

# IPv6, Next Generation Internet Protocol



**Khalid Raza Distinguished Engineer** 

## **Agenda**

- Why IPv6
- Refresher

Header

Types of address

Extension

ISP Deployment

MPLS network

**Dual Stack** 

# Why IPv6

## **Basic goals of internet**

- Universal connectivity anyone can send packets to anyone.
- Applications run at the edge so anyone can install and offer services.
- Cheap and cheerful" core technology so transmission is cheap.
- Natural selection no grand plan, but good technology survives and the rest dies

**Brian Carpenter (IBM)** 

# Why IPv6

#### Evolution of network

- Production: During 80's to early 90's it was all about office automation, manufacturing, supply chain
- Transaction: From mid to late 90's it was about the web, Google search, on line orders, check status of an order
- 3. Interaction/Experience: Rich multi-media application, voice, file sharing, youtube etc

## Why IPv6

- Rapid adoption of Broadband changing traffic patterns
- Digitization of traffic types driving convergence of diverse networks
- Mobile Technology dramatically changing the network access model



# IPv6 - Key drivers for Next Generation Ubiquitous Networking



IP Mobility



Devices, Mobile Networks, mobile wireless

The Ubiquitous Internet



Edge's appliances & services



Higher Ed./Research

**Government Public Sector** 







### Refresher



# **IPv6 Addressing**

#### Representation

- 16-bit hexadecimal numbers
- Numbers are separated by (:)
- Hex numbers are not case-sensitive
- Example:

2003:0000:130F:0000:0000:087C:876B:140B

## **IPv6 Address Representation**

 16-bit fields in case-insensitive colon hexadecimal representation

2031:0000:130F:0000:0000:09C0:876A:130B

Leading zeros in a field are optional

2031:0:130F:0:0:9C0:876A:130B

 Successive fields of 0 represented as (::), but only once in an address

2031:0:130F::9C0:876A:130B

2031::130F::9C0:876A:130B not valid!

## **IPv6 Addressing**

### **Prefix Representation**

- Representation of prefix is just like CIDR
- In this representation you attach the prefix length
- IPv4 address: 198.10.0.0/16
- IPv6 address: 3ef8:ca62:12FE::/48

# IPv6 Address Range Reserved or Assigned

#### Of the Full Address Space

- 2000::/3 (001) is for aggregatable global unicast addresses
- FE80::/10 (1111 1110 10) for link-local
- FEC0::/10 (1111 1110 11) for site-local
- FC00::7 (1111 110x) for unique-local
- FF00::/8 (1111 1111) is for multicast

**Site-Local Address Deprecated** in RFC 3879

### **Unicast**

- Unicast addresses are used in a one-to-one context
- IPv6 unicast addresses are:
  - Unspecified, loopback, and IPv4 mapped
  - Link-local
  - Site-local (deprecated)
  - **Unique-local**
  - Aggregatable global unicast

## **IPv6 Address Representation**

IPv4 mapped

0:0:0:0:0::FFFF:IPv4 = ::FFFF:IPv4

0:0:0:0:0:FFFF:192.168.30.1 = ::FFFF:C0A8:1E01

13

# **IPv6 Addressing**

- IPv6 addressing rules are covered by multiple RFCs
- Address types are:

Unicast: one to one (global, link local, compatible)

Anycast: one to nearest (allocated from unicast)

Multicast: one to many

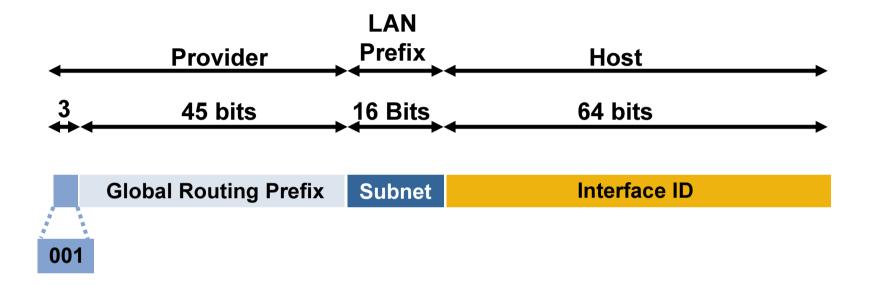
Reserved

 A single interface may be assigned multiple IPv6 addresses of any type (unicast, anycast, multicast)

No broadcast address use multicast



## **Aggregatable Global Unicast Addresses**

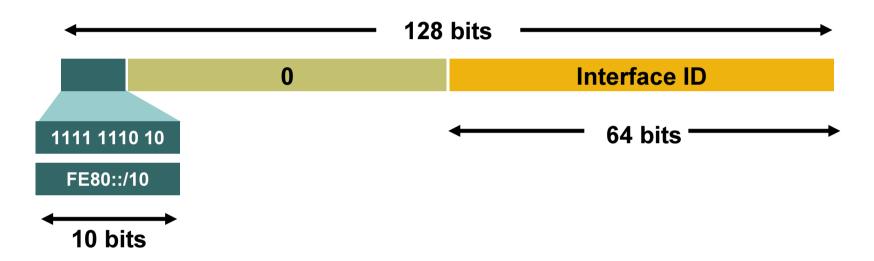


Aggregatable global unicast addresses are:

Addresses for generic use of IPv6
Structured as a hierarchy to keep the aggregation

See RFC 3513

### **Link-Local**



Link-local addresses:

Have a limited scope of the link

Are automatically configured with the interface ID

## **Link-Local**

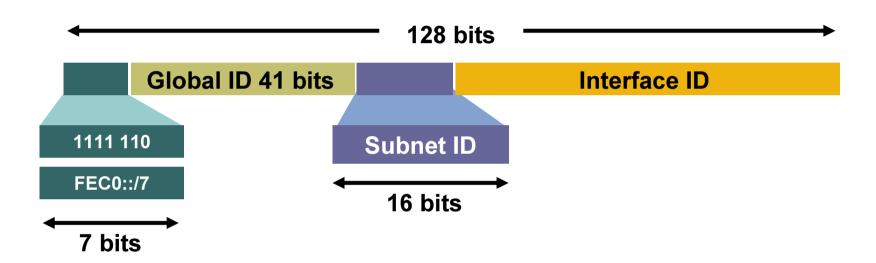
### **Aggregatable Address**

2001::4: 204:9AFF:FEAC:7D80

#### **Link-Local Address**

FE80:0:0:0 204:9AFF:FEAC:7D80

## **Unique-Local**



• Unique-local addresses used for:

Local communications

Inter-site VPNs

Not routable on the Internet

## Aggregatable Global Unicast Addresses

Lowest-order 64-bit field of unicast addresses may be assigned in several different ways:

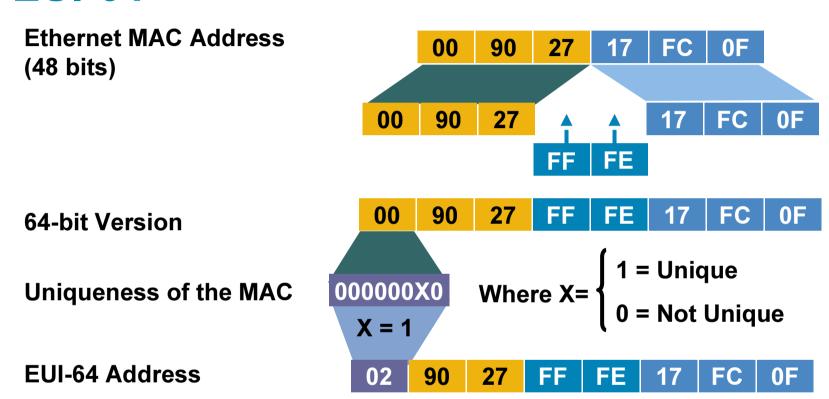
Auto-configured from a 64-bit EUI-64, or expanded from a 48-bit MAC address (e.g. Ethernet address)

Auto-generated pseudo-random number (to address privacy concerns)

Assigned via DHCP

Manually configured

### **EUI-64**

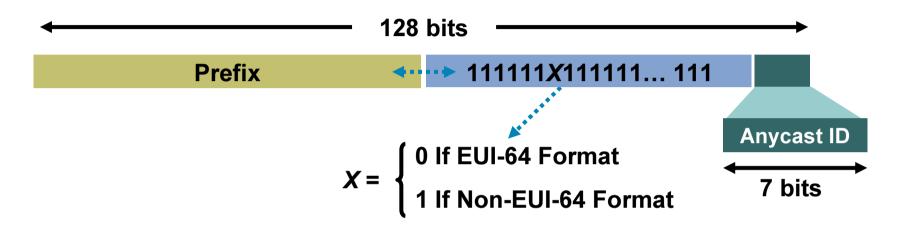


 EUI-64 address is formed by inserting "FFFE" and ORing a bit identifying the uniqueness of the MAC address

## **Anycast**

 Anycast allows a source node to transmit IP datagrams to a single destination node out of a group of destination nodes with same subnet ID based on the routing metrics

## **Anycast Address**

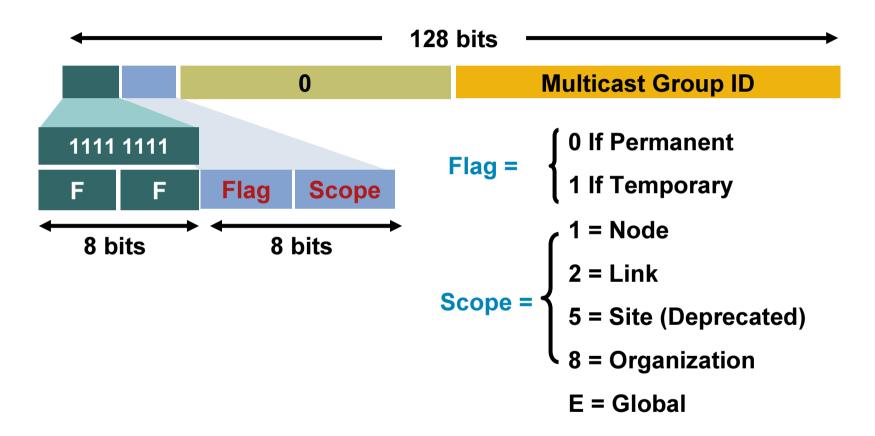


#### Anycast:

Is one-to-nearest type of address

Has a current limited use

#### **Multicast**



Multicast is used in the context of one-to-many;
 a multicast scope is new in IPv6

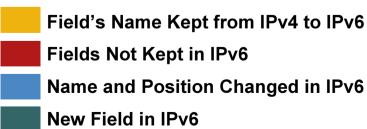
## **Expanded Address Space** Multicast Assigned Addresses (RFC 3306)

Address	Scope	Meaning
FF01::1	Node-Local	All Nodes
FF02::1	Link-Local	All Nodes
FF01::2	Node-Local	All Routers
FF02::2	Link-Local	All Routers
FF05::2	Site-Local (Deprecated)	All Routers
FF02::1:FFXX:XXXX	Link-Local	Solicited-Node

## **IPv4** and **IPv6** Header Comparison

#### **IPv4** Header

#### Type of **Total Length** Version Service Fragment Identification Flags Offset Time to Live Header Checksum Protocol Source Address **Destination Address Options** Padding Field's Name Kept from IPv4 to IPv6

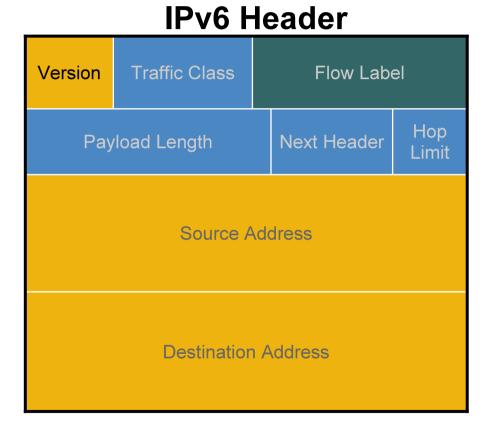


#### **IPv6** Header



# **IPv4** and **IPv6** Header Comparison

 Version: a 4-bit field that contains the number 6 instead of 4



## **IPv4** and **IPv6** Header Comparison **Fields Renamed**

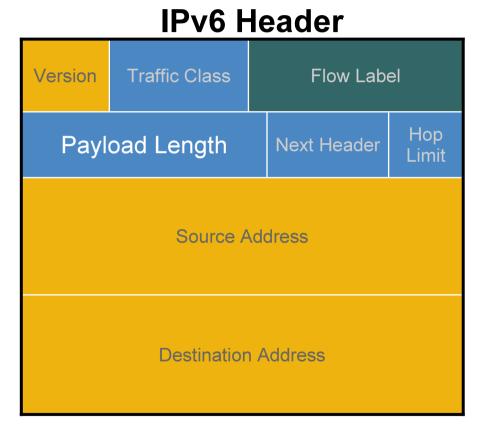
- Traffic class: an 8-bit field that is similar to the TOS field in IPv4
- It tags the packet with a traffic class that can be used in differentiated services
- These functionalities are the same as in IPv4

#### **IPv6** Header



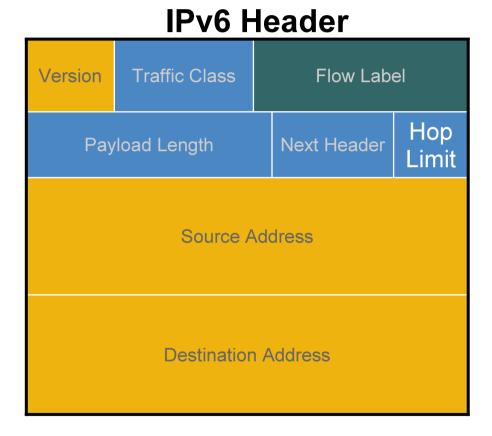
# IPv4 and IPv6 Header Comparison Fields Renamed

 Payload length: this is similar to the total length in IPv4, except it does not include the 40-byte header



# IPv4 and IPv6 Header Comparison Fields Renamed

 Hop limit: like TTL field, decrements by one for each router



## **IPv4** and **IPv6** Header Comparison **Fields Renamed**

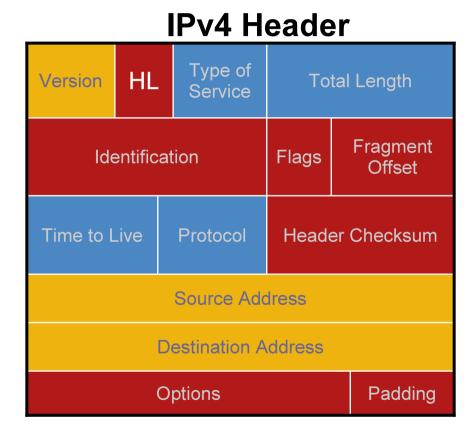
- Next header: similar to the protocol field in IPv4
- The value in this field tells you what type of information follows

E.g. TCP, UDP, extension header

### **IPv6** Header **Traffic Class** Version Flow Label Next qoH Payload Length Limit Header Source Address **Destination Address**

# IPv4 and IPv6 Header Comparison Fields Removed

 Header length: IPv6 has a fixed header length (40 bytes)



# IPv4 and IPv6 Header Comparison Fields Removed

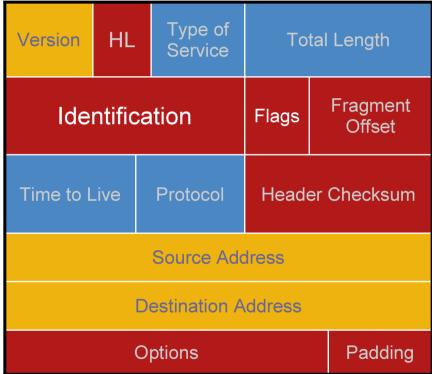
- Fragmentation: IPv6 does not do fragmentation
- If a sending host wants to do fragmentation, it will do it through extension headers

#### **IPv4** Header Type of HL Version **Total Length** Service Fragment Identification Flags Offset **Header Checksum** Time to Live Protocol Source Address **Destination Address Options Padding**

## **IPv4** and **IPv6** Header Comparison **Fields Removed**

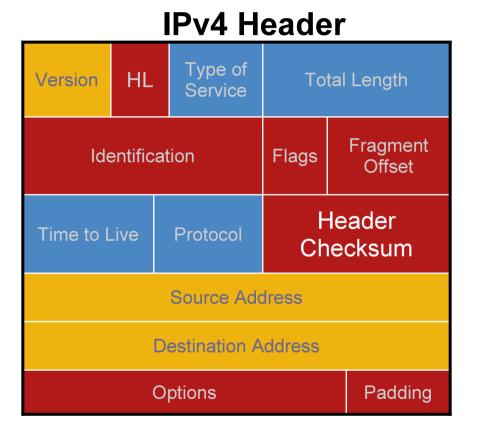
- Identification: used to identify the datagram from the source
- No fragmentation is done in IPv6 so no need for identification, also no need for flags

#### **IPv4** Header Type of HL



# IPv4 and IPv6 Header Comparison Fields Removed

 Checksum not needed because both media access and upper layer protocol (UDP and TCP) have the checksum; IP is best-effort, plus removing checksum helps expedite Packet processing



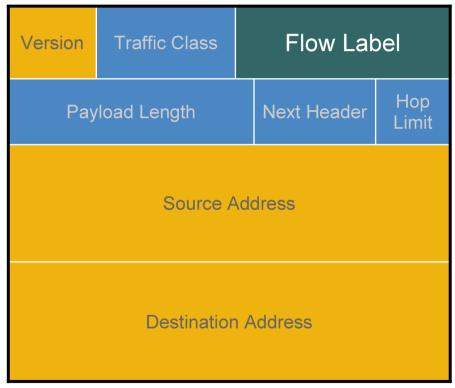
## **IPv4** and **IPv6** Header Comparison Fields Added

20-bit flow label field to identify specific flows needing special QoS

> Each source chooses its own flow label values; routers use source/dist addr + flow label to identify distinct flows

Flow label value of 0 used when no special QoS requested (the common case today)

#### **IPv6** Header



**RFC 3697** 

## **Extension Headers**

**IPv6 Header Next Header** = TCP

**TCP Header** + Data

**IPv6 Header Next Header** = Routing

**Routing Header Next Header = TCP**  **TCP Header** + Data

**IPv6 Header Next Header** = Routing

**Routing Header** Next Header = **Destination** 

**Destination Header Next Header = TCP** 

**Fragment of TCP Header** + Data

### **Extension Headers Are Daisy Chained**



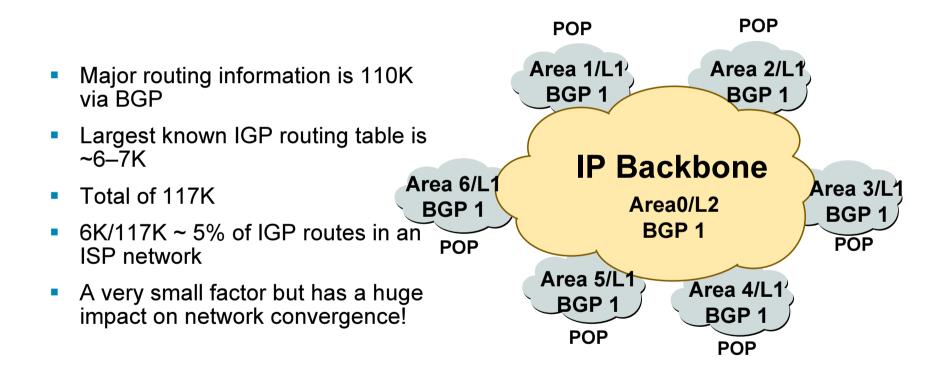
# **Dual Stack**



#### **Dual Stack IPv4-IPv6 Infrastructure**

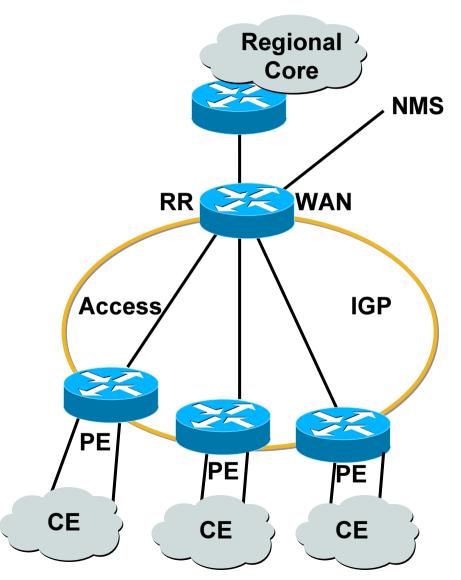
- It is generally the goal when IPv6 traffic and users will be rapidly increasing
- May not necessarily apply to the overall infrastructure. One may begin on network's portion such as Campus or Access or core networks
- Network design must be well planned Memory size to handle the growth for both IPv4 & IPv6 routing tables IGP options & its management: Integrated versus "Ships in the Night" Full network upgrade impact
- IPv4 and IPv6 Control & Data planes should not impact each other Feedback, requirements & experiments are welcome

#### **Service Provider networks**



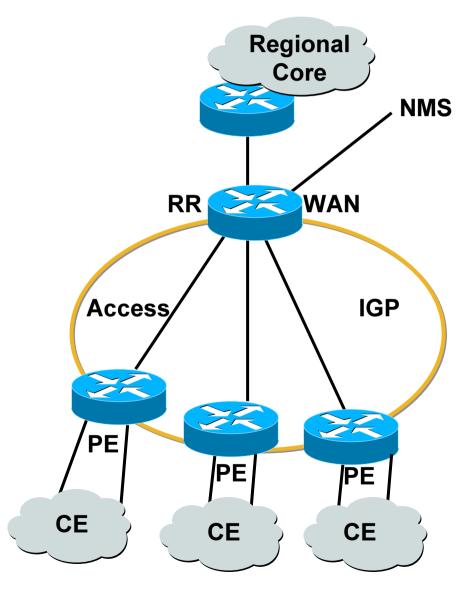
#### **Service Provider networks**

- You can reduce the IGP size from 6K to approx the number of routers in your network
- This will bring really fast convergence
- Optimized where you must and summarize where you can
- Stops unnecessary flapping



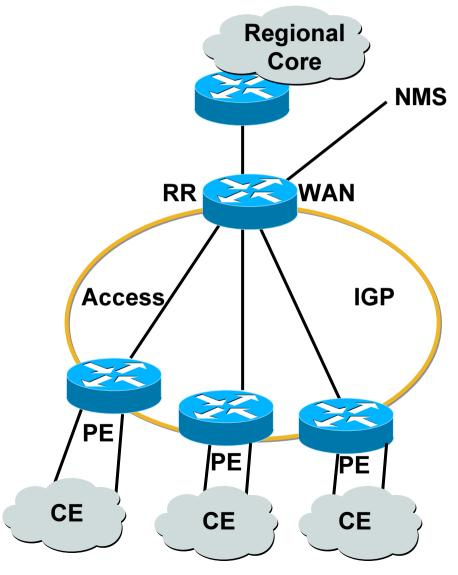
## Addressing

- The link between PE-CE needs to be known for management purpose
- BGP next-hop-self should be done on all access routers—unless PE-CE are on shared media (rare case)
- This will cut down the size of the IGP
- For PE-CE link do redistributed connected in BGP
- These connected subnets should ONLY be sent through RR to NMS for management purpose; this can be done through BGP communities



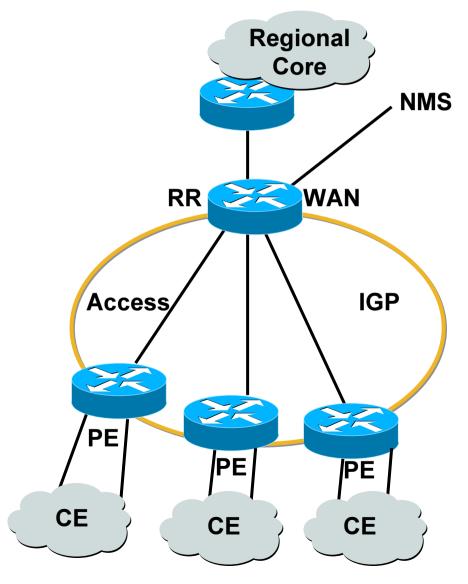
#### Addressing

- Divide the address into two parts
  - 1.Physical links
  - 2. Loopback interfaces
- Physical address should be a contagious
- Loopback should be from Aggregatable IPv6 Unicast address
- Optimal path to the next hop is necessary



## Addressing

- Assign ::/56 per pop for physical links
- Once out grow add another contiguous ::/56
- When assigning address to another POP keep few contiguous address open
- Summarize pop address at the **WAN** routers
- Leak loopback as specific



#### **Dual Stack IGP**

- IGP only carries next hop in SP network
- OSPF v3 introduces some great concepts
  - -Protocol processing per-link, not per-subnet
  - -Removal of addressing semantics
  - -Addition of Flooding scope
  - -Explicit support for multiple instances per link
  - -Use of link-local addresses
  - -Handling unknown LSA types

# **Dual Stack IGP (OSPF)**

- Protocol processing per-link, not per-subnet
  - -IPv6 uses the term "link" to indicate communication
  - -Interfaces connect to links
  - -Multiple IP subnets can be assigned to a single link, and two nodes can talk directly over a single link, even if they do not share a common IP subnet

#### **IPv6** multicast address

- The multicast address AllSPFRouters is FF02::5 note that 02 means that this is a permanent address and has link scope.
- The multcast address ALLDRouters is FF02::6

#### **OSPF**

#### Flooding

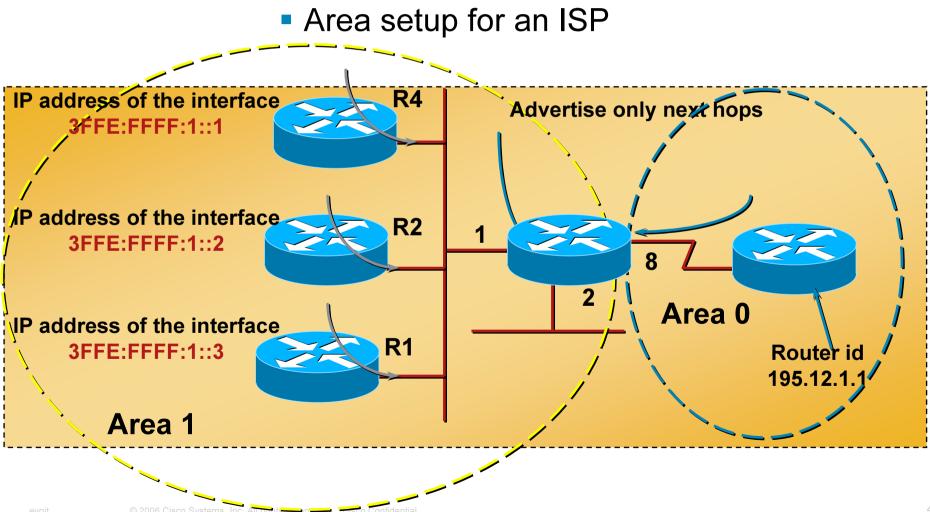
OSPFv2 had two flooding scope, AS wide and area wide. In OSPFv3 there are three flooding scope

AS scope, LSA is flooded through out the AS

Area scope, LSA is flooded only within an area

Link-local scope, LSA is flooded only on the local link.

# **Advertisement of Next hop**



# **Dual stack IGP (OSPF)**

- Much more scalable
- Infrastructure routes within an area not part of entire OSPF domain
- Provides much more scalability
- Unnecessary flooding/summaries eliminated

# **BGP Deployment**

- Two possible techniques
  - 1. Run BGP on the same hierarchy
  - 2. Run BGP with different hierarchy

If you are running VPN's then its better for scalability that you separate Route Reflection based on services.

#### Routing architecture -Theory

#### In Theory:

- The similarity between the IPv6 and IPv4 routing protocols leads to similar behaviour and expectations
- To select the IPv6 IGP, start by using the IPv4 IGP rules of thumb



**MPLS** 



#### **IPv6 over MPLS**

Many ways to deliver IPv6 services to end users

Most important is end-to-end IPv6 traffic forwarding

- Many service providers have already deployed MPLS in their IPv4 backbone for various reasons
- MPLS can be used to facilitate IPv6 integration
- Multiple approaches for IPv6 over MPLS:

IPv6 over L2TPv3

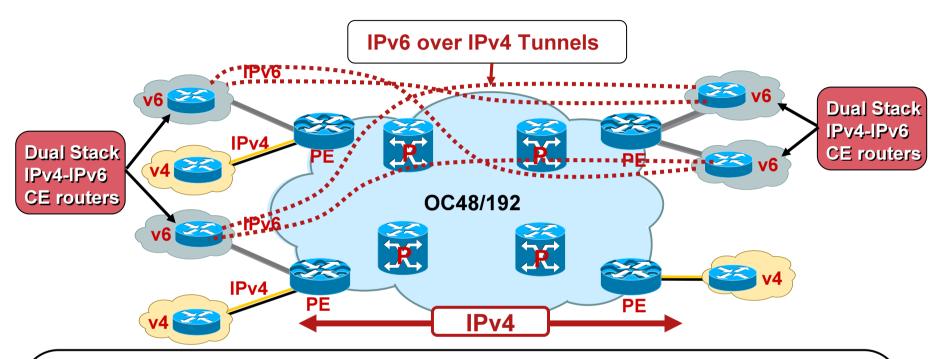
IPv6 over EoMPLS/AToM

IPv6 CE-to-CE IPv6 over IPv4 tunnels

Native IPv6 MPLS (Future)

IPv6 provider edge router (6PE) over MPLS

IPv6 VPN provider edge (6VPE) over MPLS



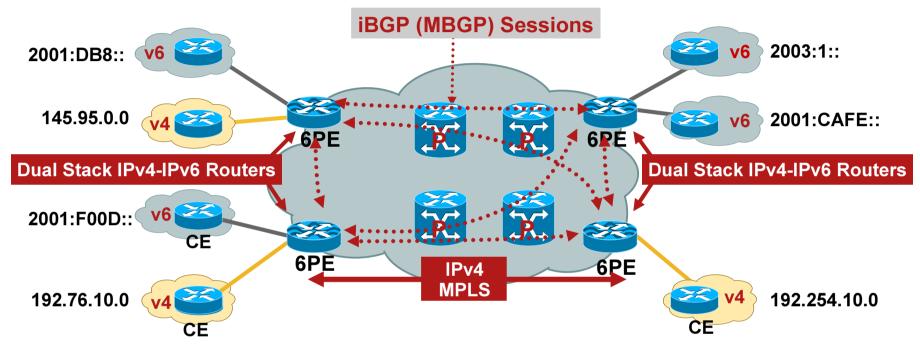
No impact on existing IPv4 or MPLS Core (IPv6 unaware)
Only CEs have to be IPv6-aware (Dual stack)
Mesh of IPv6 over IPv4 Tunnels CE-to-CE
Overhead: IPv4 header + MPLS header
MPLS/VPN support IPv4-native and IPv6 tunnels







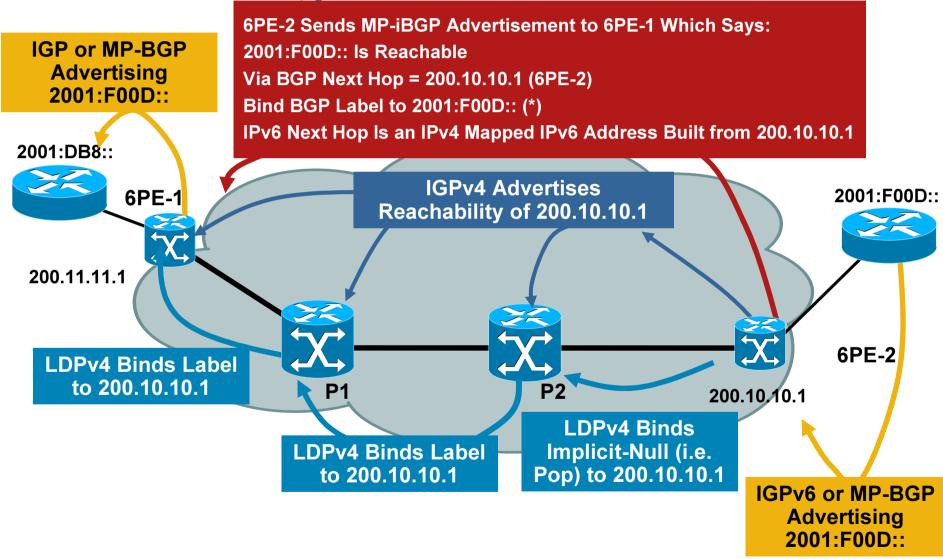
# IPv6 Provider Edge Router (6PE) over MPLS



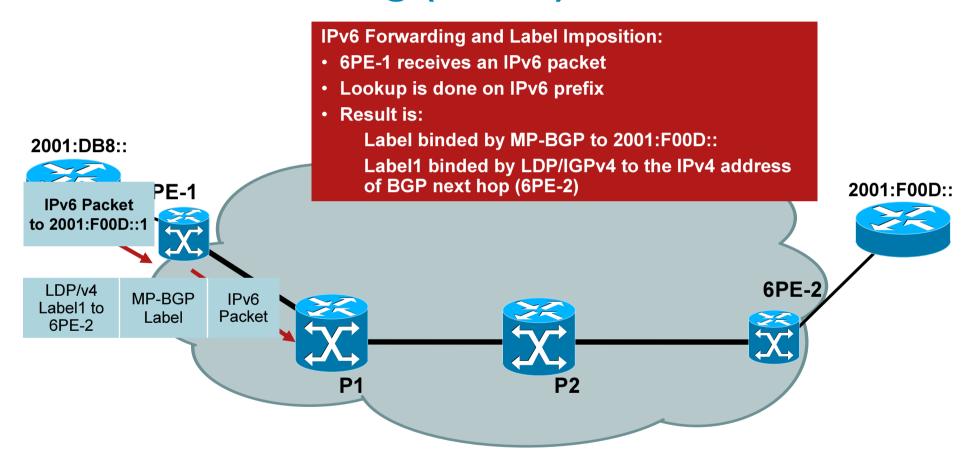
- IPv6 global connectivity over an IPv4-MPLS core
- Transitioning mechanism for providing unicast IP
- PEs are updated to support dual stack/6PE
- IPv6 reachability exchanged among 6PEs via iBGP (MBGP)
- IPv6 packets transported from 6PE to 6PE inside MPLS

http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/iosip\_an.htm

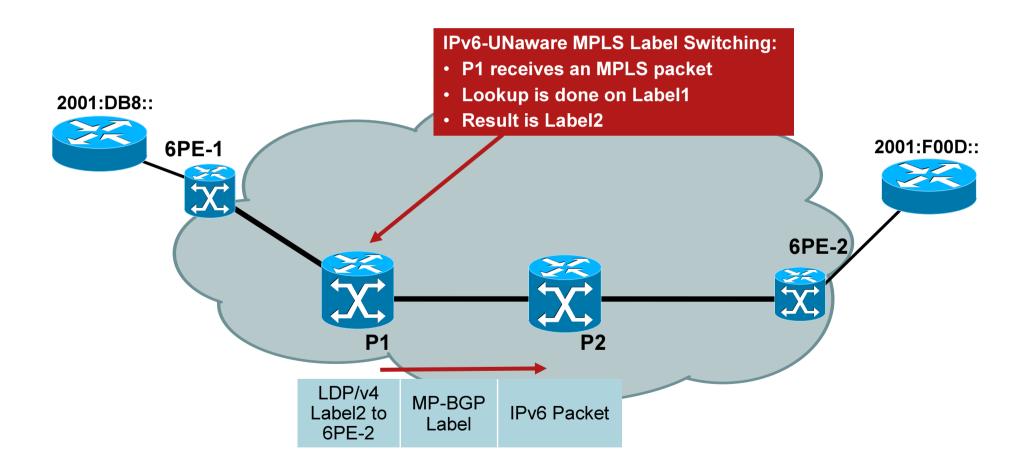
#### 6PE Routing/Label Distribution



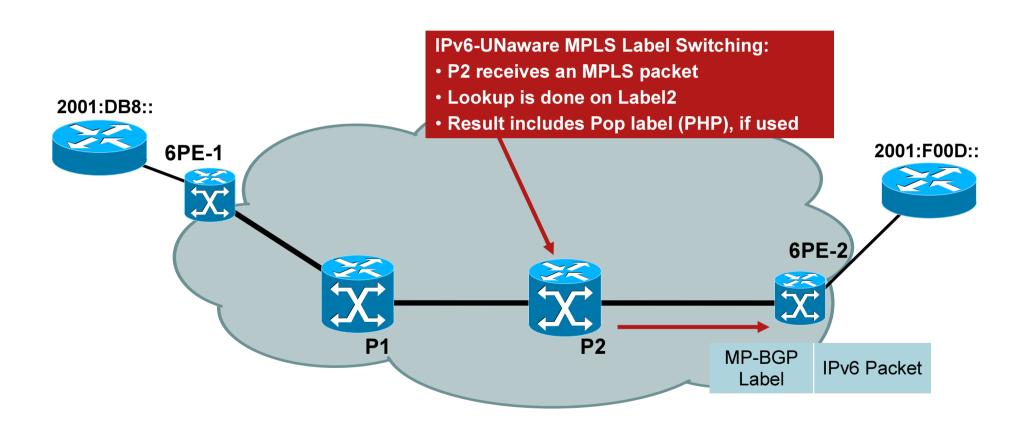
## **6PE Forwarding (6PE-1)**



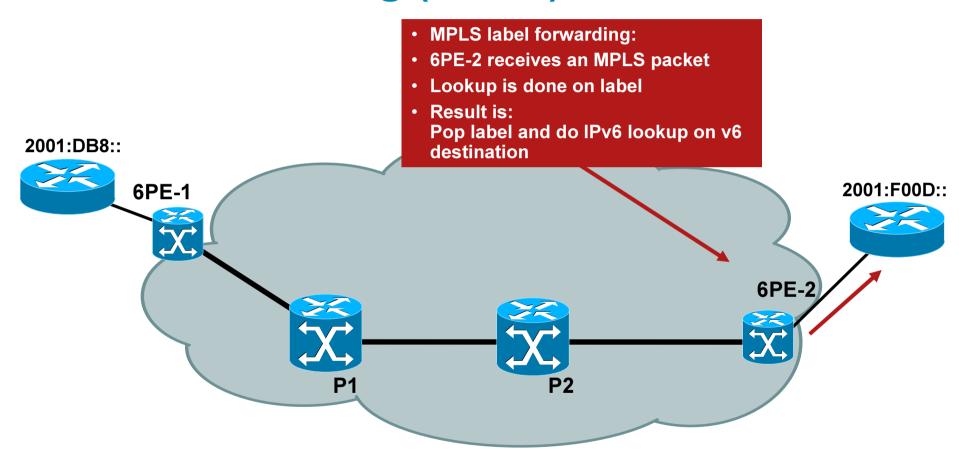
# **6PE Forwarding (P1)**



# **6PE Forwarding (P2)**

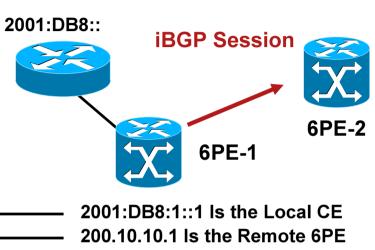


## **6PE Forwarding (6PE-2)**



# **6PE-1 Configuration**

```
ipv6 cef
mpls label protocol ldp
router bgp 100
no synchronization
no bgp default ipv4 unicast
 neighbor 2001:DB8:1::1 remote-as 65014
neighbor 200.10.10.1 remote-as 100
neighbor 200.10.10.1 update-source Loopback0
 address-family ipv6
 neighbor 200.10.10.1 activate
 neighbor 200.10.10.1 send-label
 neighbor 2001:DB8:1::1 activate
 redistribute connected
no synchronization
 exit-address-family
```



Send Labels Along with IPv6 Prefixes by Means of MP-BGP Note: Will Cause Session to Flap

#### **6PE Show Output**

```
6PE-1#show ip route 200.10.10.1
Routing entry for 200.10.10.1/32
Known via "isis", distance 115, metric 20, type level-2
[snip]
  * 10.12.0.1, from 200.10.10.1, via FastEthernet1/0
   Route metric is 20, traffic share count is 1
6PE-1#show ipv6 route
B 2001:F00D::/64 [200/0]
  via :: FFFF: 200.10.10.1, IPv6-mpls
6PE-1#show ipv6 cef internal #hidden command
.. OUTPUT TRUNCATED ..
2001:F00D::/64,
  nexthop :: FFFF: 200.10.10.1
fast tag rewrite with F0/1, 10.12.0.1, tags imposed \{17.28\}
```

#### **Other Useful Output:**

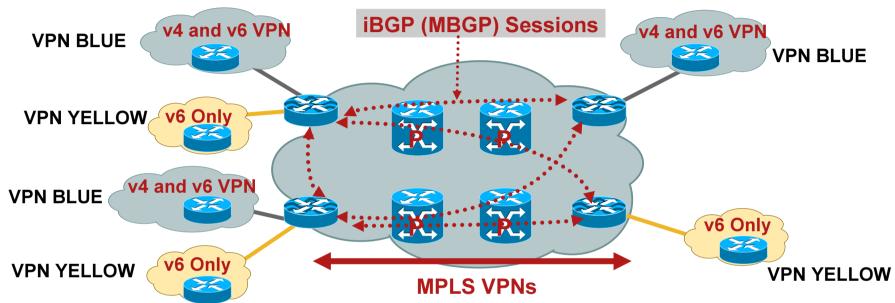
```
show bgp ipv6 neighbors
show bgp ipv6 unicast
show mpls forwarding #more on this later
```







#### **6VPE Deployment**



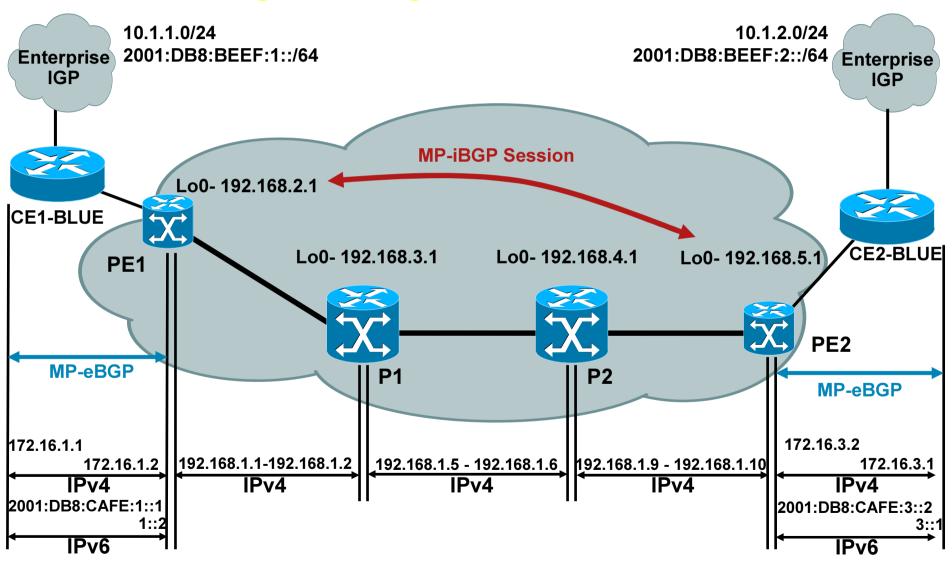
- 6VPE ~ IPv6 + BGP-MPLS IPv4 VPN + 6PE
- Cisco 6VPE is an implementation of RFC4659
- VPNv6 address:

Address including the 64 bits route distinguisher and the 128 bits IPv6 address

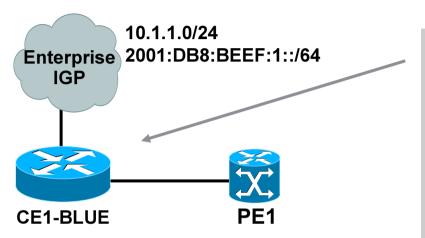
- MP-BGP VPNv6 address-family: AFI "IPv6" (2), SAFI "VPN" (128)
- VPN IPv6 MP REACH NLRI With VPNv6 next-hop and NLRI in the form of <length, IPv6-prefix, label>
- Encoding of the BGP next-hop

#### **6VPE Example Design**

#### Addressing/Routing



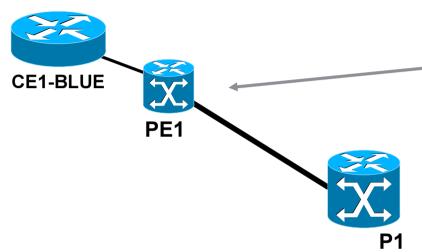
#### **6VPE Configuration Example CE1-BLUE to PE1**



```
ipv6 unicast-routing
ipv6 cef
interface Ethernet0/0
description to PE1
 ip address 172.16.1.1 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1::1/64
interface Ethernet1/0
description to BLUE LAN
 ip address 10.1.1.1 255.255.255.0
 ipv6 address 2001:DB8:BEEF:1::1/64
 ipv6 rip BLUE enable
```

```
router bgp 500
bgp log-neighbor-changes
neighbor 2001:DB8:CAFE:1::2 remote-as 100
neighbor 172.16.1.2 remote-as 100
 address-family ipv4
redistribute connected
redistribute eigrp 100
no neighbor 2001:DB8:CAFE:1::2 activate
neighbor 172.16.1.2 activate
no auto-summary
no synchronization
 exit-address-family
 address-family ipv6
neighbor 2001:DB8:CAFE:1::2 activate
redistribute connected
redistribute rip BLUE
no synchronization
exit-address-family
ipv6 router rip BLUE
redistribute bgp 500
```

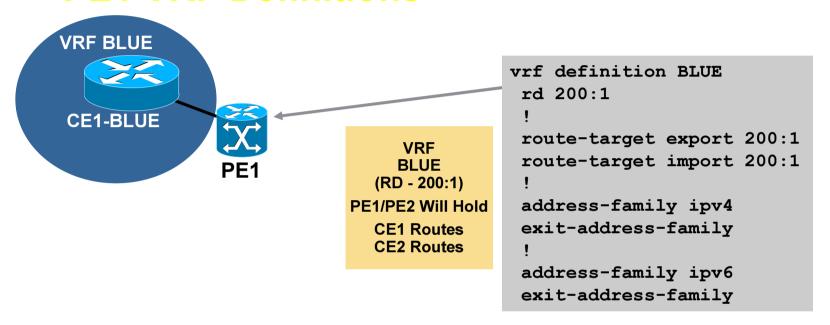
#### **6VPE Configuration Example PE1 Connections**



- Standard MPLS configuration between PE-P
- Running IGP in the cloud

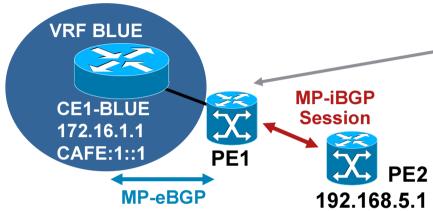
```
ipv6 unicast-routing
ipv6 cef
mpls ldp router-id Loopback0
interface Loopback0
 ip address 192.168.2.1 255.255.255.255
interface Ethernet0/0
 description to CE1-BLUE
vrf forwarding BLUE
 ip address 172.16.1.2 255.255.255.0
 ipv6 address 2001:DB8:CAFE:1::2/64
interface Ethernet2/0
 description to P1
 ip address 192.168.1.1 255.255.255.252
mpls ip
router ospf 1
 log-adjacency-changes
 redistribute connected subnets
passive-interface Loopback0
 network 192.168.1.0 0.0.0.255 area 0
```

# **6VPE Configuration Example PE1 VRF Definitions**



# Define VRF Name Route-Distinguisher Route-Targets

#### **6VPE Configuration Example** PE1 BGP Setup



```
router bgp 100
bgp log-neighbor-changes
 neighbor 192.168.5.1 remote-as 100
 neighbor 192.168.5.1 update-source
Loopback0
 address-family ipv4
 neighbor 192.168.5.1 activate
 no auto-summary
 no synchronization
 exit-address-family
 address-family vpnv4
 neighbor 192.168.5.1 activate
 neighbor 192.168.5.1 send-community
extended
 exit-address-family
```

```
address-family vpnv6
neighbor 192.168.5.1 activate
neighbor 192.168.5.1 send-community
extended
exit-address-family
 address-family ipv4 vrf BLUE
 redistribute connected
neighbor 172.16.1.1 remote-as 500
neighbor 172.16.1.1 activate
no auto-summary
 no synchronization
 exit-address-family
 address-family ipv6 vrf BLUE
neighbor 2001:DB8:CAFE:1::1 remote-as
500
neighbor 2001:DB8:CAFE:1::1 activate
 redistribute connected
no synchronization
 exit-address-family
```

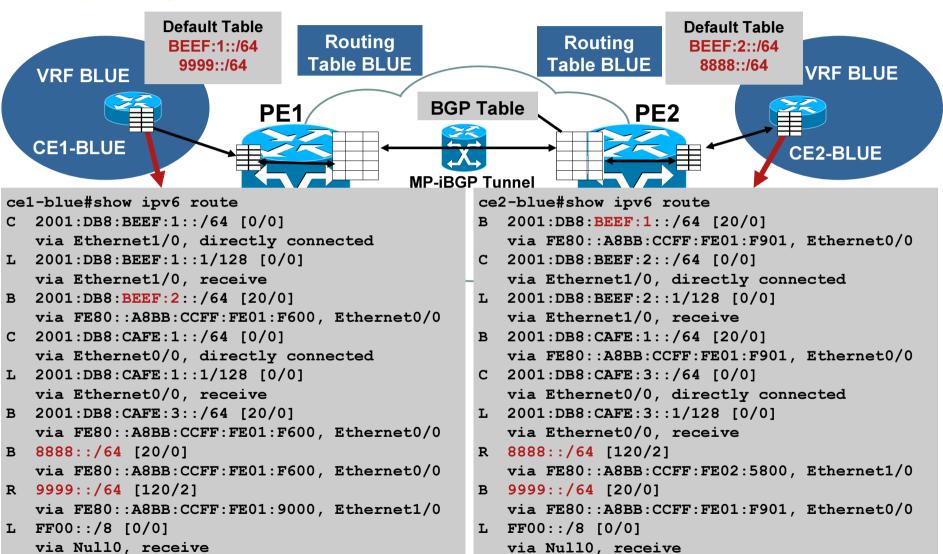
#### **6VPE Configuration Example** P Connections

```
mpls ldp router-id Loopback0
interface Loopback0
 ip address 192.168.3.1 255.255.255.255
interface Ethernet0/0
 description to PE1
 ip address 192.168.1.2 255.255.255.252
mpls ip
interface Ethernet1/0
 description to P2
 ip address 192.168.1.5 255.255.255.252
mpls ip
router ospf 1
 log-adjacency-changes
 redistribute connected subnets
 passive-interface Loopback0
 network 192.168.1.0 0.0.0.255 area 0
```

```
mpls ldp router-id Loopback0
interface Loopback0
 ip address 192.168.4.1 255.255.255.255
interface Ethernet0/0
 description to P1
 ip address 192.168.1.6 255.255.255.252
mpls ip
interface Ethernet1/0
 description to PE2
 ip address 192.168.1.9 255.255.255.252
mpls ip
router ospf 1
 log-adjacency-changes
 redistribute connected subnets
passive-interface Loopback0
 network 192.168.1.0 0.0.0.255 area 0
```

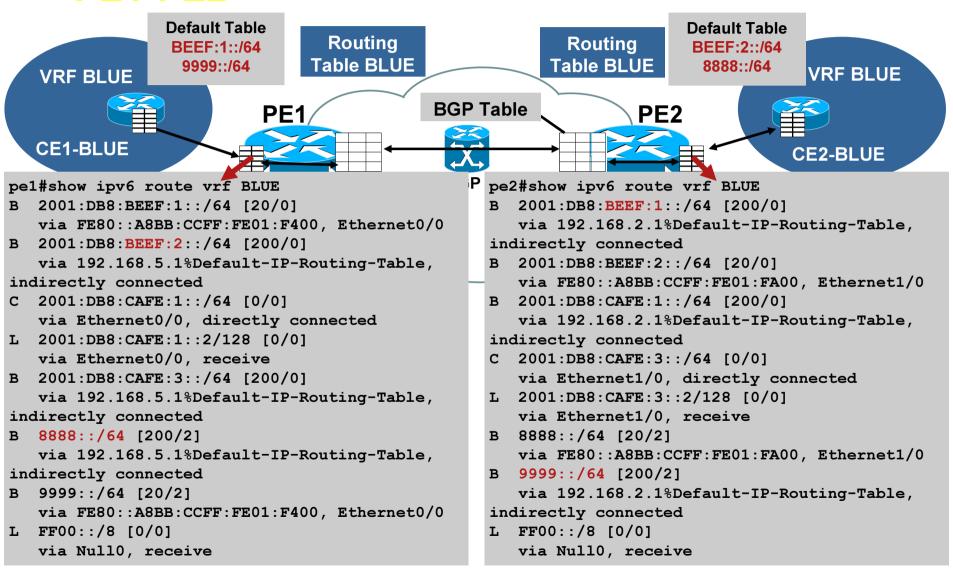
# **IPv6 Routing Tables**

CE1-CE2

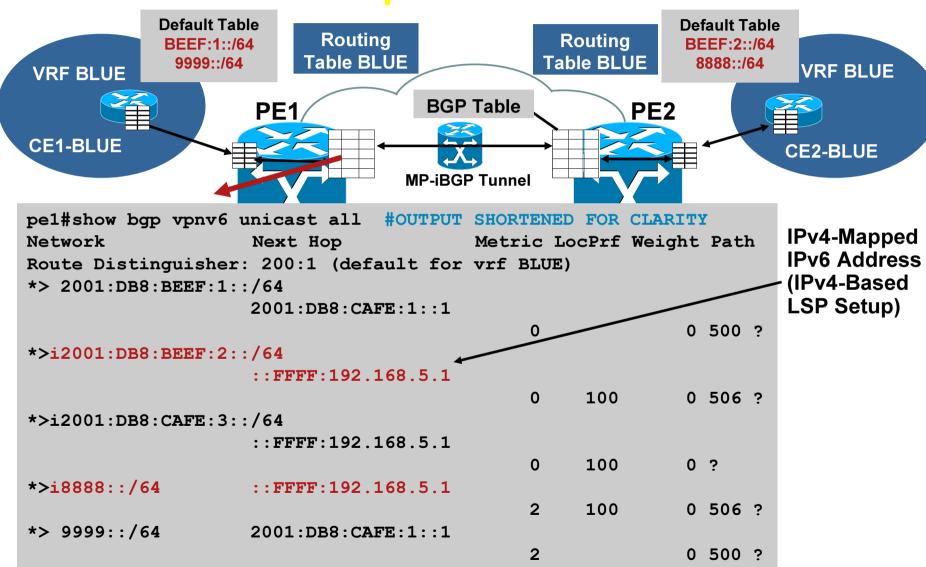


# **IPv6 Routing Tables**

PE1-PE2

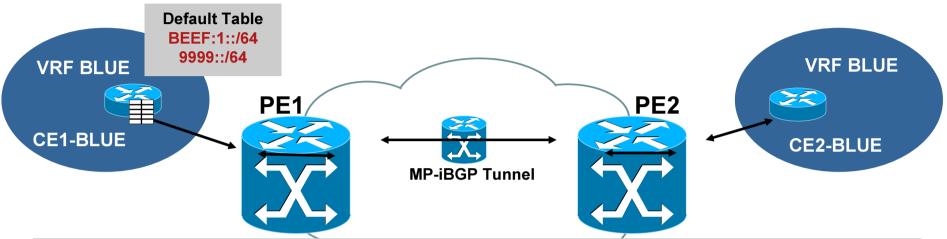


## **IPv6 Routing Tables PE1 BGP Next-Hop**



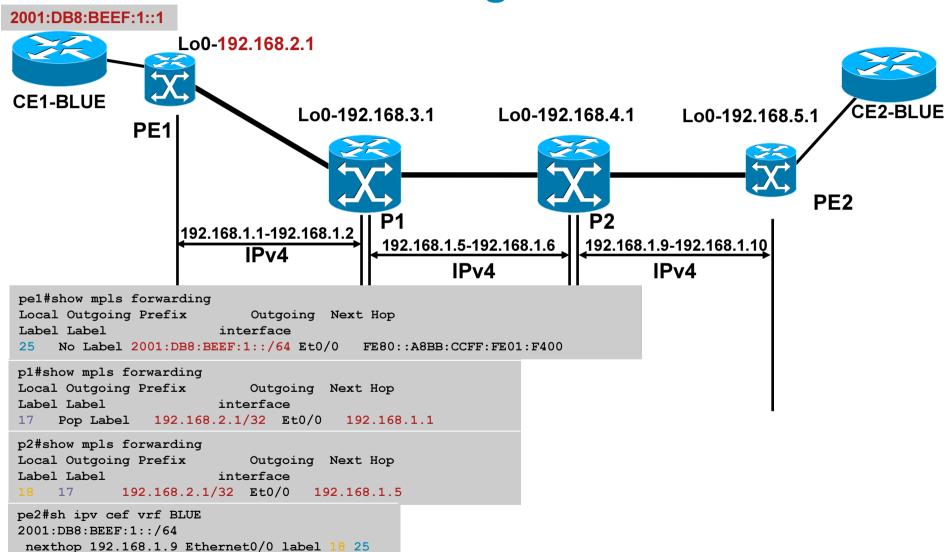
# **MPLS Forwarding**

PE<sub>1</sub>



pel#show mpls forwarding					
Local	Outgoing	Prefix	Bytes Label	Outgoing	Next Hop
Label	Label or VC	or Tunnel Id	Switched	interface	
16	Pop Label	192.168.1.4/30	0	Et2/0	192.168.1.2
17	16	192.168.1.8/30	0	Et2/0	192.168.1.2
18	Pop Label	192.168.3.1/32	0	Et2/0	192.168.1.2
19	18	192.168.4.1/32	0	Et2/0	192.168.1.2
20	19	192.168.5.1/32	0	Et2/0	192.168.1.2
21	No Label	10.1.1.0/24[V]	0	Et0/0	172.16.1.1
22	Aggregate	172.16.1.0/24[V]	570	BLUE	
25	No Label	2001:DB8:BEEF:1::	/64[V] \		
			570	Et0/0	FE80::A8BB:CCFF:FE01:F400
26	Aggregate	2001:DB8:CAFE:1::	/64[V] \		
			35456	BLUE	
27	No Label	9999::/64[V]	570	Et0/0	FE80::A8BB:CCFF:FE01:F400

## A Look at Forwarding



#