IPv6 **over** MPLS (Cisco 6PE)

Introduction

There are multiple techniques available to integrate IPv6 services over Service Providers core backbones: dedicated IPv6 network running over various data link layers, dual stack IPv4-IPv6 backbone, or leveraging of an existing MPLS backbone. These solutions [IPv6_Deploy] are deployed on Service Providers backbones when the amount of IPv6 traffic and the revenue generated are in line with the necessary investments and the risks consented.

Conditions are favorable for the introduction of native IPv6 service, from the edge, in a scalable way, without any IPv6 addressing restrictions and without putting a well-controlled IPv4 backbone in jeopardy. Backbone stability is key for Service Providers, which recently stabilized their IPv4 infrastructure.

Service Providers running an MPLS/IPv4 infrastructure follow the same trends, as several integration scenarios are possible to offer IPv6 services on an MPLS network. Cisco Systems specially developed IPv6 Provider Edge Router over MPLS—(Cisco 6PE) to meet all of those requirements.

Service Providers that already deploy MPLS, or plan to do so, can garner the following benefits from Cisco 6PE:

- Minimal operational cost and risk
  - No impact on existing IPv4 and MPLS services
- Provider Edge routers upgrade only
  - 6PE router can be an existing PE router or a new one dedicated to IPv6 traffic
- No impact on IPv6 customer edge routers
  - The ISP can connect to any Customer CE running Static, IGP or EGP
- Ready for production services
  - An ISP can delegate IPv6 prefixes from its own to the customers
- Non-disruptive IPv6 introduction into an existing MPLS service
  - 6PE routers can be added at any time
- Switching up to OC-192 speed is possible in the core

Deploying IPv6 over MPLS Backbones

Cisco 6PE (IPv6 over MPLS)-enabled backbones allow IPv6 domains to communicate with each other over an MPLS IPv4 core network. This implementation requires no backbone infrastructure upgrades and no reconfiguration of core routers, because forwarding is based on labels rather than on the IP header itself. This provides a very cost-effective strategy for Ipv6 deployment.

Additionally, the inherent Virtual Private Network (VPN) and Traffic Engineering (TE) services available within an MPLS environment allow IPv6 networks to be combined into VPNs or extranets over an infrastructure that supports IPv4 VPNs and MPLS-TE.

In addition to Cisco 6PE, additional deployment strategies are available:

- IPv6 using tunnels on the customer edge (CE) routers
- IPv6 over a circuit transport over MPLS

Table 1 summarizes the primary use, benefits, and limitations for each MPLS mechanism.

Table 1 MPLS Mechanisms: Primary Uses, Benefits and Limitations

MPLS Mechanism	Primary Use	Benefits	Limitations	Requirements
IPv6 Using Tunnels on CE Routers	Enterprise customers who want to use IPv6 over existing MPLS IPv4 services	VPN-like IPv6 clouds isolation	Routers use IPv4-compatible or 6to4 addresses	
Difficult to deploy global IPv6 addresses	Dual-stack CE routers			
IPv6 over a Circuit Transport over MPLS	Service Providers with ATM or Ethernet links to CE routers	Fully transparent IPv6 communication	Layer 2 VPN complexity equivalent	Any router with AtoM feature set support
IPv6 on PE Routers	Internet and mobile service providers wanting to offer an IPv6 service	No software upgrade or reconfiguration of the MPLS core		
Layer 3 VPN complexity equivalent	Requires BGP understanding as well as MPLS	IPv6 software upgrade for PE routers		

Cisco IOS IPv6 currently supports all of these strategies. Another conceptual strategy would be to run a native IPv6 MPLS core, but this strategy requires a full network upgrade of all P and PE routers, with dual control planes for IPv4 and IPv6. The core IGP and Label Distribution protocol would require an upgrade that Service Providers prefer to avoid.

The following sections describe each mechanism briefly. Cisco 6PE is extensively covered later as the main topic.

IPv6 Using Tunnels on the Customer Edge Routers

Using tunnels on the CE routers is the simplest way to deploy IPv6 over MPLS networks. It has no impact on the operation or infrastructure of MPLS, and requires no changes to either the P routers in the core or the PE routers connected to the customers. However, a difficult to scale tunnel meshing is required as the number of CEs to connect increases. It is also difficult to delegate a global IPv6 prefix for an ISP.

Communication between the remote IPv6 domains uses standard tunnelling mechanisms: running IPv6 over IPv4 tunnels on the CEs, similarly to the way that MPLS VPNs support native IPv4 tunnels. The CE routers must be upgraded in order to be dual-stack, and configured for IPv4-compatible or 6to4 tunnels, but communication with the PE routers is IPv4. The MPLS domain perceives the traffic to be IPv4. The dual-stack routers use the IPv4-compatible or 6to4 address, rather than an IPv6 address supplied by the Service Provider. Figure 1 depicts the network architecture using tunnels on the CE routers.

Figure 1. IPv6 Using Tunnels on the CE Routers
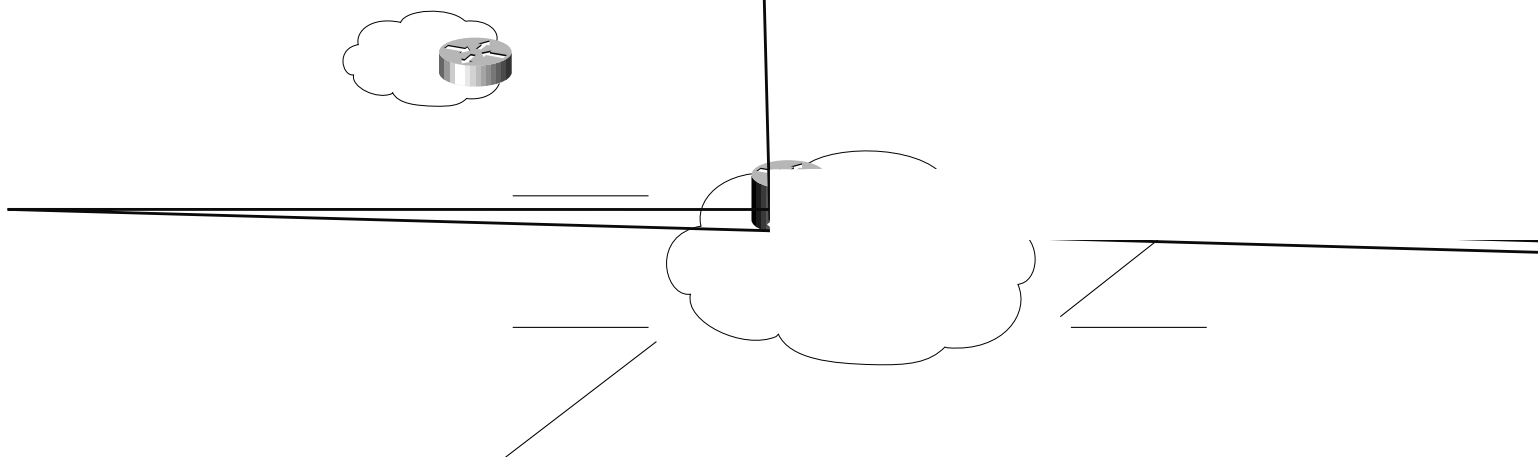
IPv6 over a Circuit Transport over MPLS

The use of Any Transport over MPLS (AtoM) for deploying IPv6 over MPLS networks does not impact the operation or infrastructure of MPLS (neither P router software version nor configuration is changed). However, PE routers connected to customers must be upgraded with AToM. It is also difficult to delegate a global IPv6 prefix for an ISP.

Communication between the remote IPv6 domains runs native IPv6 protocols over a dedicated link, where the underlying mechanisms are fully transparent to IPv6. The IPv6 traffic is tunneled via AToM, with the IPv6 routers connected through an ATM or Ethernet interface, respectively. This potential heavy tunnel meshing is a strong drawback of the solution. Figure 2 depicts the network architecture for IPv6 with AToM.

Figure 2. IPv6 over a Circuit Transport over MPLS
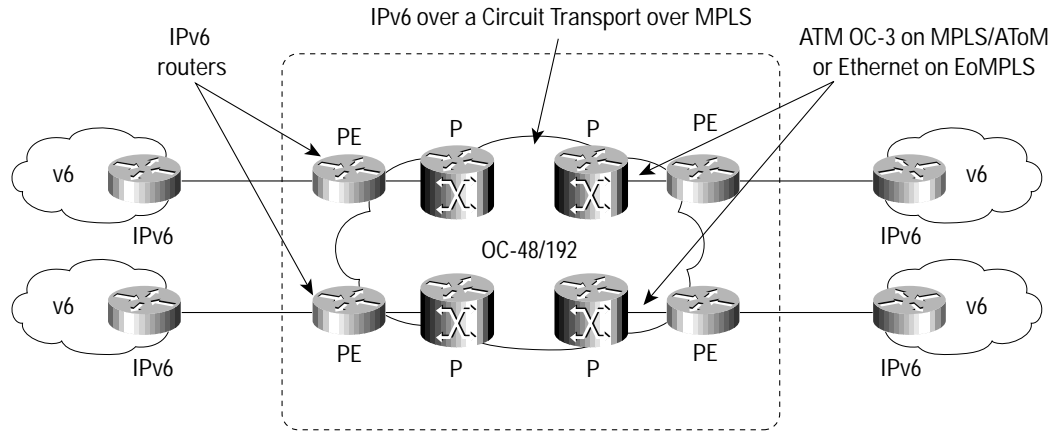
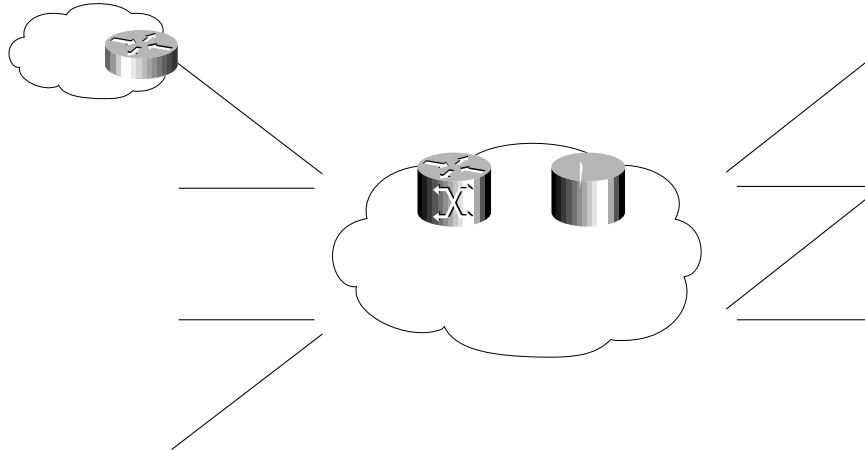## IPv6 on the Provider Edge Routers (Cisco 6PE)

The Cisco 6PE feature is particularly applicable to Service Providers who already run an MPLS network or plan to do it. One of the Cisco 6PE advantages is that there is no need to upgrade the hardware, software or configuration of the core network. Thus it eliminates the impact on the operations and the revenues generated by the existing IPv4 traffic. MPLS has been chosen by many Service Providers as a vehicle to deliver services to customers. MPLS as a multi-service infrastructure technology is able to provide layer 3 VPN, QoS, traffic engineering, fast re-routing and integration of ATM and IP switching. It is in a very natural manner that MPLS is put to contribution to ease IPv6 introduction in existing production networks.

MPLS de-coupling of the Control Plane and Data Plane provide an interesting alternative to the integration and co-existence of IPv4, IPv6 and ATM over a single infrastructure, thus fulfilling environments such as 3G networks where UMTS Release 5 needs in terms of transport: Cisco 6PE for IPv6 traffic, ATM over MPLS and regular IPv4 switching with its VPN, traffic engineering and QoS extensions. From an operational standpoint, new CEs introduction is straightforward and painless as it leverages the Layer 3 VPN scalability.

Figure 3 depicts the network architecture for the Cisco 6PE solution.

Figure 3. IPv6 on the Provider Edge Routers
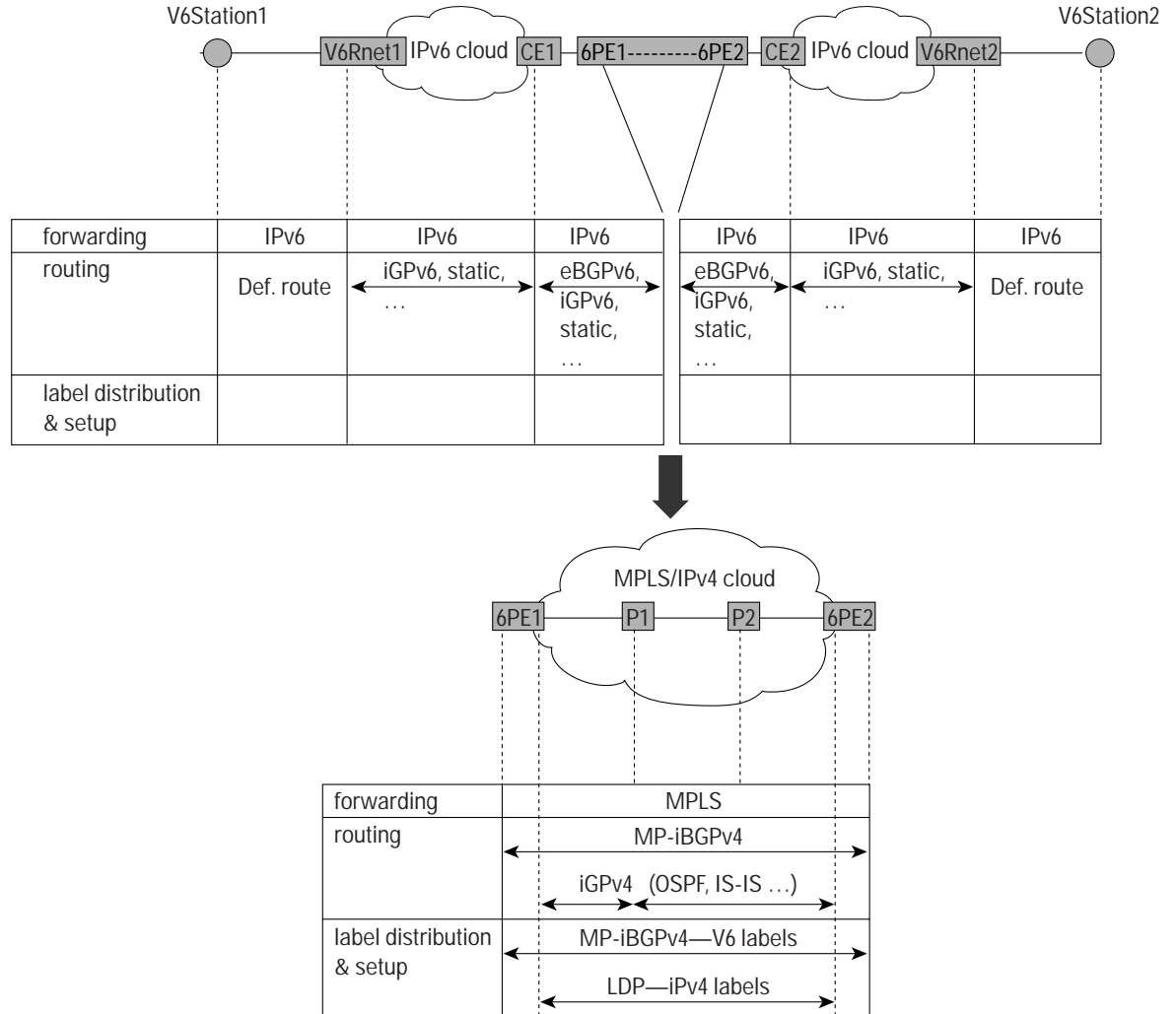

Cisco 6PE Detailed Architecture

Cisco 6PE enables IPv6 islands to communicate with each other over an MPLS/IPv4 core network using MPLS LSPs (Label Switched Path). The method relies on BGP extensions in the IPv4 network Provider Edge Routers (6PE) to exchange IPv6 reachability information along with an MPLS label for each IPv6 address prefix announced. 6PE routers are dual-stack (IPv4 and IPv6).

In order for the IPv6 transport to be transparent to all but 6PE routers, it is necessary to impose a hierarchy of labels at the 6PE ingress router. The top label provides connectivity inside the IPv4 MPLS core network: LDP (Label Distribution Protocol), TDP (Tag Distribution Protocol) or RSVP (Resource Reservation Protocol) distributes it. The next label is used at each 6PE egress router for IPv6 forwarding: it is distributed by MP-BGP (Multi-protocol BGP) in the "IPv6+label" address family.

Figure 4 depicts the routing/forwarding mechanisms used end-to-end when two IPv6 capable stations communicate over IPv6 clouds, themselves connected over an MPLS cloud.

Figure 4. Global Cisco 6PE architecture

V6Station1 ... V6Rnet1 | IPv6 cloud | CE1 | 6PE1---------6PE2 | CE2 | IPv6 cloud | V6Rnet2 ... V6Station2

forwarding	IPv6	IPv6	IPv6	IPv6	IPv6	IPv6
routing	Def. route	←— iGPv6, static, … —→	eBGPv6, iGPv6, static, …	eBGPv6, iGPv6, static, …	←— iGPv6, static, … —→	Def. route
label distribution & setup						

MPLS/IPv4 cloud

6PE1 — P1 — P2 — 6PE2

forwarding	MPLS		
routing	←— MP-iBGPv4 —→		
		←— iGPv4 (OSPF, IS-IS …) —→	
label distribution & setup	←— MP-iBGPv4—V6 labels —→		
	←— LDP—IPv4 labels —→		

The following devices can be found in Figure 4:

- *V6Station*: client stations running IPv6 application and an IPv6 stack.
- *V6R*: IPv6 router on the customer premises.
- *V6CE*: Customer Edge equipment, connected to the Provider Edge equipment and running an IPv6 stack.
- *6PE*: Provider Edge equipment, connected to V6CEs as well as entry points to the MPLS cloud and running in a dual stack mode.
- *P*: Provider routers, core of the MPLS backbone running MPLS and an IPv4 stack.

Four types of routing interactions can be found on the path to or from the IPv6 stations:

1. IPv6 clouds running an IPv6 Internal Gateway Protocol (RIPng, IS-ISv6, OSPFv3)
2. V6CE and 6PE router exchanging IPv6 routing information through an IPv6 External Gateway Protocol (MP-BGP), Interior Gateway Protocol (IGP) or using static routes

3. 6PE router peering together through MP-BGP4 for exchanging IPv6 reachability with the MPLS cloud and performing IPv6 forwarding.

4. MPLS/IPv4 cloud, running an IPv4 Internal Gateway Protocol (OSPF or IS-IS) to establish reachability between 6PEs, an IPv4 label distribution protocol (LDP, TDP, RSVP) to distribute IPv4 labels.

V6stations, V6R routers and V6CE routers are not aware that forwarding occurs over MPLS clouds. They operate in their regular IPv6 way. Therefore there is no impact on these boxes because of the Cisco 6PE feature. ISP who delegates /48 or /64 from their registered IPv6 prefixes can do it over a Cisco 6PE infrastructure.

P routers inside the MPLS clouds are not aware that they are switching IPv6 packets, as they only use MPLS forwarding, LDP, TDP or RSVP for binding IPv4 labels and an IPv4 IGP (OSPF or IS-IS) to establish internal reachability inside the MPLS cloud. Therefore, the Cisco 6PE feature does not impact these MPLS core devices.

Finally the 6PE routers have to:

1. Participate in IPv4 IGP to establish internal reachability inside the MPLS cloud

2. Participate in LDP or TDP for binding IPv4 labels

3. Run Multi-Protocol iBGP(MP-iBGP) to advertise IPv6 reachability and distribute aggregate[1] IPv6 labels among them

4. Run MP-eBGP, an IPv6 IGP or static routing with CE routers to advertise IPv6 reachability learnt from their peers over the MPLS cloud.

Cisco is working at IETF to ensure progress of corresponding standardization. In particular, the Cisco 6PE scheme is compliant with the more recent versions of the IETF draft on "Connecting IPv6 Domains across IPv4 Clouds with BGP" [IPv6_BGP].

Cisco 6PE Control Plane

From a control plane perspective, the main principles of the solution are:

1. The 6PE MP-BGP routers are dual-stack: IPv6 toward the CE and IPv4 toward the MPLS core.

2. Multi-Protocol iBGP is used between 6PE routers to exchange IPv6 reachability information. A label is bounded to each advertised destination IPv6 prefix. This label, the Aggregate IPv6 Label, is allocated by the egress 6PE from a local pool of 16 labels (there is a single pool of 16 labels used for all IPv6 prefixes). The 6PE router creates an entry in its MPLS table for every "Aggregate IPv6 label," which points to a "pop label + IPv6 lookup[2]" operation in the IPv6 FIB. Through MP-BGP, an ingress 6PE router knows the IPv4 address of the remote egress 6PE router (IPv4 mapped IPv6 address) to reach a destination IPv6 prefix.

3. At recursion time, the ingress 6PE extracts the IPv4 address contained in the IPv4 mapped IPv6 address. Then the ingress 6PE resolves this IPv4 address (using the IPv4 routing table) in order to get the label associated to the LSP for this destination. This "IPv4 label" has been populated in the IPv4 tables through regular IPv4 MPLS control plane procedures using IPv4 IGP and IPv4 label distribution protocols. This "IPv4 label" is then stored along with the BGP label for the destination IPv6 subnet in the IPv6 forwarding table of the ingress PE router.

---

1. On GSR and some Cisco IOS Software releases, labels are not aggregated but unique per IPv6 prefixes.
2. When labels are not aggregated, the packet is MPLS switched, no IPv6 lookup is performed.

Multi-Protocol BGP Extension for "labeled IPv6 prefixes"

Multi-protocol BGP already supports the IPv6 address family via the Multi-protocol Reachable NLRI (MP_REACH_NLRI) and Multi-protocol Unreachable NLRI (MP_UNREACH_NLRI) attributes.

Cisco 6PE requires that Multi-protocol BGP be further extended to be able to bind an MPLS label to the IPv6 route as per [MPLS_BGP]. The main aspects of this extension for support of [MPLS_BGP] are presented below:

NLRI (Network Layer Reachability Information) Format

The label binding information is piggybacked along the prefix advertisement in the same MP_REACH_NLRI attribute. The fact that the MP_REACH_NLRI contains a label is indicated by SAFI (Sub-Address Family) value of 4.

The NLRI for labeled IPv6 routes (used in "NLRI" field of MP_REACH_NLRI and in "Withdrawn Routes" field of MP_UNREACH_NLRI) contains one or more triple <Length, Label, Prefix>:

> Length: total length of the label plus prefix
>
> Label: In the 6PE case, this field carries one label, where:
>
>> Label value: 20 high order bits
>>
>> Unused: 3 following bits set to zero
>>
>> Bottom of stack bit: low order bit
>
> Prefix: IPv6 prefix of destination

Capability Negotiation

Cisco IOS Software already supports Capability Negotiation as specified in [BGP_CAP], but the Cisco 6PE feature extends BGP capability negotiation for supporting "IPv6+label" capability in the following way:

- 6PE advertises capability for "IPv6+label" to a neighbor when configured to do so for this neighbor via the new command (see Command Line Interface section in this document).

- 6PE also advertises capability for unlabeled IPv6 since there is a separate Capabilities Optional Parameters for each SAFI (Subsequent Address Family Identifier).

- If a neighbor has advertised "IPv6+label" capability, the 6PE advertises all IPv6 routes as labeled routes.

- If a neighbor has not advertised "IPv6+label" capability but has advertised "IPv6" capability, the 6PE advertises all IPv6 routes as IPv6 (unlabelled) routes to this neighbor. Note that if a 6PE receives unlabelled IPv6 routes, then the 6PE does not resolve the recursion and marks these prefixes as unreachable in the IPv6 routing table so that packets to this destination get dropped and not sent into the MPLS cloud. This behavior avoids having a penultimate MPLS/IPv4 P router dropping an IPv6 packet because of Penultimate Hop Popping (PHP).

Control Plane Function at Egress 6PE

In terms of routing function, the egress 6PE is responsible to learn by usual means and to advertise through MP-BGP the IPv6 prefixes in its reach:

1. 6PE2[3] includes, in its IPv6 routing table, prefix P2 reachable through connected interfaces or connected CEs. Any IPv6 routing protocol can run between CEs and 6PEs (RIPng, OSPFv3, ISISv6, MP-BGP) or static routing to reduce configuration complexity.

---

3. See figure 4 for router names.

2. 6PE2 redistribute P2 prefixes in MP-BGP under the "IPv6+label" address family.

3. MP-BGP assigns a local label to prefix P2. This label is an aggregate label[4]. It is not specific to a given prefix, it is chosen out of a sixteen-label pool in a round robin fashion. Consequently, egress 6PE is not able to take a forwarding decision solely based on this label. The advantage of using more than one label value is that it results in better load sharing in the core of the network, since the P routers hash the bottom label value because the encapsulated packet is not IPv4.

4. MP-BGP create a BGP update including a MP_REACH_NLRI attribute with AFI=2 (IPv6), SAFI=4 (label), Prefix = P2, Label =L2, next-hop = ::ffff:<ipv4addr of the BGP connection>. The IPv6 next-hop address is expressed as an IPv4-mapped IPv6 address.

Example

On 6PE2 router (egress 6PE), prefix 2001:33::/64 is locally connected via interface loopback0:

```
6PE2#show ipv6 route
IPv6 Routing Table - 5 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
 I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea
L 2001:33::1/128 [0/0]
 via ::, Loopback0
C 2001:33::/64 [0/0]
 via ::, Loopback0
```

And then injected in MP-BGP:

```
6PE2#show bgp ipv6
BGP table version is 9, local router ID is 192.168.0.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
 Network Next Hop Metric LocPrf Weight Path
*> 2001:33::/64 :: 32768 i
```

Aggregate label 17 is associated to this route:

```
6PE2#show bgp ipv6 tags
 Network Next Hop            In tag/Out tag
 2001:33::/64 ::             17/notag
6PE2#show mpls forwarding-table
Local Outgoing Prefix        Bytes tag          Outgoing    Next Hop
tag tag or VC or Tunnel Id   switched           interface
17 Aggregate IPv6            1040
```

Control Plane Function at Ingress 6PE

1. 6PE1 extract the <prefix, label , next-hop=6PE2> triple from MP_REACH_NLRI attributes included in MP-BGP updates coming from 6PE2. In turn the <prefix, label> couple is installed in the MP-BGP table for AFI=2(IPv6)/ SAFI=4(Label) and the IPv4 address is extracted from the "IPv4 mapped" IPv6 address which was advertised as next-hop.

---

4. When the labels are not aggregated (one label per prefix is assigned) load sharing is performed as well.

2. This IPv4 address must be present in the IPv4 routing table and a LSP must exist to this destination. If the above conditions are not both met, the IPv6 prefix is declared inaccessible.

3. A lookup for the extracted IPv4 next-hop is performed in the routing table to return the metric to this IPv4 address.

4. If a metric is found, it is associated to the IPv6 prefix learnt from 6PE2. Prefix P2 is then inserted in IPv6 RIB (Routing Information Base). In case a metric is not found the BGP scanner mechanism re-runs step (2) on a regular basis.

5. 6PE1 triggers the 6PE resolution algorithm for prefix P2. This leads to complete a two label MPLS stack for prefix P2. This algorithm is explained in the next section.

Example

Prefix 2001:33::/64 is learnt by 6PE1 router (ingress 6PE) from 6PE2 router.

```
6PE1#show bgp
BGP table version is 5, local router ID is 2.6.0.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
 Network Next Hop Metric LocPrf        Weight        Path
*>i2001:33::/64 ::FFFF:192.168.0.2 100   0                 I
6PE1#show ipv6 route
IPv6 Routing Table - 6 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
 I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea
B 2001:33::/64 [200/0]
 via ::FFFF:192.168.0.2, IPv6-mpls
6PE1#show ipv6 cef
2001:2::2/128
 Receive
2001:33::/64
 nexthop ::FFFF:192.168.0.2
 fast tag rewrite with Et1/2, 172.16.0.49, tags imposed: {16 17}
```

Cisco 6PE Resolution

For each IPv6 prefix learnt by the way of the Cisco 6PE mechanism and present in the routing table of the ingress 6PE, the following steps are rolled out by 6PE to perform the IPv6 next-hop resolution. Each steps refers to figure 5.

1. For each IPv6 prefix the IPv6 next-hop is transmitted by MP-BGP within the "IPv6+label" updates. The associated label is the bottom of the stack MPLS label for the new adjacency associated with this prefix in the CEFv6 table.

2. The IPv4 next-hop is extracted from the IPv6 next-hop: it is encoded in an IPv4-mapped IPv6 address.

3. In the IPv4 CEF table, the entry built for the IPv4 next-hop includes an IPv4 MPLS label. This label is the top of the stack label for the new layer 2 header associated with this prefix in the CEFv6 table. It is a "sine qua non" condition for Cisco 6PE to work. If there is no LSP to reach the IPv4 next-hop, it cannot work.
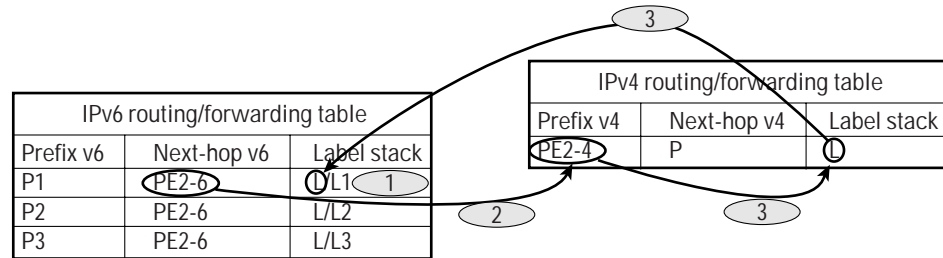
Figure 5. Cisco 6PE resolution at ingress 6PE

IPv6 routing/forwarding table		
Prefix v6	Next-hop v6	Label stack
P1	PE2-6	L/L1
P2	PE2-6	L/L2
P3	PE2-6	L/L3

IPv4 routing/forwarding table		
Prefix v4	Next-hop v4	Label stack
PE2-4	P	L

Figure 6. Resolution for prefix 2001:33::/64

```
6PE1#show ipv6 cef
2001:33::/64
     nexthop ::FFFF:192.168.0.2
     fast tag rewrite with Et1/2, 172.16.0.49,
tags imposed: {16 17}
```

```
6PE1#show ip cef 192.168.0.2
192.168.0.3/32, version 46, cached adjacency 172.16.0.49
0 packets, 0 bytes
 tag information set
   local tag: 17
   fast tag rewrite with Et1/2, 172.16.0.49, tags imposed: {16}
 via 172.16.0.49, Ethernet1/2, 0 dependencies
 next hop 172.16.0.49, Ethernet1/2
 valid cached adjacency
 tag rewrite with Et1/2, 172.16.0.49, tags imposed: {16}
```

```
6PE1#show bgp tags
   Network              Next Hop         In tag/Out tag
   2001:33::/64      ::FFFF:192.168.0.2    notag/17
```

This Cisco 6PE resolution algorithm generally occurs in the control plane operation, when the router computes the routing, forwarding and adjacency table.

In the common case, the resolution algorithm is triggered by:

- BGP on receipt of an update containing an MP_REACH_NLRI to insert a prefix in the MP-BGP table
- The BGP scanner to insert a prefix in the routing table
- Change in the MPLS table

An exception is the multi-path case, during which the next- hop is resolved at forwarding time when the router assigns the next hop based on the source, destination address couple. This case is addressed in a later section of this document.

Cisco 6PE Forwarding Plane

Packet forwarding can take place after the control plane actions are performed, and the routing table and forwarding table have been populated. The main principles of the solution are:

- Normal IPv6 forwarding is used in the IPv6 clouds
- When receiving an IPv6 packet, the ingress PE does a lookup in the IPv6 FIB. It finds the MPLS header to impose on the IPv6 packet (top label is "IPv4 label" to reach egress 6PE; bottom label is "Aggregate IPv6 label").
- Normal MPLS label switching is used in the MPLS cloud.
- On the egress 6PE, the "Aggregate IPv6 label" is used to do a lookup in the LFIB that instructs the egress 6PE to do a pop and a lookup in the IPv6 FIB[5].

Figure 7 shows the forwarding header used on the path between two IPv6 stations communicating via the MPLS cloud when Penultimate Hop Popping (PHP) is used.

Figure 7. The life of an IPv6 packet crossing a Cisco 6PE enabled infrastructure
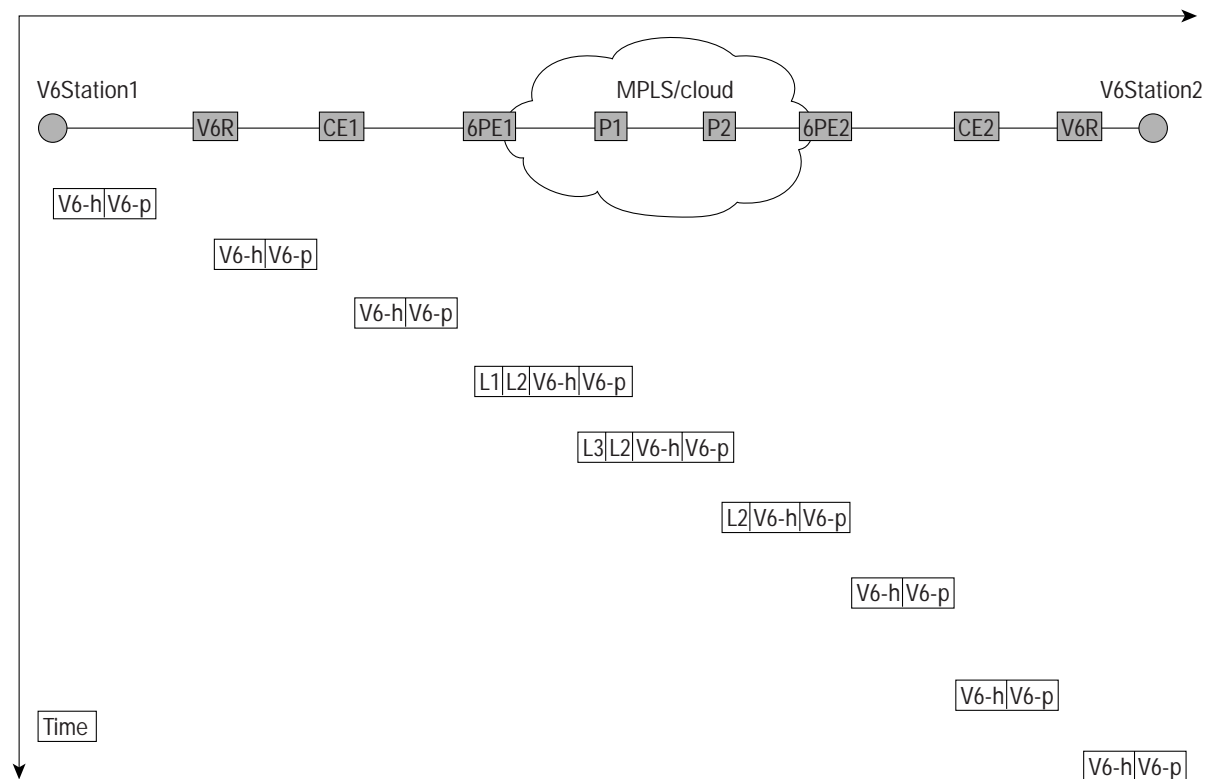
V6-h = IPv6 header.

V6-p= IPv6 payload.

L1[6] and L3 = V4 labels. L1 allows 6PE1 to reach P1 and L3 to reach P2

L2 = Aggregate IPv6 label bound to V6station2 prefix by 6PE2.

5. When labels are not aggregated, the packet is MPLS switched, no IPv6 lookup is performed.
6. See figure 1 for routers' names.

Forwarding Plane Function at Ingress 6PE Router

A packet entering the 6PE router by one of its IPv6 interface is going through the followings steps, which are CEF switched. Consequently, this provides the typical CEF performances.

1. An IPv6 lookup is completed for the IPv6 destination address. By configuration, in the particular case of a self originated packet, a source IPv6 address is assigned to this packet.

2. The forwarding information (two-label MPLS stack and MAC rewrite) is used to create the new Layer 2 and MPLS shim header of the incoming IPv6 packet.

3. At this stage, load balancing can occur at the IGP level:
   – IGP multi-path occurs when multiple paths co-exist to reach the IPv4 next-hop corresponding the IPv6 MP-BGP next-hop announced by the remote 6PE (6PE2). The hashing is then based on the 256 bits of the IPv6 source / destination address couple.
   – CEFv6 will support per-packet load-balancing in the future.

4. The new IPv6 packet is placed in the output queue of the outgoing interface for forwarding.

Forwarding Plane Function at P Router

Only regular MPLS switching occurs in the P router. Switching of MPLS-encapsulated IPv6 packets is thus supported at line rate, including OC-192 speed. A P router is only required to support MPLS, an IPv4 Label Distribution Protocol, and the IPv4 IGP of choice.

1. IPv6 packet is switched through the P routers, based on the MPLS stack top label.
   – The assumption is that P routers remain IPv6 unaware, and are thus unable to perform a hashing based on the IPv6 address (this is performed for IPv4 MPLS load balancing).
   – If load balancing occurs on a P router because there are multiple LSPs to the remote 6PE (6PE2), the hashing is based on the MPLS stack bottom label.

2. By the use of a two-label stack, Cisco 6PE becomes compatible with penultimate hop popping (PHP). The penultimate router, a P router, pops up the top label, which is the forwarding label so far. The MPLS stack stays with a one-label depth. The IPv6 packet is then forwarded to the remote 6PE with this one-label stack. The second label is mandatory when PHP is enabled. Its absence would cause the IPv6 packet to appear "nude" on the penultimate P router after the PHP is performed. As P routers are IPv6 unaware, this packet would be dropped.

Forwarding Plane Function at Egress 6PE Router

All the following steps are performed by CEF, once again providing generally expected CEF performances.

1. The IPv6 packet is encapsulated in a 2 or 1 label depth MPLS stack, depending whether or not PHP is enabled.
   – If PHP is enabled, then the label stack is one label deep. A lookup in the LFIB leads to a one-label pop-up operation.
   – If PHP is not enabled, then the label stack has a depth of two labels. A lookup on the top label leads to pop up the "explicit-null" label; a second lookup in the LFIB lead to a one-label pop up operation.

2. On the de-capsulated IPv6 packet, a lookup is performed on the CEFv6 table based on the IPv6 destination address. The new MAC header is based on the CEFv6 adjacency table; it is written and the packet is placed in the output queue of the outgoing interface for forwarding.

Command Line Interface

Cisco 6PE mainly requires the network manager to be familiar with Cisco IOS MPLS and BGP4 configuration and troubleshooting. The impact of configuring the 6PE feature through Cisco IOS CLI is minimal.

Cisco 6PE—The 2 Magic Commands

Essentially, one new command is added to advertise the "IPv6+label" capability and to start 6PE operations.

In address-family IPv6 configuration mode, specify that "IPv6+label" capability is advertised to a neighbor:

```
router(config-router-af)#neighbor <ip-address> send-label
```

An additional command is added in global configuration mode, to specify the interface to take source address from, for self-generated packets:

```
router(config)#mpls ipv6 source-interface <interface>
```

This command applies only to locally generated packets and is mandatory in that case.

A Configuration Example

The following piece of configuration is a recommended minimal 6PE router configuration:

```
6PE2#show running config
ip cef
!
ipv6 cef
!
ipv6 unicast-routing
mpls ipv6 source-interface Loopback0
!
router bgp 100
 no synchronization
 no bgp default ipv4-unicast
 neighbor 192.168.0.1 remote-as 100
 neighbor 192.168.0.1 update-source Loopback0
 !
 address-family ipv6
 neighbor 192.168.0.1 activate
 neighbor 192.168.0.1 send-label
 exit-address-family
```

TTL Processing

At imposition time, the 6PE supports the same TTL processing options for IPv6 as currently supported for IPv4:

Propagate: IP TTL used to set initial MPLS TTL

Do not propagate: MPLS TTL set to 255

Note that it is the same single configuration knob (i.e. existing CLI propagate/do not propagate command) that controls the option for both IPv4 imposition and 6PE IPv6 imposition.

By default the IPv6 TTL value is copied to MPLS TTL and you can see how many hops, there are in the network.

```
router(config)# mpls ip propagate-ttl
```

In most cases, disabling the TTL propagation makes the IPv4 core network completely transparent. At the LSP tail-end, the minimum of the MPLS TTL and IP TTL is copied into the IP TTL.

```
router(config)#no mpls ip propagate-ttl
```

Monitoring Cisco 6PE Configuration and Traffic

There are no specific Cisco 6PE "show commands", but the output of the following components has been enhanced to display additional outputs when 6PE is enabled. Monitor the following component when running Cisco 6PE:

- MP-BGP
- MPLS
- CEFv6
- IPv6 routing table

The output of the following commands is relative to the following network architecture:

```
2001:33::/64
```

Figure 8. Sample 6PE network

6PE-1   P   P   6PE-2
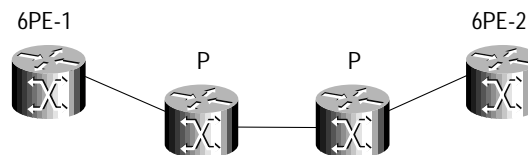
show bgp ipv6 <ipv6-prefix>

This command displays the details of Cisco 6PE BGP routes, which the ingress 6PE (6PE-1) learns from the egress 6PE (6PE-2):

```
6PE1#show bgp ipv6 2001:33::/64
BGP routing table entry for 2001:33::/64, version 3
Paths: (1 available, best #1, table Global-IPv6-Table)
 Not advertised to any peer
 Local
 ::FFFF:192.168.0.2 (metric 30) from 192.168.0.2 (192.168.0.2)
 Origin IGP, localpref 100, valid, internal, best
```

show bgp ipv6 neighbor

This command displays, among other information, the capabilities of the BGP peer. Since a new capability "IPv6+label" is added, it must appear in the list of capabilities returned by this command.

```
6PE2#sh bgp ipv6 neighbors 192.168.0.1
BGP neighbor is 192.168.0.1, remote AS 100, internal link
 BGP version 4, remote router ID 192.168.0.1
 BGP state = Established, up for 00:03:28
 Last read 00:00:28, hold time is 180, keepalive interval is 60 seconds
 Neighbor capabilities:
 Route refresh: advertised and received(old & new)
 Address family IPv6 Unicast: advertised and received
 ipv6 MPLS Label capability: advertised and received
 Received 8 messages, 0 notifications, 0 in queue
 Sent 7 messages, 0 notifications, 0 in queue
 Default minimum time between advertisement runs is 5 seconds
 For address family: IPv6 Unicast
 BGP table version 3, neighbor version 3
 Index 1, Offset 0, Mask 0x2
 Route refresh request: received 0, sent 0
 Sending Prefix & Label
 2 accepted prefixes consume 144 bytes
 Prefix advertised 0, suppressed 0, withdrawn 0
 Number of NLRIs in the update sent: max 0, min 0
```

show mpls forwarding-table

This command displays correspondence between labels and prefix. In the case of 6PE, labels are "aggregate label". This indicates that there are several prefixes for one local label. The command now shows "Aggregate", rather than target prefix. In our case 6PE2 is the egress 6PE generating the aggregate label for 2001:33/64 prefix. The MPLS forwarding table has a label to reach the BGP next-hop (192.168.0.3).

```
6PE2#show mpls forwarding-table
Local Outgoing     Prefix           Bytes tag   Outgoing        Next Hop
tag tag or VC      or Tunnel Id     switched    interface
17  16            192.168.0.1/32   0           Et1/2           172.16.0.49
19  Aggregate     IPv6             0
```

show bgp ipv6 tags

On the ingress 6PE, the top of the stack label is shown by the following command, which describes the label switching operations for prefix 2001:33::/64:

```
6PE1#show bgp ipv6 tags
 Network Next Hop   In tag/Out tag
 2001:33::/64 ::FFFF:192.168.0.2       notag/17
```

show ipv6 cef <prefix>

The CEF tables contain a two-label stack entry for this prefix:

```
6PE1#sh ipv6 cef 2001:33::/64
2001:33::/64
 nexthop :FFFF:192.168.0.2
 fast tag rewrite with Et1/2, 172.16.0.49, tags imposed: {16 17}
```

show ipv6 route <prefix>

The following command specifies, with the "IPv6-mpls" keyword, the Cisco 6PE origin of the route.

```
6PE1#show ipv6 route 2001:33::/64
IPv6 Routing Table - 6 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
 I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea
B 2001:33::/64 [200/0]
 via :FFFF:192.168.0.2, IPv6-mpls
```

Debug Commands

The following three commands have seen messages added for 6PE debugging control. All MPLS VPNv4 resolution messages are also applicable to 6PE.

debug bgp ipv6

The following debug messages have been added:

```
BGP: resolving bestpath with tag %d for network %P
BGP: reflected - forwarding received tag %d for prefix %P
BGP: allocate local label %d for network %P
BGP: sending label %d for network %P
BGP: receiving tag %d from peer for network %P
BGP: %s insert tag %d for network %P into path
BGP: deletes path with tag %d for network %P
BGP: next-hop %P is not a compatible v4
BGP: no parent fib for %P
BGP: no tag rewrite for parent %i
BGP: parent %i not resolved
BGP: no label binding for parent %i
BGP: parent %i has no adjacency
BGP: multipath for parent %i hash %d not resolved
BGP: multipath for parent %i hash %d rewrite %d has no label binding
```

debug ipv6 packet

The following debug messages have been added:

```
6PE: multicast packet dropped. dest: %P
6PE: null route or tag_rewrite. dest: %P
6PE: tagsw_ipv6_get_shared_label allocating label %d at egress
6PE: unknown route while installing RIB entry for prefix %P-%d
6PE: hwidb is a tunnel but not output interface was found
6PE: setting %s vif output interface (tunnel) to %s
```

debug mpls lfib struct

The following debug messages have been added:

```
6PE: tfib_pe6_install_label_egress cannot allocate tag_rewrite at egress
6PE: tfib_pe6_res_install: prefix %P --> %i
6PE: no parent fib entry - next hop %i
6PE: no tag rewrite for parent fib entry - next hop %i
6PE: no label binding for parent %i
6PE: multipath for parent %i hash %d not resolved
6PE: multipath for parent %i hash %d rewrite %d has no label binding
6PE: TFIB entry installed for prefix %P
6PE: TFIB entry uninstalled for prefix %P
```

Deployment Scenarios

IPv6 services that run through Cisco 6PE can be deployed on any MPLS topology as it only involves the addition or reconfiguration of some PE devices.

Nevertheless, as always, the size of the network and attached sites may impact topology. Topology examples follow in the next section.

Full Mesh Structure

This configuration is following the well-known BGP rule to have all the BGP speakers in an Autonomous System peering together. It is the vanilla Cisco 6PE configuration, which conforms to BGP and MP-BGP specifications. One of the new rules it must follow is that all the Cisco 6PE speakers must support the new "IPv6+label" capability. On the MPLS side, a LSP exists from each 6PE router to all the others 6PE routers.

Figure 9. A full mesh BGP topology for Cisco 6PE
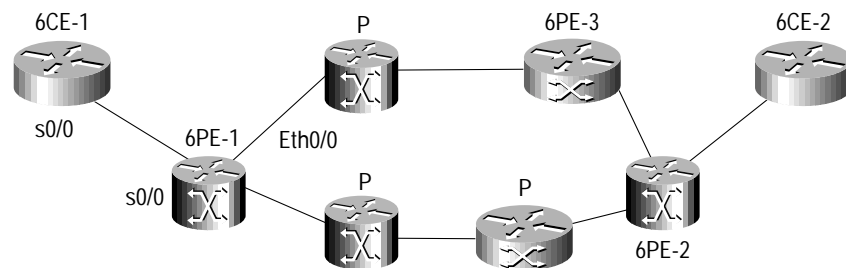
6PE-1, 6PE-2, 6PE-3: router with Cisco 6PE capability.

6CE1, 6CE-2: simple CE router supporting IPv6.

P: provider router running MPLS and IPv4.

Configurations

6PE-1 router configuration

In this example, the IPv4 IGP is ISIS. P and 6PE router are participating in a level-1/level-2 area. LDP is the Label Distribution Protocol. A BGP peering is created for each remote 6PEs (6PE-2, 6PE-3).

```
ip cef
!
ipv6 cef
!
ipv6 unicast-routing
!
mpls ipv6 source-interface Loopback0
tag-switching tdp router-id Loopback0
!
interface Loopback0
 ip address 172.16.0.1 255.255.255.255
 ipv6 address 2001:1::1/64
!
```

```
interface Serial0/0
 description to_CE_router
 no ip address
 ipv6 address 2001:2::2/64
!
interface Ethernet0/0
 description to_MPLS_backbone
 ip address 172.16.0.121 255.255.255.252
 ip router isis
 mpls label protocol ldp
 tag-switching ip
!
router isis
 passive-interface Loopback0
 net 49.0001.0000.0000.0006.00
!
router bgp 100
 no synchronization
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor 192.168.0.2 remote-as 100
 neighbor 192.168.0.2 description to_6PE2
 neighbor 192.168.0.2 update-source Loopback0
 neighbor 192.168.0.3 remote-as 100
 neighbor 192.168.0.3 description to_6PE3
 neighbor 192.168.0.3 update-source Loopback0
 !
 address-family ipv6
 neighbor 192.168.0.2 activate
 neighbor 192.168.0.2 send-label
 neighbor 192.168.0.3 activate
 neighbor 192.168.0.3 send-label
 network 2001:21::/64
 exit-address-family
!
ipv6 route 2001:21::/64 Serial0/0
```

P router configuration

The P router configuration is given here as a reference. It is a simple MPLS switching configuration. Nothing is specific to Cisco 6PE. The Label Distribution Protocol is common to P and 6PE routers.

```
ip cef
!
tag-switching tdp router-id Loopback0
!
interface Loopback0
 ip address 192.168.0.4 255.255.255.255
!
interface POS1/0
description to_MPLS backbone
 ip address 172.16.0.246 255.255.255.252
 ip router isis
 mpls label protocol ldp
 tag-switching ip
!
```

```
router isis
 passive-interface Loopback0
 net 49.0001.0000.0000.0008.00
CE-PE configuration with static routing
In case of static routing between CE and 6PE, use the one of the following configurations:
```

6CE-1 router configuration

6CE-1 presents a default route to the 6PE-1 router.

```
ip cef
!
ipv6 unicast-routing
!
interface Serial0/0
 description to_6PE1_router
 no ip address
 ipv6 address 2001:2::1/64
!
interface FastEthernet0/0
 no ip address
 ipv6 address 2001:21::1/64
!
ipv6 route ::/0 Serial0/0
```

6PE-1 router configuration

See the above 6PE-1 reference configuration. 6PE-1 presents a route to the CE-1 router Ethernet interface prefix "2001:21::1/64".

CE-PE configuration with eBGP

6CE-1 router configuration

An eMP-BGP peering is established towards the 6PE-1 router, and local prefixes are announced through "network" statements. A default route is pointing to the 6PE-1 router.

```
router bgp 1722
 no synchronization
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor 2001:2::2 remote-as 100
 !
 address-family ipv6
 neighbor 2001:2::2 activate
 network 2001:21::/64
 exit-address-family
!
ipv6 route ::/0 Serial0/0
```

6PE-1 router configuration

An eMP-BGP peering is established toward the CE-1 router.

```
router bgp 100
 no synchronization
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor 2001:2::2 remote-as 1722
 !
 address-family ipv6
 neighbor 2001:2::2 activate
exit-address-family
```

Several options are available to establish an ISISv6 adjacency between a 6CE router and a 6PE router. The following configurations represent one of the possible solutions. A default route is pointing to the 6PE-1 router.

6CE-1 router configuration
```
ipv6 unicast-routing
!
interface Serial0/0
 description to_6PE1_router
 no ip address
 ipv6 address 2001:2::1/64
 ipv6 router isis
!
interface FastEthernet0/0
 no ip address
 ipv6 address 2001:21::1/64
!
ipv6 route ::/0 Serial0/0
!
router isis
 !
 address-family ipv6
 redistribute connected
 exit-address-family
 net 49.0001.0000.0000.0005.00
!
ipv6 route ::/0 Serial0/0
```
6PE-1 router configuration
```
interface Serial0/0
 description to_CE_router
 no ip address
 ipv6 address 2001:2::2/64
 ipv6 router isis
!
router isis
 net 49.0001.0000.0000.0006.00
!
ipv6 route ::/0 Serial0/0
```
CE-PE configuration with RIPng

6PE-1 router configuration

A RIPng process is created on the router, and enabled on the interface that connects the 6PE-1 and 6CE-1 routers.
```
interface Serial0/O
ipv6 address 2001:2::2/64
ipv6 rip ce1 enable
!ipv6 router rip ce1
```
6CE-1 router configuration

A RIPng process is created on the router, and enabled on the interface that connects the 6PE-1 and 6CE-1 routers. A "redistribute connected" statement contributes to announce the locally reachable prefixes. A default route points to the 6PE-1 router.interface Serial0/0

```
description to_6PE1_router
ipv6 address 2001:2::1/64
 ipv6 rip ce1 enable
!
interface FastEthernet0/0
 no ip address
 ipv6 address 2001:21::1/64
!
ipv6 route ::/0 Serial0/0
!
ipv6 router rip ce1
 redistribute connected!
ipv6 route ::/0 Serial0/0
```

Route Reflector

Cisco 6PE uses MP-BGP to exchange IPv6 prefixes reachability, so it is subject to BGP characteristics and features. One of these characteristics is a mandatory full meshing. Fortunately, the Route Reflector concept alleviates this constraint and is also applicable to Cisco 6PE. It is possible to have groups of 6PEs peering with a small number of router instead of establishing a full mesh of MP-BGP sessions. This Router Reflector 6PE (6PE-RR) must support the "IPv6+label" address family. In practice, it is a router supporting the Cisco 6PE feature.

Route Reflector clients PEs send updates containing MP_REACH_NLRI attributes. These updates are reflected, unchanged, to the others Route Reflector client PEs. However the 6PE-RR performs a best path calculation on routes presented to it by the Route Reflector client PEs. Depending on the result of this best path calculation, the MP-REACH_NLRI attribute is reflected to other Route Reflector clients or not.

Figure 10. A Route Reflector topology for Cisco 6PE
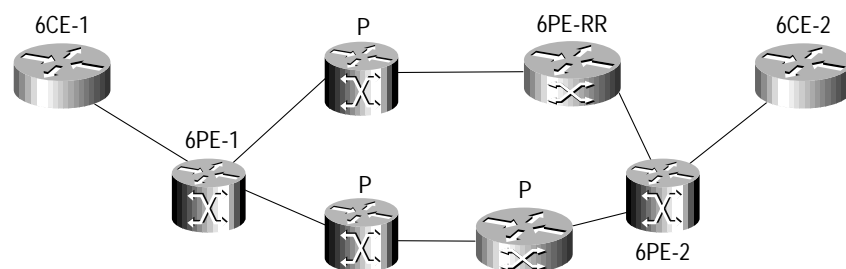
6PE-1, 6PE-2: Route Reflector client router with Cisco 6PE capability.

6PE-RR: Route Reflector server router with Cisco 6PE capability.

Configuration and Monitoring

Route Reflector server 6PE-RR router configuration

An MP-BGP session with Route Reflector capacity is maintained to each Route Reflector clients (6PE-1, 6PE-2).

```
router bgp 100
 no synchronization
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor 192.168.0.1 remote-as 100
 neighbor 192.168.0.1 update-source Loopback0
 neighbor 192.168.0.1 description to_6PE1
 neighbor 192.168.0.2 remote-as 100
 neighbor 192.168.0.2 update-source Loopback0
 neighbor 192.168.0.2 description to_6PE2
 !
 address-family ipv6
 neighbor 192.168.0.1 activate
 neighbor 192.168.0.1 route-reflector-client
 neighbor 192.168.0.1 send-label
 neighbor 192.168.0.2 activate
 neighbor 192.168.0.2 route-reflector-client
 neighbor 192.168.0.2 send-label
 exit-address-family
```

Prefix 2001:39::/64 is learnt from 6PE-2 with "::FFFF:192.168.0.2" as a next-hop and "19" as the aggregate label.

```
6PE-RR# show bgp ipv6 ipv6 2001:39::/64
sh bgp ipv6 2001:39::/64
BGP routing table entry for 2001:39::/64, version 8
Paths: (1 available, best #1, table Global-IPv6-Table)
 Advertised to non peer-group peers:
 192.168.0.6
 Local, (Received from a RR-client)
 ::FFFF:192.168.0.2 (metric 10) from 192.168.0.2 (192.168.0.2)
 Origin IGP, localpref 100, valid, internal, best
```

As the Route Reflector 6PE is in the data path the MPLS label stack is two label deep: "19" the Aggregate label and "23" a local forwarding label.

```
6PE-RR#show ipv6 cef 2001:39::/64
2001:39::/64
 nexthop ::FFFF:192.168.0.2
 fast tag rewrite with Et1/3, 172.16.0.89, tags imposed: {23 19}
```

Route Reflector client 6PE-1 router configuration

A single MP-BGP session to the Route Reflector server is maintained as in a classical BGP configuration.

```
router bgp 100
 no synchronization
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor 192.168.0.6 remote-as 100
 neighbor 192.168.0.6 description to_6PE_RR
 neighbor 192.168.0.6 update-source Loopback0
 !
 address-family ipv6
 neighbor 192.168.0.6 activate
 neighbor 192.168.0.6 send-label
 exit-address-family
```

On 6PE-1, label "19" is kept as the aggregate label and next-hop "::FFFF:192.168.0.2" for prefix 2001:39::/64. The Route Reflector role is correctly fulfilled by the Route Reflector server router.

```
6PE1#show ipv6 cef 2001:39::/64
2001:39::/64
 nexthop ::FFFF:192.168.0.2
 fast tag rewrite with Et1/0, 172.16.0.82, tags imposed: {27 19}
6PE-1# show bgp ipv6 2001:39::/64
BGP routing table entry for 2001:39::/64, version 6
Paths: (1 available, best #1, table Global-IPv6-Table)
 Not advertised to any peer
 Local
 ::FFFF:192.168.0.2 (metric 20) from 192.168.0.6 (192.168.0.6)
 Origin IGP, localpref 100, valid, internal, best
 Originator: 192.168.0.2, Cluster list: 192.168.0.6
```

Confederations

Another way to avoid the full meshing of all 6PE routers in an Autonomous System is to take advantage of the confederation scheme that applies to 6PE as well. All routers included in a peering with a 6PE router must support the Cisco 6PE feature in order to participate to the "IPv6+label" route exchange.

Figure 11. A BGP Confederation topology for Cisco 6PE
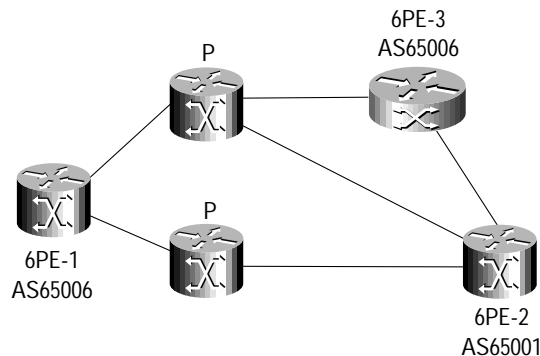
6PE-1, 6PE-2, 6PE-3: 6PR router with a BGP confederation configuration.

Configuration

6PE3 router configuration

6PE3 router establishes 2 MP-BGP sessions with 2 different sub-AS in AS100.

```
router bgp 65006
 no synchronization
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 bgp confederation identifier 100
 bgp confederation peers 65001
 neighbor 192.168.0.2 remote-as 65001
 neighbor 192.168.0.2 ebgp-multihop 10
 neighbor 192.168.0.2 update-source Loopback0
 neighbor 192.168.0.2 description to_6PE2
 neighbor 192.168.0.1 remote-as 65006
 neighbor 192.168.0.1 update-source Loopback0
 neighbor 192.168.0.1 description to_6PE1
 !
 address-family ipv6
 neighbor 192.168.0.1 activate
 neighbor 192.168.0.1 send-label
 neighbor 192.168.0.2 activate
 neighbor 192.168.0.2 send-label
 exit-address-family
!
```

Prefix 2001:39::/64 is learnt from 6PE-2 which is in a different sub-AS. Consequently the AS path to reach it is: (65001). The next-hop "::FFFF:192.168.0.2" stays unchanged following the confederation scheme.

```
6PE3#show bgp ipv6
*> 2001:39::/64 ::FFFF:192.168.0.2
 100 0 (65001) i
6PE1#show bgp ipv6 2001:39::/64
BGP routing table entry for 2001:39::/64, version 3
Paths: (1 available, best #1, table Global-IPv6-Table)
 Not advertised to any peer
 (65001)
 ::FFFF:192.168.0.2 (metric 30) from 192.168.0.2 (192.168.0.2)
 Origin IGP, localpref 100, valid, confed-internal, best
```

Traffic Engineering

The current MPLS backbones offer a Traffic Engineering (TE) feature to provide more control on the path taken by the traffic crossing the backbone. This feature is also available for the Cisco 6PE traffic. Instead of distributing a top-of-the-stack label with TDP or LDP, the forwarding label can be distributed by RSVP-TE, which offering a TE function for the IPv6 traffic. In the following diagram, the IPv6 traffic carried by MPLS between 6PE1 and 6PE-3 is following an MPLS-TE tunnel (6PE-1, P2, P1, 6PE2, 6PE-3). The IPv4 traffic is following the shortest path calculated by the IGP (6PE-1, P3, 6PE-3).

TE can be also deployed from P router to P router; however, IPv4 and IPv6 traffic is following the same path in this case.

Figure 12. A full-mesh BGP topology with traffic engineering tunnels for Cisco 6PE
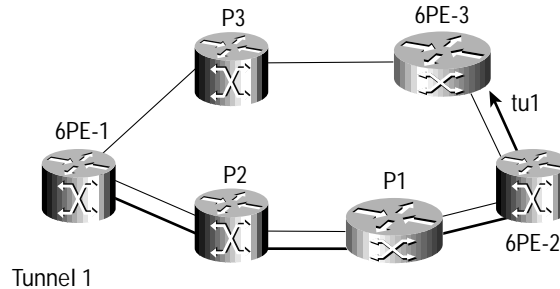
6PE-1, 6PE-3: tunnel end-points.

Configuration and Monitoring

6PE-1 router configuration

MPLS TE is enabled globally. RSVP-TE is running on the backbone interfaces. An MPLS-TE tunnel is created and an explicit path is associated to this tunnel. IGP traffic engineering extensions are enabled. The Cisco 6PE and IPv4 BGP sessions are peering to different IPv4 addresses of the same remote router (6PE-3). As the traffic to the 6PE BGP remote peer address "192.168.0.3" is traffic-engineered, the path to the IPv4 BGP remote peer address "192.168.0.103" is let to the ISIS decision.

```
ip cef
!
ipv6 unicast-routing
ipv6 cef
mpls traffic-eng tunnels
tag-switching tdp router-id Loopback0
!
interface Loopback0
 ip address 192.168.0.1 255.255.255.255
 ipv6 address 2001:55::1/64
!
interface Loopback1
 ip address 192.168.0.101 255.255.255.255
!
interface Tunnel1
 ip unnumbered Loopback0
 tag-switching ip
 tunnel destination 192.168.0.3
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng priority 7 7
 tunnel mpls traffic-eng bandwidth 128
 tunnel mpls traffic-eng path-option 1 explicit identifier 33
!
interface FastEthernet0/0
 description to_MPLS_backbone
 ip address 172.16.0.81 255.255.255.252
 ip router isis
 mpls traffic-eng tunnels
 tag-switching ip
 ip rsvp bandwidth 256 256
!
```

```
router isis
 passive-interface Loopback0
 passive-interface Loopback1
 net 49.0001.0000.0000.0006.00
 metric-style wide
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng level-1
!
ip explicit-path identifier 33 enable
 next-address 172.16.0.82
 next-address 172.16.0.86
 next-address 172.16.0.90
!
router bgp 100
 no synchronization
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor 192.168.0.2 remote-as 100
 neighbor 192.168.0.2 description to_6PE2
 neighbor 192.168.0.2 update-source Loopback0
 neighbor 192.168.0.3 remote-as 100
 neighbor 192.168.0.3 description to_6PE3
 neighbor 192.168.0.3 update-source Loopback0
 neighbor 192.168.0.103 remote-as 100
 neighbor 192.168.0.103 update-source Loopback1
 neighbor 192.168.0.103 activate
 !
 address-family ipv6
 neighbor 192.168.0.2 activate
 neighbor 192.168.0.2 send-label
 neighbor 192.168.0.3 activate
 neighbor 192.168.0.3 send-label
 network 2001:21::/64
 network 2001:55::1/64
 exit-address-family
!
ip route 192.168.0.3 255.255.255.255 tu1
!
```

P router configuration

MPLS TE is enabled globally on the P routers. RSVP is running on the interfaces. IGP traffic engineering extensions are enabled.

```
mpls traffic-eng tunnels
!
interface FastEthernet0/0
 description to_2652
 ip address 172.16.0.82 255.255.255.252
 ip router isis
 duplex auto
 speed auto
 mpls traffic-eng tunnels
 tag-switching ip
 ip rsvp bandwidth 256 256
!
```

```
router isis
 passive-interface Loopback0
 net 49.0001.0000.0000.0003.00
 metric-style wide
 mpls traffic-eng router-id Loopback0
 mpls traffic-eng level-1
```

IGP Load-Balancing in a 6PE Router

A load-balancing scenario happens in a 6PE router when there are multiple IGP paths from a 6PE to reach another 6PE. There are several equal cost paths to the MP-BGP next-hop (a remote 6PE) in the IGP routing table. In the following example, there are two equal cost paths to reach 6PE-2 from 6PE-1.

Figure 13. IGP Load-sharing in the ingress 6PE
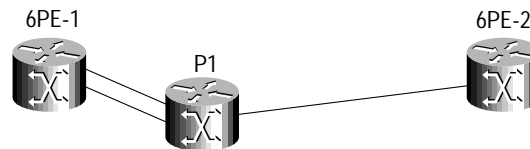
Monitoring

A single next-hop exists in the MP-BGP table "::FFFF:192.168.0.2" to "2001:33::/64".

```
6PE-1#sh bgp ipv6 2001:33::/64
BGP routing table entry for 2001:33::/64, version 3
Paths: (1 available, best #1, table Global-IPv6-Table)
 Not advertised to any peer
 Local
 ::FFFF:192.168.0.2 (metric 30) from 192.168.0.2 (192.168.0.2)
 Origin IGP, localpref 100, valid, internal, best
```

The CEF table shows a load-sharing structure attached to "192.168.0.2" because 2 next-hops exist to reach the IPv4 next-hop "192.168.0.2". The traffic split is based on the IPv6 source/destination couple.

```
6PE-1#show ip cef 192.168.0.2
192.168.0.7/32, version 69, per-destination sharing
0 packets, 0 bytes
 tag information set
 local tag: 25
 via 172.16.0.53, Ethernet1/3, 0 dependencies
 traffic share 1
 next hop 172.16.0.53, Ethernet1/3
 valid adjacency
 tag rewrite with Et1/3, 172.16.0.53, tags imposed: {22}
 via 172.16.0.122, Ethernet1/1, 0 dependencies
 traffic share 1
 next hop 172.16.0.122, Ethernet1/1
 valid adjacency
 tag rewrite with Et1/1, 172.16.0.122, tags imposed: {22}
 0 packets, 0 bytes switched through the prefix
 tmstats: external 0 packets, 0 bytes
 internal 0 packets, 0 bytes
```

```
6PE1#show ipv6 cef 2001:33::/64
2001:34::/64
 nexthop ::FFFF: 192.168.0.2
 fast tag rewrite with
 Recursive rewrite via 192.168.0.2/32, tags imposed {22 23}
```

IGP Load-Balancing in a P Router

Load-sharing on a P router occurs independently of 6PE. As the MPLS-encapsulated packet is not an IPv4 packet, the traffic split is based on the bottom of the stack label. The pool of sixteen aggregate labels allows a certain level of load sharing to occur in the P router. The prefixes originated by the same egress 6PE carry an aggregate label value chosen among 16 possible values. Thus, if the number of prefixes is big enough, a correct load balancing can happen on P routers.

Figure 14. MPLS Load-sharing happening in a P router
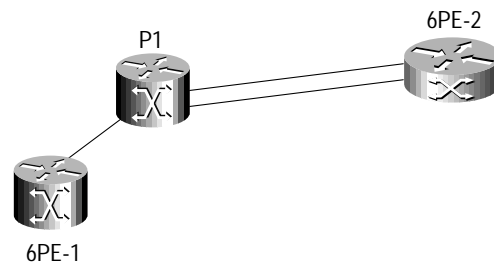
Monitoring

On P1 the IGP is reporting 2 equal cost paths to reach the IPv4 BGP next-hop (192.168.0.7) for the prefixes advertised by 6PE-2.

```
P1#show ip cef 192.168.0.2 255.255.255.255
192.168.0.7/32, version 710, per-destination sharing
0 packets, 0 bytes
 tag information set
 local tag: 22
 via 172.16.0.242, POS0/0, 0 dependencies
 traffic share 1
 next hop 172.16.0.242, POS0/0
 valid adjacency
 tag rewrite with POS0/0, point2point, tags imposed {21}
 via 172.16.0.245, POS1/0, 0 dependencies
 traffic share 1
 next hop 172.16.0.245, POS1/0
 valid adjacency
 tag rewrite with PO1/0, point2point, tags imposed {22}
 0 packets, 0 bytes switched through the prefix
 tmstats: external 0 packets, 0 byte
 internal 0 packets, 0 bytes
P1#show mpls forwarding-table
Local      Outgoing Prefix Bytes      Outgoing            Next Hop
tag        tag or VC or Tunnel Id switched                interface
21         21        192.168.0.7/32 234 POS0/0
     point2point
           22        192.168.0.7/32 485                   POS1/0
               point2point
```

Conclusion

IPv6 services can be integrated without MPLS dependency. Cisco 6PE leverages MPLS and MP-BGP, if both have been deployed in the network infrastructure. MPLS enables an attractive approach for IPv6 migration when it is deployed to offer services such as Layer 2 and Layer 3 VPN, TE, DS-TE, Fast Reroute.

Amongst the various approaches to deploy IPv6 traffic over an MPLS infrastructure, Cisco 6PE offers IPv6 deployment at marginal cost and risk. Additionally, it offers traditional Layer 3 network scalability benefits, and does not impact the IPv4/MPLS core. IPv6 is activated on the edge, on top of the existing MPLS infrastructure, which is not impacted, for support of existing revenue-generating services. IPv6-enabled devices are added gradually where required.

Furthermore, training and operational costs are minimized because only a small part of the network is upgraded, and Cisco 6PE is based on well-known MPLS-VPN technology.

References

[IPv6_Deploy] Cisco IOS IPv6 Deployment scenario guide

http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/ipv6_sol/ipv6dswp.pdf

[IPv6_BGP] "Connecting IPv6 Domains across IPv4 Clouds with BGP", De Clerq et al, draft-ietf-ngtrans-bgp-tunnel-04, January 2002, Work in progress.

[MPLS_BGP] Carrying Label Information in BGP-4, RFC3107, May 2001. Y. Rekhter and E. Rosen.

[BGP_CAP] Capabilities Advertisement with BGP-4, RFC-2842, May 2000.

[v6_ADDRESS] "IPv6 Addressing Architecture", R. Hinden and S. Deering, draft-ietf-ipngwg-addr-arch-v3-10, September 2002, Work in Progress.

CISCO SYSTEMS