BGP Enhancements for IPv6

ISP/IXP Workshops

Adding IPv6 to BGP…

- RFC4760

  Defines Multi-protocol Extensions for BGP4

  Enables BGP to carry routing information of protocols other than IPv4

  e.g. MPLS, IPv6, Multicast etc

  Exchange of multiprotocol NLRI must be negotiated at session startup

- RFC2545

  Use of BGP Multiprotocol Extensions for IPv6 Inter-Domain Routing

RFC4760

- New optional and non-transitive BGP attributes:

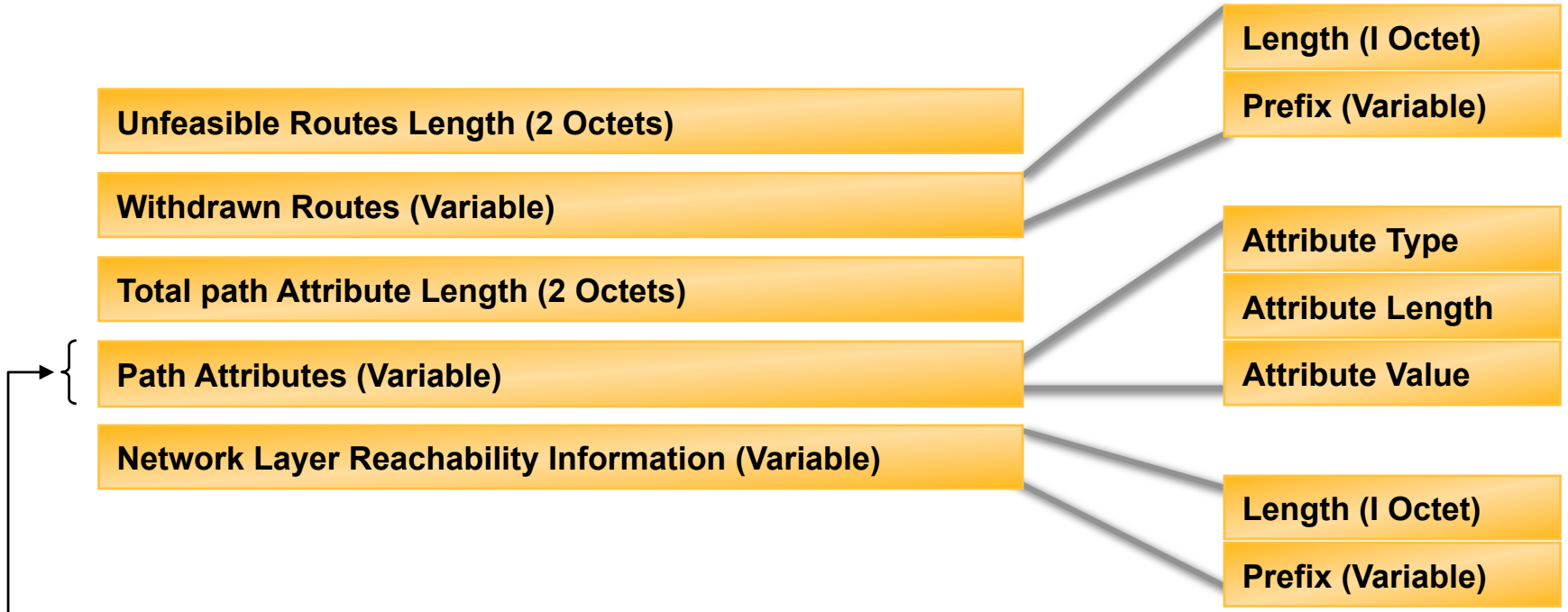
  MP_REACH_ NLRI (Attribute code: 14)

  Carry the set of reachable destinations together with the next-hop information to be used for forwarding to these destinations (RFC2858)

  MP_UNREACH_NLRI (Attribute code: 15)

  Carry the set of unreachable destinations

MPBGP Update Message



New Multiprotocol Attributes added to Path Attributes:

MP_REACH_NLRI

MP_UNREACH_NLRI

MPBGP New Attributes

- Attribute contains one or more Triples:
  - AFI Address Family Information
  - Next-Hop Information (must be of the same address family)
  - NLRI Network Layer Reachability Information

AFI/SAFI Information

- Address Family Information (AFI)
  - Identifies Address Type (see RFC1700)
    - AFI = 1 (IPv4)
    - AFI = 2 (IPv6)
- Subsequent Address Family Information (Sub-AFI)
  - Sub category for AFI Field
  - Address Family Information (AFI) = 2 (IPv6)
    - Sub-AFI = 1 (NLRI is used for unicast)
    - Sub-AFI = 2 (NLRI is used for multicast RPF check)
    - Sub-AFI = 3 (NLRI is used for both unicast and multicast RPF check)
    - Sub-AFI = 4 (label)

RFC2545

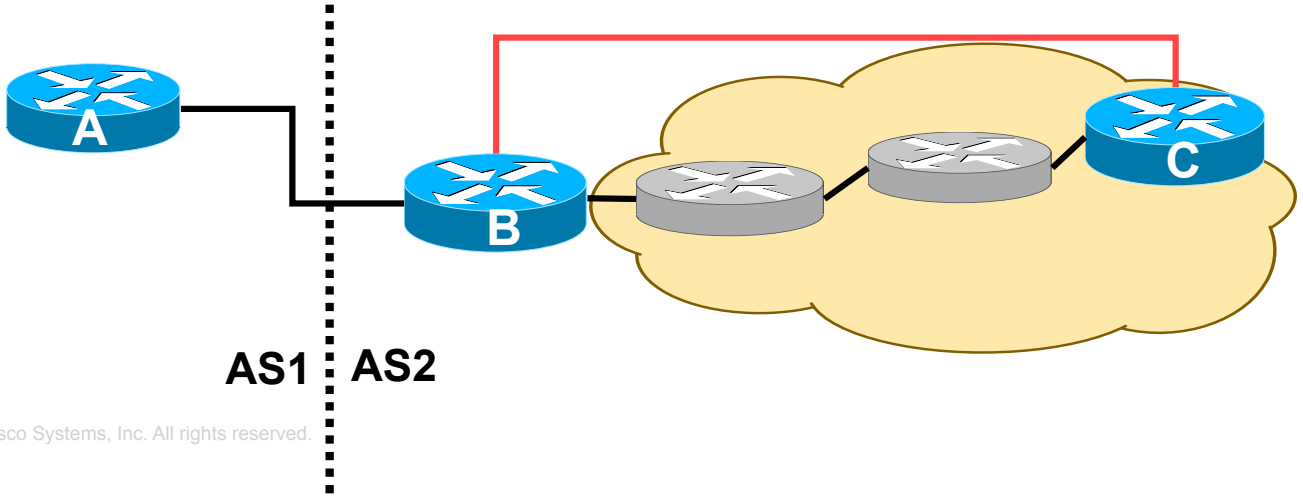
- IPv6 specific extensions

  IPv6 AFI/SAFI definition

  NEXT_HOP and NLRI are expressed as IPv6 addresses and prefix

  Scoped addresses: Next-hop contains a global IPv6 address and/or potentially a link-local address

    Link local address as a next-hop is only set if the BGP peer shares the subnet with both routers (advertising and advertised)



MP-BGP Capabilities Negotiation

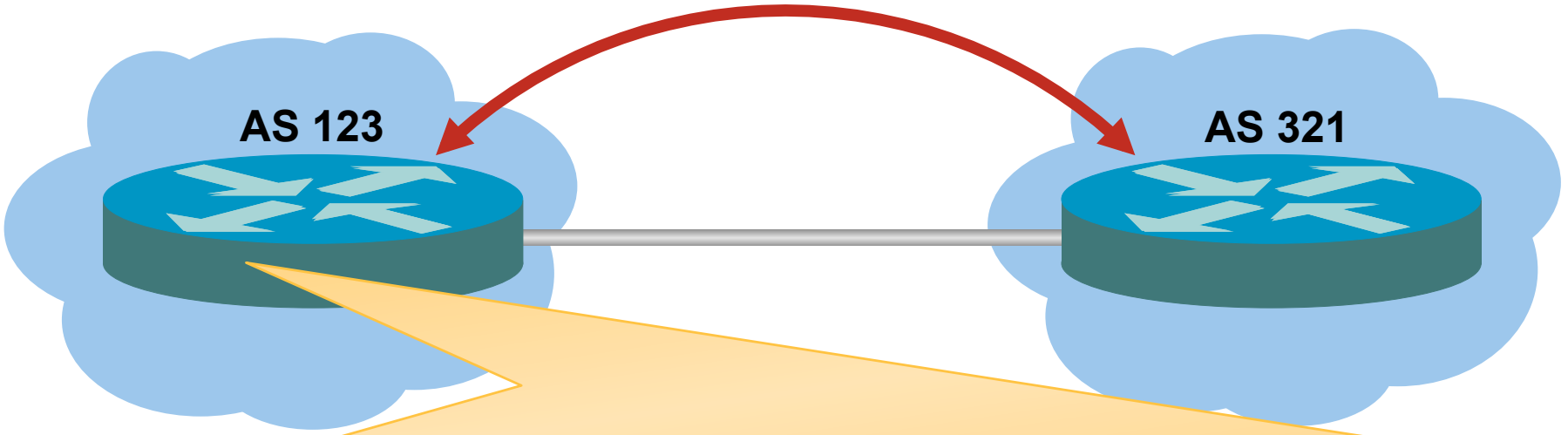
- BGP routers establish BGP sessions through the OPEN message
- OPEN message contains optional parameters
- If OPEN parameters are not recognized, BGP session is terminated
- A new optional parameter:

  CAPABILITIES

MP-BGP Capabilities Negotiation (Cont.)

- BGP router sends an OPEN message with CAPABILITIES parameter containing its capabilities :
  - Multi-Protocol extensions (AFI/SAFI)
  - Route Refresh
  - Outbound Route Filtering

MP-BGP Session Establishment




```
BGP: 3FFE:B00:C18:2:1::1 sending OPEN, version 4, my as: 100
BGP: 3FFE:B00:C18:2:1::1 rcv OPEN, version 4
BGP: 3FFE:B00:C18:2:1::1 rcv OPEN w/ OPTION parameter len: 16
BGP: 3FFE:B00:C18:2:1::1 rcvd OPEN w/ optional parameter type 2 (Capability) len 6
BGP: 3FFE:B00:C18:2:1::1 OPEN has CAPABILITY code: 1, length 4
BGP: 3FFE:B00:C18:2:1::1 OPEN has MP_EXT CAP for afi/safi: 2/1
BGP: 3FFE:B00:C18:2:1::1 rcvd OPEN w/ optional parameter type 2 (Capability) len 2
BGP: 3FFE:B00:C18:2:1::1 went from OpenSent to OpenConfirm
BGP: 3FFE:B00:C18:2:1::1 went from OpenConfirm to Established
%BGP-5-ADJCHANGE: neighbor 3FFE:B00:C18:2:1::1 Up
```

BGP Configuration

- New ‘address-family’ structure

```
router bgp <asn>
 address-family <afi> [<sub-afi>]
                .
                .
 exit-address-family
```

Example:

```
address-family ipv4
address-family ipv4 multicast
address-family vpnv4
address-family ipv6
address-family vpnv6
```

- Separates configurations by address family

BGP Configuration

- IOS assumes by default that all BGP neighbours will be IPv4 unicast neighbours

  We need to remove this assumption

```
router bgp 100
 no bgp default ipv4-unicast
```

- Failing to do this will result in all neighbours being defined as IPv4 unicast neighbours

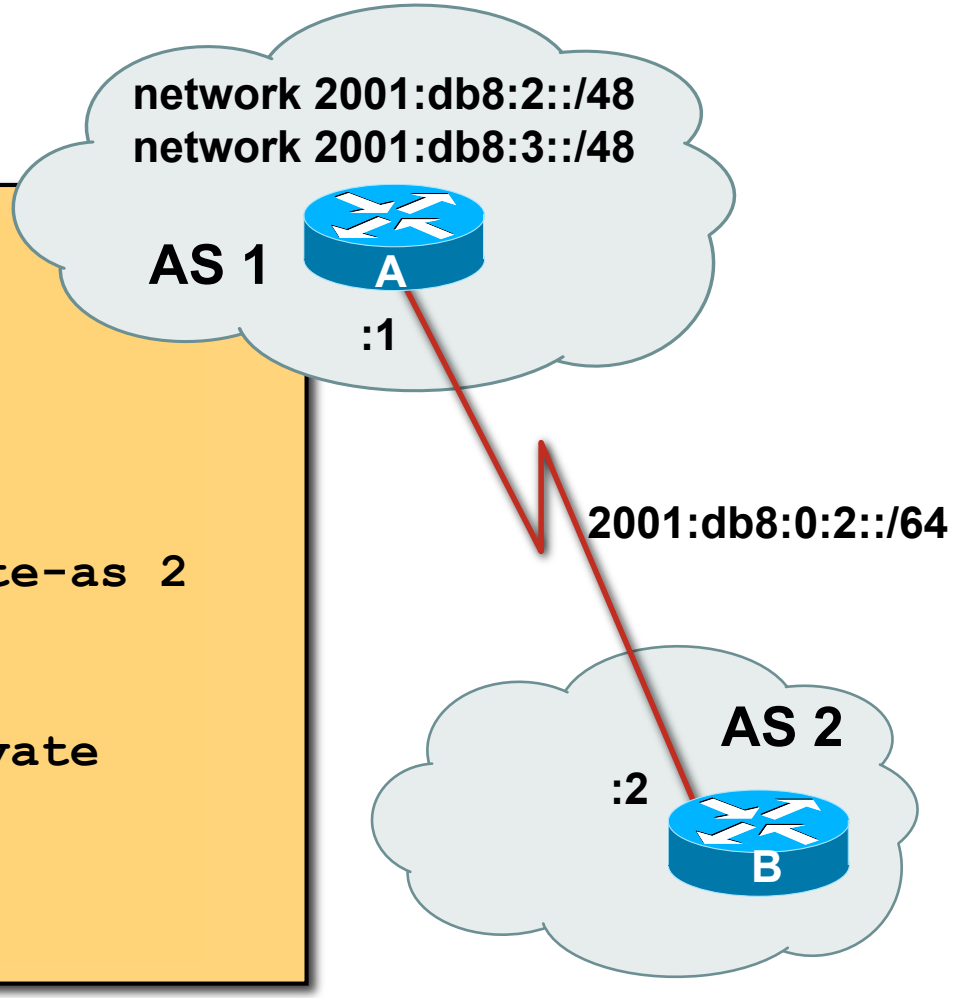
  Non-IPv4 neighbours will have no specific unicast IPv4 configuration

  Cluttered configuration, confusing troubleshooting and diagnosis

BGP Configurations
Regular Peering

Router A

```
router bgp 1
 no bgp default ipv4-unicast
 bgp router-id 1.1.1.1
 neighbor 2001:db8:0:2::2 remote-as 2
!
 address-family ipv6
 neighbor 2001:db8:0:2::2 activate
 network 2001:db8:2::/48
 network 2001:db8:3::/48
!
```



activate means that the BGP peering is activated for this particular address family

Routing Information

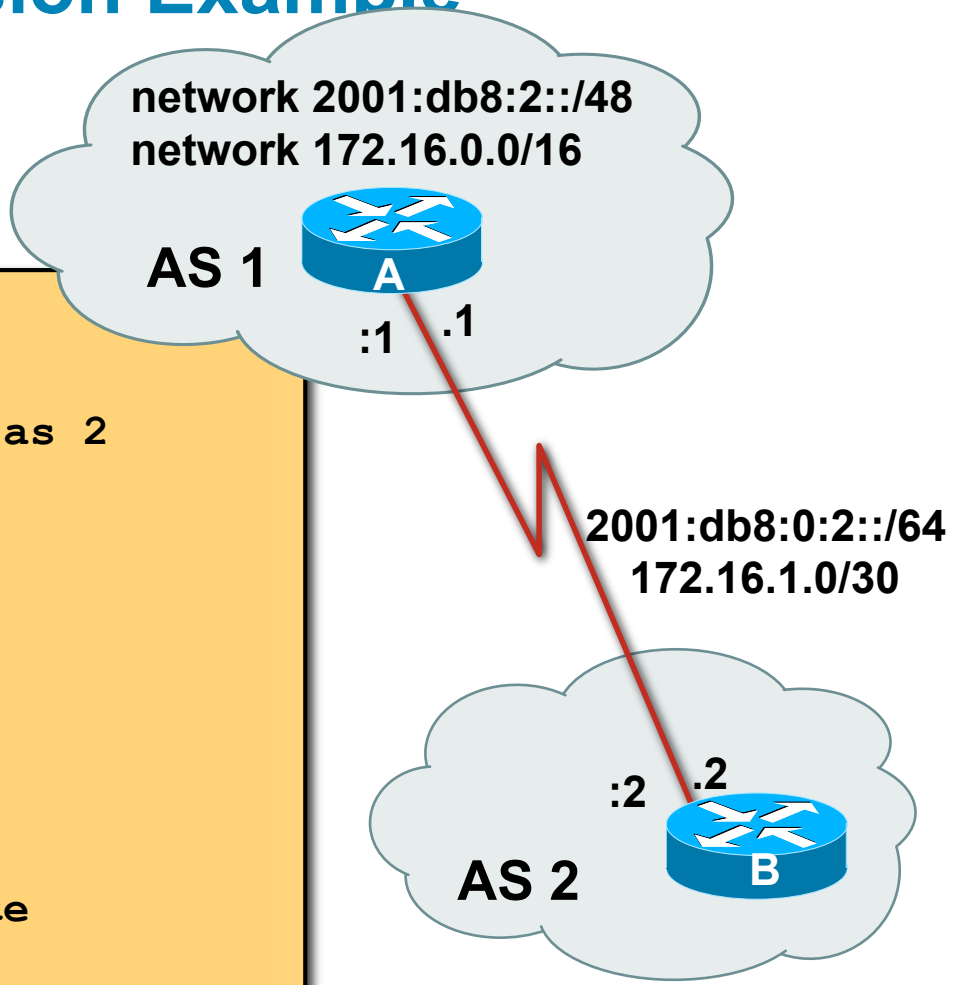
- Independent operation
  - One RIB per protocol
    - e.g. IPv6 has its own BGP table
  - Distinct policies per protocol

BGP Configuration

- TCP Interaction
  - BGP runs on top of TCP
  - This connection could be set up either over IPv4 or IPv6
- Peering sessions can be shared when the topology is congruent

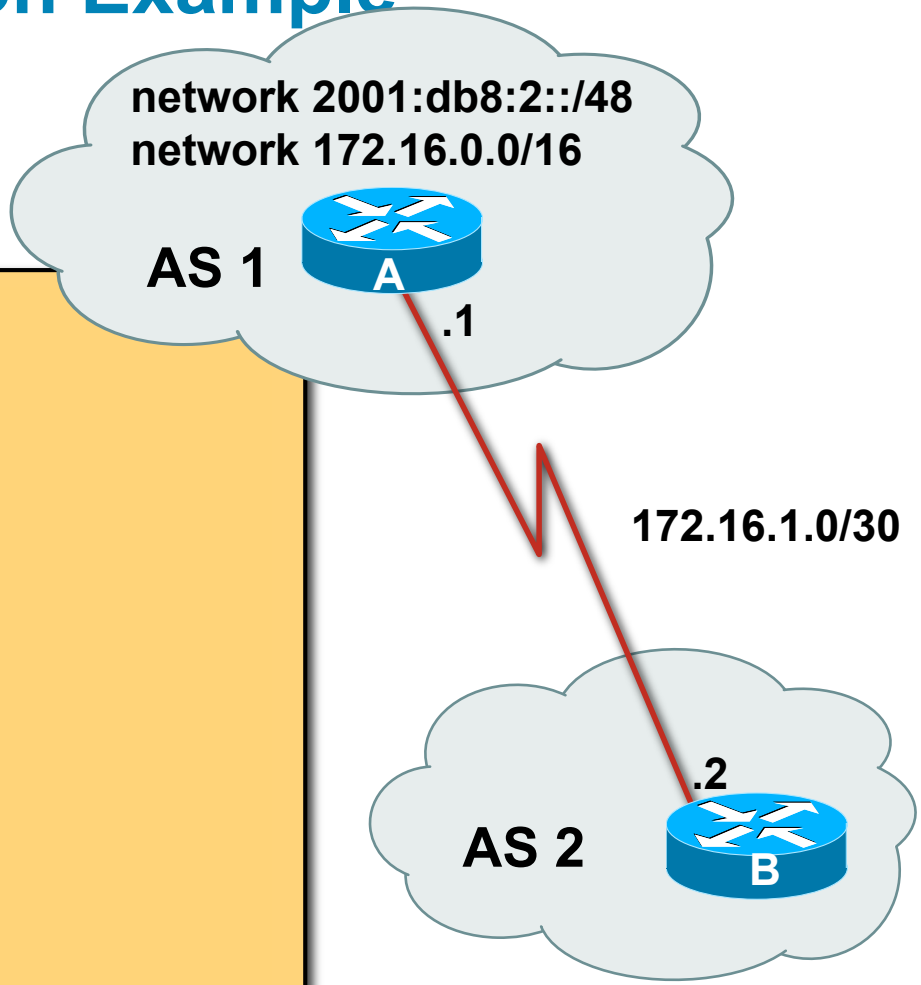
BGP Configuration
Separate Peering Session Example

```
router bgp 1
 no bgp default ipv4-unicast
 neighbor 2001:db8:0:2::2 remote-as 2
 neighbor 172.16.1.2 remote-as 2
!
 address-family ipv4
 neighbor 172.16.1.2 activate
 network 172.16.0.0
 exit-address-family
!
 address-family ipv6
 neighbor 2001:db8:0:2::2 activate
 network 2001:db8:2::/48
 exit-address-family
!
```



BGP Configuration
Shared Peering Session Example

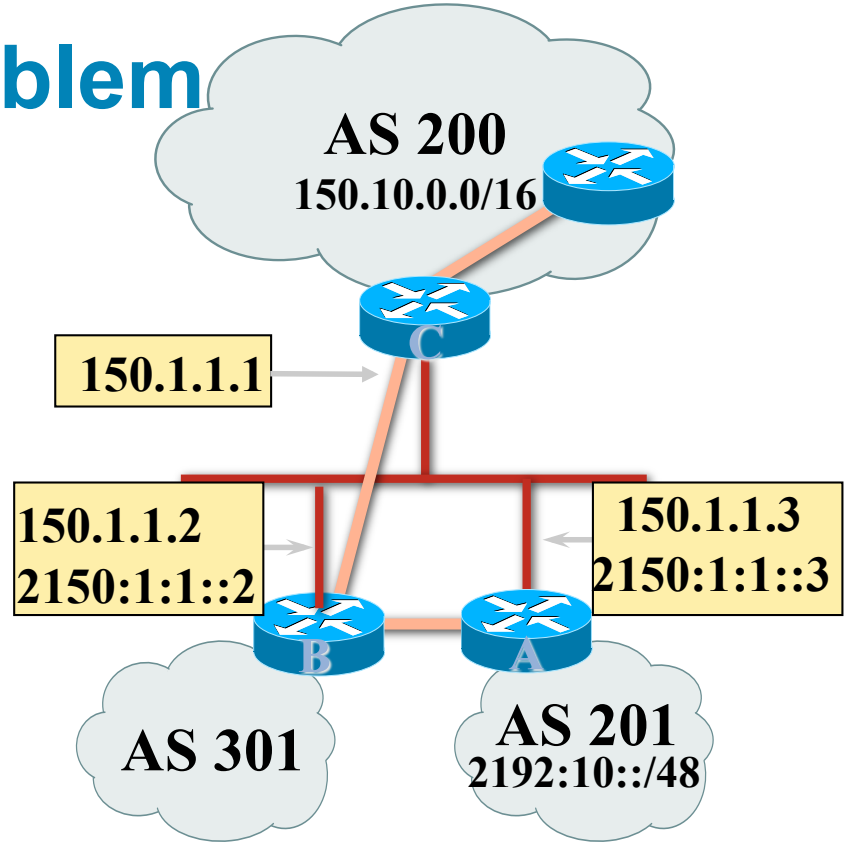
```
router bgp 1
 no bgp default ipv4-unicast
 neighbor 172.16.1.2 remote-as 2
!
 address-family ipv4
 neighbor 172.16.1.2 activate
 network 172.16.0.0
 exit-address-family
!
 address-family ipv6
 neighbor 172.16.1.2 activate
 network 2001:db8:2::/48
 exit-address-family
!
```



IPv6 NLRI in IPv4 – Problem

Router A

```
router bgp 201
 bgp router-id 192.168.30.1
 neighbor 150.1.1.2 remote-as 301
 !
 address-family ipv6
 neighbor 150.1.1.2 activate
 network 2192:10::/48
!
```



Router A:

BGP(1): 150.1.1.2 send UPDATE (format) 2192:10::/48, next ::FFFF:150.1.1.3, metric 0, path Local

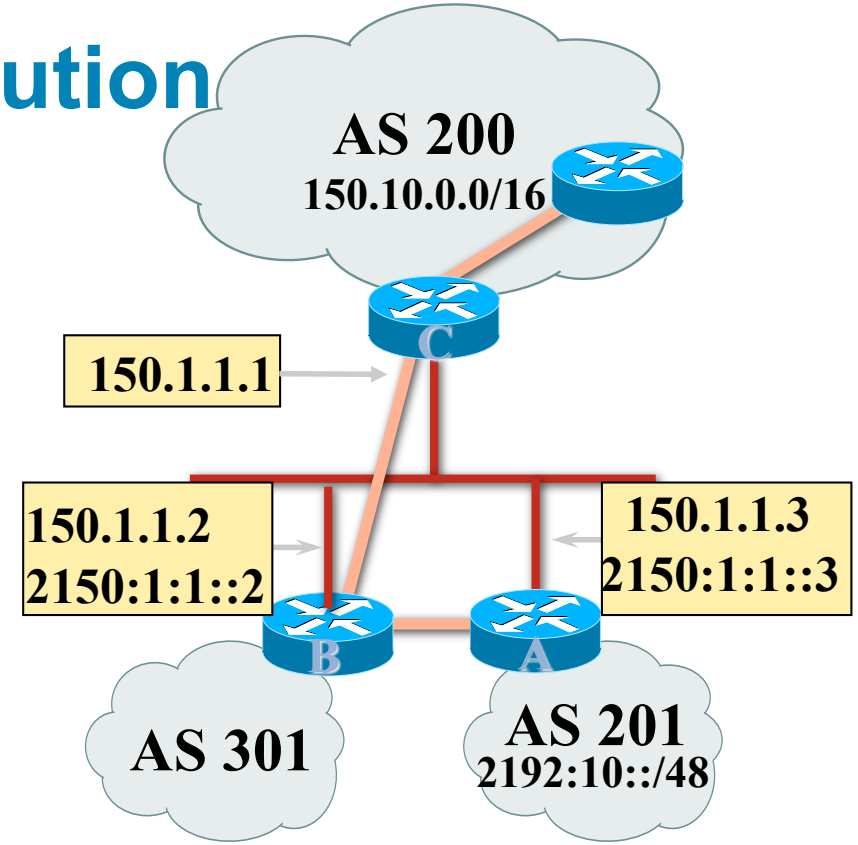
Router B:

BGP(1): 150.1.1.3 rcvd UPDATE w/ attr: nexthop ::FFFF:150.1.1.3, origin i, localpref 100, metric 0
BGP(1): 150.1.1.3 rcvd 2192:10::/48
BGP(1): no valid path for 2192:10::/48

IPv6 NLRI In IPv4 – Solution

Router A

```
router bgp 201
 bgp router-id 192.168.30.1
 neighbor 150.1.1.2 remote-as 301
 !
 address-family ipv6
 neighbor 150.1.1.2 activate
 neighbor 150.1.1.2 route-map SETNH out
 network 2192:10::/48
!
route-map SETNH permit 10
 set ipv6 next-hop 2150:1:1::3
```



Router A:

BGP(1): 150.1.1.2 send UPDATE (prepend, chgflags: 0x820) 2192:10::/48, next 2150:1:1::3, metric 0, path Local

Router B:

BGP(1): 2150:1:1::3 rcvd UPDATE w/ attr: nexthop 2150:1:1::3, origin i, localpref 100, metric 0
BGP(1): 2150:1:1::3 rcvd 2192:10::/48
BGP(1): Revise route installing 2192:10::/48 -> 2150:1:1::3 (::) to main IPv6 table

BGP next-hop Behavior

IPv6 Prefix Advertisement over an IPv4 BGP Session

- **Setting IPv6 next-hop explicitly is usually necessary**
- **IOS-XR 3.8, 12.0(32)SY09, 12.0(33)S4 and later:**

  Tries to use automatically a valid IPv6 address as next-hop

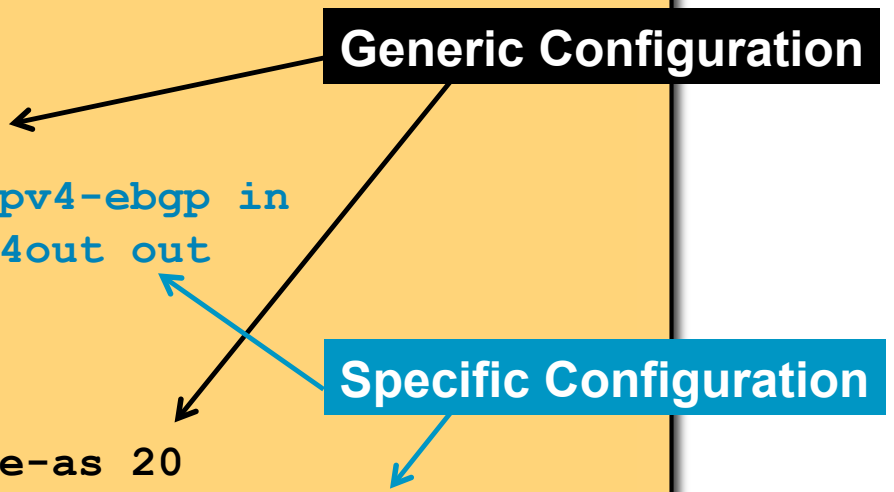
  - **if this is a directly connected peering session, pick up local interface IPv6 address (global address and link-local address)**
  - **if this is a loopback peering session (loopback interface IPv4 address configured as update-source), pick up a IPv6 address from that loopback interface (global address and link-local address)**
  - **otherwise, fall back to the default behavior of a v4-mapped-v6 address**

BGP Configuration IPv4 and IPv6

- When configuring the router, recommendation is:
  - Put **all** IPv6 configuration directly into IPv6 address family
  - Put **all** IPv4 configuration directly into IPv4 address family
- Router will sort generic from specific address family configuration when the configuration is saved to NVRAM or displayed on the console
- Example follows…
  - Notice how **`activate`** is added by the router to indicate that the peering is activated for the particular address family

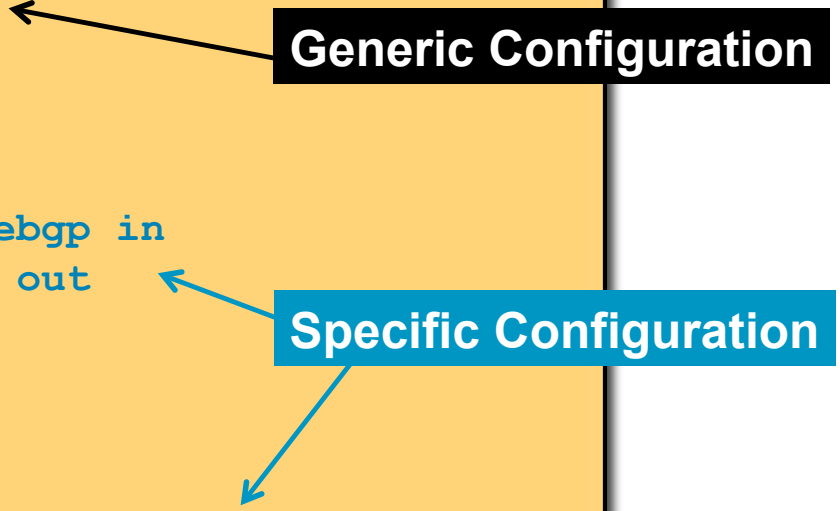
BGP Address Families Applied Configuration

```
router bgp 10
 no bgp default ipv4-unicast
!
address family ipv4
 neighbor 172.16.1.2 remote-as 30
 neighbor 172.16.1.2 prefix-list ipv4-ebgp in
 neighbor 172.16.1.2 prefix-list v4out out
 network 172.16.0.0
!
 address-family ipv6
 neighbor 2001:db8:1:1019::1 remote-as 20
 neighbor 2001:db8:1:1019::1 prefix-list ipv6-ebgp in
 neighbor 2001:db8:1:1019::1 prefix-list v6out out
 network 2001:db8::/32
!
ip prefix-list ipv4-ebgp permit 0.0.0.0/0 le 32
ip prefix-list v4out permit 172.16.0.0/16
ipv6 prefix-list ipv6-ebgp permit ::/0 le 128
ipv6 prefix-list v6out permit 2001:db8::/32
```



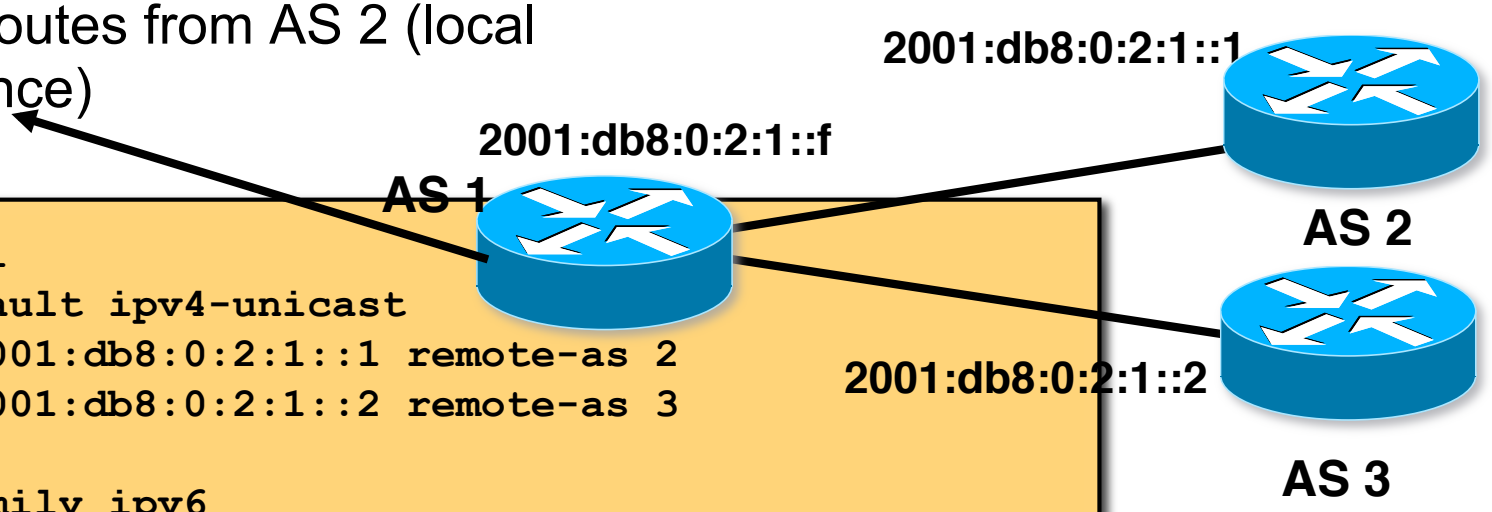
BGP Address Families
End result

```
router bgp 10
 no bgp default ipv4-unicast
 neighbor 2001:db8:1:1019::1 remote-as 20
 neighbor 172.16.1.2 remote-as 30
!
 address-family ipv4
 neighbor 172.16.1.2 activate
 neighbor 172.16.1.2 prefix-list ipv4-ebgp in
 neighbor 172.16.1.2 prefix-list v4out out
 network 172.16.0.0
 exit-address-family
!
 address-family ipv6
 neighbor 2001:db8:1:1019::1 activate
 neighbor 2001:db8:1:1019::1 prefix-list ipv6-ebgp in
 neighbor 2001:db8:1:1019::1 prefix-list v6out out
 network 2001:db8::/32
 exit-address-family
!
ip prefix-list ipv4-ebgp permit 0.0.0.0/0 le 32
ip prefix-list v4out permit 172.16.0.0/16
ipv6 prefix-list ipv6-ebgp permit ::/0 le 128
ipv6 prefix-list v6out permit 2001:db8::/32
```



BGP Configuration
Manipulating Attributes

- Prefer routes from AS 2 (local preference)




```
router bgp 1
 no bgp default ipv4-unicast
 neighbor 2001:db8:0:2:1::1 remote-as 2
 neighbor 2001:db8:0:2:1::2 remote-as 3
!
 address-family ipv6
 neighbor 2001:db8:0:2:1::1 activate
 neighbor 2001:db8:0:2:1::1 prefix-list in-filter in
 neighbor 2001:db8:0:2:1::1 route-map fromAS2 in
 neighbor 2001:db8:0:2:1::2 activate
 neighbor 2001:db8:0:2:1::2 prefix-list in-filter in
 network 2001:db8::/32
 exit-address-family
!
route-map fromAS2 permit 10
 set local-preference 120
```

BGP Configuration
Prefix List

```
ip prefix-list ipv4-ebgp permit 0.0.0.0/0 le 32
!
ip prefix-list v4out permit 172.16.0.0/16
!
ipv6 prefix-list ipv6-ebgp permit ::/0 le 128
!
ipv6 prefix-list v6out permit 2001:db8::/32
!
```

- Compare IPv4 prefix filters with IPv6 prefix filters

```
ip prefix-list <name> permit|deny <ipv4 address>
ipv6 prefix-list <name> permit|deny <ipv6 address>
```

BGP Configuration
Carrying IPv4 inside IPv6 peering

- IPv4 prefixes can be carried inside an IPv6 peering

  Note that the next-hop for received prefixes needs to be “fixed”

- Example

```
router bgp 1
 neighbor 2001:db8:0:2::2 remote-as 2
!
address-family ipv4
 neighbor 2001:db8:0:2::2 activate
 neighbor 2001:db8:0:2::2 route-map ipv4 in
!
route-map ipv4 permit 10
 set ip next-hop 131.108.1.1
```

BGP Status Commands

- IPv6 BGP show commands take ipv6 as argument

```
show bgp ipv6 unicast <parameter>
```

```
Router1#show bgp ipv6 unicast 2017::/32
BGP routing table entry for 2017::/32, version 11
Paths: (1 available, best #1)
Local
2001:db8:c18:2:1::1 from 2001:db8:c18:2:1::1 (10.10.20.2)
Origin incomplete, localpref 100, valid, internal, best
```

- IPv4 BGP show commands can also use this format:

```
show bgp ipv4 unicast <parameter>
```

BGP Status Commands

Display summary information regarding the state of the BGP neighbours

show bgp ipv6 unicast summary

```
BGP router identifier 128.107.240.254, local AS number 109
BGP table version is 400386, main routing table version 400386
585 network entries using 78390 bytes of memory
9365 path entries using 674280 bytes of memory
16604 BGP path attribute entries using 930384 bytes of memory
8238 BGP AS-PATH entries using 228072 bytes of memory
42 BGP community entries using 1008 bytes of memory
9451 BGP route-map cache entries using 302432 bytes of memory
584 BGP filter-list cache entries using 7008 bytes of memory
BGP using 2221574 total bytes of memory
2 received paths for inbound soft reconfiguration
BGP activity 63094/62437 prefixes, 1887496/1878059 paths, scan interval 60secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
2001:1458:C000::64B:4:1
                4   513 1294728  460213   400386    0    0 3d11h           498
```

Neighbour Information

BGP Messages Activity

Conclusion

- BGP extended to support multiple protocols
  - IPv6 is but one more address family
- Operators experienced with IPv4 BGP should have no trouble adapting
  - Configuration concepts and CLI is familiar format

BGP Enhancements for IPv6

ISP/IXP Workshops