set the Tag field to the BGP AS-number when redistributing non-OSPF routes from MP-BGP into OSPF**
- **Tag field is propagated between OSPF domains when the external OSPF routes are redistributed between OSPF domains**
- **PE routers never redistribute OSPF routes with Tag field equal to their BGP AS-number into MP-BGP**

www.cisco.com MPLS v2.1 -24

The routing loops introduced by route redistribution between OSPF domains can be solved with the help of the *tag* field, using standard BGP - OSPF redistribution rules.

In standard BGP - OSPF or OSPF - OSPF redistribution, the following rules apply:

■ Whenever a router redistributes a BGP route into OSPF, the *tag* field in the type-5 (or type-7) LSA is set to the AS-number of the redistributing router

■ The *tag* field from an external OSPF route is propagated across OSPF domains when the external OSPF route is redistributed into another OSPF domain

In addition to these standard mechanisms, PE-routers filter external OSPF routes based on their *tag* field and do not redistribute routes with *tag* field equal to the BGP AS-number into MP-BGP.
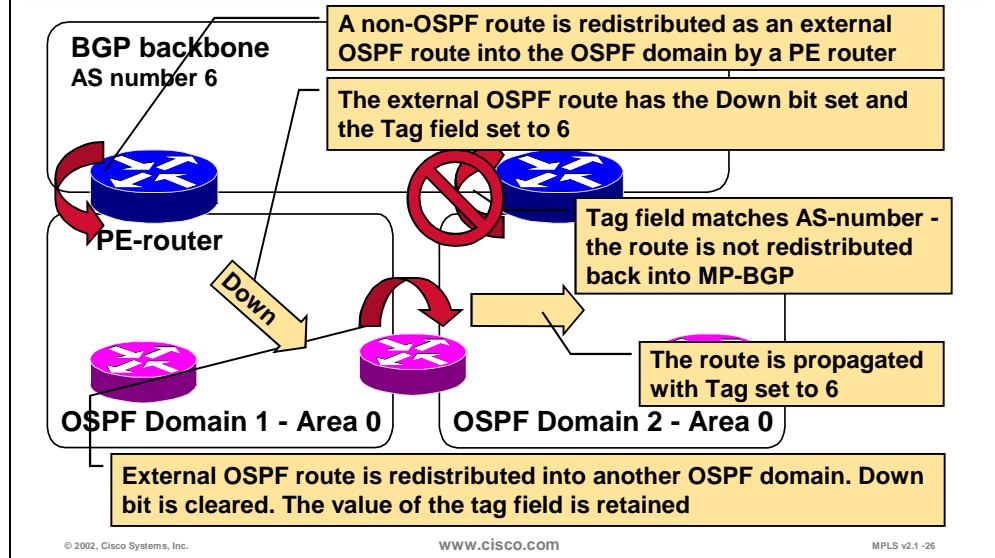
The OSPF *tag* field is only present in the external OSPF routes (type-5 LSA or type-7 LSA). This technique therefore cannot detect cross-domain loops involving internal OSPF routes. There are two manual methods that can be used to overcome this OSPF limitation:

■ The tag field can be set manually on the router redistributing routes between OSPF domains using the **redistribute ospf** *source-process-id* **tag** *value* command.

■ The PE-router can be configured to redistribute only internal OSPF routes into MP-BGP.

Routing Loop Prevention with OSPF Tag Field

BGP backbone
AS number 6

A non-OSPF route is redistributed as an external OSPF route into the OSPF domain by a PE router

The external OSPF route has the Down bit set and the Tag field set to 6

PE-router

Down

Tag field matches AS-number - the route is not redistributed back into MP-BGP

The route is propagated with Tag set to 6

OSPF Domain 1 - Area 0

OSPF Domain 2 - Area 0

External OSPF route is redistributed into another OSPF domain. Down bit is cleared. The value of the tag field is retained

www.cisco.com
MPLS v2.1 -26

The figure illustrates how the OSPF *tag* field could be used to prevent routing loops when the redistribution is done between OSPF domains.

Step 1 A non-OSPF route is redistributed as an external OSPF route by a PE-router. The *tag* field is set to the BGP AS-number; the *down* bit is set.

Step 2 The redistributed route is propagated across the OSPF domain.

Step 3 When the route is redistributed into another OSPF domain, the *tag* field is propagated, but the *down* bit is cleared.

Step 4 Another PE-router receives the external OSPF route and filters the route based on the *tag* field. The route is not redistributed into MP-BGP.

Packet Forwarding through MPLS VPN Backbone

© 2002, Cisco Systems, Inc.   www.cisco.com   MPLS v2.1 -27
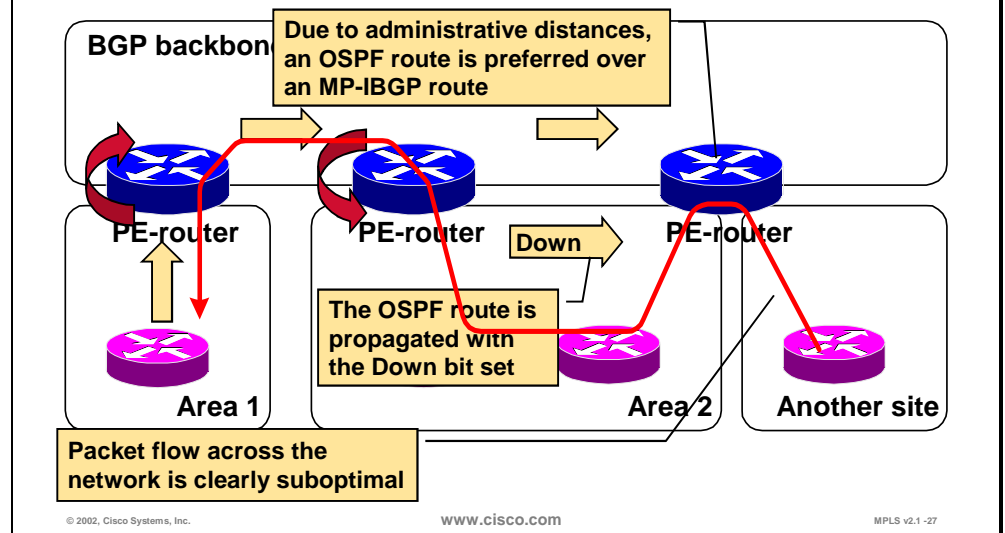
The OSPF super-backbone implementation with MP-BGP has other implications beyond the potential for routing loops between OSPF and BGP. Consider, for example, the network in the figure:

Step 1   The PE-router redistributes the OSPF route into MP-BGP. The route is propagated to other PE-routers as an MP-BGP route. It is also redistributed into other OSPF areas.

Step 2   The redistributed OSPF route is propagated across the OSPF area with the *down* bit set.

Step 3   The ingress PE-router receives an MP-IBGP route with an administrative distance of 200 and an OSPF route with an administrative distance of 110. The OSPF route is preferred over the MP-IBGP route and the data packets flow across customer sites, not directly over the MPLS VPN backbone.

Optimizing Packet Forwarding Across MPLS VPN Backbone

PE-routers ignore OSPF routes with Down bit set for routing purposes:

- **These routes originated at other sites, therefore the traffic toward them should go via MP-BGP backbone**

Routing bit is not set on OSPF routes with Down bit set:

- **These routes do not enter IP routing table even when they are selected as the best routes using the SPF algorithm**

www.cisco.com

To prevent the customer sites from acting as transit parts of the MPLS VPN network, the OSPF route selection rules in PE-routers need to be changed. The PE-routers have to ignore all OSPF routes with the *down* bit set, as these routes originated in the MP-BGP backbone and the MP-BGP route should be used as the optimum route toward the destination.

This rule is implemented with the *routing* bit in the OSPF LSA. For routes with the *down* bit set, the *routing* bit is cleared and these routes never enter the IP routing table, even if they are selected as the best routes by the Shortest Path First (SPF) algorithm.

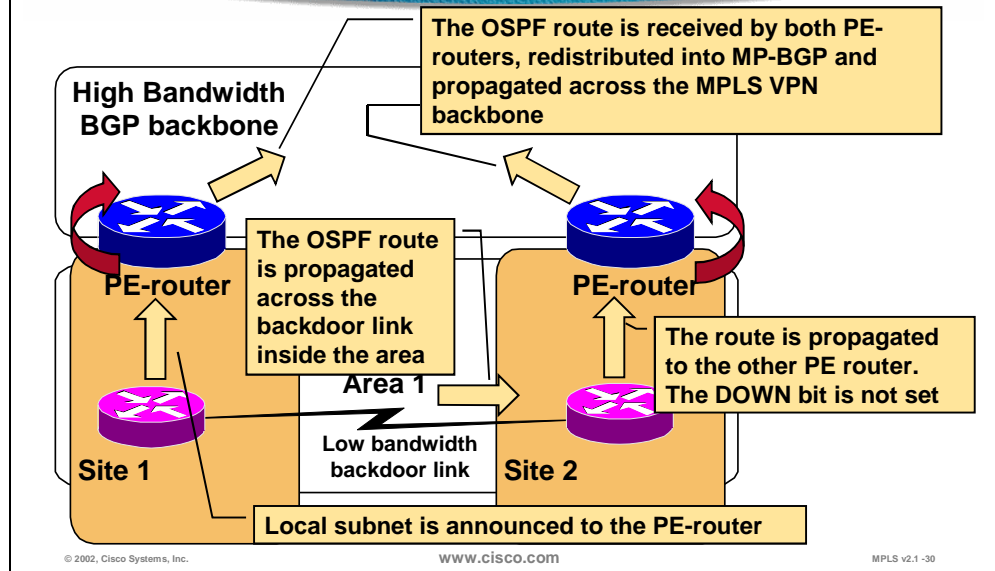|	
Note	The *routing* bit is Cisco's extension to OSPF and is used only internally in the router. It is never propagated between routers in LSA updates.

Packet Forwarding with Down Bit Processing
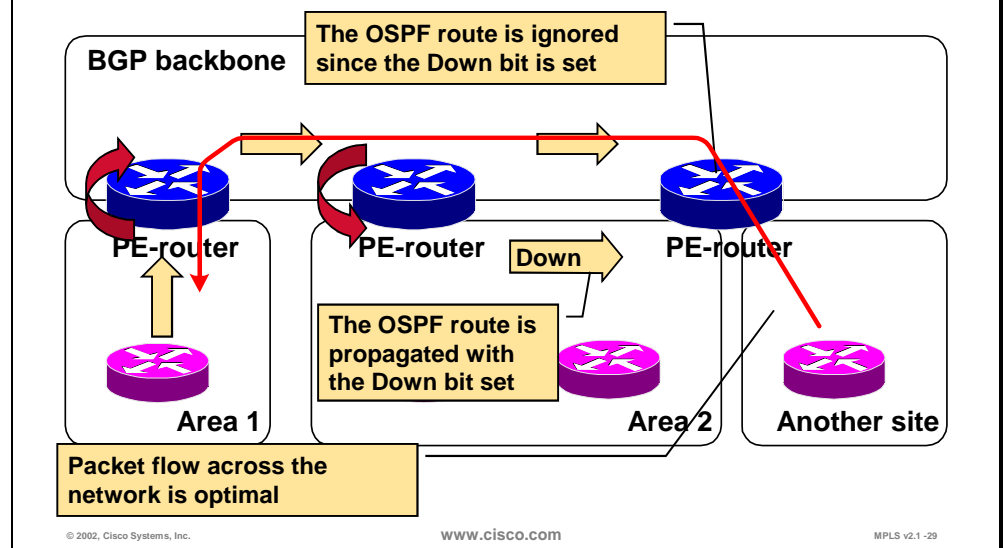
With the new route OSPF selection rules in place, the packet forwarding in the network shown in the figure follows the desired path:

Step 1   The OSPF route is redistributed into MP-BGP by a PE-router and propagated to other PE-routers.

Step 2   The receiving PE-routers redistribute the MP-BGP route into OSPF.

Step 3   Other PE-routers might receive the MP-BGP and OSPF routes, but will ignore the OSPF route for routing purposes because it has the *down* bit set. The data packets will flow across the MPLS VPN backbone following only the MP-BGP routes, not the OSPF routes derived from the MP-BGP routes.

OSPF Backdoor Routing

The OSPF route is received by both PE-routers, redistributed into MP-BGP and propagated across the MPLS VPN backbone

High Bandwidth BGP backbone

PE-router

The OSPF route is propagated across the backdoor link inside the area

Area 1

PE-router

The route is propagated to the other PE router. The DOWN bit is not set

Site 1

Low bandwidth backdoor link

Site 2

Local subnet is announced to the PE-router

© 2002, Cisco Systems, Inc.   www.cisco.com   MPLS v2.1 -30
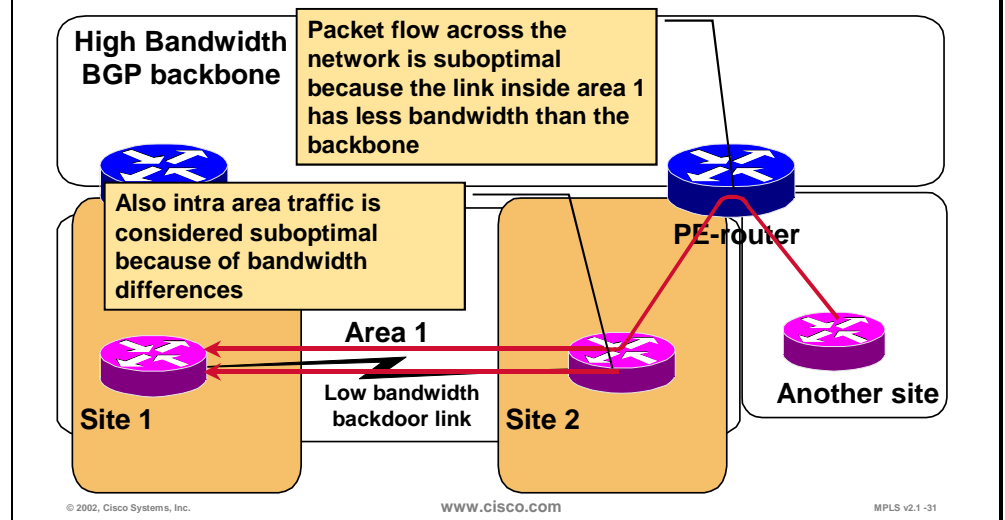
The figure illustrates a network scenario where two sites are connected to a high bandwidth BGP based MPLS/VPN backbone. But in addition to this connection the two sites have a backup link. This is a low bandwidth backdoor link. The objective is that this link should only be used in the case where any of the two sites (or both) are disconnected from the MPLS/VPN backbone.

Both of the sites belong to area 1. This means that the backdoor link is an intra-area link. None of the CE routers in the figure belong to the backbone. This is a plausible design when migrating from an overlay VPN and OSPF area re-design was not possible.

In this case, the local subnet announced by the CE router at Site 1 will be flooded across Area 1. Both PE routers will receive it and the *down* bit is not set. Both PE router will redistribute it into BGP and advertise it to the MPLS/VPN backbone.

Packet Forwarding with Backdoor

High Bandwidth BGP backbone

Packet flow across the network is suboptimal because the link inside area 1 has less bandwidth than the backbone

Also intra area traffic is considered suboptimal because of bandwidth differences

PE-router

Area 1

Low bandwidth backdoor link

Site 1

Site 2

Another site

www.cisco.com

MPLS v2.1 -31

Packet forwarding across the network with the backdoor link could be suboptimal. The right-most PE router will install and use the OSPF route in the vrf since the *down* bit was not set. The access link to the other site is connected to the same vrf since the sites share the same connectivity requirements.

Packets from the other site will therefore be routed to the CE router in Site 2, then propagated across the backdoor link to Site 1. Also traffic from Site 2 to Site 1 will use the backdoor link, because any alternative routes from the MPLS/VPN backbone would be inter-area routes and intra-area routes are preferred.

The scenario described here is normally the intended behavior. Normally traffic inside an area should stay in the area, as in the case in this example. The traffic forwarding is considered suboptimal because the backdoor link has low bandwidth and is intended for backup.

Sham Link

- A logical intra-area link
- Carried by the superbackbone
- Required only between two VPN sites that belong to the same area AND have a backdoor link for backup purposes
- OSPF adjacency is established across the sham link
- LSA flooding occurs across the sham link

© 2002, Cisco Systems, Inc.   www.cisco.com   MPLS v2.1 -32

Two VPN sites that are connected to an MPLS/VPN backbone and belong to the same area and have a backdoor link between them for backup purposes can use a *sham-link* to avoid suboptimal routing.

A *sham-link* is a logical link that belong to the area (intra-area) but is carried by the BGP based superbackbone. The two PE routers will be the end-points of the *sham-link*. They will form an OSPF adjacency across it and flood intra-area LSAs via this link.

When a *sham-link* is in use, the PE routers will receive intra-area OSPF routes two ways. One way is via the backdoor link. The other way is via the *sham-link*.

When a sham-link route is preferred by OSPF:

- **The OSPF route is not redistributed to MP-BGP**

- **Instead, the router on the other end of the sham link performs the redistribution**

- **The forwarding information from the MP-BGP route is used**
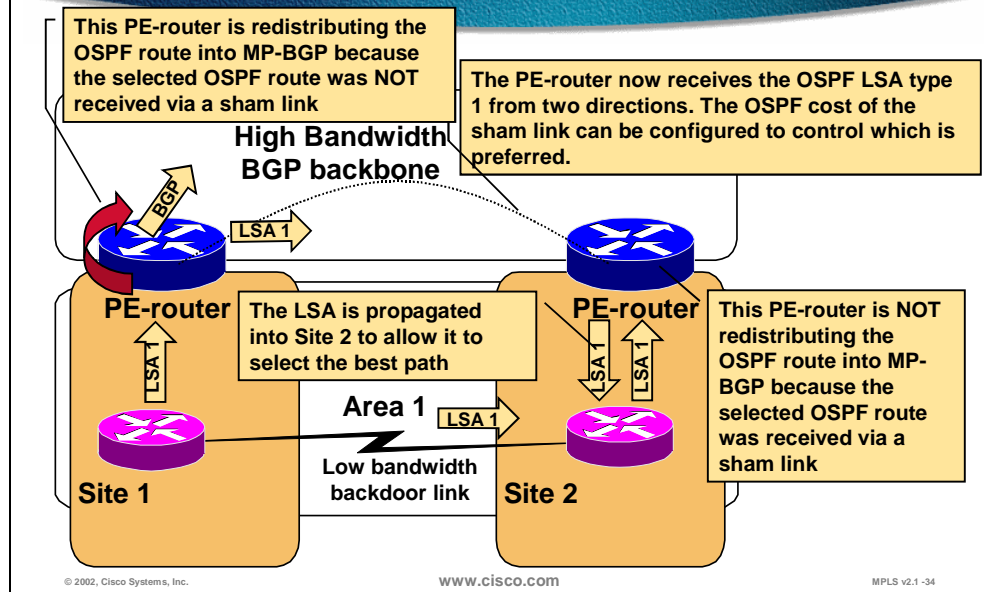
www.cisco.com
MPLS v2.1 -33

A router that receives an intra-area route two ways will use the standard SPF algorithm to determine which is the best. The accumulated cost to cross all links in one path will be compared with the accumulated cost to cross all links in the other path. The OSPF cost associated with the *sham-link* can be configured to a suitable value to get the intended behavior.

When a PE router decides to prefer an OSPF route received via a *sham-link*, then that PE router does not redistribute that route into MP-BGP. It is assumed that the PE router in the other end of the *sham-link* has already done so.

The OSPF information received over the *sham-link* is not enough to forward packets. Packets across the MPLS/VPN backbone must be label switched and the route must be associated with a label stack. The forwarding information, including the labels to use, is instead derived from MP-BGP.

Sham Link (Cont.)

This PE-router is redistributing the OSPF route into MP-BGP because the selected OSPF route was NOT received via a sham link

The PE-router now receives the OSPF LSA type 1 from two directions. The OSPF cost of the sham link can be configured to control which is preferred.

High Bandwidth BGP backbone

BGP

LSA 1

PE-router

The LSA is propagated into Site 2 to allow it to select the best path

PE-router

This PE-router is NOT redistributing the OSPF route into MP-BGP because the selected OSPF route was received via a sham link

LSA 1

LSA 1

LSA 1

Area 1

LSA 1

Low bandwidth backdoor link

Site 1

Site 2

www.cisco.com

MPLS v2.1 -34

In this figure, the two sites that belong to area 1, now have a *sham-link* between them. The local route from the CE router at Site 1 is propagated as an LSA type 1 (intra-area information). The LSA reaches the left-most PE router, which redistributes it into MP-BGP and advertises it to the MPLS/VPN backbone. But the PE router also floods the LSA across the *sham-link* to the right-most PE router.

The LSA type 1 is of course also flooded across the backdoor link to the CE router at site 2 which propagates the LSA to the right-most PE router.

The two PE routers have formed an adjacency across the *sham-link*. This means that the link state topology information includes an intra-area point-to-point link between them. All routers inside area 1 will have this information and take that link into account when determining the best path.

The right-most PE router now sees two alternative intra-area paths to reach the destination. It will perform the SPF algorithm and determine what the best path is. By tweaking the OSPF cost parameter of the *sham-link* the decision of the PE router may be influenced to select the *sham-link* route as the best.

That route will be installed in the vrf routing table, but the forwarding information (next-hop, label stack etc) will be derived from MP-BGP.

The CE router at site 2 also has the topology information about the *sham-link*. It is included in the router LSAs that both PE routers have flooded across area 1. That means that the CE router now also sees two alternatives. The OSPF cost of the *sham-link* and the cost of the backdoor link can be set to appropriate values to influence the router to select the MPLS/VPN backbone as the best route.

Packet Forwarding with Sham Link
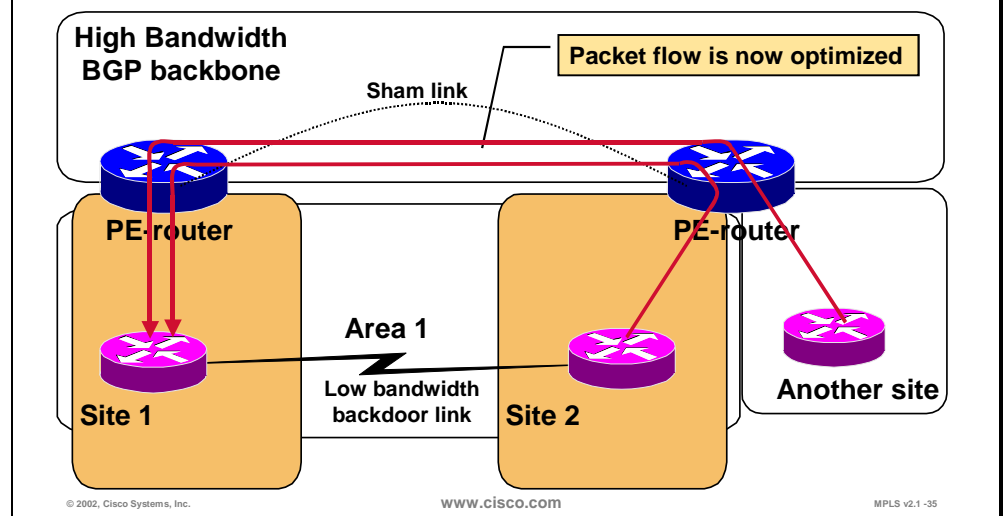
The right-most PE router has installed the OSPF route received over the *sham-link* in the vrf routing table. The forwarding information (label stack) has been derived from MP-BGP.

Packets received from the other site will therefore be routed across the MPLS/VPN backbone and use the high bandwidth links. Also, the CE router at site 2 has decided to select the *sham-link* as the best path. Therefore it routes the traffic to site 1 up to the PE router, which in turn forwards the traffic across the MPLS/VPN.

Traffic flow has now been optimized to utilize the highest bandwidth links in this example network.

Summary

The MPLS VPN architecture introduces a routing model where a BGP backbone is inserted into the customer network. Traditional OSPF - BGP interactions would imply that the OSPF routes received from one customer site would be inserted as external OSPF routes into other customer sites. As the external OSPF routes are treated differently from internal OSPF routes and the customer OSPF routing often relies on properties of various OSPF route types, this option is not acceptable. A different model is needed where the MPLS VPN backbone would be implemented transparently to the customer.

The OSPF *super-backbone* was introduced in MPLS VPN architecture to support the transparency requirements. The OSPF super-backbone, although implemented with MP-BGP, looks like a regular area to the CE-routers and the PE-routers look like Area Border Routers (ABR) even though they are in reality redistributing routes between MP-BGP and OSPF.

Additional extended BGP communities are used to propagate the OSPF route type across an MP-BGP backbone. The OSPF route type carried in the MP-BGP update received by the PE-router is used to generate a summary LSA in the OSPF topology database. An additional bit (called the *down* bit) is used in the Options field of the OSPF header to prevent routing loops between MP-BGP and OSPF. The same bit is also used on the PE-routers to prefer MP-BGP routes over OSPF routes derived from MP-BGP routes through redistribution.

Traffic may be suboptimally forwarded in some special network scenarios, where two sites connected to the superbackbone belong to the same area and also have a backdoor link for backup purposes. To handle this special case, the concept of *sham-links* was introduced. The *sham-link* is opened between PE routers and is used to propagate intra-area OSPF routes.

Lesson Review

Answer the following questions:

1. Why is the OSPF super-backbone needed in MPLS VPN environments?

2. What is the interaction between Area 0 and super-backbone?

3. What is the interaction between super-backbone and other areas?

4. How are OSPF route attributes propagated across MPLS VPN backbone?

5. What is the purpose of the Down bit in LSA header?

6. What is the influence of the Down bit on route selection process? Why is this influence needed?

7. When should a sham link be used?

Configuring and Monitoring OSPF in an MPLS VPN Environment

Objectives

Upon completion of this lesson, the learner will be able to perform the following tasks:

- Configure OSPF in a customer VPN

- Configure sham links

- Monitor MPLS VPN-specific attributes in an OSPF topology database

- Monitor OSPF-specific extended communities in an MP-BGP table

Follow these steps to configure OSPF as the PE-CE routing protocol:

- **Configure per-VRF copy of OSPF**
- **Configure redistribution of MP-BGP into OSPF**
- **Configure redistribution of OSPF into MP-BGP**

www.cisco.com

MPLS v2.1 -40

Configuring OSPF as a PE-CE routing protocol is performed in three steps:

Step 1 Configure per-VRF copy of OSPF process and define all usual OSPF parameters (networks, areas, neighbors).

Step 2 Configure redistribution of MP-BGP into OSPF.

Step 3 Configure redistribution of OSPF into MP-BGP.

Note Contrary to conventional wisdom, two-way redistribution between OSPF and MP-BGP is safe in MPLS VPN environments because of additional mechanisms that prevent routing loops or suboptimal routing.

Per-VRF OSPF Configuration

router(config)#

```
router ospf process-id vrf name
 ... Standard OSPF parameters ...
```

- **This command starts per-VRF OSPF routing process**
- **The total number of routing processes per router is limited to 32**

router(config-router)#

```
redistribute bgp as-number subnets
```

- **This command redistributes MP-BGP routes into OSPF. The Subnets keyword is mandatory for proper operation**

www.cisco.com

MPLS v2.1 –41

The per-VRF OSPF process is started with the **router ospf** *process-id* **vrf** *name* command.

Note	A separate OSPF process is needed for every VRF in the router, even if the VRFs participate in the same VPN. As the number of routing processes in Cisco IOS is limited to 32, the number of OSPF customers that can be supported by any single PE-router is limited.

The redistribution of MP-BGP routes into OSPF is configured with the **redistribute bgp** *as-number* **subnets** command. The **subnets** keyword is mandatory for proper MPLS VPN operation; otherwise only the major networks would be redistributed from MP-BGP into OSPF.

Instead of route redistribution from MP-BGP, the default route could be announced into OSPF sites with the **default-information originate always** command.

Configuring Route Redistribution

```
router(config)#

router bgp as-number
  address-family ipv4 vrf vrf-name
   redistribute ospf process-id [match [internal] [external-1] [external-2]]
```

- **OSPF to BGP route redistribution is configured with the redistribute command under the proper address-family**
- **Without the OSPF match keyword specified, only internal OSPF routes are redistributed into OSPF**

© 2002, Cisco Systems, Inc.   www.cisco.com   MPLS v2.1 -42

The OSPF routes are redistributed into MP-BGP with the **redistribute ospf process-id** command, which needs to be configured in the proper VRF address family. The VRF address family is selected with the **address-family ipv4 vrf name** command.

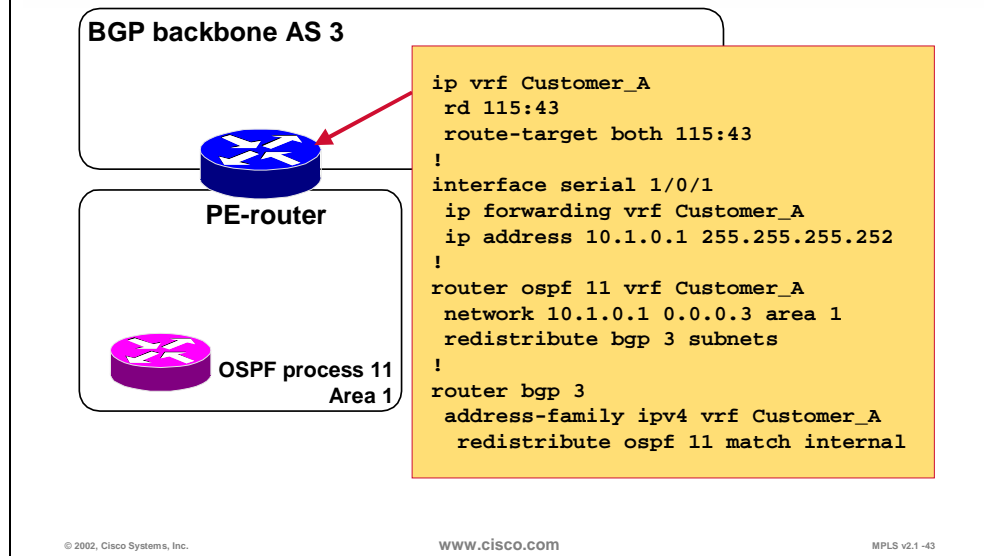| Note	Please refer to the **MPLS VPN Implementation** module for more information on BGP address families.

The **redistribute** command with no addition parameters will redistribute only internal OSPF routes into MP-BGP. Redistribution of external OSPF routes into MP-BGP must be configured manually with the **match** option of the **redistribute** command. The command **redistribute ospf process-id match internal external 1 external 2** can be used to redistribute all OSPF routes into MP-BGP.

Sample Super-Backbone Router Configuration

BGP backbone AS 3

PE-router

OSPF process 11 Area 1

```
ip vrf Customer_A
 rd 115:43
 route-target both 115:43
!
interface serial 1/0/1
 ip forwarding vrf Customer_A
 ip address 10.1.0.1 255.255.255.252
!
router ospf 11 vrf Customer_A
 network 10.1.0.1 0.0.0.3 area 1
 redistribute bgp 3 subnets
!
router bgp 3
 address-family ipv4 vrf Customer_A
  redistribute ospf 11 match internal
```

www.cisco.com MPLS v2.1 –43

The figure illustrates a site connected to the MPLS/VPN backbone. OSPF is used as the CE to PE routing protocol at the site. The MPLS/VPN backbone is the BGP based superbackbone for OSPF.

The PE router has the ordinary vrf and interface configurations with **ip vrf forwarding** as described in the MPLS VPN Implementation module. The sample configuration then continues with the OSPF configuration. A dedicated OSPF process is created for the vrf. The keyword **vrf** in the **router OSPF** configuration line indicates that this OSPF process will be used inside the vrf and that the PE router should take the role as ABR connected to the OSPF superbackbone.

The **network** configuration command refers to ip addresses inside the vrf. In this case the network statement will match the ip address assigned to the vrf interface Serial 1/0/1. Interface Serial 1/0/1 will be in area 1.

Redistribution from bgp into ospf will be done in the vrf routing context because the OSPF process is totally inside the vrf.

Redistribution from OSPF into BGP is configured in the BGP routing protocol context of the vrf.

Sham Link End-Point Addresses

- **A separate /32 address space is required in each PE router for the sham link**
- **This /32 address space:**
  - **Is required so that OSPF packets can be sent over the VPN backbone to the remote end of the sham-link**
  - **Must belong to the vrf**
  - **Must not be advertised by OSPF**
  - **Must be advertised by BGP**

www.cisco.com

MPLS v2.1 -44

A sham link is required only if there are two (or more) sites which belong to the same area and have a backdoor link between them which should be used for backup purposes only. The MPLS/VPN backbone is the primary means of communication between the two sites.

The sham link must have two distinct end points. One end point is in the vrf of one PE router where the first site is connected. The other is in the vrf of the other PE router where the second site is connected. These two end points are identified by IP addresses. IP packets will be exchanged between the sham link end points across the MPLS/VPN backbone. Therefore a separate /32 address space is required in each PE router.

The /32 address space must belong to the vrf since the sham link is used by the OSPF process inside the vrf. But it should be reachable only across the MPLS/VPN backbone. The IP packets that are exchanged between the end points should be forwarded via the backbone. The /32 address space should not be advertised via OSPF. If they were, the end points may reach each other across the site rather than across the backbone.

Configuring Sham Link

router(config-router)#

area *area-id* sham-link *source-address destination-address* cost *number*

- **This command was introduced in IOS 12.2(8)T**
- **The sham link belongs to the specified area**
- **Sham link packets sent across the MPLS/VPN backbone will have the specified source and destination addresses**
- **When the SPF algorithm is executed, the sham link will have the specified cost**

www.cisco.com

MPLS v2.1 –45

Support for sham links was added in Cisco IOS 12.2(8)T

area sham-link cost

To configure a sham-link interface on a provider edge (PE) router in a Multiprotocol Label Switching (MPLS) Virtual Private Network (VPN) backbone, use the area sham-link cost command in global configuration mode. To remove the sham-link, use the no form of this command.

area *area-id* sham-link *source-address destination-address* cost *number*

no area *area-id* sham-link *source-address destination-address* cost *number*

Syntax Description

area-id	ID number of the Open Shortest Path First (OSPF) area assigned to the sham-link. Valid values: numeric value or valid IP address. There is no default.
source-address	IP address of the source PE router.
destination-address	IP address of the destination PE router.
Number	OSPF cost to send IP packets over the sham-link interface. Valid values are from 1 to 65535.

Sample Sham Link Configuration

BGP backbone AS 3

PE-router

OSPF process 11 Area 1

```
ip vrf Customer_A
 rd 115:43
 route-target both 115:43
!
interface Loopback11
 ip forwarding vrf Customer_A
 ip address 10.2.1.1 255.255.255.255
!
interface serial 1/0/1
 ip forwarding vrf Customer_A
 ip address 10.1.0.1 255.255.255.252
!
router ospf 11 vrf Customer_A
 network 10.1.0.1 0.0.0.3 area 1
 redistribute bgp 3 subnets
 area 1 sham-link 10.2.1.1 10.2.1.2 cost 10
!
router bgp 3
 address-family ipv4 vrf Customer_A
  network 10.2.1.1 mask 255.255.255.255
  redistribute ospf 11 match internal
```

www.cisco.com MPLS v2.1 -46

This figure illustrates one of the PE router configurations. The other end of the sham link is another PE router with corresponding configurations.

This end of the sham link is configured using the **area sham-link** configuration command in the router OSPF configuration mode. The sham link belongs to area 1. It is an unnumbered intra-area link. The source address must be an address space that belongs to the vrf but is only advertised by BGP. A loopback interface is configured for this purpose.

The router OSPF configuration does not include a network statement for the address 10.2.1.1. This means that the loopback interface will not be advertised by OSPF into the customer site. This ensures that the other end of the sham link will not try to reach this end via the customer site. Instead the address space is advertised by BGP. The router BGP configuration does include a network statement for the address. The PE routers will exchange the packets that constitute the sham link only across the MPLS/VPN backbone.

The majority of OSPF-related **show** commands will display MPLS VPN-specific OSPF parameters on the PE-routers. The **show ip bgp vpnv4 vrf** *name prefix mask* command will also display detailed information on the MP-BGP route including the extended BGP route communities.

show ip ospf

```
Router#show ip ospf
Routing Process "ospf 250" with ID 10.2.3.4
 Supports only single TOS(TOS0) routes
 Supports opaque LSA
 Connected to MPLS VPN Superbackbone
 It is an area border and autonomous system boundary
      router
 Redistributing External Routes from,
    bgp 1, includes subnets in redistribution
```

www.cisco.com

MPLS v2.1 -48

The **show ip ospf** command displays whether the router is a PE-router (and thus connected to the MPLS VPN super-backbone). A PE-router is always an area border router (ABR) or an AS boundary router (ASBR).

show ip ospf *type prefix*

```
Router#show ip ospf database summary 10.0.1.0
OSPF Router with ID (10.4.3.2) (Process ID 250)

                Summary Net Link States (Area 0)

  LS age: 1298
  Options: (No TOS-capability, DC, Downward)
  LS Type: Summary Links(Network)
  Link State ID: 10.0.1.0 (summary Network Number)
  Advertising Router: 10.2.3.4
  LS Seq Number: 80000002
  Checksum: 0x5C2F
  Length: 28
  Network Mask: /24
        TOS: 0  Metric: 7435
```

The **show ip ospf database** *type prefix* command displays the status of the *down* bit. If the *down* bit is set, you will see the keyword *Downward* displayed in the Options field of the LSA. If the bit is not set, the keyword *Upward* will be displayed.

show ip bgp vpnv4 vrf *name prefix*

```
Router#show ip bgp vpnv4 vrf Customer_A 10.0.1.0
BGP routing table entry for 1:10:10.0.1.0/24, version 64
Paths: (1 available, best #1, table Customer_A)
  Advertised to non peer-group peers:
  10.2.3.6
  Local
    10.2.3.4 from 0.0.0.0 (10.2.3.4)
      Origin incomplete, metric 2, localpref 100, weight
      32768, valid, sourced, best
      Extended Community: RT:100:27 OSPF DOMAIN ID:0.0.0.11
      OSPF RT:0:3:0
```

The **show ip bgp vpnv4 vrf** *customer prefix* command displays all details of a MP-BGP route, including the OSPF extended BGP community. In the printout above, the route redistributed into MP-BGP from OSPF process number 11 was a summary route (LSA type 3) redistributed from OSPF area 0.

Summary

The OSPF process in a VRF is started with the **router ospf** *process* **vrf** *name* command. As the overall number of routing processes per router is limited to 32, a single PE-router can serve only a small number of VRFs.

Two-way redistribution between BGP and OSPF is usually configured. The redistribution is safe because of additional attributes introduced with the super-backbone architecture.

By default, only major networks are redistributed into OSPF. Redistribution of subnets needs to be configured with the **redistribute … subnets** command.

By default, only internal OSPF routes are redistributed from OSPF into MP-BGP. Redistribution of external routes has to be configured with the **redistribute … match** *route-type-list* command.

The **show ip ospf** command will display whether a router is a PE-router connected to the MPLS VPN backbone. The detailed printouts from the **show ip ospf database** command will display the value of the *down* bit. The detailed printouts from the **show ip bgp** command will display the OSPF-specific extended BGP community.

Sham links must be manually configured in both end points. The configuration is done in the PE routers, in the vrf OSPF configuration mode. Source and destination IP addresses is configured and these addresses must belong to the vrf and be advertised across the BGP backbone but not be advertised into the customer sites.

Lesson Review

1. How can you verify if an OSPF route was received from a local OSPF router or through MPLS VPN backbone?

2. How can you verify if your router is participating in OSPF super-backbone?

3. How can you display OSPF-related extended communities attached to a route?

Summary

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

■ Describe OSPF operation inside a VPN

■ Describe the enhanced OSPF hierarchical model

■ Describe the interactions between OSPF and MP-BGP

■ Use OSPF as the PE-CE routing protocol in complex MPLS VPN environment