IPv6 Routing Protocols

IPv6 Routing Protocols

- 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 (RFC 5308)
  - OSPFv3 (RFC 5340 Obsolètes RFC 2740)
- EGP : MP-BGP4 (RFC 2858 and RFC 2545)
- Cisco IOS supports all of them
  - Pick one that meets your objectives

RIPng
(RFC 2080)

Enhanced Routing Protocol Support
RIPng Overview RFC 2080

command	version	must be zero
Address Family Identifier		Route Tag
IPv4 Address		
Subnet Mask		
Next Hop		
Metric		

command	version	must be zero
IPv6 prefix		
route tag	prefix len	metric

- Similar characteristics as IPv4

  Distance-vector, hop limit of 15, split-horizon, multicast based (FF02::9), UDP port (521) etc.

- Updated features for IPv6

  IPv6 prefix & prefix len.

- Special Handling for the NH

  Route tag and prefix len for NH is all 0. Metric will have 0xFF. NH must be link local

EIGRP for IPv6

EIGRP for IPv6 Overview

- Just another protocol module (IP, IPX, AppleTalk) with three new TLVs:

  IPv6_REQUEST_TYPE (0X0401)

  IPv6_METRIC_TYPE (0X0402)

  IPv6_EXTERIOR_TYPE (0X0403)

- Other similarities

  Same protocol number 88

  Router ID stays 32 bits (must be configured explicitly if there is no IPv4 interface on the router)

  Uses MD5 like for IPv4
  (IPSec authentication will be available soon)

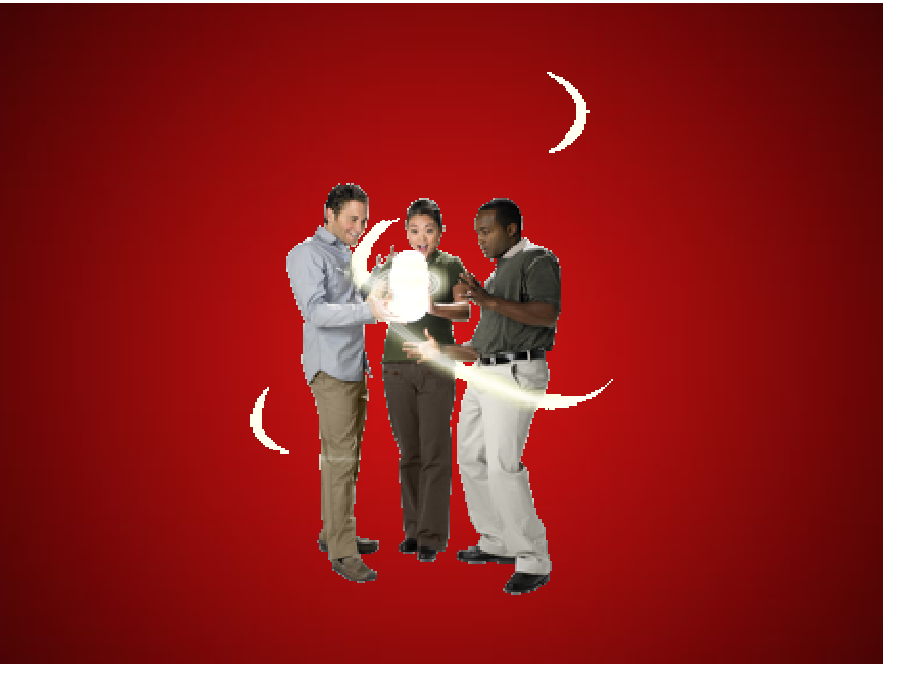
  Same metrics

EIGRP for IPv6 Specific Features

Several IPv6 Specific Differences with Respect to IPv4:

- 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)
- By default EIGRP starts in shutdown mode & needs no shut cmd.

OSPFv3 (RFC 5340) Obsoletes 2740



OSPFv3 and v2 Differences

OSPFv3 and v2 Differences

- Changes made to OSPFv2 to accommodate increased address size of IPv6
- OSPF now runs on per-link, not per-subnet
- Removal of addressing semantics from OSPF packets and LSAs making it network-protocol-independent
- New LSAs has been created to carry IPv6 addresses and prefixes
- Addition of Flooding scope (similar to RFC2370)
- Explicit support for multiple instances per link
- Use of IPv6 link-local addresses for protocol processing and providing next hop information during packet forwarding
- Authentication method changes
- Packet format & LSA's header format changes
- Handling of unknown LSA types

OSPFv3 and v2 Similarities

packet type	Description
1	Hello
2	Database Description
3	Link State Request
4	Link State Update
5	Link State Acknowledgment

- **OSPFv3 has the same 5 packet type but some fields have been changed.**
- **Mechanisms for neighbor discovery and adjacency formation**
- **Interface types**

  P2P, P2MP, Broadcast, NBMA, Virtual
- **LSA flooding and aging**
- **DR, BDR election, area support, SPF**
- **Nearly identical LSA types**

OSPFv3 and v2 header comparison

Version	Type	Packet Length
Router ID		
Area ID		
Checksum	Autype	
Authentication		
Authentication		

Version	Type	Packet Length	
Router ID			
Area ID			
Checksum	Instance ID	0	

- **Size of the header is reduced from 24 bytes to 16**
- **Router ID & Area ID are still a 32 bit numbers, The Router ID of 0.0.0.0 is reserved and SHOULD NOT be used.**
- **Instance ID is a new field that is used to run multiple OSPF protocol instances per link. In order for 2 instances talk to each other they need to have the same instance ID. By default it is 0 and for any additional instance it is increased. Affects the reception of OSPF packets**
- **Two usage of instance ID mentioned in the RFC**

  1) NAP

  2) single link in more than one area
- **Authentication fields have been suppressed – RFC 4552 talks about the authentication implementation in OSPFv3 using IPv6 AH and ESP**

OSPFv3 and v2 hello packet comparison

<table>
<tr><td colspan="3">Network Mask</td></tr>
<tr><td>HelloInterval</td><td>Options</td><td>Rtr Pri</td></tr>
<tr><td colspan="3">RouterDeadInterval</td></tr>
<tr><td colspan="3">Designated Router</td></tr>
<tr><td colspan="3">Backup Designated Router</td></tr>
<tr><td colspan="3">Neighbor ID</td></tr>
</table>

<table>
<tr><td colspan="2">Interface ID</td></tr>
<tr><td>Rtr Pri</td><td>Options</td></tr>
<tr><td>HelloInterval</td><td>RouterDeadInterval</td></tr>
<tr><td colspan="2">Designated Router</td></tr>
<tr><td colspan="2">Backup Designated Router</td></tr>
<tr><td colspan="2">Neighbor ID</td></tr>
</table>

- **Network Mask field has been removed**
- **Interface ID: a 32-bit number the originating router assigned to uniquely identify (among its own interfaces) its interface to the link. Virtual link gets its own interface ID**
- **Router Priority – Change in location**
- **Option field has been increased to 24-bit from 8-bits**
- **Hello Interval – Only change in location**
- **Dead interval have been reduced to 16 bits from 32**
- **DR and BDR are still 32-bit field and contain the Router ID of DR /BDR instead of IP address. Router ID along with the Link ID uniquely identify the DR on an interface**

OSPFv3 and v2 DBD packet comparison

Interface MTU	Options	00000IMMS
DD Sequence number		
LSA Header		

0	Options	
Interface MTU	0	00000IMMS
DD Sequence number		
LSA Header		

- DBD packets are exchanged when an adjacency is being initialized.
- It describe the contents of the link-state database
- The DD packet has been increased by 4 byte compared to OSPFv2
- The Option field is now 24-bit
- All the other field are the same and behave the same way as for OSPFv2

OSPFv3 and v2 Link State Request Packet comparison

LS Type
Link State ID
Advertising Router

0	LS type
Link State ID	
Advertising Router	

- **Every LSA is uniquely identified by { LS type, Link State ID, Advertising router }**
- **OSPFv3 has the same field as OSPFv2. Note that LS Type field is now 2 bytes.**
- **LS type has now different coding compare to OSPFv2.**
- **There are two bits within LS Type which indicates the flooding scope. (later on flooding scope)**

OSPFv3 and v2 LS Update packet comparison

LSAs
LSA (header + body)

LSAs
LSA (header + body)

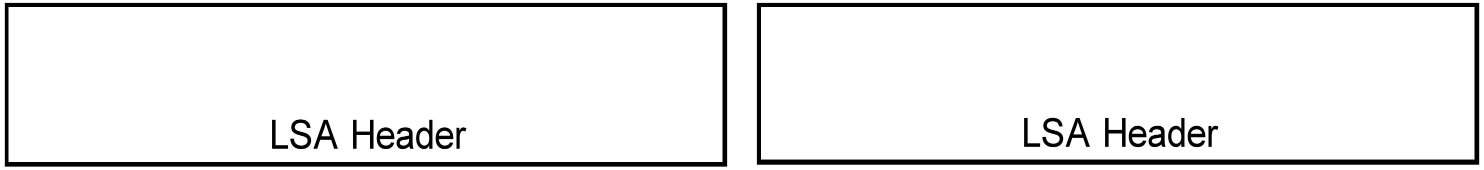
- **Both in IPv4 and IPv6, the steps for sending a Link State Update packet are the same**
- **However the list of eligible interfaces out which to flood the LSA is different**
- **For IPv6, the eligible interfaces are selected based on the following factors:**

  The LSA's flooding scope (will talk more later)

  Whether the LSA has a recognized LS type.

  The setting of the U-bit in the LS type if the LSA is not recognized

OSPFv3 & v2 LS Acknowledgment packet comparison



- **Each newly received LSA must be acknowledged.**
- **This is usually done by sending Link State Acknowledgment packets.**
- **Acknowledgments can also be accomplished implicitly by sending Link State Update packets**
- **No changes from OSPFv2**

OSPFv3 & v2 option comparison

7							0
*	O	DC	EA	N/P	MC	E	*

							15								7							0	
																	DC	R	N	X	E	V6	

- **In OSPFv3 option field has been increased from 8 to 24-bits and moved to the body of certain LSA (details later)**
- **The option field in Hello and DD packet has been also increased to 24-bit.**
- **Unused bits have been suppressed and two new bits have been introduced, R and V6 bit. V6 bit specializes the R bit.**
- **MC bit has been reclaimed in the latest RFC**

R	V6	Router behavior
0	0	Can not be a trasit
0	1	Ambigous
1	0	Transit for non-V6 only
1	1	Normal

OSPFv3 LSA Details

OSPFv3 & v2 LSA Header Comparison

LS age	Options	LS type
Link State ID		
Advertising Router		
LS sequence numer		
LS checksum	Length	

LS age	LS type
Link State ID	
Advertising Router	
LS sequence numer	
LS checksum	Length

- **Option field is removed from the header to the body of the LSA**
- **LS type field in the LSA header has increased from 1 byte to 2 bytes. It indicates the function performed by the LSA (more on next slide)**
- **The behavior of assigning Link State ID value has changed from v4 to v6, we will talk about the change of behavior as we go to each of the LSA**
- **Advertising router still contain the RID of the LSA originator**

OSPFv3 Flooding Scope

LS age	Options	LS type

LS age	U	S2	S1	LSA Function Code

- **The high-order three bits of LS type {1 bit (U) for handling unrecognized LSA and two bits (S2, S1) for flooding scope} encode generic properties of the LSA, while the remainder, (called LSA function code) indicate the LSA's specific functionality**
- **OSPFv2 had two flooding scope, AS wide and area wide**
- **OSPFv3 has three 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.

OSPFv3 Flooding Scope

- U (unrecognized) bit is used to indicate a router how to handle an LSA if it is unrecognized

U-bit	LSA Handling
0	Treat this LSA as if it has link-local Scope
1	Store and flood this LSA as if type understood

- S2 / S1 bit indicates the three flooding scopes

S2	S1	Flooding scope
0	0	Link-Local flooding scope
0	1	Area flodding scope
1	0	AS flooding scope
1	1	Reserved

- Unrecognized LS type with flooding scope set to link local or area local can be flooded into stub area or NSSA with U bit set to 1.

OSPFv3 LSA Types

- Here is the list of LSA in OSPFv3

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 Router LSA

OSPFv3 and v2 Router LSA comparison

- **Function Code 0x2001 – Area wide flooding scope**
- **Link State ID has shed addressing semantics.**
- **Link State ID for OSPFv2 in case of the router LSA is the router ID of the originating router. Since router could only originate a single router LSA in OSPFv2, fragmentation was up to the IP process so the link state ID was set to a single value**
- **In OSPFv3, a single router MAY originate one or more Router LSAs, distinguished by their Link-State IDs (which are chosen arbitrarily by the originating router).**

OSPFv3 and v2 Router LSA comparison

00000VEB	0	# Links
Link ID		
Link Data		
Type	# TOS	metric
:::::::::::		
TOS	0	TOS metric

000NtxVEB	Options	
Type	0	Metric
Interface ID		
Neighbor Interface ID		
Neighbor Router ID		

- **New bit W (removed in latest specs)**

bit Nt: This is from RFC 3101, when set, the router is an NSSA ABR

bit x: This used to be W bit (wild-card multicast receiver) for MOSPF and has been reclaimed in the RFC 5340

Options: we talked about earlier. Currently V6, E, R, DC, N bits. V6 and R bit should always be set normally(1 1)

The Options field and Nt, V, E and B bits should be the same in all Router LSAs from a single originator. In case of a mismatch the values in the LSA with the lowest Link State ID take precedence.

Router interface information may be spread across multiple Router LSAs. Receivers must concatenate all the Router-LSAs originated by a given router when running the SPF calculation.

Next five fields is what you call the link description fields, taking 16 bytes. These include Type, Metric, Interface ID, Neighbor Interface ID and Neighbor Router ID

OSPFv3 Network LSA

OSPFv3 and vs Network LSA comparison

Network Mask
Attached Router
:::::::::::::

0	Options
Attached Router	
:::::::::::::	

- **Function code 0x2002 – Area Wide flooding scope**
- **In OSPFv2, Link State ID in the header contained DR IP address of the Transit link and the mask was inside the type 2 LSA. This combination was used to install the IP address of the transit link.**
- **In OSPFv3 IP address are carried in intra-area-prefix-LSA (FC 9) therefore the mask field has been removed from network LSA. Also, Link State ID in the LSA header contain DR's Interface ID**

OSPFv3 Link LSA

OSPFv3 (Link LSA Details) - NEW

- Function Code of 0x0008 - Link local flooding scope.
- This is a new LSA in OSPFv3 and it is used to advertise one or more IPv6 prefixes on a given link
- Link LSA's are generated for every link that has 2 or more routers
- Note, in OSPFv2 link address information (routers' interface ip address) was carried in Router LSA's "Link Data" field. This information was used for the next hop calculation
- Link LSA MUST not be originated for Virtual links.
- LinkLSASupression (added in the latest specs)

OSPFv3 (Link LSA Details)

- Link LSA has the following three purposes:

  I- Announce the IPv6 link local address to all router(s) attached to the link. This is needed for the Next Hop calculation

  II- 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. Duplicates will be ignored

- III- Announce the options capability of a given router to the DR. The DR will then perform an "OR" operation on the options received from all the attached routers. This will be the final option field set in the Network LSA

OSPFv3 (Link LSA Details)

Rtr Pri	Options	
Link-local Interface address (128-bit)		
# prefixes		
PrefixLength	PrefixOptions	0
Adress prefix (128-bit)		

- **Rtr Pri is The Router Priority of the interface attaching the originating router to the link.**
- **# prefixes is the number of prefix advertised**
- **Link-local interface address is used for next hop calculation.**
- **Link-LSA also advertise a list of IPv6 prefixes identified by {Address prefix, PrefixLength, PrefixOptions} to other attached router. For example a DR will include this list of IPv6 prefix advertised by a router in its intra-area-prefix-LSA**
- **Link State ID in the header of the Link-LSA is set to router's Interface ID on the link therefore, size of this LSA is not limited.**

OSPFv3 Intra-area-prefix LSA

OSPFv3 (Intra-area-prefix LSA Details) - NEW

- **Function code 0x2009 – Area wide scope**
- **This is a new LSA in OSPFv3 to advertise one or more IPv6 prefixes. The prefixes are associated with router segments(p2p), stub network segment or transit network segment.**
- **Intra-area prefix LSA is ONLY generated by the DR for all the prefixe(s) on a broadcast segment**
- **In OSPFv2 the intra area prefix information was carried in the Router and Network LSA's**

OSPFv3 (Intra-area-prefix LSA Details)

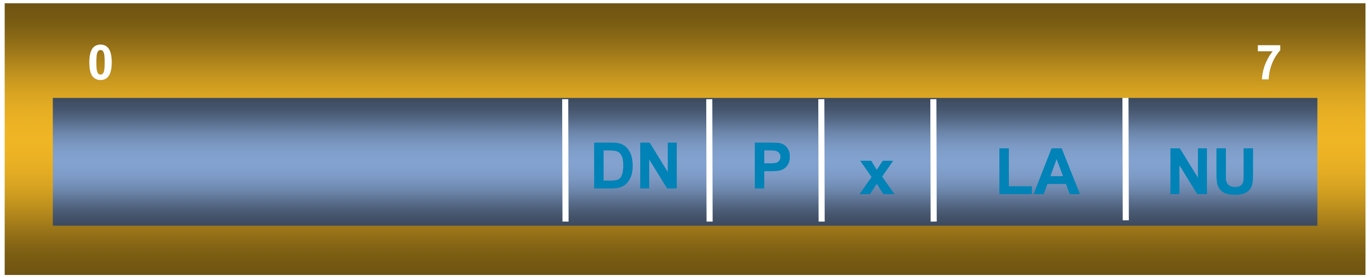
<table>
<tr><td colspan="2"># Prefixes</td><td colspan="2">Referenced LS type</td></tr>
<tr><td colspan="4">Referenced Link State ID</td></tr>
<tr><td colspan="4">Referenced Advertsing Router</td></tr>
<tr><td>PrefixLength</td><td>PrefixOptions</td><td colspan="2">Metric</td></tr>
<tr><td colspan="4">Address Prefix (128-bit)</td></tr>
</table>

- **# prefixes is the number of prefixes advertised**
- **Each IPv6 address is associate with { Address prefix, PrefixLength, PrefixOptions }**
- **The three field { Referenced LS type, Referencd Link State ID, Referenced Advertising Router } identifies the Router LSA or Network LSA that the Intra-Area-Prefix-LSA should be associated with.**

OSPFv3 (Intra-area-prefix LSA Details)

	Referenced LSA Type	Referenced LS ID	Referenced Advertising Router
Router LSA	2001	0	Originating router’s RID
Network LSA	2002	Interface ID of the DR	DR’s RID

OSPFv3 (Intra-area-prefix LSA Details)



- Prefix Options in Intra-area-prefix-LSA is an 8 bit field serve as input to the various routing calculations
- NU-bit: The "no unicast" capability bit. If set, the prefix should be excluded from IPv6 unicast calculations, otherwise it should be included.
- LA: "local address" capability bit. If set, the /128 prefix is actually an IPv6 interface address of the advertising router
- x: the "multicast" capability bit. MOSPF has be deprecated in the latest spec
- P: The "propagate" bit. Set on an NSSA area prefixes then needs be re-advertised at the NSSA area border as an external LSA
- DN: The down bit, used for MPLS VPN loop prevention

OSPFv3 Inter-area-prefix LSA

OSPFv3 (Inter-area-prefix LSA Details)

Network Mask	
0	metric
TOS	TOS metric

0	metric	
PrefixLength	PrefixOptions	0
Adress Prefix 128-bit		

- **Function code 0x2003 – Area wide flooding scope**
- **Inter-Area-Prefix-LSA announce destinations outside of the area (type 3 in OSPFv2)**
- **All TOS field have been suppressed**
- **In OSPFv2 Link State ID in the LSA header contain IP destination out side of the area and the mask is in the body of the LSA**
- **In OSPFv3 Link State ID is just a fragment number and the prefix is moved into the body of the LSA**
- **All Prefix in OSPFv3 is defined by 3 fields {Address Prefix, PrefixLength, PrefixOptions}**

OSPFv3 Inter-area-router LSA

OSPFv3 (Inter-Area-Router-LSA Details)

Network Mask	
0	metric
TOS	TOS metric

0	Options
0	Metric
Destination Router ID	

- **Function code 0x2004 – Area wide flooding scope**
- **Inter-Area-Router-LSA announce the location of ASBR (type 4 in OSPFv2)**
- **In OSPFv2 the mask field is not used for type 4 and contains zero so suppressed in OSPFv3**
- **All TOS related fields are suppressed**
- **In OSPFv2 Link State ID in the header contain the Router ID of the ASBR. In OSPFv3 Link State ID is just a fragment number and ASBR Router ID is inside the body of LSA.**
- **Note, Cisco implementation converts the ASBR RID in decimal format and put it under Link State ID (This may change in future)**

OSPFv3 External LSA

OSPFv3 & v2 External LSA Difference

Network Mask	
E0000000	Metric
Forwarding Adress	
External Route Tag	
E| TOS	TOS metric

0 0 0 0 0 E FT	metric	
PrefixLength	PrefixOptions	Referenced LS Type(opt)
Adress Prefix 128-bit		
Forwarding address (optional) 128-bit		
External Route Tag (optional)		
Referenced Link State ID (optonal)		

- **Function Type code 0x4005 – AS wide flooding scope**
- **The Link State ID of an AS-external-LSA has lost all of its addressing semantics, it is used just to distinguish between multiple external LSA originated by the same ASBR**
- **There are two new fields in OSPFv3, Referenced LS type and Referenced Link State ID. The Referenced Link State ID field would present ONLY if Referenced LS Type is non-zero**
- **If a router advertising an As-External-LSA wants to announce additional information regarding external route that is not used by OSPF itself (for example BGP external route attribute) it sets Referenced LS type and Referenced Link State ID in order to announce additional information.**
- **Fwd address is now optional – must NOT be set to :: or link-local (currently no fwd address implementation done in IOS). The forwarding address is present in the AS-external-LSA if and only if the AS-external-LSA's bit F is set.**
- **The external route tag is present in the AS-external-LSA if and only if the AS-external-LSA's bit T is set**
- **The prefix is described by the PrefixLength, PrefixOptions and Address Prefix fields embedded within the LSA body. Network Mask is no longer specified**

OSPFv3 LSAs summary

- We need router and network LSA for the topology information irrespective of what address family OSPFv3 is being used for. Stub link are excluded from Router LSA because 1. it does not have any topology info 2. To carry just the prefix info would increase the size of the router LSA since it will have to carry 128 bits address
- We need Link LSA for the 1. NH calculation (link local). 2. Prefix advertisement on broadcast segments so every router won't have to generate separate intra area prefix LSA 3. Optional capabilities awareness to the DR. Link LSA could be suppressed on p2p link in the latest RFC
- We need intra area prefix lsa for prefix propagation area wide. Link LSA can not go area wide. Stub link prefix, p2p & transit prefixes will be generated by this LSA. On broadcast segments, ONLY DR will generate this LSA including ALL prefixes sent to the DR through link LSA. P2P link prefix needs to be generated by this LSA so the info about the prefix could be propagated within an area.
- Inter-area prefix is a summary LSA just like Type 3 in OSPFv2
- Inter-area router is a info about the ASBR just like Type 4 in OSPFv2
- External and NSSA is also the same as OSPFv2. Main difference is the FWD address & Tags are now optional.

OSPFv3 Fast Convergence

- Following Techniques/tools are available for fast convergence in OSPFv3
  - Carrier Delays **Detect**
  - Hello/dead timers (Fast Hellos) **Detect**
  - Bi-Directional Forwarding Detection—(BFD) **Detect**
  - LSA packet pacing **Propagate**
  - Interface event dampening - **Propagate**
  - Exponential throttle timers for LSA & SPF **Process**
  - MinLSArrival Interval **Process**
  - Incremental SPF **Process**
- Techniques/tools for Resiliency
  - Stub router (e.g., max-metric)
  - Cisco NSF (RFC 4811,4812,4813)
  - Graceful Restart (ONLY RFC 3623)

ISIS for IPv6

Integrated IS-IS for IPv6

- 2 Tag/Length/Values 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 TLV's 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 & OSI uses different values)

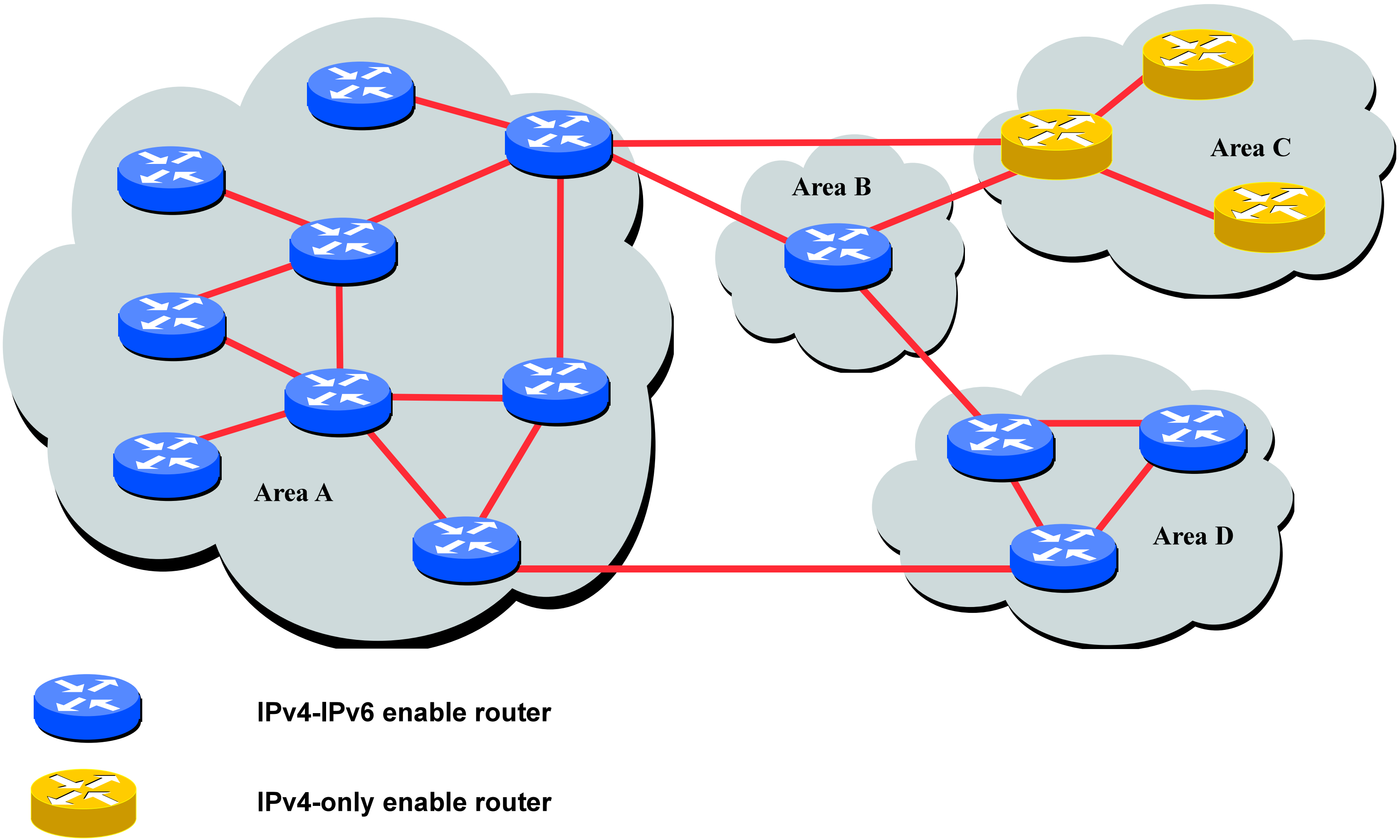
Single SPF IS-IS for IPv6 Restrictions

- IS-IS for IPv6 uses the same SPF for both IPv4 and IPv6.

Therefore:

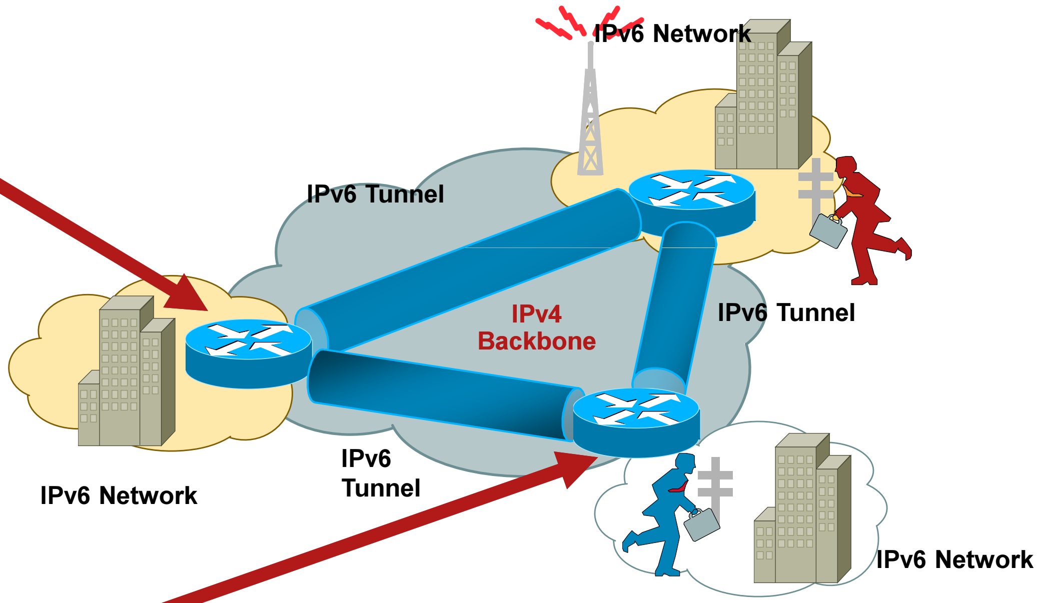
- Not really suitable for an existing IPv4 IS-IS network where customer wants to turn on scattered IPv6 support.
- If using IS-IS for both IPv4 and IPv6 then the IPv4 and IPv6 topologies MUST match exactly. Cannot run IS-IS IPv6 on some interfaces, IS-IS IPv4 on others.
- Will only form adjacencies with similarly-configured routers. E.g. An IS-IS IPv4-only router will not form an adjacency with an IS-IS IPv4/IPv6 router. (Exception is over L2-only interface)
- 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 Hierarchy & IPv6 Example



IS-IS for 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

Multi-Topology IS-IS
(RFC 5120)

Introduction

- Mechanism that allows IS-IS, used within a single domain, to maintain a set of independent IP topologies
- Multi-Topologies extension can be used to maintain separate topologies for:
  - IPv4
  - IPv6
  - Multicast
- Topologies need not to be congruent (of course)
- Multiple topologies for same address family is allowed
  - Think about QBR…
  - The multicast dimension …

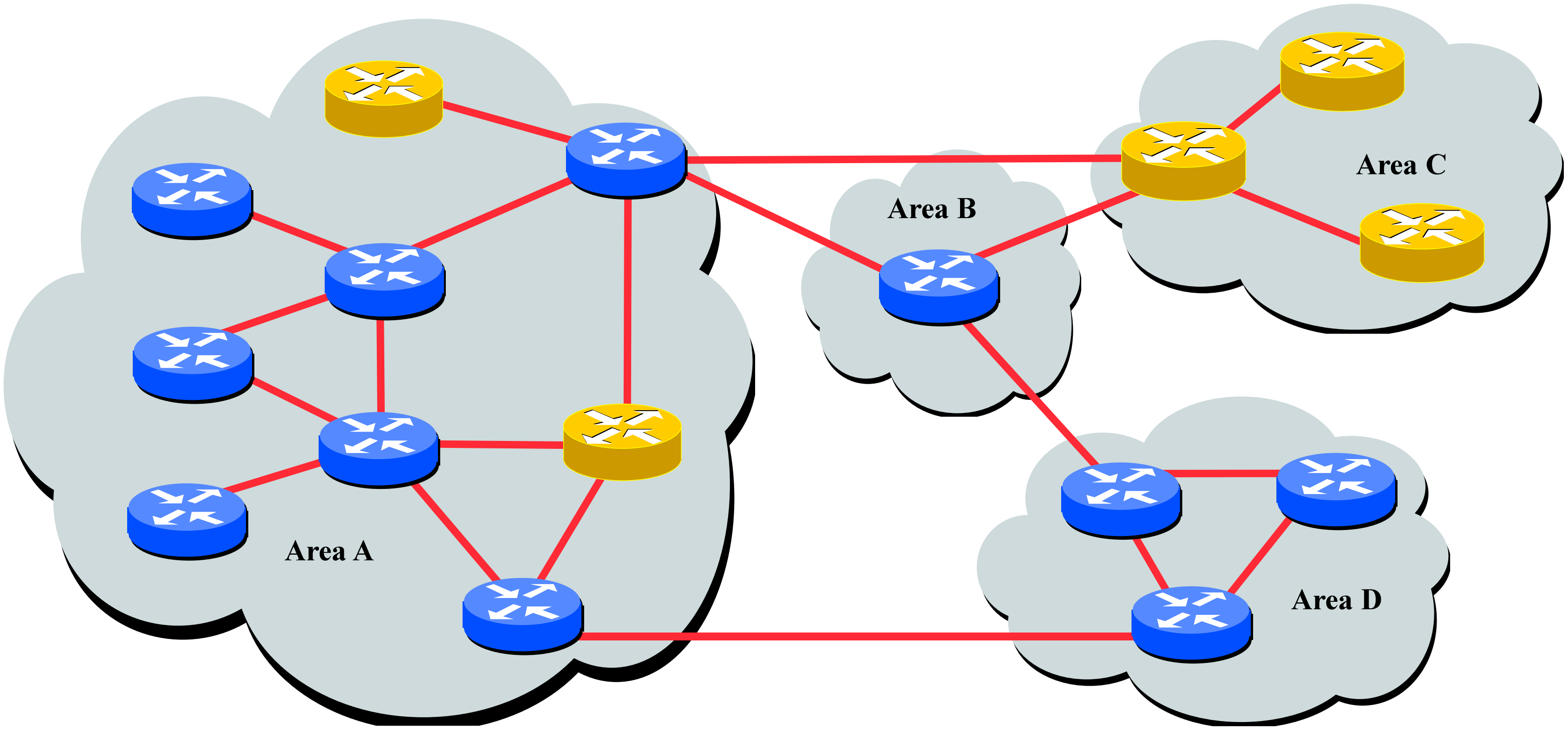
The problem

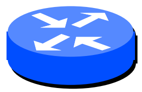
- Current IS-IS spec and implementation forces all protocols carried by IS-IS to agree on a common Shortest Path Tree
  - Single SPF run for all protocols
- Single SPT means congruent topologies
- Single SPT means all links need to understand all address families present in the domain

IS-IS Multi-Topologies Architecture

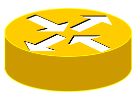
- Each router knows on which topologies it will establish adjacencies and build SPTs
  - Through configuration
- During adjacency establishment, peers need to agree on topologies
  - Topologies identifiers are exchanged in IIH packets

Multi-Topology IS-IS





IPv4-IPv6 enable router



IPv4-only enable router

The Multi-Topology software will create two topologies inside Area for IPv4 and IPv6.
IPv4-only routers will be excluded from the IPv6 topology

Two methods

- Multi-Topology
  - Single ISIS domain with set of independent IP topologies
  - Common flooding and resource associated with both router and network
  - Multiple SPF
  - Large Database
- Multi-instance
  - Multiple instance of protocol on a given link
  - Enhances the ability to isolate the resources associated with both router and network
  - Instance specific prioritization for PDUs and routing calculations

Two methods

- OSPF currently is based on multi-instance
  - Adding multi topology is very easy for OSPFv3
  - Multiple address family support is already there just minor extension for multi-topology needs to be added
- ISIS
  - Multi-topology support has been there for a while
  - Multi-instance draft is there for ISIS now
- Which one is better
  - Depends who you talk to
    - Operation (Multi-instance is better)
    - Development (Multi-Topology is better)

ISIS Fast Convergence

- Following Techniques/tools are available for fast convergence in ISIS
  - Carrier Delays **Detect**
  - Hello/dead timers (Fast Hellos) **Detect**
  - Bi-Directional Forwarding Detection—(BFD) **Detect**
  - LSP pacing **Propagate**
  - Interface event dampening - **Propagate**
  - Exponential throttle timers for LSA & SPF **Process**
  - PRC-interval **Process**
  - Incremental SPF **Process**
- Techniques/tools for Resiliency
  - Cisco NSF
  - Graceful Restart

BGP-4 Extensions for IPv6 (RFC 2545)

BGP-4 Extensions for IPv6

- BGP-4 carries only 3 pieces of information which is 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 in the OPEN message & AGGREGATOR attribute
- To make BGP-4 available for other network layer protocols, RFC 2858 (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 contains one or more Triples:
  - Address Family Information (AFI) – (next slide..)
  - 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 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 for BGP MUST be a valid IPv4 address. For e.g. 0.0.0.1 could NOT be use as a valid RID for BGP (Note: 0.0.0.1 CAN be used as a valid RID for OSPF & EIGRP)

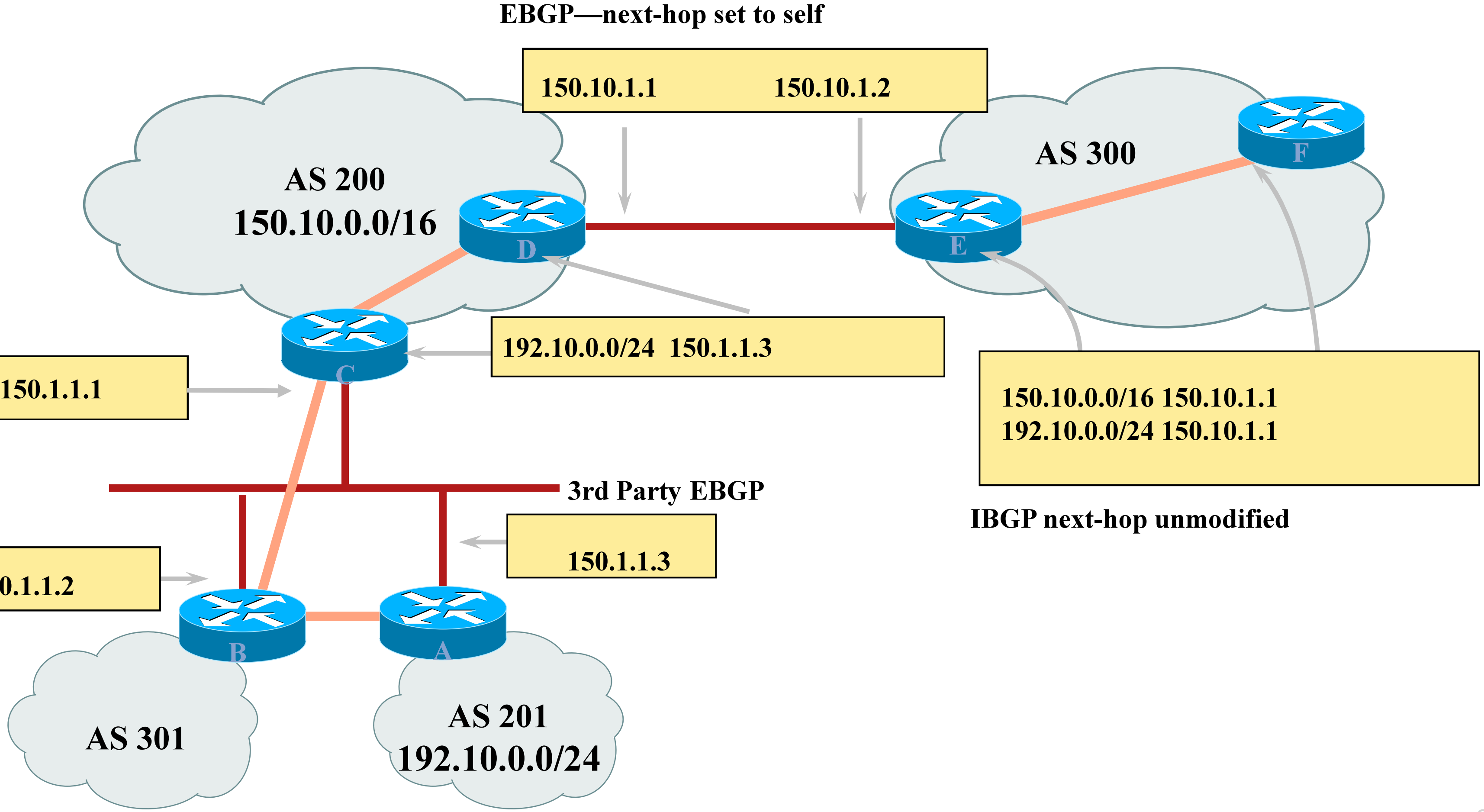
  The sole purpose of RID is for identification

  In BGP, it is used as a tie breaker, and is send within the OPEN message

IPv4 NLRI in IPv4

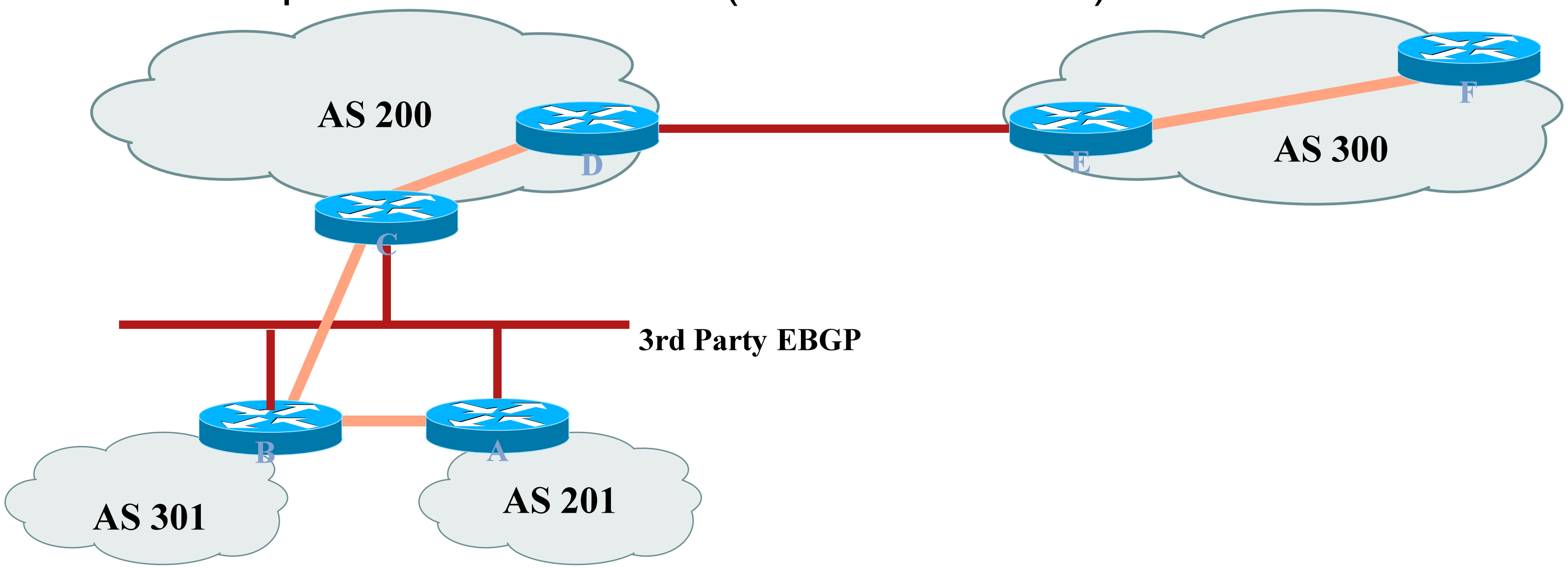
Router A

```
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

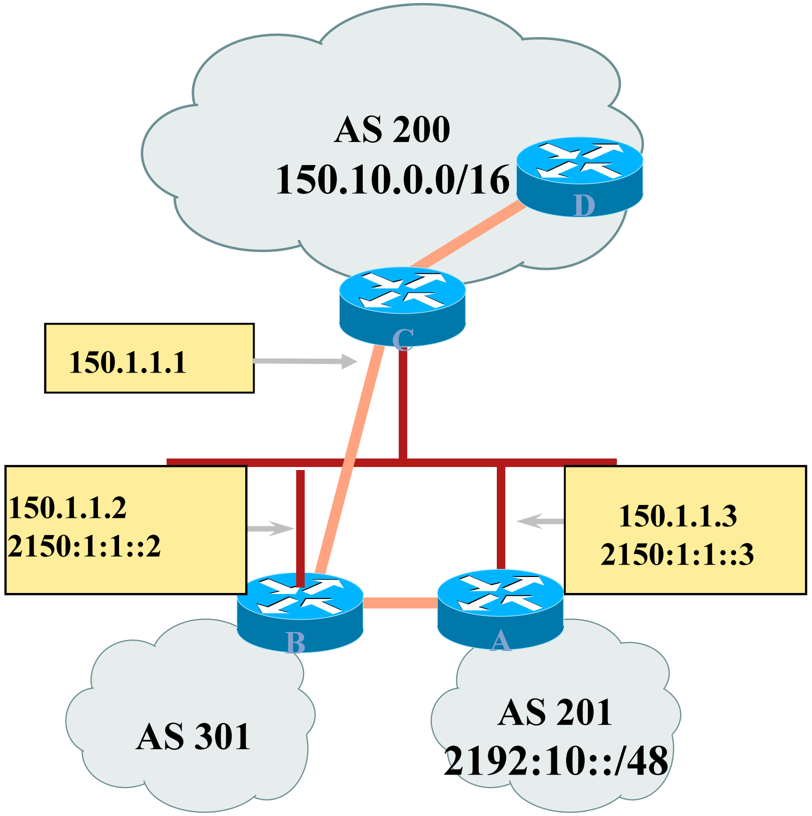
- Next hop reachability is VERY important in BGP
- If the Next Hop is inaccessible, the routes learned via BGP will not be installed in the routing table (In some case the routes will be rejected by BGP)
- The value of the length of the next hop field on MP_REACH_NLRI attribute is set to 16 when only global is present and is set to 32 if link local is present as well
- 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 (next few slides..)



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

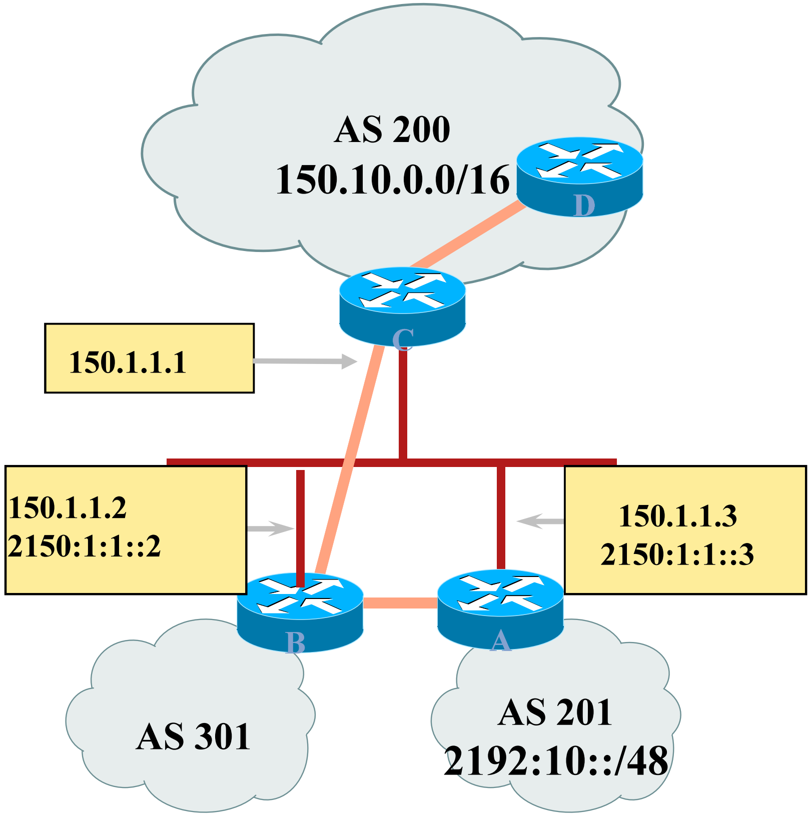
Router B:

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

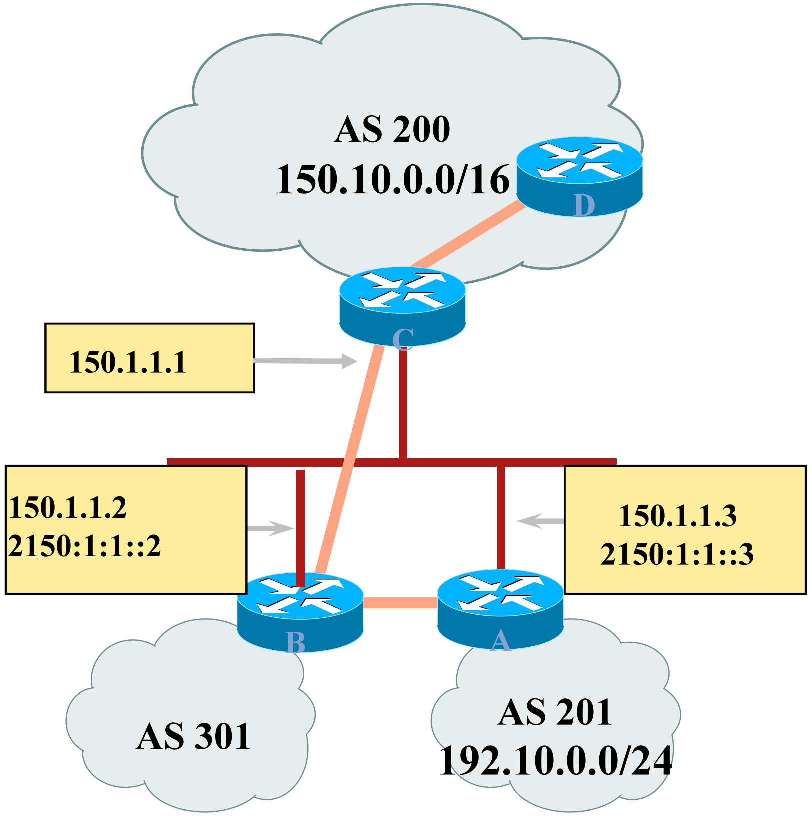
Router B:

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

IPv4 NLRI in IPv6(Global) - Problem

Router A

```
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 192.10.0.0
!
```



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

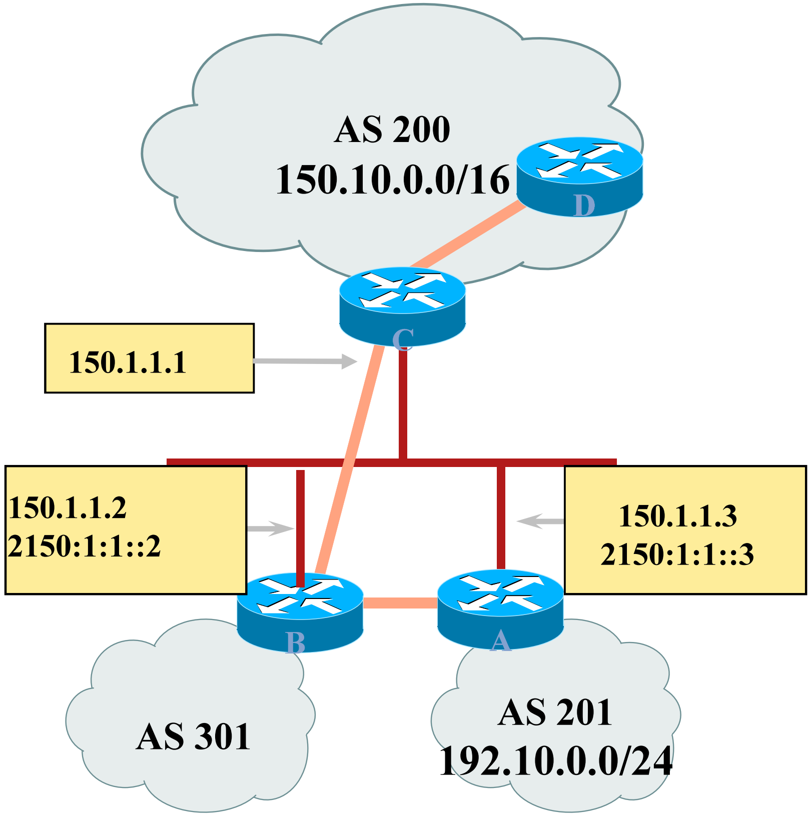
Router B:

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(Global) - Solution

Router A

```
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 192.10.0.0
!
route-map SETNH 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

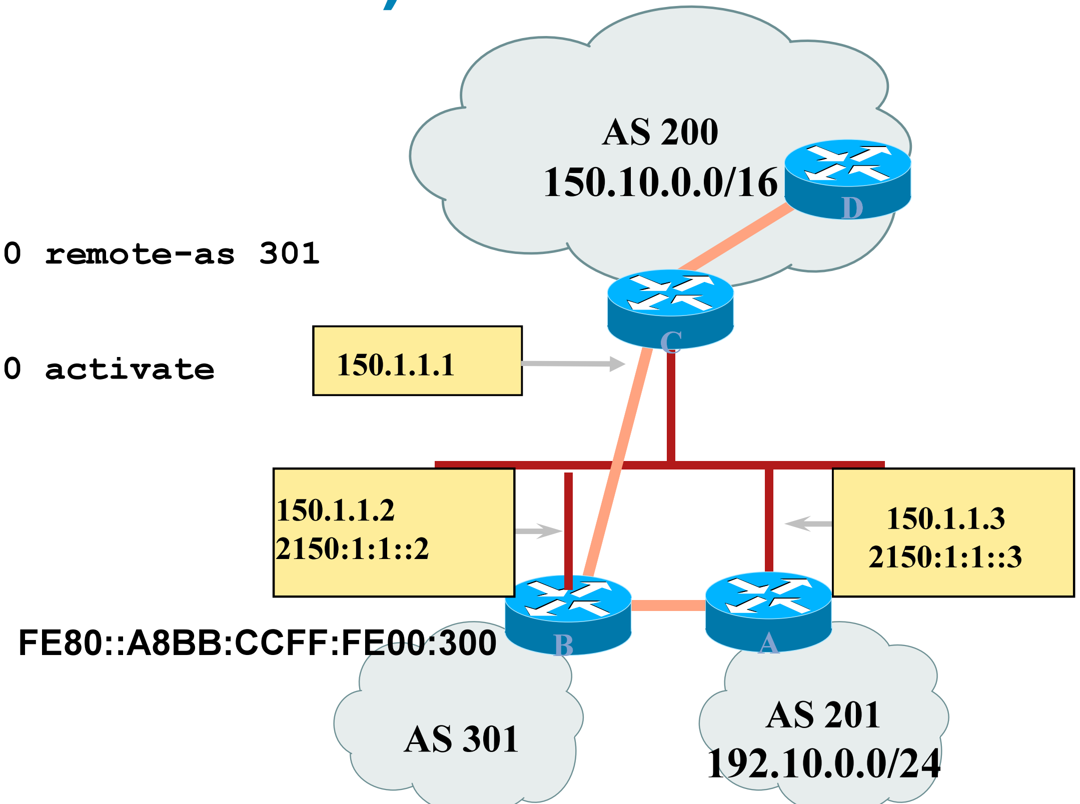
Router B:

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(Link Local) - Problem

Router A

```
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 192.10.0.0
!
```



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

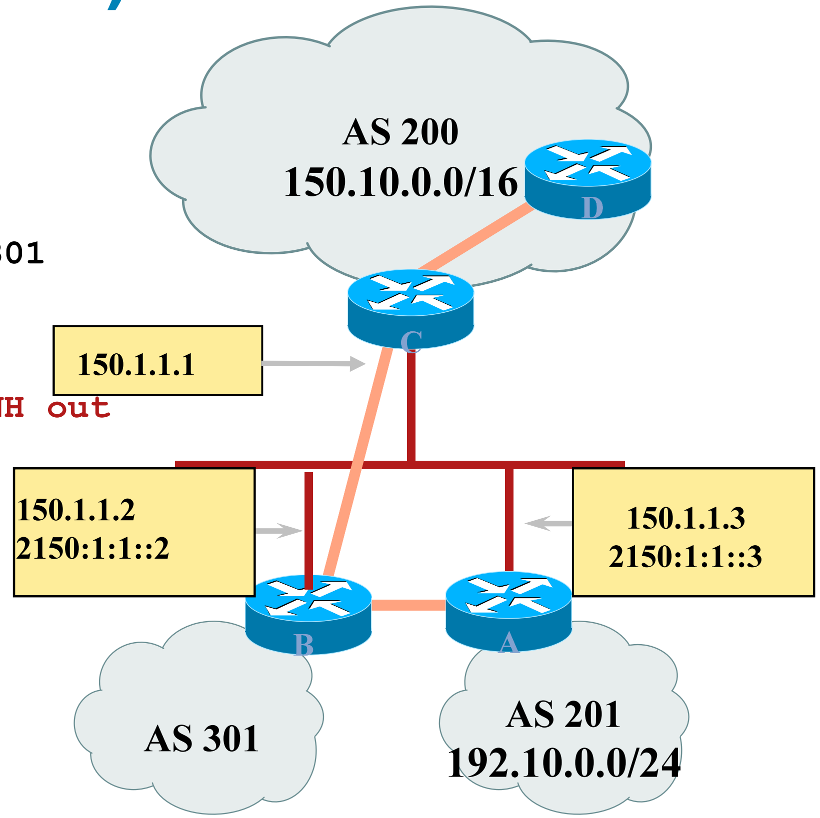
Router B:

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(Link Local) - Solution

Router A

```
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 route-map SETNH out
 network 192.10.0.0
!
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

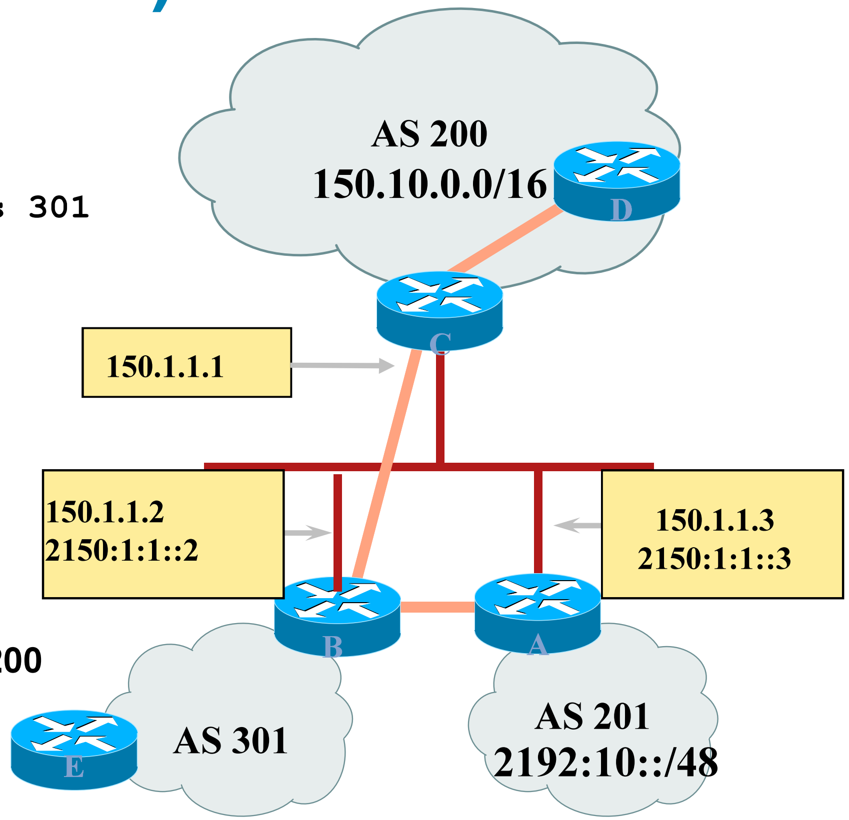
Router B:

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 NLRI in IPv6(Link Local) - Router A & B

Router A

```
router bgp 201
 bgp router-id 192.168.30.1
 neighbor FE80::A8BB:CCFF:FE00:200%E0 remote-as 301
 !
 address-family ipv6
 neighbor FE80::A8BB:CCFF:FE00:200%E0 activate
 network 2192:10::/48
!
```



Router A:

BGP(1): Can't advertise 2192:10::/64 to FE80::A8BB:CCFF:FE00:200%Ethernet0/0 session 1 with NEXT_HOP FE80::A8BB:CCFF:FE00:100
BGP(1): FE80::A8BB:CCFF:FE00:200%Ethernet0/0 send UPDATE (format) 2192:10::/64, next ::, metric 0, path Local

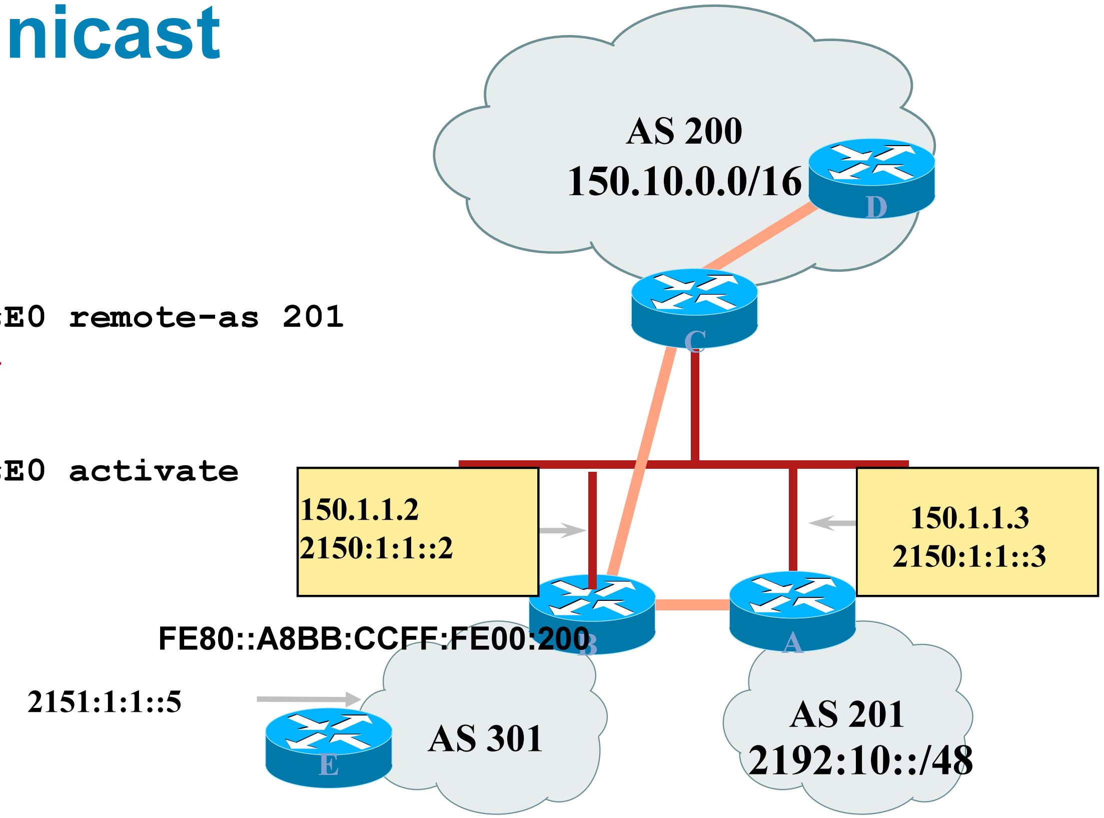
Router B:

BGP(1): FE80::A8BB:CCFF:FE00:100%Ethernet0/0 rcvd UPDATE w/ attr: nexthop FE80::A8BB:CCFF:FE00:100 (FE80::A8BB:CCFF:FE00:100), origin i, metric 0, path 201
BGP(1): FE80::A8BB:CCFF:FE00:100%Ethernet0/0 rcvd 2192:10::/64
BGP(1): Revise route installing 2192:10::/64 -> FE80::A8BB:CCFF:FE00:100 (FE80::A8BB:CCFF:FE00:100) to main IPv6 table

IPv6 NLRI in IPv6(Link Local) & iBGP between Router B & E using Global Unicast

Router B

```
router bgp 301
 bgp router-id 192.168.30.2
 neighbor FE80::A8BB:CCFF:FE00:100%E0 remote-as 201
 neighbor 2151:1:1::5 remote-as 301
 !
 address-family ipv6
 neighbor FE80::A8BB:CCFF:FE00:300%E0 activate
 neighbor 2151:1:1::5 activate
 !
```



Router B:

BGP(1): Can't advertise 2192:10::/64 to 2151:1:1::5 with NEXT_HOP FE80::A8BB:CCFF:FE00:200
BGP(1): 2151:1:1::5 send UPDATE (format) 2192:10::/64, next 2151:1:1::2, metric 0, path 10

Router E:

BGP(1): 2151:1:1::2 rcvd UPDATE w/ attr: nexthop 2151:1:1::2, origin i, localpref 100, metric 0, path 10
BGP(1): 2151:1:1::2 rcvd 2192:10::/64
BGP(1): Revise route installing 2192:10::/64 -> 2151:1:1::2 (::) to main IPv6 table

BGP-4 Configurations for IPv6
Link Local Peering 12.2.33SRC and above IOS

Router A

```
router bgp 200
 neighbor FE80::A8BB:CCFF:FE01:F600%Ethernet0/0 remote-as 100
 !
 address-family ipv6
 neighbor FE80::A8BB:CCFF:FE01:F600%Ethernet0/0 activate
 neighbor FE80::A8BB:CCFF:FE01:F600%Ethernet0/0 route-map SETNH out
 redistribute connected
 no synchronization
!
route-map SETNH permit 10
 set ipv6 next-hop 2001:100:1:1::2
```

New CLI per RFC 4007