IPv6 Routing Protocols

Vladimir Settey

Agenda

- Static Routing
- RIPng
- EIGRP for IPv6
- OSPFv3
- IS-IS for IPv6
- BGP-4 for IPv6

Routing in IPv6

- As in IPv4, IPv6 has 2 families of routing protocols: IGP and EGP, and still uses the longest-prefix match routing algorithm
- IGP
  - RIPng (RFC 2080)
  - Cisco EIGRP for IPv6
  - Integrated IS-ISv6 (draft-ietf-isis-ipv6-07)
  - OSPFv3 (RFC 5340)
- EGP → MP-BGP4 (RFC 2858 and RFC 2545)
- Cisco IOS supports all of them
  - Pick one that meets your objectives

IPv4 and IPv6 Parallels

RIP	RIPv2 for IPv4 RIPng for IPv6 Distinct but similar protocols with RIPng taking advantage of IPv6 specificities
OSPF	OSPFv2 for IPv4 OSPFv3 for IPv6 Distinct but similar protocols with OSPFv3 being a cleaner implementation that takes advantage of IPv6 specificities
IS-IS	Extended to support IPv6 Natural fit to some of the IPv6 foundational concepts Supports Single and Multi Topology operation
EIGRP	Extended to support IPv6 (IPv6_REQUEST_TYPE, IPv6_METRIC_TYPE, IPv6_EXTERIOR_TYPE) Some changes reflecting IPv6 characteristics
BGP	New MP_REACH_NLRI, MP_UNREACH_NLRI, AFI=2 with SAFI for Unicast/Multicast /Label/VPN Peering over IPv6 or IPv4 (route maps)

- For all intents and purposes, IPv6 IGPs are similar to their IPv4 counterparts
- IPv6 IGPs have additional features that could lead to new designs

IPv6 Default & Static Routing

Default and Static Routing

- Similar to IPv4. Need to define the next hop / interface.
- Default route denoted as ::/0

  ipv6 route *ipv6-prefix/prefix-length {ipv6-address | interface-type interface-number [ipv6-address]} [administrative-distance] [administrative-multicast-distance* | unicast | multicast] [tag *tag*]

- Examples:

  Forward packets for network 2001:DB8::/32 through 2001:DB8:1:1::1 with an administrative distance of 10

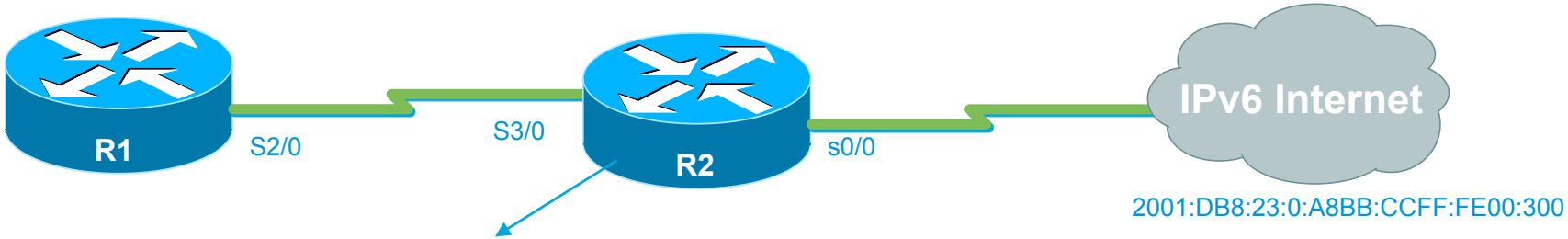
```
Router(config)# ipv6 route 2001:DB8::/32 2001:DB8:1:1::1 10
```

  Default route to 2001:DB8:1:1::1

```
Router(config)# ipv6 route ::/0 2001:DB8:1:1::1
```

IPv6 Routing – Default Route

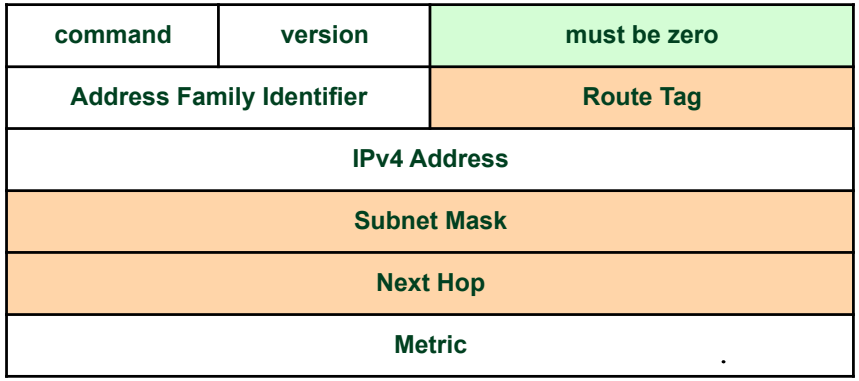
Routing table and testing connectivity

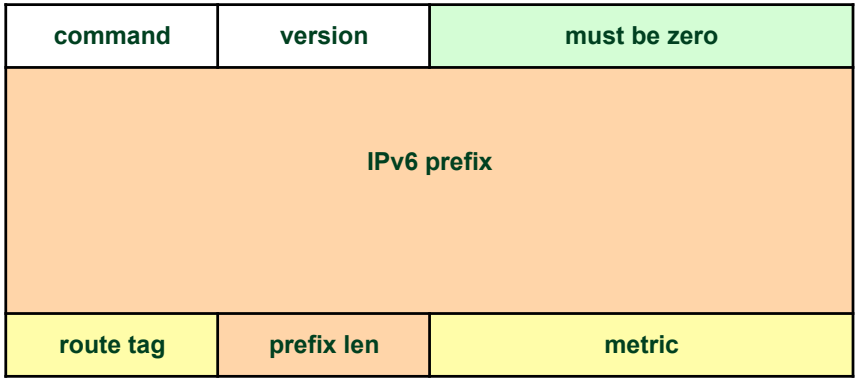



```
R1#sh run | b ipv6 route
ipv6 route ::/0 2001:DB8:12:0:A8BB:CCFF:FE00:300

R1#show ipv6 route ::/0
IPv6 Routing Table - 11 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
       U - Per-user Static route
       I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary
       O - OSPF intra, OI - OSPF inter, OE1 - OSPF ext 1, OE2 - OSPF ext 2
       ON1 - OSPF NSSA ext 1, ON2 - OSPF NSSA ext 2
       D - EIGRP, EX - EIGRP external
S   ::/0 [1/0]
     via 2001:DB8:12:0:A8BB:CCFF:FE00:300

R1#ping 2001:DB8:23:0:A8BB:CCFF:FE00:300

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 2001:DB8:23:0:A8BB:CCFF:FE00:300, timeout is
 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 20/23/36 ms
```

RIPng (RFC 2080)

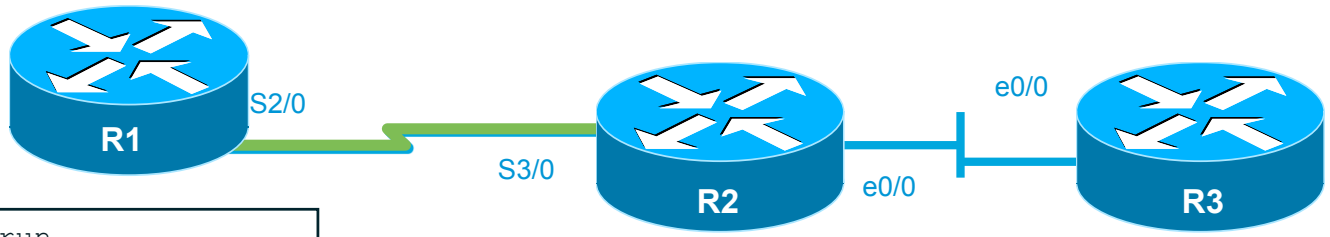
RIPng Overview (RFC 2080)





- Similar characteristics as IPv4
  - Distance-vector, hop limit of 15, split-horizon, multicast based (FF02::9), UDP port (521) etc.
  - Based on RIPv2 (RFC 2453) and extended to handle IPv6
- Updated features for IPv6
  - IPv6 prefix & prefix length
- Special Handling for the NH
  - Route tag and prefix length for NH is all 0. Metric will have 0xFF. NH must be link local.

RIPng – Configuration and Display

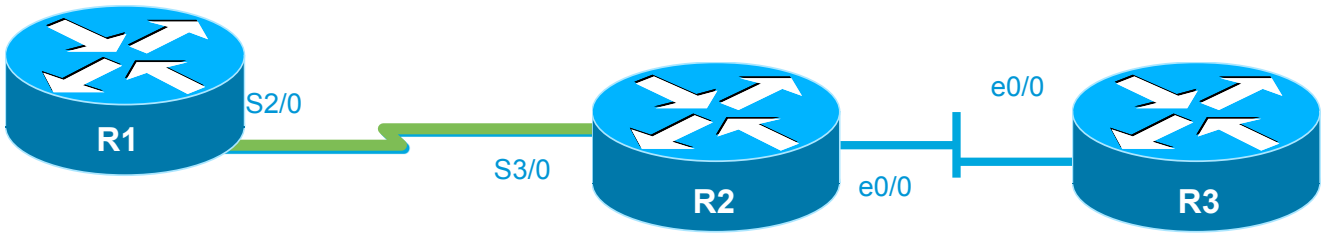



```
R1#show run

!
hostname R1
ipv6 unicast-routing

interface Loopback1
 no ip address
 ipv6 address
 2001:DB8:1::/64 eui-64
 ipv6 rip TEST enable
!
interface Serial2/0
 no ip address
 ipv6 address
 2001:DB8:12::/64 eui-64
 ipv6 rip TEST enable
 serial restart-delay 0
```

```
R2#sh run
!
hostname R2
interface Loopback2
 no ip address
 ipv6 address 2001:DB8:2::/64 eui-64
 ipv6 rip TEST enable
!
interface Ethernet0/0
 no ip address
 ipv6 address 2001:DB8:23::/64 eui-64
 ipv6 rip TEST enable
interface Serial3/0
 no ip address
 ipv6 address 2001:DB8:12::/64 eui-64
 ipv6 rip TEST enable
 serial restart-delay 0
```

```
R3#sh run

!
hostname R3
interface Loopback3
 no ip address
 ipv6 address 2001:DB8:3::/64 eui-64
 ipv6 rip TEST enable
!
interface Ethernet0/0
 no ip address
 ipv6 address 2001:DB8:23::/64 eui-64
 ipv6 rip TEST enable

ipv6 router rip TEST
```

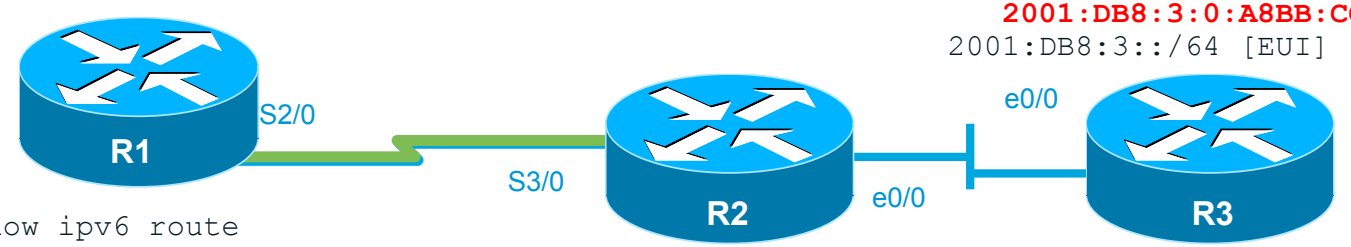
RIPng – ‘debug ipv6 rip’




```
R1#debug ipv6 rip
RIP Routing Protocol debugging is on
R1#
*Oct  1 02:40:10.673: RIPng: Sending multicast update on Serial2/0 for TEST
*Oct  1 02:40:10.673:        src=FE80::A8BB:CCFF:FE00:100
*Oct  1 02:40:10.673:        dst=FF02::9 (Serial2/0)
*Oct  1 02:40:10.673:        sport=521, dport=521, length=52
*Oct  1 02:40:11.985: RIPng: response received from FE80::A8BB:CCFF:FE00:200 on Serial2/0 for TEST
*Oct  1 02:40:11.985:        src=FE80::A8BB:CCFF:FE00:200 (Serial2/0)
*Oct  1 02:40:11.985:        dst=FF02::9
*Oct  1 02:40:11.985:        sport=521, dport=521, length=92
```

RIPng – 'show ipv6 route'

```
Loopback3 is up, line protocol is up
  IPv6 is enabled, link-local address is
 FE80::A8BB:CCFF:FE00:300
  No Virtual link-local address(es):
  Global unicast address(es):
    2001:DB8:3:0:A8BB:CCFF:FE00:300, subnet is
 2001:DB8:3::/64 [EUI]
```




```
R1#show ipv6 route
IPv6 Routing Table - 10 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
       U - Per-user Static route
       I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea, IS - ISIS summary
       O - OSPF intra, OI - OSPF inter, OE1 - OSPF ext 1, OE2 - OSPF ext 2
       ON1 - OSPF NSSA ext 1, ON2 - OSPF NSSA ext 2
       D - EIGRP, EX - EIGRP external
C    2001:DB8:1::/64 [0/0]
     via ::, Loopback1
L    2001:DB8:1:0:A8BB:CCFF:FE00:100/128 [0/0]
     via ::, Loopback1
R    2001:DB8:2::/64 [120/2]
     via FE80::A8BB:CCFF:FE00:200, Serial2/0
R    2001:DB8:3::/64 [120/3]
     via FE80::A8BB:CCFF:FE00:200, Serial2/0
C    2001:DB8:12::/64 [0/0]
     via ::, Serial2/0
L    2001:DB8:12:0:A8BB:CCFF:FE00:100/128 [0/0]
     via ::, Serial2/0
R    2001:DB8:23::/64 [120/2]
     via FE80::A8BB:CCFF:FE00:200, Serial2/0
L    FF00::/8 [0/0]
     via ::, Null0
```

```
R1# ping 2001:DB8:3:0:A8BB:CCFF:FE00:300

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to
 2001:DB8:3:0:A8BB:CCFF:FE00:300, timeout is 2
 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip
 min/avg/max = 20/23/36 ms
```

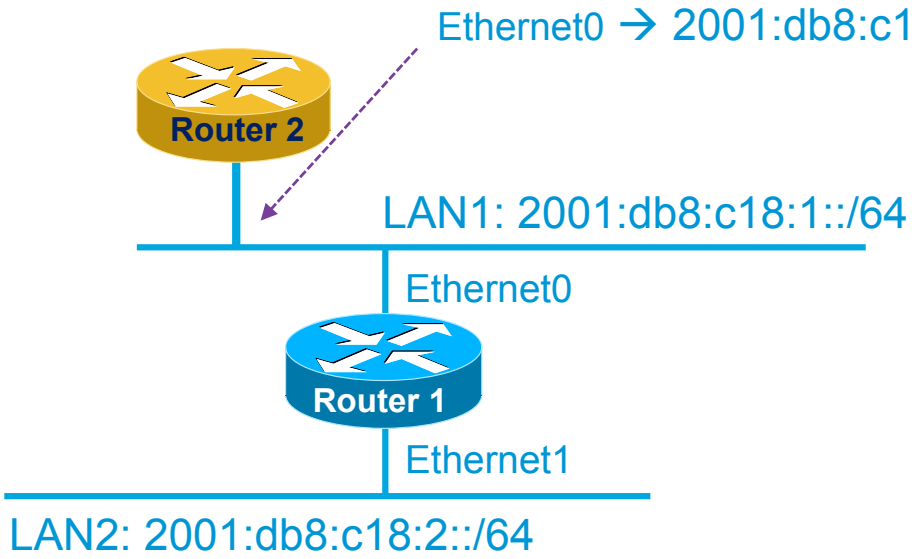
EIGRP for IPv6

EIGRP for IPv6 – Specific Features

Several IPv6 Specific Differences With Respect to IPv4:

- Three new TLVs:
  - IPv6_REQUEST_TYPE (0X0401)
  - IPv6_METRIC_TYPE (0X0402)
  - IPv6_EXTERIOR_TYPE (0X0403)
- Hellos are sourced from the link-local address and destined to FF02::A (all EIGRP routers). This means that neighbors do not have to share the same global prefix (with the exception of explicitly specified neighbors where traffic is unicasted).
- Automatic summarization is disabled by default for IPv6 (unlike IPv4)
- No split-horizon in the case of EIGRP for IPv6 (because IPv6 supports multiple prefixes per interface)
- RID stays 32 bits

Configuration and Display




```
Router2#
ipv6 router eigrp 100
  eigrp router-id 10.10.10.2

interface Ethernet0
 ipv6 address 2001:db8:c18:1::/64 eui-64
 ipv6 enable
 ipv6 eigrp 100
```

```
Router1#show ipv6 eigrp neighbor
IPv6-EIGRP neighbors for process 100
H   Address                 Interface       Hold Uptime(sec)  SRTT(ms)  RTO  Q  Seq Cnt Num
0   FE80::260:3eff:fe47:1530  E0              14 00:01:43       1      4500  0    1
```

Neighbor Identified by Link-Local Address

```
Router1#show ipv6 eigrp topology all-links
IPv6-EIGRP Topology Table for AS(100)/ID(10.10.10.1)
Codes: P - Passive, A - Active, U - Update, Q - Query, R - Reply,
       r - reply Status, s - sia Status
P 2001:db8:c18:1::/64, 1 successors, FD is 28160, serno 1
        via Connected, Ethernet0
        via FE80::260:3eff:fe47:1530 (30720/28160), Ethernet0
```

OSPFv3

OSPFv3 – Changes from OSPFv2

- Per Link Processing
- New Link LSA
- Handling of unknown LSA types
- Addition of flooding scope
- Virtual Link Changes
- Authentication changes

OSPFv3 - Per-Link Processing

- IPv6 uses the term “link” instead of network or subnet to indicate communication
- Interfaces connect to links
- Adjacencies are formed on link local addresses
- Multiple IPv6 subnets can be assigned to a single link
- Two nodes can talk directly over a single link, even if they do not share a common IPv6 subnet
- Network address and mask do not impact the formation of adjacencies

OSPFv3 – Link LSA

- Announces the IPv6 link local address to all the router(s) attached to the link
  - This is needed for the next hop calculation
- Announce a list of IPv6 prefixes associated with the link
  - This is used for a router attached to a LAN to announce its prefix to the DR so DR can include this IPv6 address in its intra-area-prefix-LSA
- Announce the router's options capability router to the DR
  - The DR will then perform an "OR" operation on the options received from all the attached routers
  - The final option field set in the network LSA
- Generated for every link that has two or more routers
- Not be originated for virtual links
- May be suppressed

OSPFv3 – Handling Unknown LSA Types

- Each LSA now contains an “unknown LSA” bit
  - 0: Treat this LSA as a link local
  - 1: Store and flood this LSA even if you don’t understand it
- This allows the deployment of new features in the future
  - Routers that don’t understand the new feature will simply store and forward the LSA
  - Features can be deployed at edges, within a flooding domain, etc., without the need to upgrade all routers

OSPFv3 – Flooding Scope

- Each LSA now contains two bits indicating the flooding scope
  - AS scope, LSA is flooded throughout the AS
  - Area scope, LSA is flooded only within an area
  - Link-local scope, LSA is flooded only on the local link
- These changes also impact the names of the LSAs
  - Type 3 (Summary LSA) is now called the inter-area-prefix-LSA
  - Type 4 (Autonomous System Border LSA) is now called the inter-area-router-LSA
  - Other new LSAs have been added

OSPFv3 – Flooding Scope

LSA Name	LS Type code	Flooding scope	LSA Function code
Router LSA	0x2001	Area scope	1
Network LSA	0x2002	Area scope	2
Inter-Area-Prefix-LSA	0x2003	Area scope	3
Inter-Area-Router-LSA	0x2004	Area scope	4
AS-External-LSA	0x4005	AS scope	5
Group-membership-LSA	0x2006	Area scope	6
Type-7-LSA	0x2007	Area scope	7
Link-LSA	0x0008	Link-local scope	8
Intra-Area-Prefix-LSA	0x2009	Area scope	9

OSPFv3 – Virtual Link Requirements

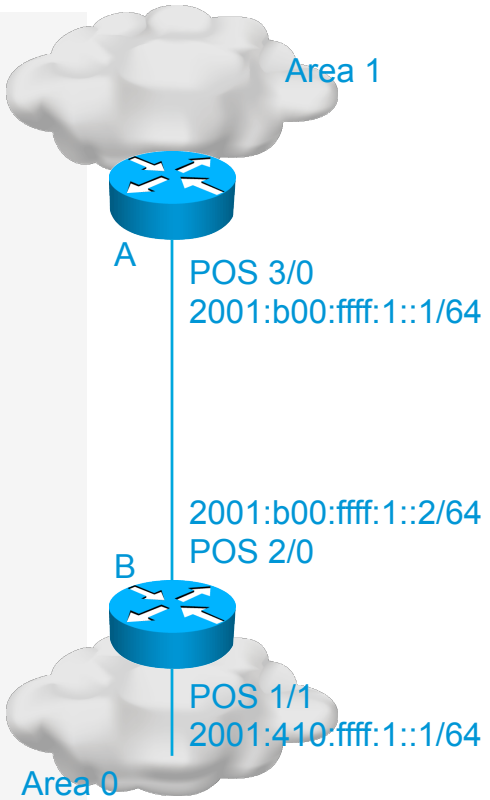
- At least one global/unique local IPv6 address in the transit area
  - OSPFv3 normally sends LSAs with a link local source address
  - This won't work over a virtual link – the packet needs to be forwarded through the intervening area
- Advertisement of a /128 prefix
  - If no /128 is available in the table, a /128 from within an existing prefix space will be used
  - This provides reachability between the endpoints of the virtual link

OSPFv3 - Authentication

- OSPFv3 currently only supports IPsec for authentication
  - Group keying is painful for IPsec
  - There is current work in GDOI and other spaces to make group keying work better for this space
- There is current work in the OSPF working group to allow HMAC-SHA and other forms of “in packet” authentication

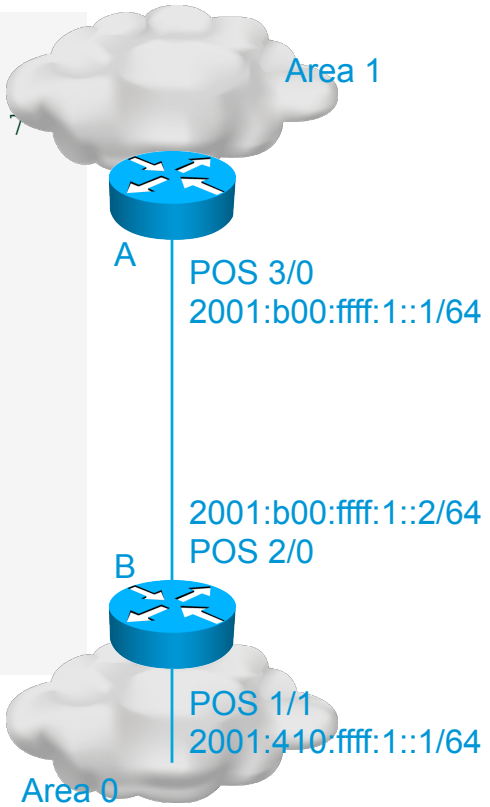
OSPFv3 – Configuration and Display

```
Router-B#
interface POS1/1
 ipv6 address 2001:410:FFFF:1::1/64
 ipv6 enable
 ipv6 ospf 100 area 0

interface POS2/0
 ipv6 address 2001:B00:FFFF:1::2/64
 ipv6 enable
 ipv6 ospf 100 area 1

 ipv6 router ospf 100
   router-id 10.1.1.3

Router-A#
interface POS3/0
 ipv6 address 2001:B00:FFFF:1::1/64
 ipv6 enable
 ipv6 ospf 100 area 1

ipv6 router ospf 100
   router-id 10.1.1.4
```



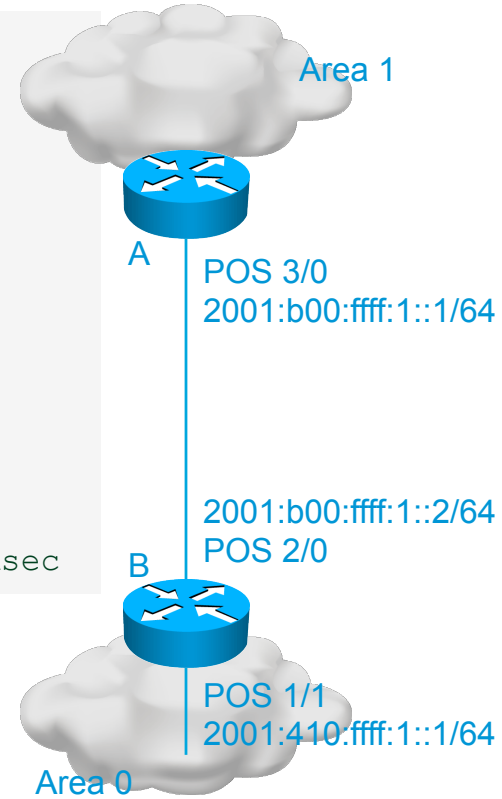
OSPFv3 – Configuration and Display

```
Router-A#sh ipv6 ospf int pos 3/0
POS3/0 is up, line protocol is up
  Link Local Address FE80::290:86FF:FE5D:A000, Interface ID 7
  Area 1, Process ID 100, Instance ID 0, Router ID 10.1.1.4
  Network Type POINT_TO_POINT, Cost: 1
  Transmit Delay is 1 sec, State POINT_TO_POINT,
  Timer intervals configured, Hello 10, Dead 40, Wait 40,
    Retransmit 5
    Hello due in 00:00:02
  Index 1/1/1, flood queue length 0
  Next 0x0(0)/0x0(0)/0x0(0)
  Last flood scan length is 3, maximum is 3
  Last flood scan time is 0 msec, maximum is 0 msec
  Neighbor Count is 1, Adjacent neighbor count is 1
    Adjacent with neighbor 10.1.1.3
  Suppress hello for 0 neighbor(s)
```



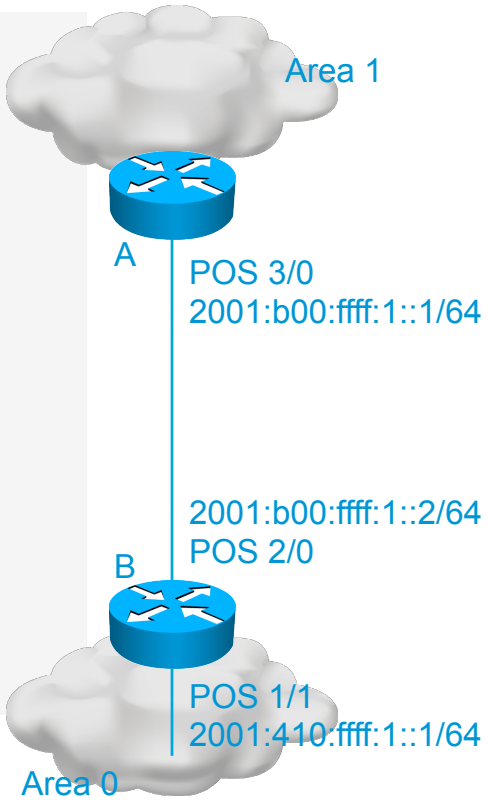
OSPFv3 – Configuration and Display

```
Router-A#sh ipv6 ospf  neighbor detail
 Neighbor 10.1.1.3
    In the area 1 via interface POS3/0
    Neighbor: interface-id 8, link-local address
FE80::2D0:FFFF:FE60:DFFF
    Neighbor priority is 1, State is FULL, 12 state changes
    Options is 0x630C34B9
    Dead timer due in 00:00:33
    Neighbor is up for 00:49:32
    Index 1/1/1, retransmission queue length 0, number of
retransmission 1
    First 0x0(0)/0x0(0)/0x0(0) Next 0x0(0)/0x0(0)/0x0(0)
    Last retransmission scan length is 2, maximum is 2
    Last retransmission scan time is 0 msec, maximum is 0 msec
```



OSPFv3 – Configuration and Display

```
Router-A#sh ipv6 route
IPv6 Routing Table - 5 entries
Codes: C - Connected, L - Local, S - Static, R - RIP,
       B - BGP, U - Per-user Static route
       I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea
       O - OSPF intra, OI - OSPF inter, OE1 - OSPF ext 1,
       OE2 - OSPF ext 2
OI  2001:410:FFFF:1::/64 [110/2]
     via FE80::2D0:FFFF:FE60:DFFF, POS3/0
C   2001:B00:FFFF:1::/64 [0/0]
     via ::, POS3/0
L   2001:B00:FFFF:1::1/128 [0/0]
     via ::, POS3/0
L   FE80::/10 [0/0]
     via ::, Null0
L   FF00::/8 [0/0]
     via ::, Null0
```



OSPFv3 – Routing Consideration

- New Routing Protocol
  - New training, show commands, troubleshooting procedures
  - Must run “dual stack” in the control plane
- Consider design carefully
  - Be intentional about flooding domain boundaries
  - Don’t just deploy “one big area” because you can, it’s simple, it’s a test, etc…
  - Probably best to place ABRs in the same places just to facilitate management and troubleshooting
- IPv6 rollout must be contiguous
  - Just like with the other IGPs…

IS-IS for IPv6

IS-IS/IPv6 – Protocol Changes

- Two TLVs (tag/length/value) added to introduce IPv6 routing
- IPv6 reachability TLV (0xEC)
  - Describes network reachability such as IPv6 routing prefix, metric information and some option bits. The option bits indicates the advertisement of IPv6 prefix from a higher level, redistribution from other routing protocols.
  - Equivalent to IP Internal/External reachability TLVs described in RFC1195
- IPv6 interface address TLV (0xE8)
  - Contains 128-bit address
  - For Hello PDUs, must contain the link-local address (FE80::/10)
  - For LSP, must only contain the non link-local address
- A new Network Layer Protocol Identifier (NLPID) is defined
  - Allowing IS-IS routers with IPv6 support to advertise IPv6 prefix payload using 0x8E value (IPv4 and OSI uses different values)

IS-IS/IPv6 – Single vs. Multi-Topology

- IS-IS supports IPv6 in two ways
- Single Topology
  - The IPv4 and IPv6 topologies must match
  - One SPF is run; IPv4 and IPv6 are mixed on the resulting SPT
- Multi-topology
  - Uses a different address family for IPv6 destinations
  - IPv4 and IPv6 topologies do not need to match

IS-IS/IPv6 – Single Topology

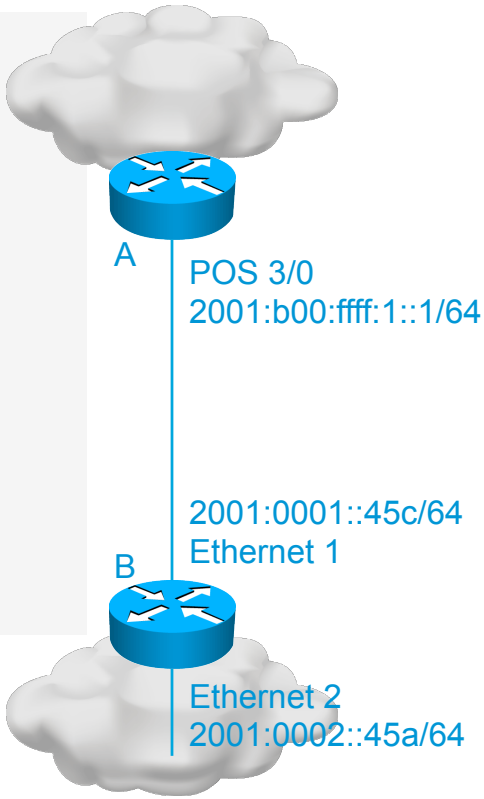
- Uses the same SPF for both IPv4 and IPv6
  - Not really suitable for overlaying pockets of IPv6 on an existing IPv4 network
  - If using both IPv4 and IPv6, topologies must match
  - Cannot run IPv4 on some interfaces, IPv6 on others
- Adjacencies on Level 1 interfaces only form when configuration is matched
- Cannot join two IPv6 areas via an IPv4-only area
  - L2 adjacencies will form OK but IPv6 traffic will black-hole in the IPv4 area.

IS-IS/IPv6 – Multi-Topology

- IPv4 and IPv6 have their own databases
- SPF is run for each topology
  - Once for IPv4, once for IPv6
- Cannot connect “islands” of IPv6 together
  - The problem here is the forwarding plane, not the control plane
  - Not really suitable for overlaying pockets of IPv6 on an existing IPv4 network
- Allows flooding domain boundaries to be in different places
  - More complex to configure and maintain

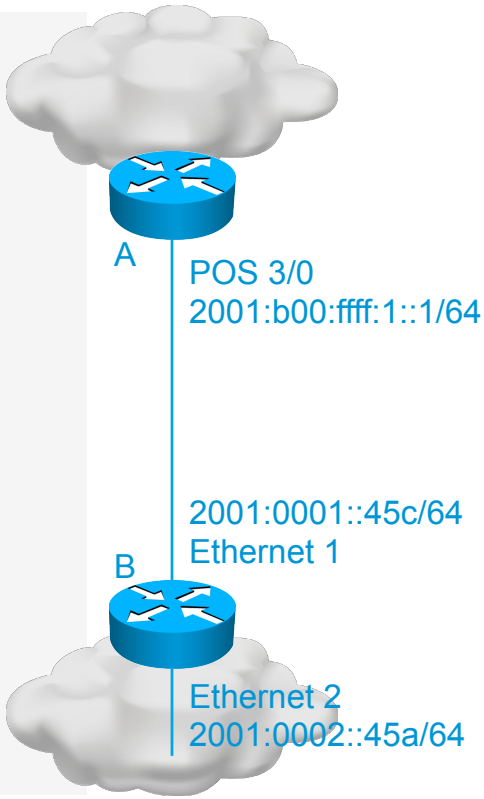
IS-IS/IPv6 – Configuration and Display

```
Router-B#
 interface ethernet-1
     ipv6 address 2001:0001::45c/64
     ipv6 router isis
     isis circuit-type level-2-only

  interface ethernet-2
     ipv6 address 2001:0002::45a/64
     ipv6 router isis

  router isis
      address-family ipv6
      redistribute static
      exit-address-family
      net 42.0001.0000.0000.072c.00
```



IS-IS/IPv6 – Configuration and Display

```
Router-B#
 interface ethernet-1
     ip address 10.1.1.1 255.255.255.0
     ipv6 address 2001:0001::45c/64
     ip router isis
     ipv6 router isis

  interface ethernet-2
     ip address 10.2.1.1 255.255.255.0
     ipv6 address 2001:0002::45a/64
     ip router isis
     ipv6 router isis

  router isis
      address-family ipv6
      redistribute static
      exit-address-family
      net 42.0001.0000.0000.072c.00
      redistribute static
```



- Dual IPv4/IPv6 Configuration
- Redistributing both IPv6 static routes and IPv4 static routes

IS-IS/IPv6 – Configuration and Display

```
brum-45c#show ipv6 route is-is
IPv6 Routing Table - 14 entries
Codes: C - Connected, L - Local, S - Static, R - RIP,
       B - BGP
       I1 - ISIS L1, I2 - ISIS L2, IA - ISIS interarea
Timers: Uptime/Expires

I1  2001:45A:1000::/64 [115/20]
     via FE80::210:7BFF:FEC2:ACCC, Ethernet1, 00:10:12/never
I1  2001:72B:2000::/64 [115/10]
     via FE80::210:7BFF:FEC2:ACCC, Ethernet1, 00:05:19/never
I1  2002:49::/64 [115/10]
     via FE80::210:7BFF:FEC2:ACCC, Ethernet1, 00:05:19/never
```

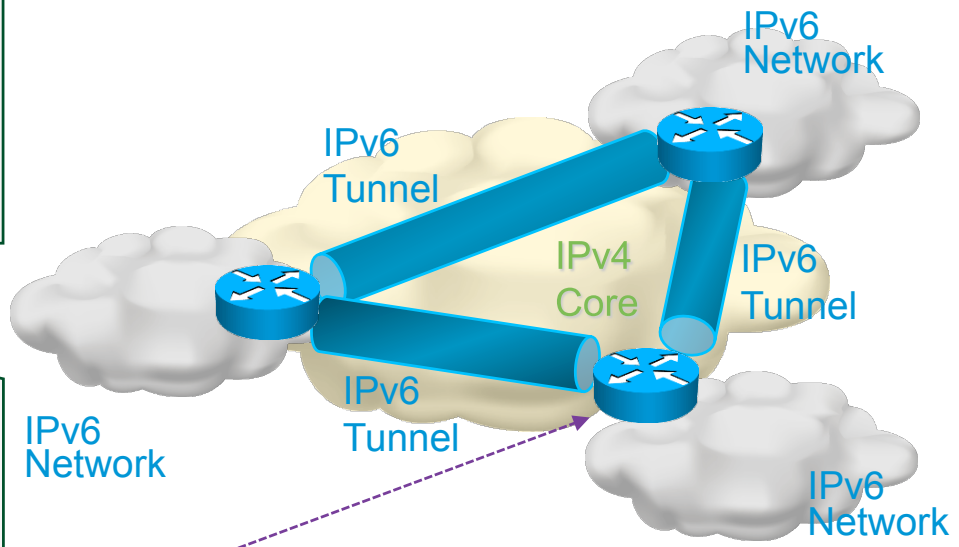
IS-IS/IPv6 – Configuration and Display

```
show clns is-neigh detail

System Id      Interface   State  Type Priority  Circuit Id         Format
brum-45a       Et1         Up     L1   64        brum-45c.01        Phase V
  Area Address(es): 47.0023.0001.0000.0001.0002.0001
  IPv6 Address(es): FE80::210:7BFF:FEC2:ACCC
  Uptime: 00:06:56

IS-IS Level-1 Link State Database:
LSPID                 LSP Seq Num  LSP Checksum  LSP Holdtime      ATT/P/OL
brum-45c.00-00      * 0x00000003   0xA745        732               0/0/0
  Area Address: 47.0023.0001.0000.0001.0002.0001
  NLPID:        0x8E
  Hostname: brum-45c
  IPv6 Address: 3F02::45C
  IPv6 Address: 2001:45C:2000::45C
  Metric: 10         IPv6 2001:45C:1000::/64
  Metric: 10         IPv6 3F02::/64
  Metric: 10         IPv6 2001:45C:2000::/64
  Metric: 10         IS brum-45c.02
  Metric: 10         IS brum-45c.01
brum-45c.01-00      * 0x00000001   0x96DB        733               0/0/0
  Metric: 0          IS brum-45c.00
  Metric: 0          IS brum-45a.00
brum-45a.00-00        0x00000005   0xDDBA        1027              0/0/0
  Area Address: 47.0023.0001.0000.0001.0002.0001
  NLPID:        0x8E
  Hostname: brum-45a
  IPv6 Address: 2001:45A:1000::45A
  Metric: 10         IPv6 2001:45A:1000::/64
  Metric: 10         IS brum-45c.01
  Metric: 0          IPv6-Ext 2001:72B:2000::/64
  Metric: 0          IPv6-Ext 2002:49::/64
```

IS-IS/IPv6 on IPv6 Tunnels over IPv4

```
interface Tunnel0
 no ip address
 ipv6 address 2001:0001::45A/64
 ipv6 address FE80::10:7BC2:ACC9:10 link-local
 ipv6 router isis
 tunnel source Ethernet1
 tunnel destination 10.42.2.1
!
router isis
 passive-interface Ethernet2
 net 42.0001.0000.0000.045a.00
```

```
interface Tunnel0
 no ip address
 ipv6 address 2001:0001::45C/64
 ipv6 address FE80::10:7BC2:B280:11 link-local
 ipv6 router isis
tunnel source Ethernet2
 tunnel destination 10.42.1.1
!
router isis
 net 42.0001.0000.0000.045c.00
```



IS-IS for IPv6 on an IPv6 tunnel requires GRE tunnel, it can't work with IPv6 configured tunnel as IS-IS runs directly over the data link layer

BGP-4 for IPv6

BGP-4 Extensions for IPv6

- TCP Interaction
  - BGP-4 runs on top of TCP
  - This connection could be setup either over IPv4 or IPv6 irrespective of what NLRI BGP is carrying
- Router ID
  - When no IPv4 is configured, an explicit BGP router-id needs to be configured in a 32 bit ipv4 type format.
  - The RID does not have to be in valid IPv4 format (e.g. 0.0.0.1 could be a valid RID)
  - The sole purpose of RID is for identification
  - In BGP, it is used as a tie breaker and is sent within the OPEN message

BGP-4 Extensions for IPv6 (RFC2545)

- BGP-4 carries only 3 pieces of information which are truly IPv4 specific:
  - NLRI in the UPDATE message contains an IPv4 prefix
  - NEXT_HOP path attribute in the UPDATE message contains a IPv4 address
  - BGP Identifier is in the OPEN message & AGGREGATOR attribute
- To make BGP-4 available for other network layer protocols, RFC2858 (obsoletes RFC 2283) defines multi-protocol extensions for BGP-4
  - Enables BGP-4 to carry information of other protocols (e.g MPLS,IPv6)
  - New BGP-4 optional and non-transitive attributes:
    - MP_REACH_NLRI
    - MP_UNREACH_NLRI
  - Protocol independent NEXT_HOP attribute
  - Protocol independent NLRI attribute

BGP-4 Extensions for IPv6

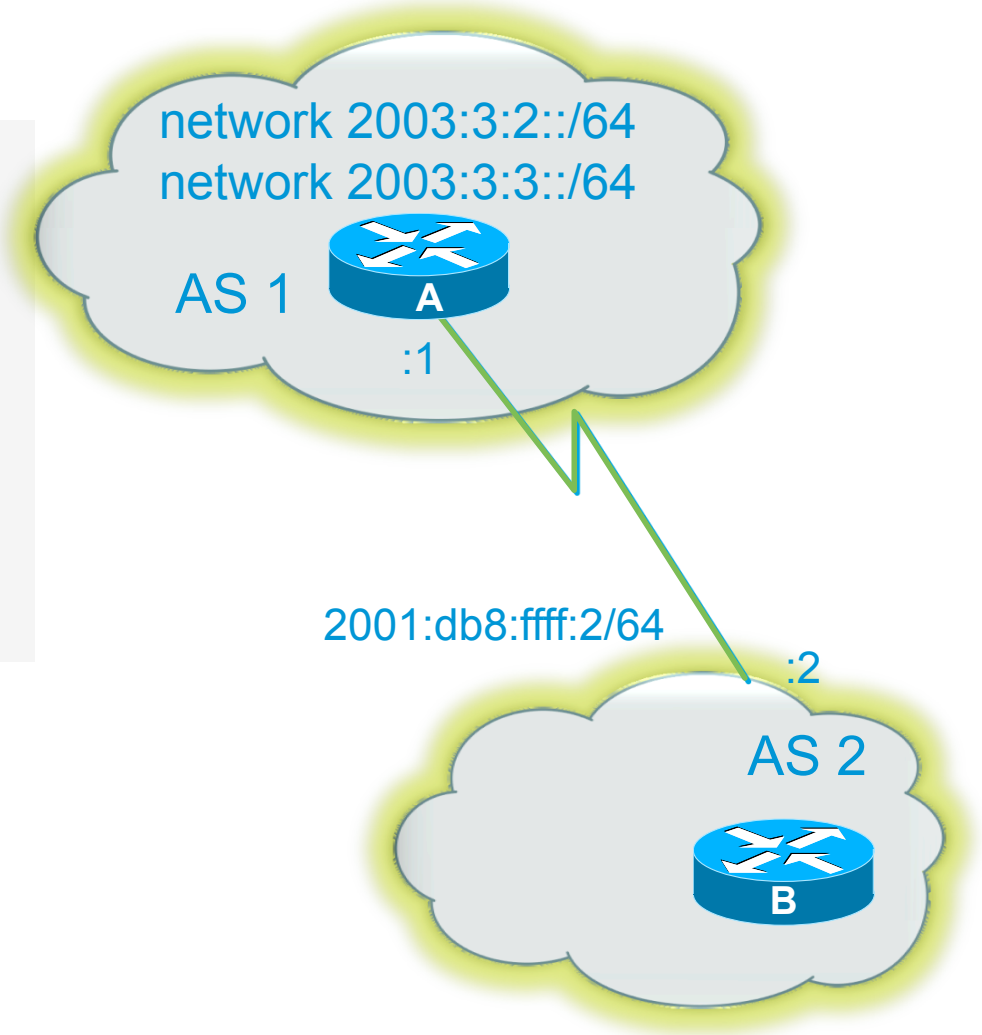
- New optional and non-transitive BGP attributes:
  - MP_REACH_NLRI (Attribute code: 14)
    - “Carry the set of reachable destinations together with the next-hop information to be used for forwarding to these destinations” (RFC2858)
  - MP_UNREACH_NLRI (Attribute code: 15)
    - Carry the set of unreachable destinations
- Attribute 14 and 15 contain one or more triples:
  - Address Family Information (AFI)
  - Next-Hop Information (must be of the same address family)
  - NLRI

BGP-4 Extensions for IPv6

- Address Family Information (AFI) for IPv6
  - AFI = 2 (RFC 1700)
  - Sub-AFI = 1 Unicast
  - Sub-AFI = 2 (Multicast for RPF check)
  - Sub-AFI = 3 for both Unicast and Multicast
  - Sub-AFI = 4 Label
  - Sub-AFI= 128 VPN

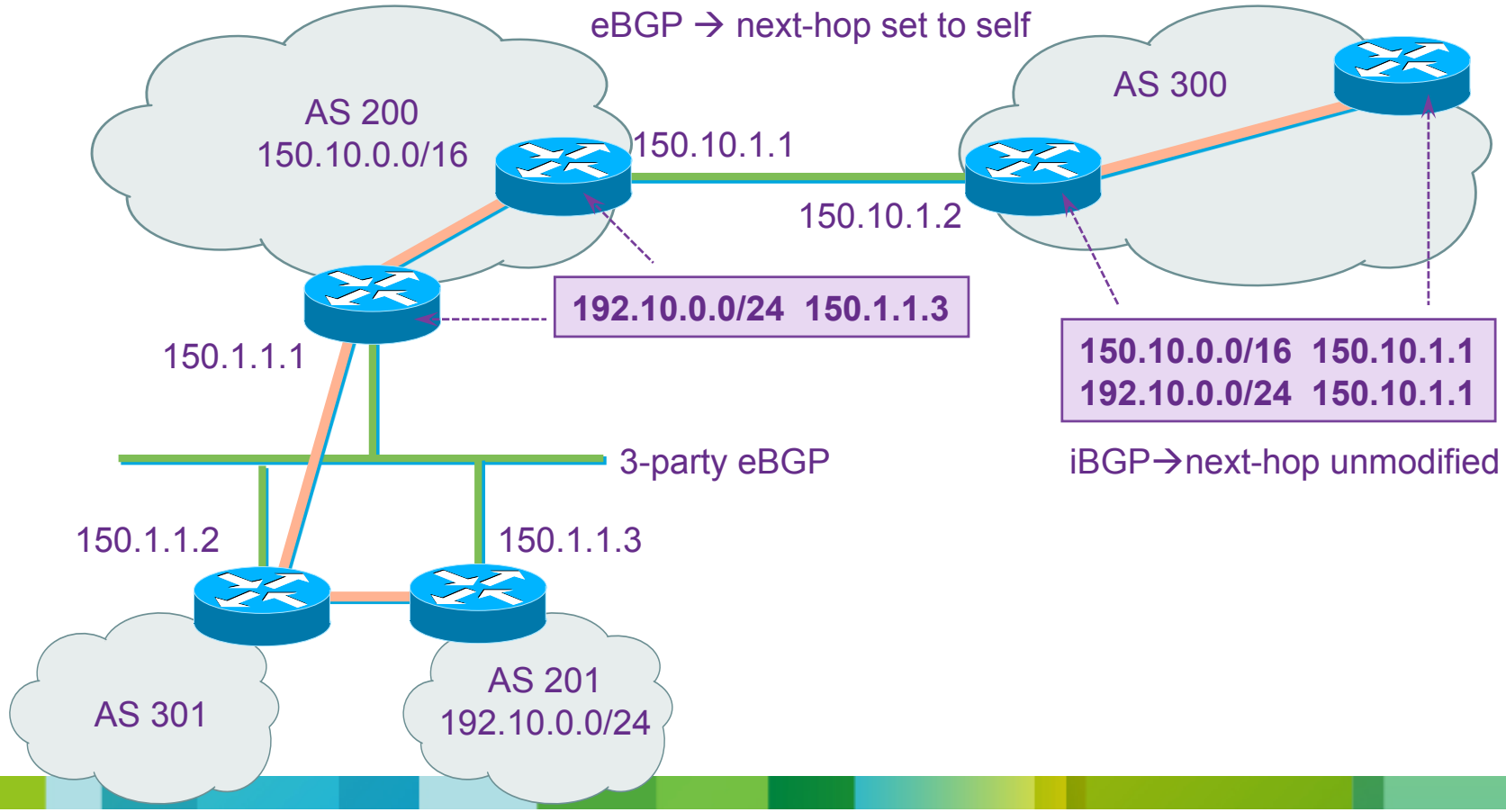
BGP-4 Configurations for IPv6
Non Link Local Peering

```
Router A# show run | b bgp

!
router bgp 1
 no bgp default ipv4 unicast
 bgp router-id 1.1.1.1
 neighbor 2001:db8:ffff:2::2 remote-as 2
 address-family ipv6
  neighbor 2001:db8:ffff:2::2 activate
  network 2003:3:2::/64
  network 2003:3:3::/64
 exit-address-family
!
```



IPv4 NLRI in IPv4

```
Router A# show run | b bgp
!
router bgp 201
 bgp router-id 192.168.30.1
 neighbor 150.1.1.2 remote-as 301
 network 192.10.0.0
!
```

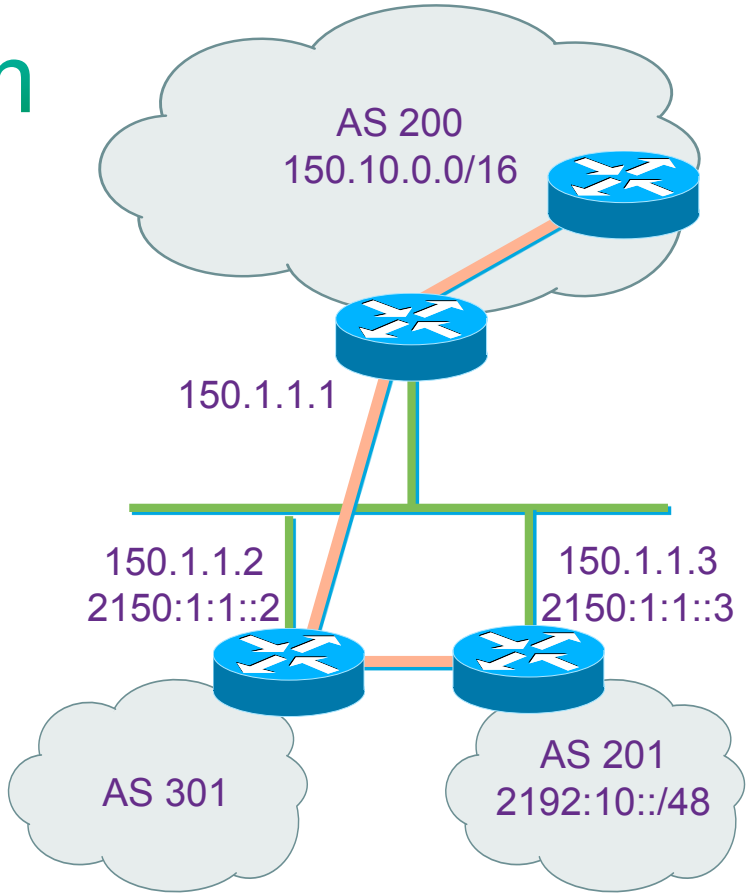


Next-Hop in BGP for IPv6

- Next-hop reachability is very important in BGP
- If the next-hop is inaccessible, the routes learned via BGP will not be installed in the routing table (In some case the routes will be rejected by BGP)
- Link-local address as a next-hop is only set if the BGP peer is also on a link-local address
- IPv6 NLRI in IPv6 (Global Unicast) works like IPv4 (3rd party NH not supported yet)
- Various next-hop behaviour in IPv6

IPv6 NLRI in IPv4 - Problem

```
Router A# show run | b bgp
!
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
!
```

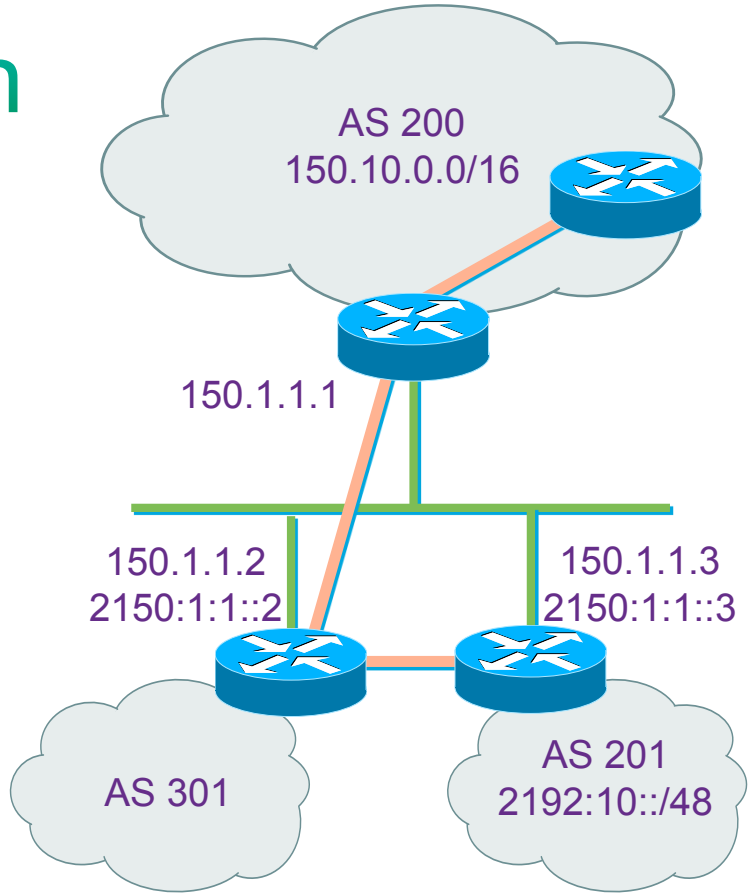


Router-A#
BGP(1): 150.1.1.2 send UPDATE (format) 2192:10::/48, next ::FFFF:150.1.1.3, metric 0, path Local

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# show run | b bgp
!
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
```

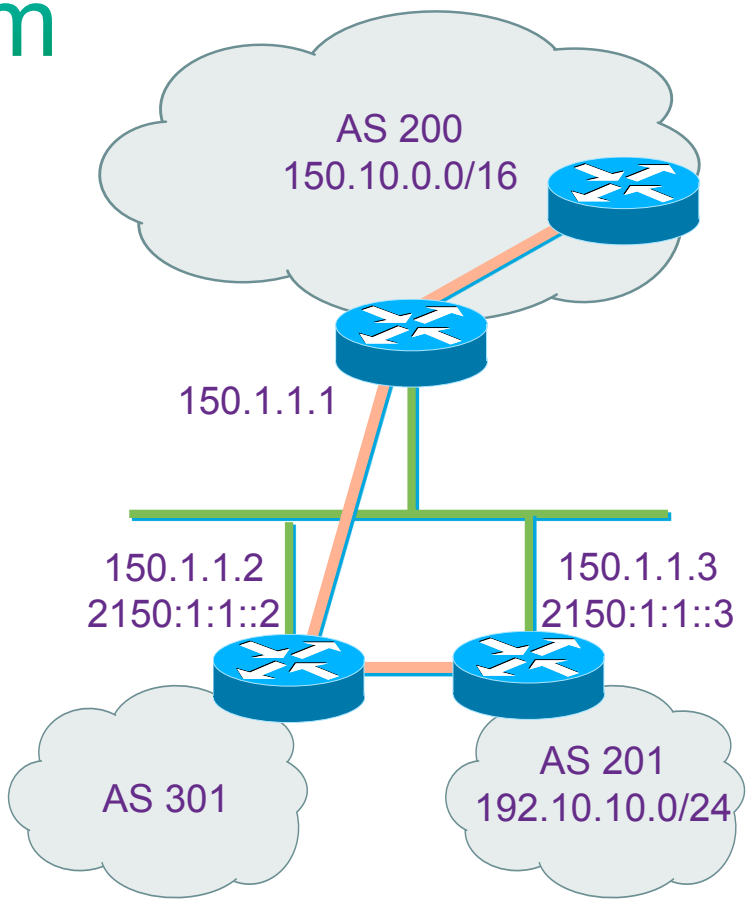


Router-A#
BGP(1): 150.1.1.2 send UPDATE (prepend, chgflags: 0x820) 2192:10::/48, next 2150:1:1::3, metric 0, path Local

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

IPv4 NLRI in IPv6 – Problem
Global Address

```
Router A# show run | b bgp
!
router bgp 201
 bgp router-id 192.168.30.1
 neighbor 2150:1:1::2 remote-as 301
!
 address-family ipv4
 neighbor 2150:1:1::2 activate
 network 192.10.10.0
!
```

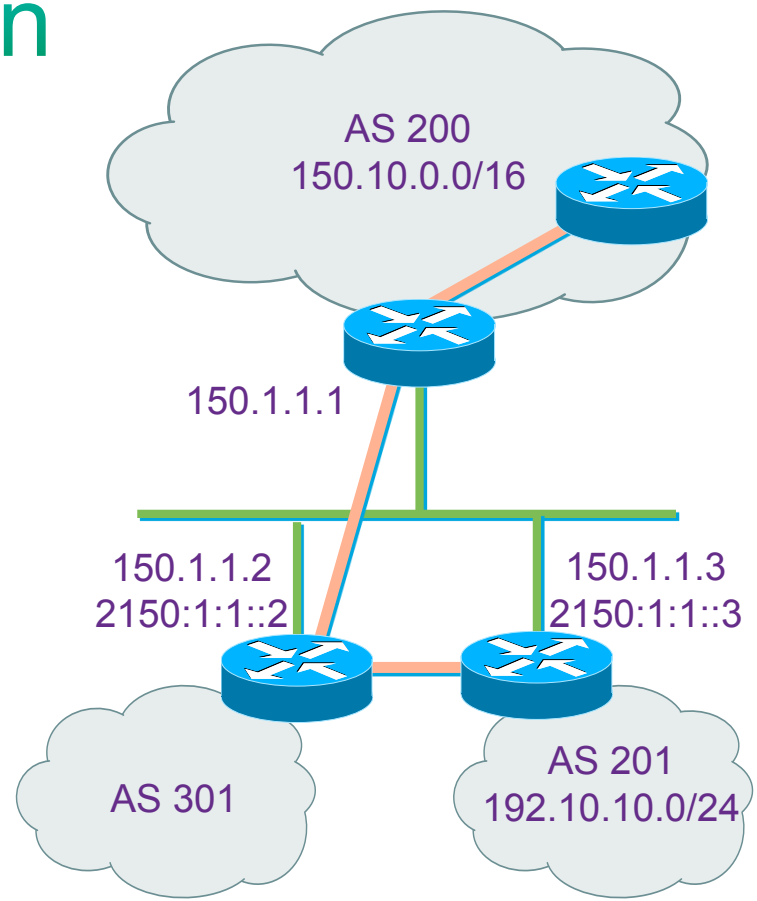


Router-A#
BGP(0): 2150:1:1::2 send UPDATE (format) 192.10.0.0/24, **next 33.80.0.1**, metric 0, path Local

Router-B#
BGP(0): 2150:1:1::3 rcvd UPDATE w/ attr: nexthop 33.80.0.1, origin i, localpref 100, metric 0
BGP(0): 2150:1:1::3 rcvd 192.10.0.0/24
BGP(0): no valid path for 192.10.0.0/24

IPv4 NLRI in IPv6 – Solution
Global Address

```
Router A# show run | b bgp
!
router bgp 201
 bgp router-id 192.168.30.1
 neighbor 2150:1:1::2 remote-as 301
!
 address-family ipv4
 neighbor 2150:1:1::2 activate
 neighbor 2150:1:1::2 route-map SETNH out
 network 192.10.10.0
!
route-map SETHN permit 10
 set ip next-hop 150.1.1.3
```

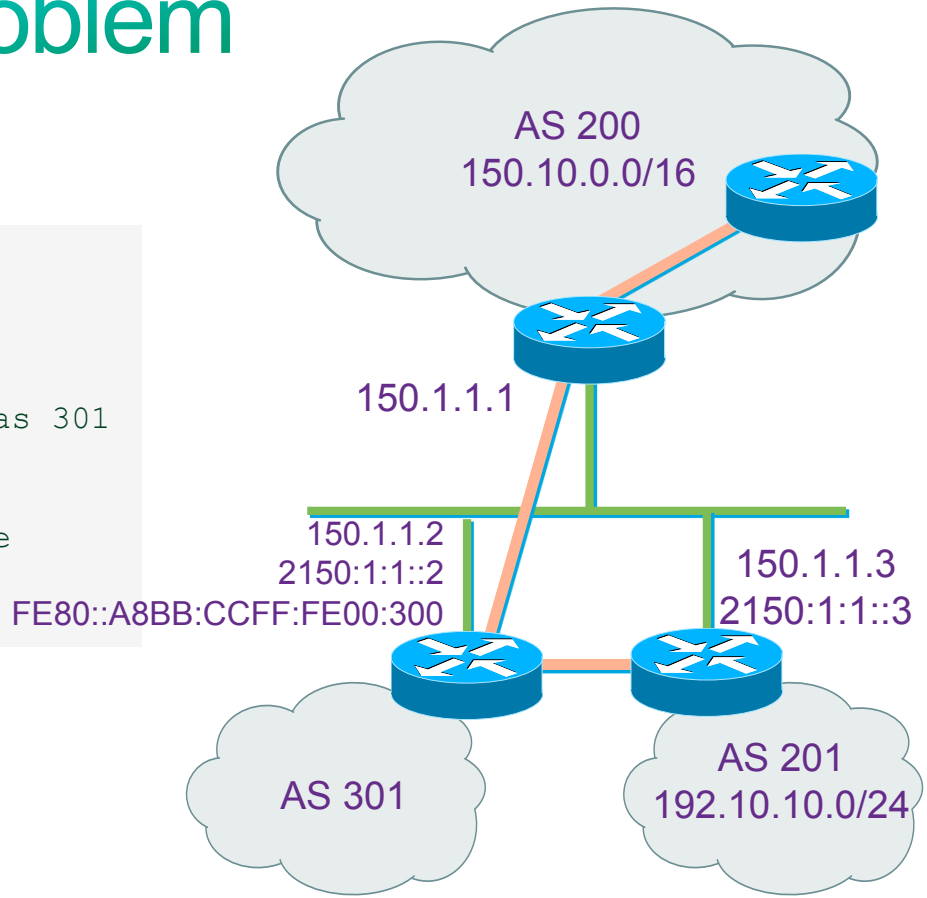


Router-A#
BGP(0): 2150:1:1::2 send UPDATE (prepend, chgflags: 0x0) 192.10.0.0/24, next 150.1.1.3, metric 0, path Local

Router-B#
BGP(0): 2150:1:1::3 rcvd UPDATE w/ attr: nexthop 150.1.1.3, origin i, metric 0, path 10
BGP(0): 2150:1:1::3 rcvd 192.10.0.0/24
BGP(0): Revise route installing 1 of 1 routes for 192.10.0.0/24 -> 150.1.1.3(main) to main IP table

IPv4 NLRI in IPv6 – Problem
Link-Local Address

```
Router A# show run | b bgp
!
router bgp 201
 bgp router-id 192.168.30.1
 neighbor FE80::A8BB:CCFF:FE00:300%E0 remote-as 301
!
 address-family ipv4
 neighbor FE80::A8BB:CCFF:FE00:300%E0 activate
 network 192.10.10.0
!
```

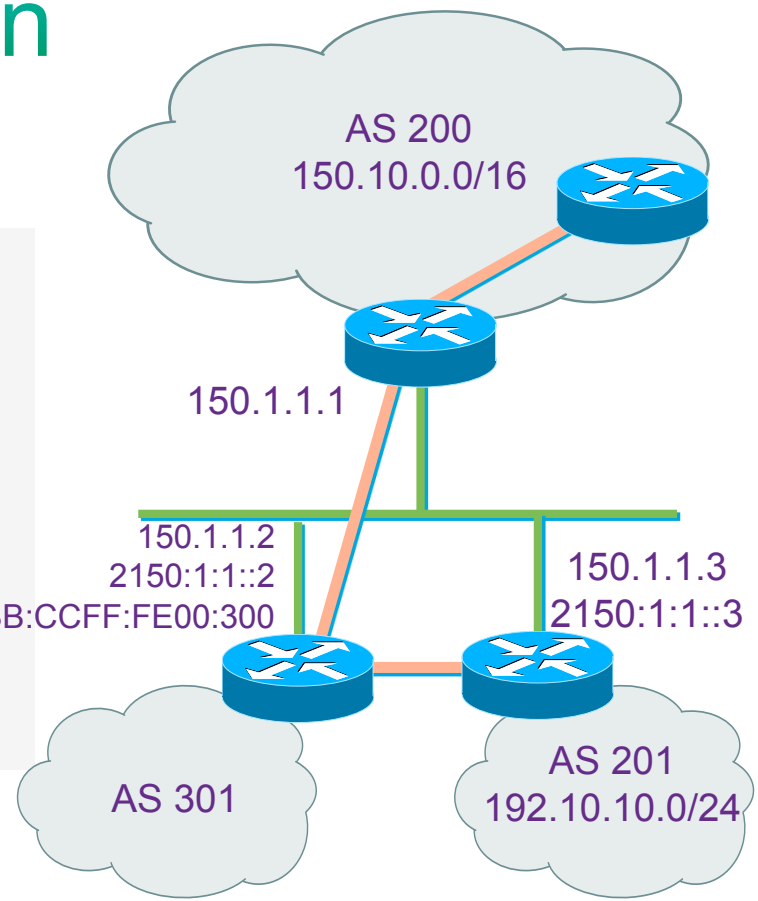


Router-A#
BGP(0): Can't advertise 192.10.0.0/24 to FE80::A8BB:CCFF:FE00:300 with NEXT_HOP 254.128.0.0
BGP(0): FE80::A8BB:CCFF:FE00:300 send UPDATE (format) 192.10.0.0/24, next 254.128.0.0, metric 0, path Local

Router-B#
BGP(0): FE80::A8BB:CCFF:FE00:200 rcv UPDATE w/ attr: nexthop 254.128.0.0, origin i, metric 0, originator 0.0.0.0, path 10, community , extended community
BGP(0): FE80::A8BB:CCFF:FE00:200 rcv UPDATE about 192.10.0.0/24 -- DENIED due to: martian NEXTHOP;

IPv4 NLRI in IPv6 – Solution
Link-Local Address

```
Router A# show run | b bgp
!
router bgp 201
 bgp router-id 192.168.30.1
 neighbor FE80::A8BB:CCFF:FE00:300%E0 remote-as 301
!
 address-family ipv4
 neighbor FE80::A8BB:CCFF:FE00:300%E0 activate
 neighbor FE80::A8BB:CCFF:FE00:300%E0 route-map SETNH
 network 192.10.10.0
!
route-map SETNH permit 10
 set ip next-hop 150.1.1.3
```



Router-A#
BGP(0): FE80::A8BB:CCFF:FE00:300 send UPDATE (format) 192.10.0.0/24, next 150.1.1.2, metric 0, path Local

Router-B#
BGP(0): FE80::A8BB:CCFF:FE00:200 rcvd UPDATE w/ attr: nexthop 150.1.1.3, origin i, metric 0, path 10
BGP(0): FE80::A8BB:CCFF:FE00:200 rcvd 192.10.0.0/24
BGP(0): Revise route installing 1 of 1 routes for 192.10.0.0/24 -> 150.1.1.3(main) to main IP table

IPv6 Routing - Summary

- BGP
  - Single protocol
  - IPv6 behavior is similar as the IPv4 behavior
  - IPv4 routes can be exchanged over an IPv6 TCP session and vice versa
  - There may be two next-hop addresses in the next-hop attribute
- OSPF
  - New protocol (OSPFv3)
  - Lots of changes and capabilities
  - IPv6 domains must be contiguous within the deployment
  - Aggregation/failure domains may not coincide

IPv6 Routing - Summary

- IS-IS/IPv6 Single Topology
  - Single protocol
  - Multiple TLVs within the single protocol
  - Topologies must be congruent
  - IPv6 domains must be contiguous within the deployment
- IS-IS/IPv6 Multi-Topology
  - Single Protocol
  - Since instance, multiple address families
  - Aggregation/failure domains may not coincide
  - IPv6 domains must be contiguous within the deployment

Thank you.