

Tutorial: IPv6 Technology Overview Part I

Speaker: Byju Pularikkal, Cisco Systems, Inc

Date: 10th October 2011

Tutorial-1: Agenda

- Structure of IPv6 Protocol
 - IPv4 and IPv6 Header Comparison
 - IPv6 Extension Headers
- IPv6 Addressing
 - Addressing Format
 - Types of IPv6 addresses
- ICMPv6 and Neighbor Discovery
 - Router Solicitation & Advertisement
 - Neighbor Solicitation & Advertisement
 - Duplicate Address Detection
- Multicast in IPv6
- DHCP & DNS for IPv6
 - DNS with IPv6
 - DHCPv6 Overview



Tutorial-2: Agenda

- Routing in IPv6
 - RIPng
 - OSPFv3
 - IS-IS support for IPv6
 - BGP-4 Extensions for IPv6

- IPv6 Transition Mechanisms
 - 6 to 4 Tunneling
 - ISATAP
 - 6RD
 - Dual-stack Lite
 - 6PE
 - 6VPE



The Structure of IPv6 Protocol



IPv4 and IPv6 Header Comparison





IPv4 Header

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

IPv6 Header

Version	Traffic Class	Flow Label		
Payload Length		Next Header	Hop Limit	
Source Address				
Destination Address				

Legend

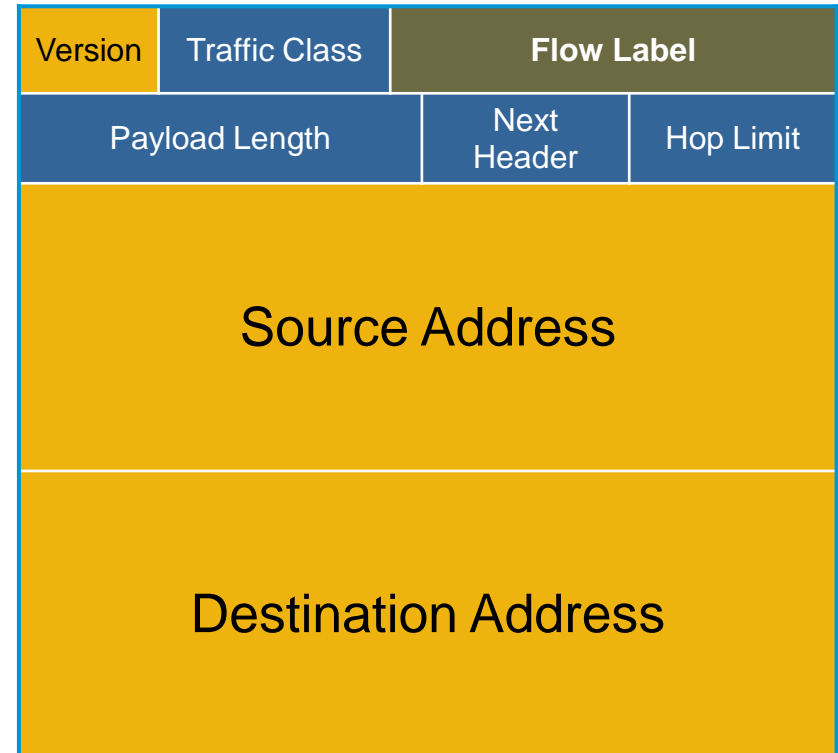
-  - Field name kept from IPv4 to IPv6
-  - Fields not kept in IPv6
-  - Name and position changed in IPv6
-  - New field in IPv6

IPv6 Header New Field: Flow Label (RFC 3697)

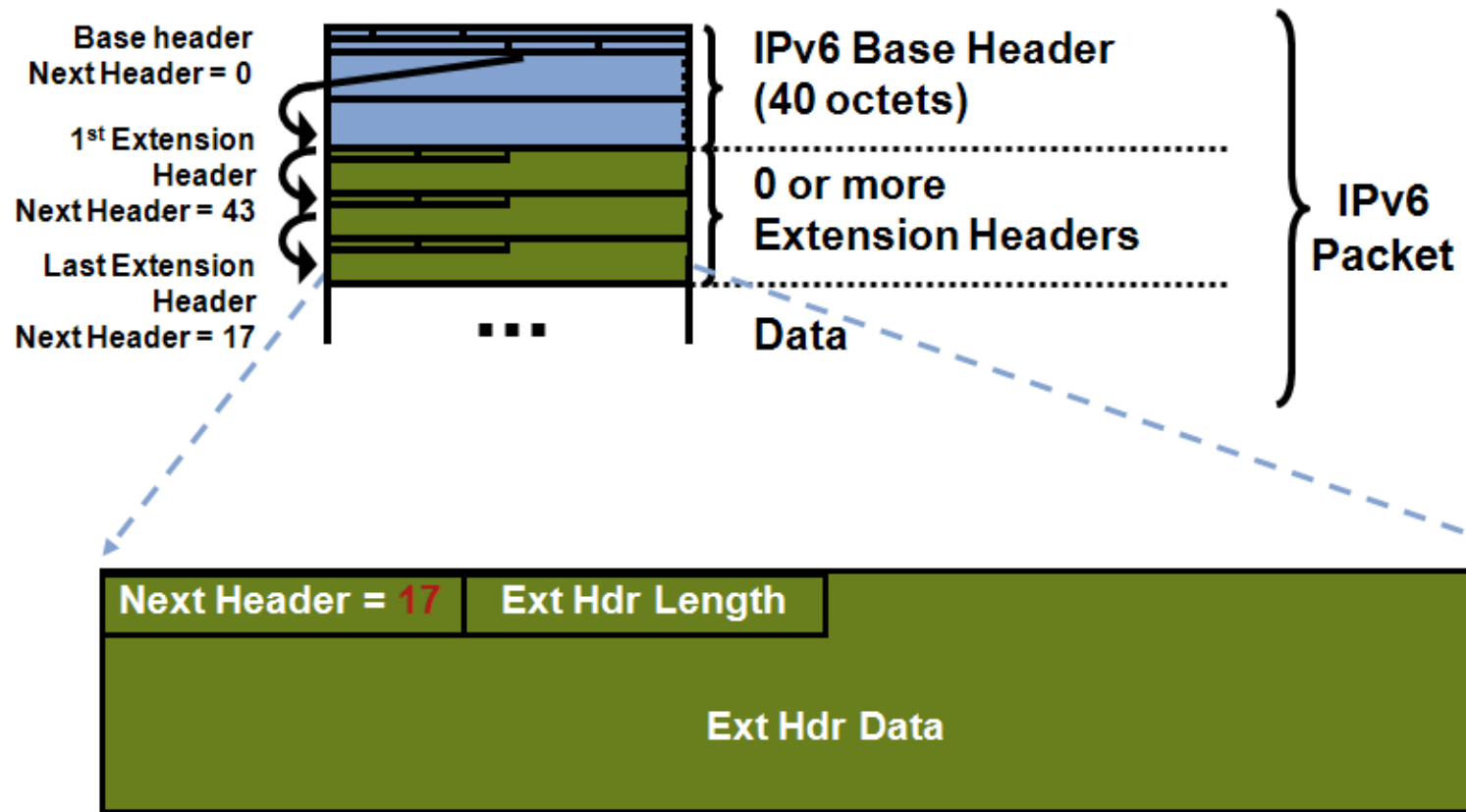
20-Bit Flow Label Field to Identify Specific Flows Needing Special QoS

- Flow classifiers had been based on 5-tuple: Source/destination address, protocol type and port numbers of transport
- Some of these fields may be unavailable due to fragmentation, encryption or locating them past extension headers
- With flow label, each source chooses its own flow label values
- Packet classifiers use the triplet of Flow Label, Source Address, and Destination Address fields to identify which flow a particular packet belongs to
- Flow label value of 0 used when no special QoS requested (the common case today)

IPv6 Header



Extension Headers



Extension Header Order

Extension Headers Should Be Constructed in the Following Sequence and Should Be Sequenced in this Order:

Hop-by-Hop header	(0)
Destination options header (w/ routing header)	(60)
Routing header	(43)
Fragment header	(44)
Authentication header	(51)
ESP header	(50)
Mobility header	(135)
Destination options header	(60)
ICMPv6	(58)
No Next header	(59)
Upper-layer header	(Varies— TCP=6, UDP=17)



MTU Issues

- Minimum link MTU for IPv6 is 1280 octets (vs. 68 octets for IPv4)
 - on links with MTU < 1280, link-specific fragmentation and reassembly must be used
- Implementations are expected to perform path MTU discovery to send packets bigger than 1280
- Minimal implementation can omit PMTU discovery as long as all packets are kept ≤ 1280 octets
- A hop-by-hop option supports transmission of “jumbograms” with up to 2^{32} octets of payload; payload is normally 2^{16}



IPv6 Addressing



IPv6 Addressing

IPv4 32-bits

IPv6 128-bits

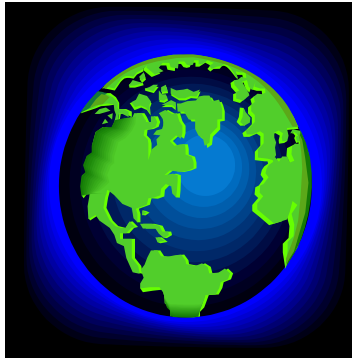
$$2^{32} = 4,294,967,296$$

$$2^{128} = 340,282,366,920,938,463,463,374,607,431,768,211,456$$

$$2^{128} = 2^{32} * 2^{96}$$

$$2^{96} = 79,228,162,514,264,337,593,543,950,336 \text{ times the number of possible IPv4 Addresses (79 trillion trillion)}$$

IPv6 Addressing



**World's population is
approximately 6.5 billion**

$$\frac{2^{128}}{6.5 \text{ Billion}} = 52 \text{ Trillion Trillion IPv6 addresses per person}$$



**Typical brain has
~100 billion brain cells
(your count may vary)**

$$\frac{52 \text{ Trillion Trillion}}{100 \text{ Billion}} = 523 \text{ Quadrillion (523 thousand trillion) IPv6 addresses for every human brain cell on the planet!}$$

Addressing Format

Representation

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

Leading zeros in contiguous block could be represented by (::)

Example:

2001:0db8:0000:130F:0000:0000:087C:140B

2001:0db8:0:130F::87C:140B

Double colon only appears once in the address



Prefix Representation in IPv6

- Representation of prefix is just like CIDR
- In this representation you attach the prefix length
- Like v4 address:
198.10.0.0/16
- V6 address is represented the same way:
2001:db8:12::/48
- Only leading zeros are omitted. Trailing zeros are not omitted

2001:0db8:0012::/48 = 2001:db8:12::/48

2001:db8:1200::/48 \neq 2001:db8:12::/48



Loopback & Unspecified Addresses

- Loopback address representation
 - 0:0:0:0:0:0:0:1=> ::1
 - Same as 127.0.0.1 in IPv4
 - Identifies self
- Unspecified address representation
 - 0:0:0:0:0:0:0:0=> ::
 - Used as a placeholder when no address available
 - Examples: Initial DHCP request, Duplicate Address Detection (DAD)



IPv6—Addressing Model

- Addresses are assigned to interfaces
- Interface “expected” to have multiple addresses
- Addresses have scope
 - Link Local
 - Unique Local
 - Global
- Addresses have lifetime
 - Valid and preferred lifetime



Address Categories

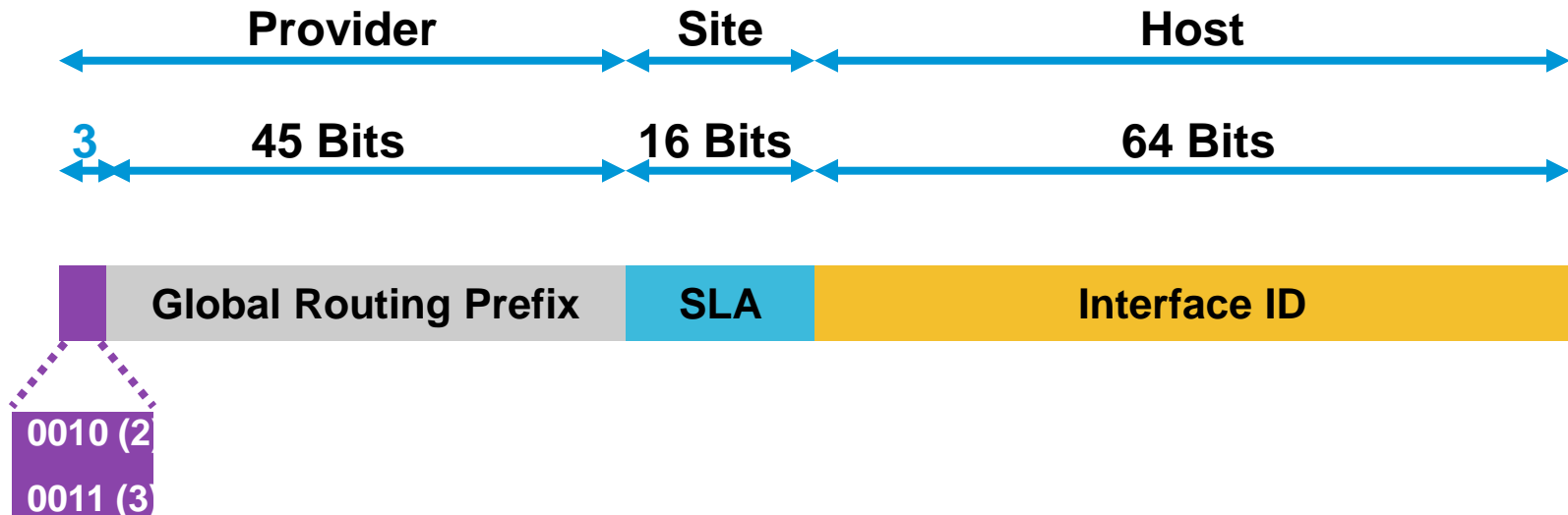
Type	Binary	Hex
Aggregatable Global Unicast Address	001	2 or 3
Link Local Unicast Address	1111 1110 10	FE80::/10
Unique Local Unicast Address	1111 1100 1111 1101	FC00::/7 FC00::/8(registry) FD00::/8 (no registry)
Multicast Address	1111 1111	FF00::/8

Types of IPv6 Addresses

- Unicast
 - Address of a single interface. One-to-one delivery to single interface
- Multicast
 - Address of a set of interfaces. One-to-many delivery to all interfaces in the set
- Anycast
 - Address of a set of interfaces. One-to-one-of-many delivery to a single interface in the set that is closest
- No more broadcast addresses

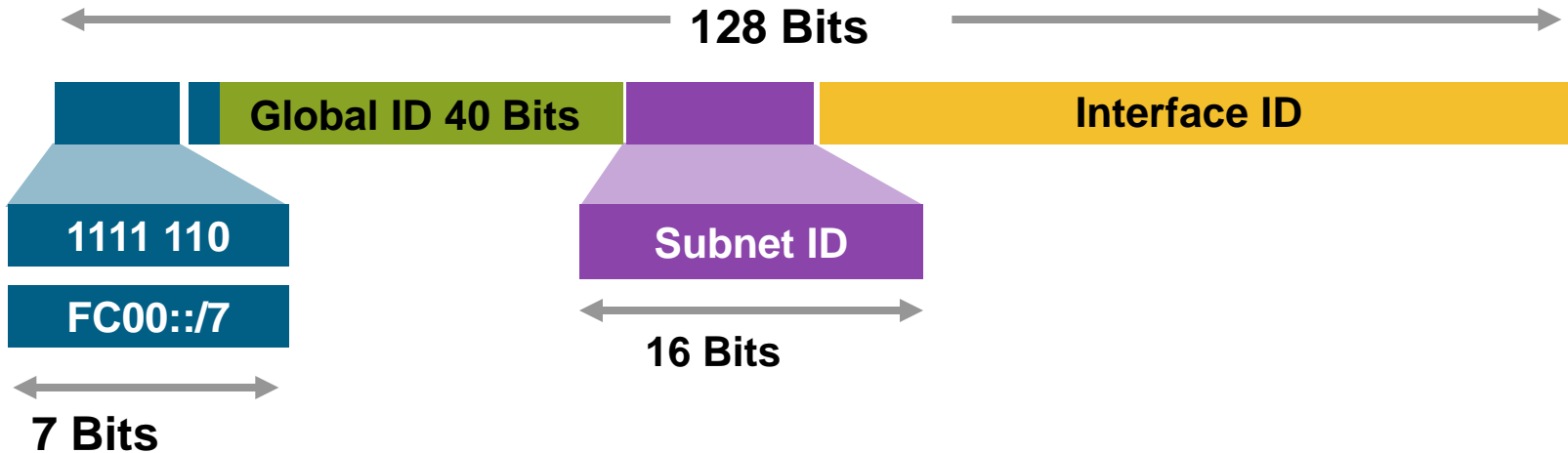


Global Unicast Addresses



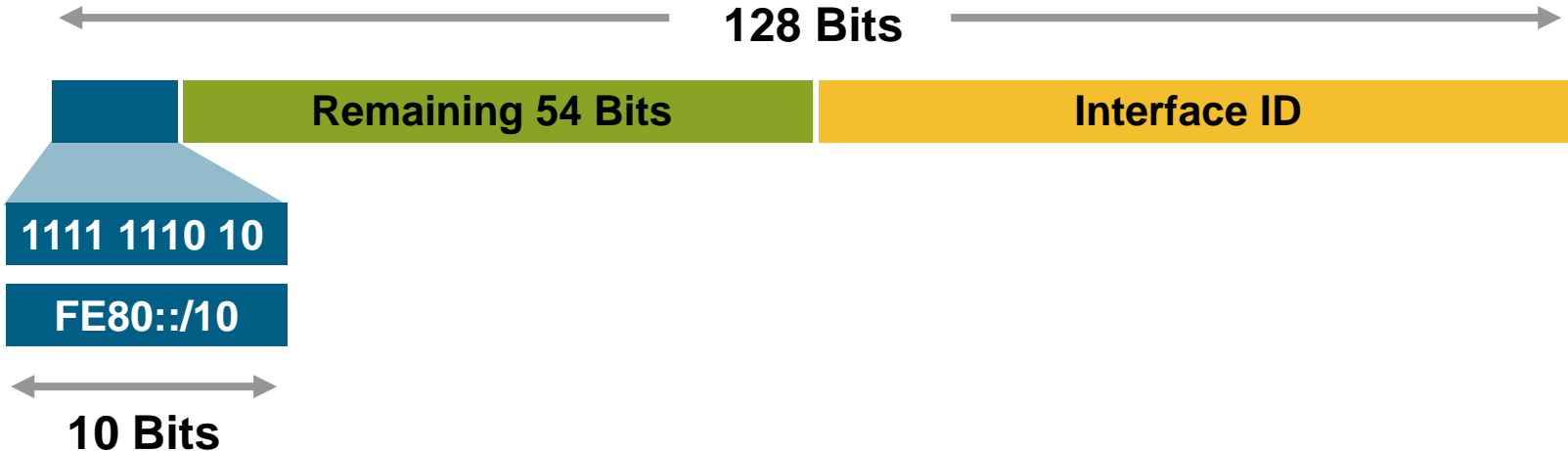
- Global Unicast Addresses Are:
 - Addresses for generic use of IPv6
 - Structured as a hierarchy to keep the aggregation

Unique-Local



- Unique-Local Addresses Used for:
 - Local communications
 - Inter-site VPNs
 - Not routable on the Internet

Link-Local



Link-Local Addresses Used for:

- Mandatory Address for Communication between two IPv6 devices (like ARP but at Layer 3)
- Automatically assigned by Router as soon as IPv6 is enabled
- Also used for Next-Hop calculation in Routing Protocols
- Only Link Specific scope
- Remaining 54 bits could be Zero or any manual configured value



IPv6 Multicast Address

- IP multicast address has a prefix FF00::/8 (1111 1111); the second octet defines the lifetime and scope of the multicast address

8-bit	4-bit	4-bit	112-bit
1111 1111	Lifetime	Scope	Group-ID

Lifetime	
0	If Permanent
1	If Temporary

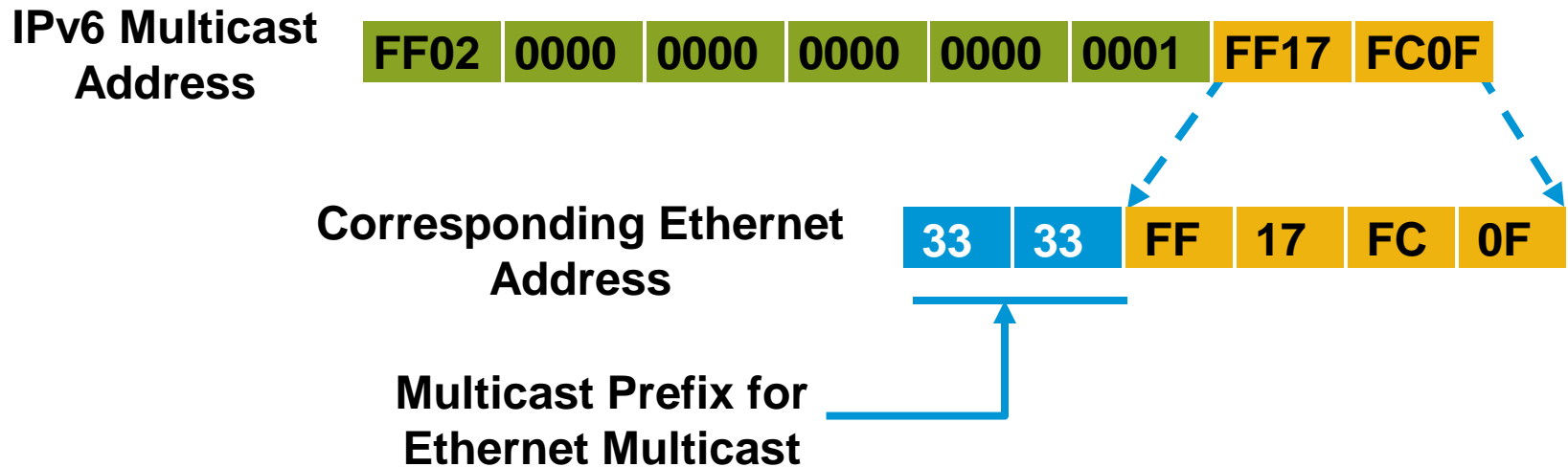
Scope	
1	Node
2	Link
5	Site
8	Organization
E	Global

Some Well Known Multicast Addresses

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	All Routers
FF02::1:FFXX:XXXX	Link-Local	Solicited-Node

- Note that 02 means that this is a permanent address and has link scope
- More details at <http://www.iana.org/assignments/ipv6-multicast-addresses>

Multicast Mapping over Ethernet



- Mapping of IPv6 multicast address to Ethernet address is:
 - 33:33:<last 32 bits of the IPv6 multicast address>

Solicited-Node Multicast Address

- For each unicast and anycast address configured there is a corresponding solicited-node multicast
- This is specially used for two purpose, for the replacement of ARP, and DAD
- Used in neighbor solicitation messages
- Multicast address with a link-local scope
- Solicited-node multicast consists of
 - FF02::1:FF & {lower 24 bits from IPv6 Unicast interface ID}

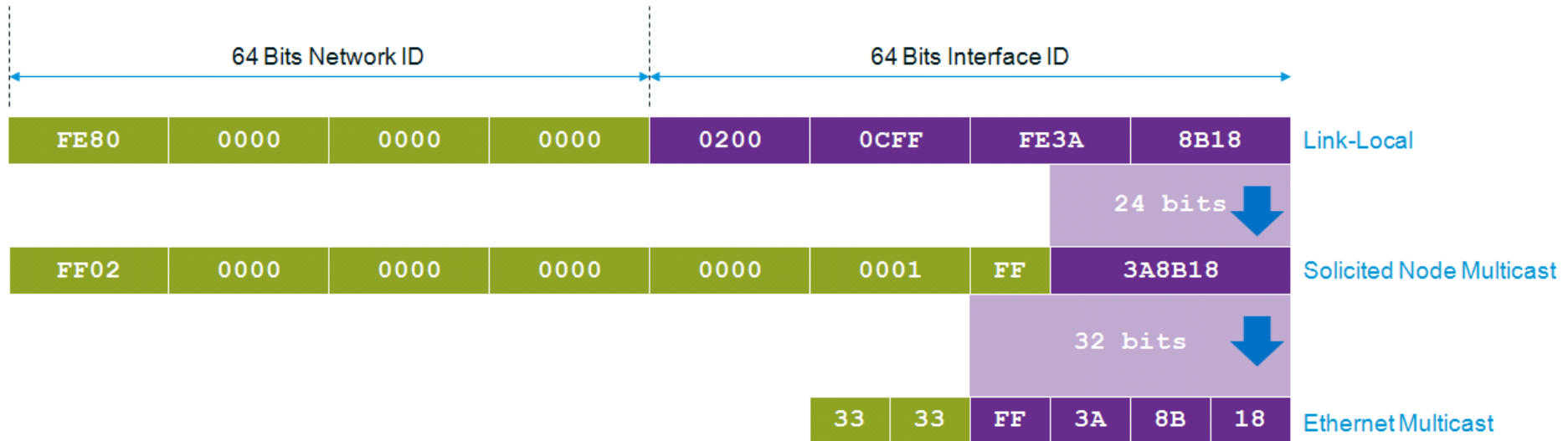
IPv6 Address



Solicited-node multicast Address



Solicited Node Multicast Address Example

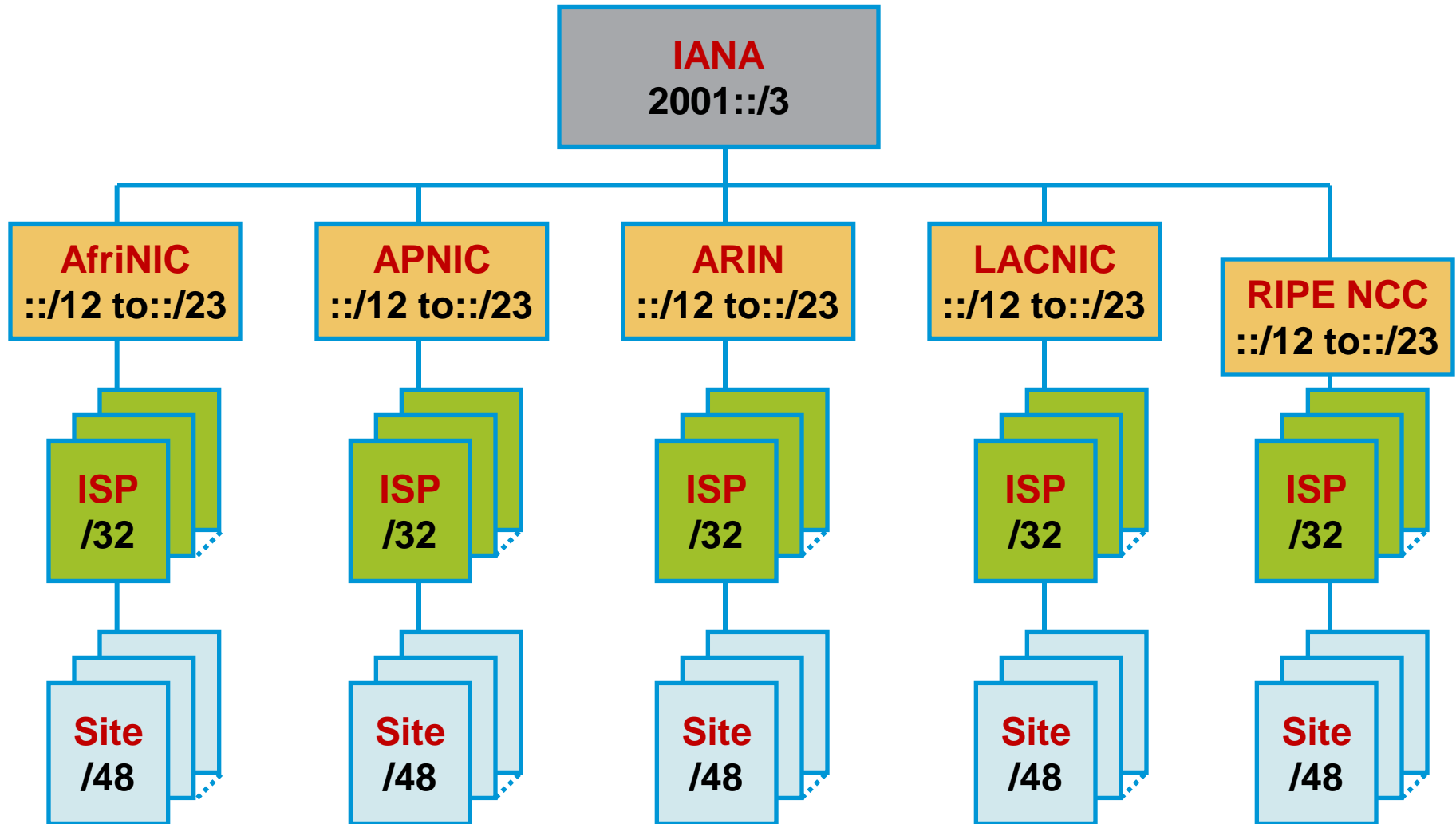


Anycast

- Anycast allows a source node to transmit IP datagrams to a single destination node out of a group destination nodes with same subnet id based on the routing metrics
- Only routers should respond to anycast addresses
- Routers along the path to the destination just process the packets based on network prefix
- Routers configured to respond to anycast packets will do so when they receive a packet send to the anycast address



IPv6 Prefix Allocation Hierarchy and Policy Example



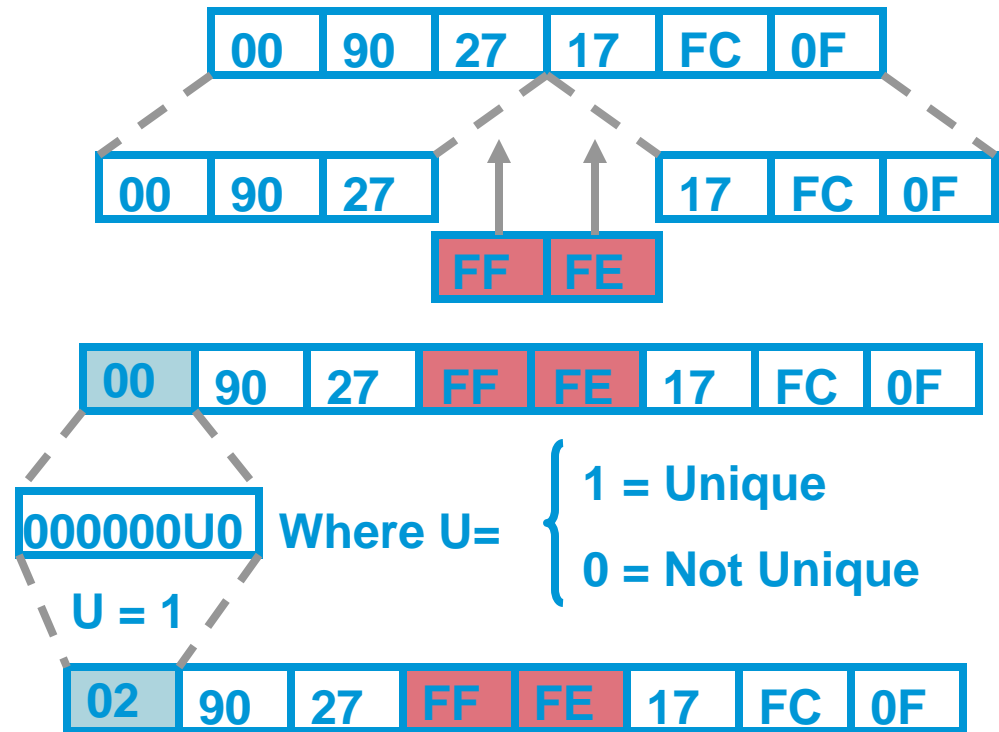
64 bit Interface Identifier

- Lowest-Order 64-bit field of unicast address 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



IPv6 Interface Identifier

- The EUI-64 format can be used to do stateless auto-configuration
- This format expands the 48 bit MAC address to 64 bits by inserting FFFE into the middle 16 bits
- To make sure that the chosen address is from a unique Ethernet MAC address, the universal/local (“u” bit) is set to 1 for global scope and 0 for local scope

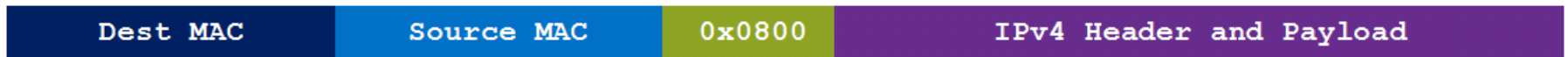


IPv6 over Ethernet

- IPv6 uses Ethernet Protocol ID (0x86DD)



- IPv4 uses Ethernet Protocol ID (0x0800)



ICMPv6 and Neighbor Discovery



ICMPv6

- Internet Control Message Protocol version 6
- RFC 2463
- Modification of ICMP from IPv4
- Message types are similar (but different types/codes)
 - Destination unreachable (type 1)
 - Packet too big (type 2)
 - Time exceeded (type 3)
 - Parameter problem (type 4)
 - Echo request/reply (type 128 and 129)



Neighbor Discovery

- Replaces ARP, ICMP (redirects, router discovery)
- Reachability of neighbors
- Hosts use it to discover routers, auto configuration of addresses
- Duplicate Address Detection (DAD)



Neighbor Discovery : Contd..

- Neighbor discovery uses ICMPv6 messages, originated from node on link local with hop limit of 255
- Consists of IPv6 header, ICMPv6 header, neighbor discovery header, and neighbor discovery options
- Five neighbor discovery messages
 1. Router solicitation (ICMPv6 type 133)
 2. Router advertisement (ICMPv6 type 134)
 3. Neighbor solicitation (ICMPv6 type 135)
 4. Neighbor advertisement (ICMPv6 type 136)
 5. Redirect (ICMPv6 type 137)



Router Solicitation and Advertisement



1—ICMP Type = 133 (RS)

Src = link-local address (FE80::1/10)

Dst = **all-routers** multicast address (FF02::2)

Query = please send RA

2—ICMP Type = 134 (RA)

Src = link-local address (FE80::2/10)

Dst = **all-nodes** multicast address (FF02::1)

Data = options, subnet prefix, lifetime, autoconfig flag

- Router solicitations (RS) are sent by booting nodes to request RAs for configuring the interfaces
- Routers send periodic Router Advertisements (RA) to the all-nodes multicast address

Neighbor Solicitation and Advertisement



Neighbor Solicitation

ICMP type = 135

Src = A

Dst = Solicited-node multicast of B

Data = link-layer address of A

Query = what is your link address?



Neighbor Advertisement

ICMP type = 136

Src = B

Dst = A

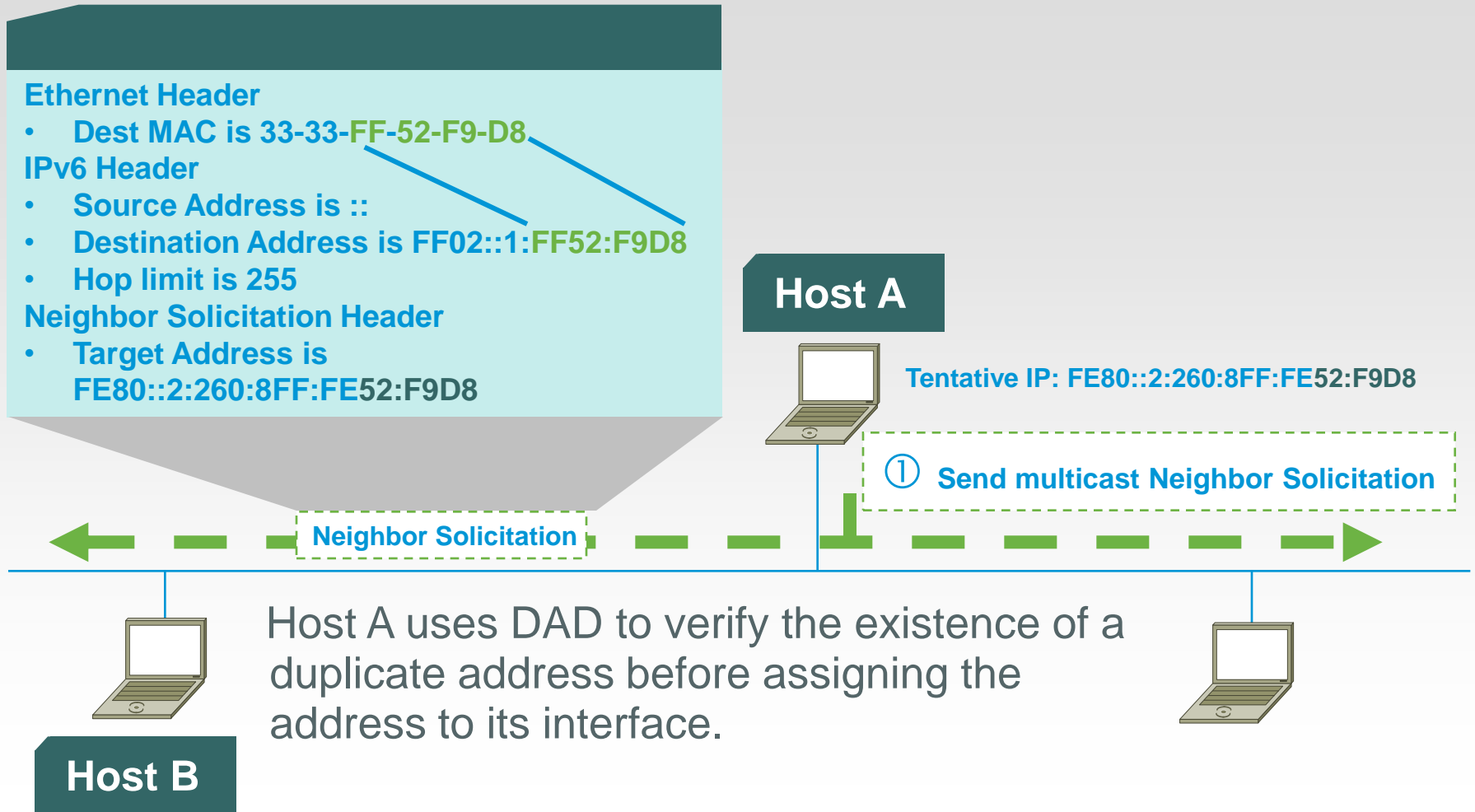
Data = link-layer address of B



A and B can now exchange
packets on this link

A blue double-headed arrow pointing left and right, indicating that A and B can now exchange packets on this link.

Multicast Neighbor Solicitation – for Duplicate Address Detection (DAD)



Multicast Neighbor Advertisement (Response)

Ethernet Header

- Destination MAC is 33-33-00-00-00-01

IPv6 Header

- Source Address is FE80::2:260:8FF:FE52:F9D8
- Destination Address is FF02::1
- Hop limit is 255

Neighbor Advertisement Header

- Target Address is FE80::2:260:8FF:FE52:F9D8

Neighbor Discovery Option

- Target Link-Layer Address is 00-60-08-52-F9-D8

Host A



Tentative IP: FE80::2:260:8FF:FE52:F9D8

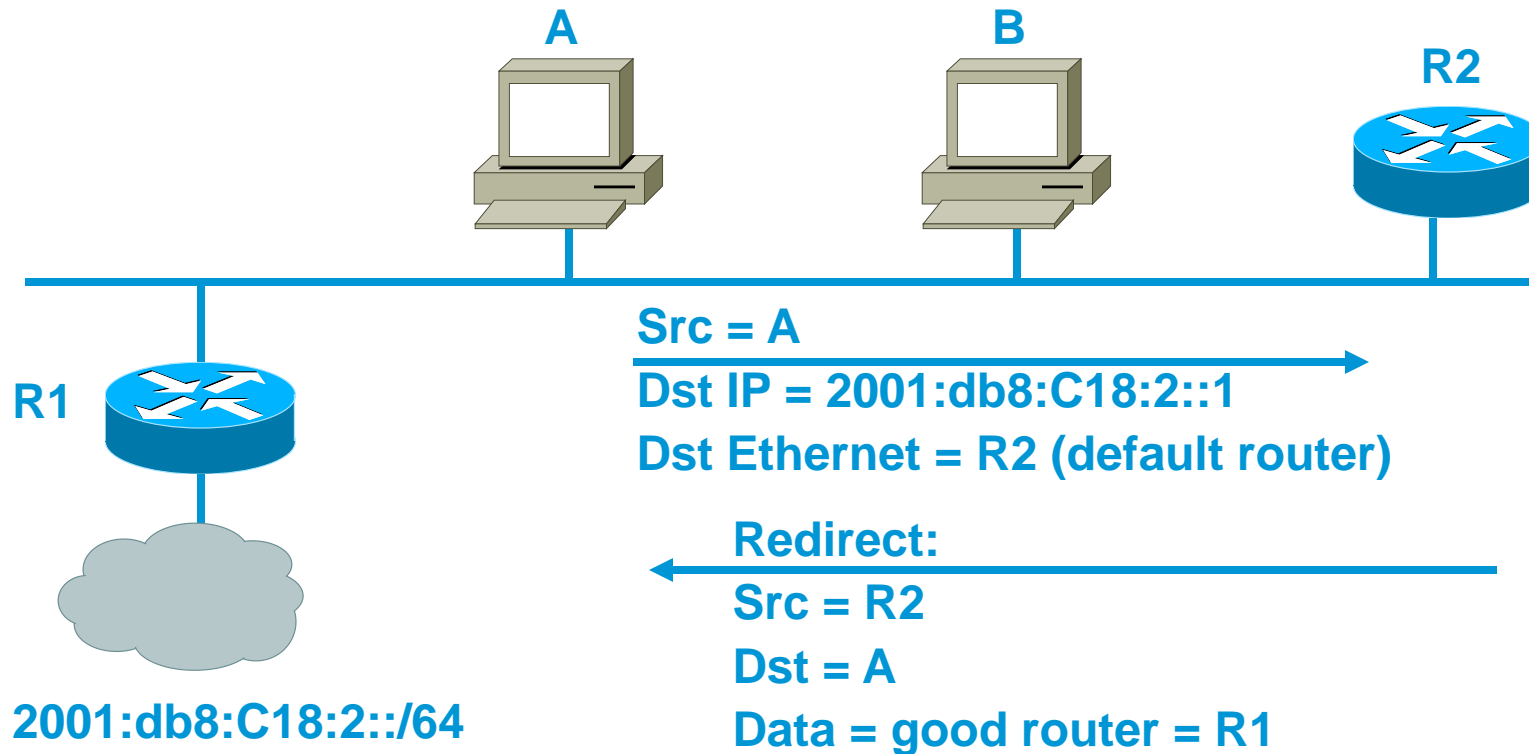
Neighbor Advertisement

MAC: 00-60-08-52-F9-D8
IP: FE80::2:260:8FF:FE52:F9D8

Host B

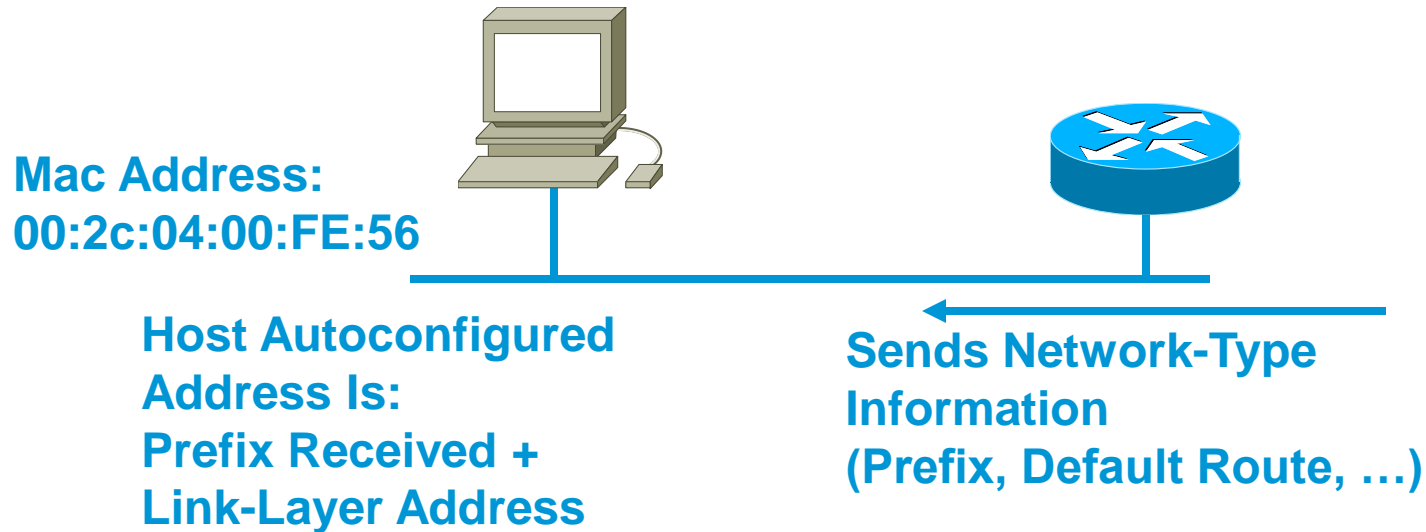
② Send multicast Neighbor Advertisement

Redirect



- Redirect is used by a router to signal the re-route of a packet to a router with better route to the destination

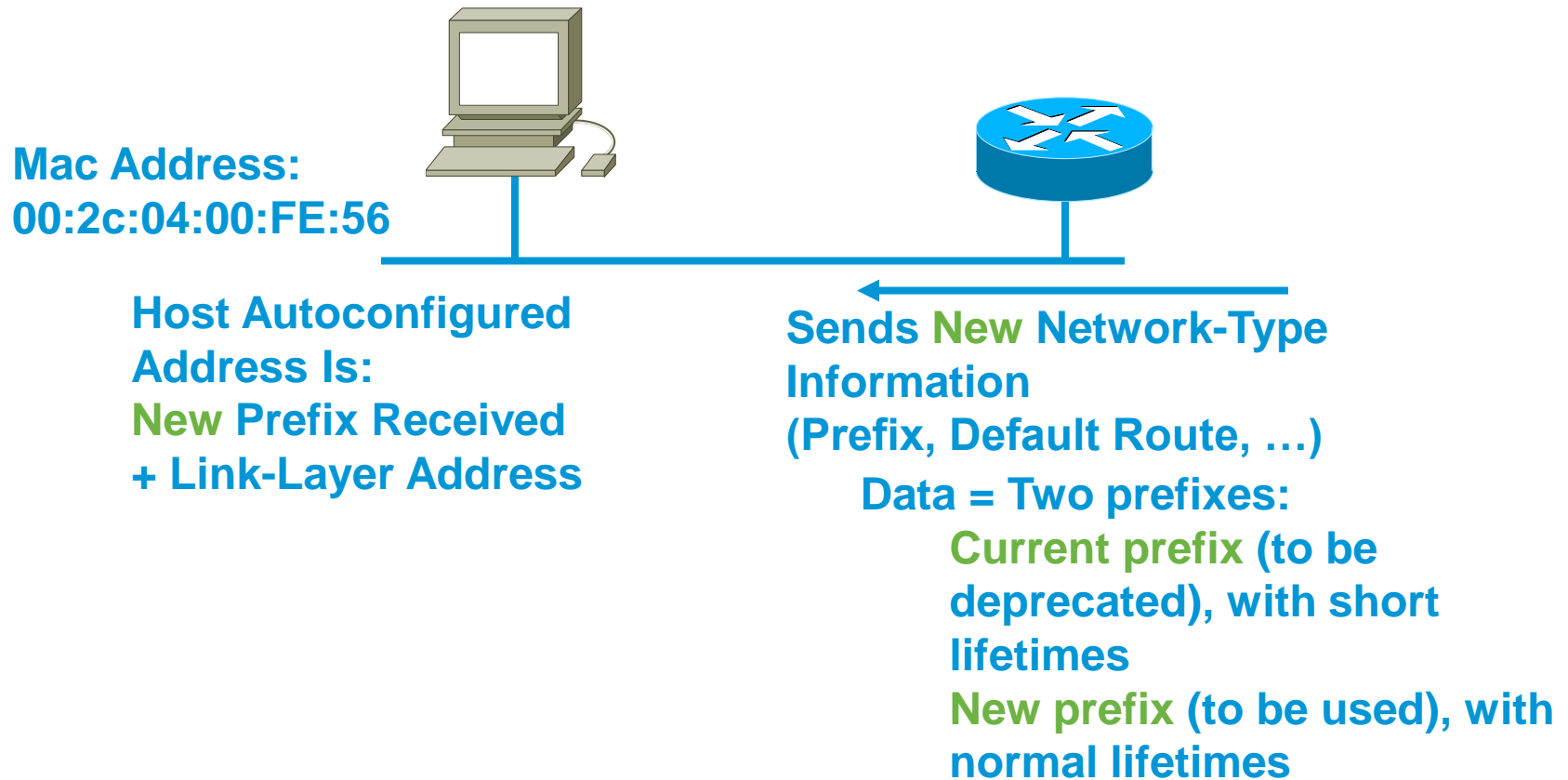
Autoconfiguration



Larger Address Space Enables:

- The use of link-layer addresses inside the address space
- Autoconfiguration with “no collisions”
- Offers “plug and play”

Renumbering



Larger Address Space Enables:

- Renumbering, using autoconfiguration and multiple addresses

Multicast Listener Discover – MLD

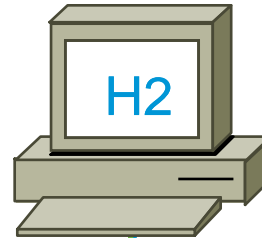
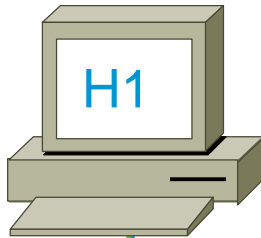
- Equivalent to IGMP in IPv4
- Messages are transported over ICMPv6
- Uses link local source addresses
- Use “Router Alert” option in header (RFC2711)
- Version number confusion:
 - MLDv1 (RFC2710) like IGMPv2 (RFC2236)
 - MLDv2 (draft-vida-mld-v2-07) like IGMPv3 (RFC3376)
 - Provides SSM support
- MLD snooping (RFC 4541)



MLD - Joining a Group (REPORT)

FE80::209:5BFF:FE08:A674

FE80::250:8BFF:FE55:78DE



Destination:
FF3E:40:3FFE:C15:C003:1109:1111:1111
ICMPv6 Type: 131



Destination:
FF3E:40:3FFE:C15:C003:1109:1111:1111
ICMPv6 Type: 131
FE80::207:85FF:FE80:692



H1 sends a REPORT for the group



H2 sends a REPORT for the group



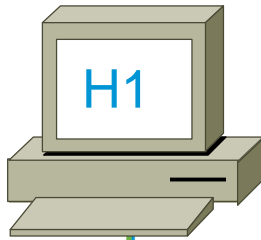
Source

Group:FF3E:40:3FFE:C15:C003:1109:1111:1111

MLD - Group-Specific Query

FE80::209:5BFF:FE08:A674

FE80::250:8BFF:FE55:78DE



REPORT to group
ICMPv6 Type: 131



Destination:
FF02::2
ICMPv6 Type: 132



Destination:
FF3E:40:3FFE:C15:C003:1109:1111:1111
ICMPv6 Type: 130



H1 sends DONE to FF02::2



RTR-A sends Group-Specific Query



H2 sends REPORT for the group



FE80::207:85FF:FE80:692

Source



Group:FF3E:40:3FFE:C15:C003:1109:1111:1111



Other MLD Operations

- Leave/DONE

- Last host leaves - Sends DONE (Type 132)
- Router responds with Group-Specific Query (Type 130)
- Router uses the Last member query response interval (Default=1 sec) for each query
- Query is sent twice and if no reports occur then entry is removed (2 seconds)

- General Query (Type 130)

- Sent to learn of listeners on the attached link
- Sets the Multicast Address Field to zero
- Sent every 125 seconds (configurable)



DNS for IPv6



DNS Basics

- DNS is a database managing Resource Records (RR)
 - Stockage of RR from various types—IPV4 and IPV6:
 - Name Server
 - Address—A and AAAA
 - Pointer—PTR
- DNS is an IP application
 - It uses either UDP or TCP on top of IPv4 or IPv6
- References
 - RFC3596: DNS Extensions to Support IP Version 6
 - RFC3363: Representing Internet Protocol Version 6 Addresses in Domain Name system (DNS)
 - RFC3364: Tradeoffs in Domain Name System (DNS) Support for Internet Protocol version 6 (IPv6)

IPv6 and DNS

	IPv4	IPv6
Hostname to IP address	A record: www.abc.test. A 192.168.30.1	AAAA record: www.abc.test AAAA 2001:db8:C18:1::2
IP address to hostname	PTR record: 1.30.168.192.in-addr.arpa. PTR www.abc.test.	PTR record: 2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.1.0.0.0.8.1.c.0.8.b.d.0.1.0.0.2.ip6.arpa PTR www.abc.test.

DHCPv6 Overview



IPv6 Address Assignment

The various IPv6 address assignment methods are as follows:

- 1. Manual Assignment**
- 2. Stateless Address Autoconfiguration (SLAAC)**
- 3. Stateless DHCPv6**
- 4. Stateful DHCPv6**
- 5. DHCPv6 Prefix Delegation(DHCPv6-PD)**



Stateless Address Auto-configuration (SLAAC)

- The network should have at least one IPv6 router configured to send periodic Router Advertisements (RA) announcements.
- IPv6 host when connected to the network sends a ICMPv6 Router Solicit (RS) message and picks up ICMPv6 RA as a response from IPv6 router.
- The IPv6 host uses a combination of IPv6 prefix received in RA message and its link layer address to form a IPv6 address.
- In ICMPv6 RA, the 'M' bit indicating managed address configuration bit would be set to zero, thus indicating IPv6 Host to perform SLAAC



Why DHCPv6 when IPv6 Stateless Auto-configuration exists ?

- Stateless auto-configuration only configures addresses; not “other configuration” information (DNS servers, domain search list, etc)
- Stateless auto-configuration is “one-size fits all”
 - Addresses can not be selectively assigned
 - Policies can not be enforced about clients allowed addresses



Stateless DHCPv6

- Stateless DHCPv6 uses SLAAC to assign one or more IPv6 addresses to an interface and utilizes DHCPv6 to receive "additional parameters" like Domain Name System (DNS) server addresses etc
- On network supporting a large number of hosts, this could mean a significant reduction in DHCPv6 messages
- Stateless DHCPv6 can be initiated by the network administrator by setting the "O" bit in RA messages sent to the client



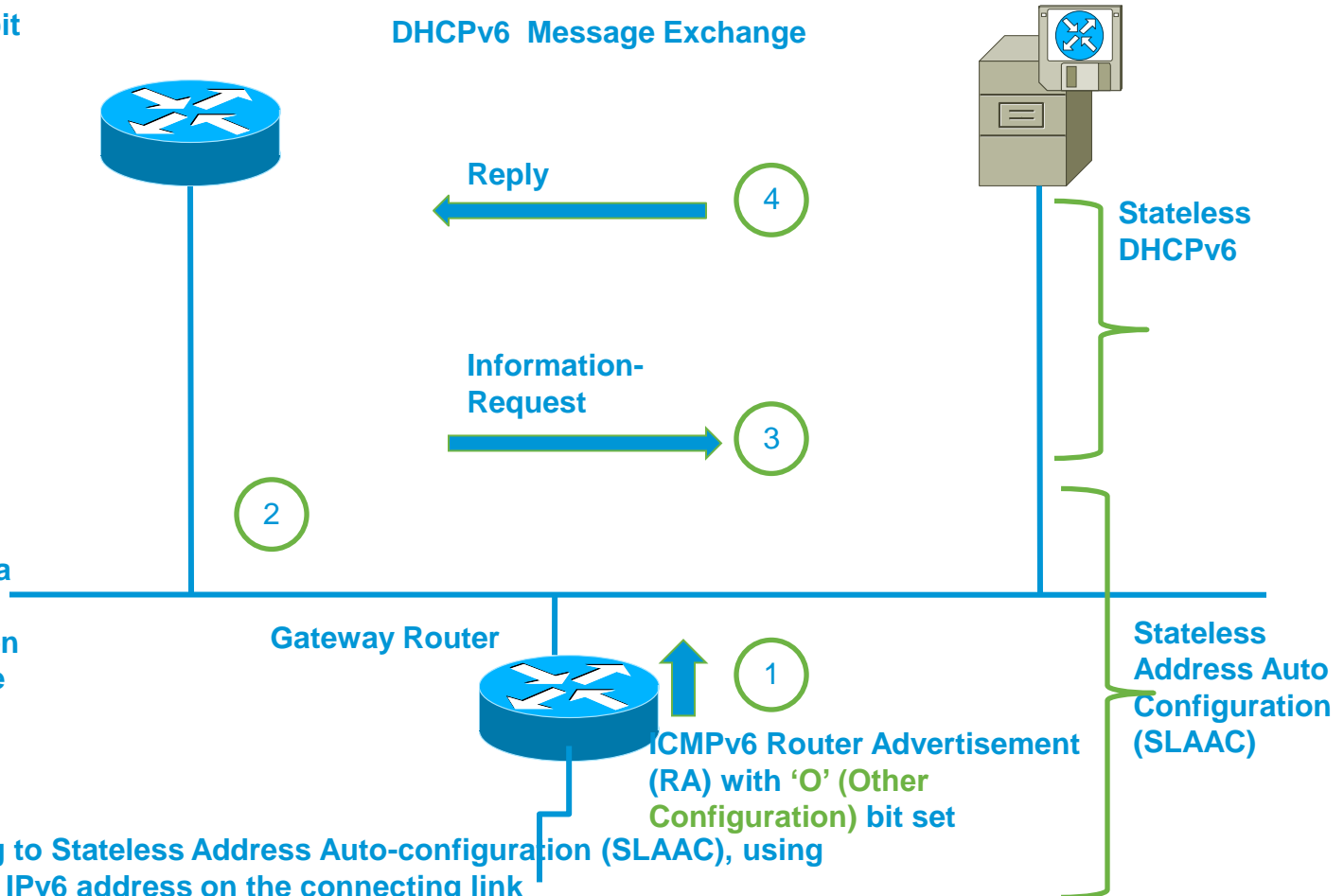
Stateless DHCPv6 – Address allocation

1. The directly connected router periodically sends out RA with 'O' bit set on the link.
2. The IPv6 host uses a combination of IPv6 prefix learned via RA message and its link layer address to assign an IPv6 address on the link.
3. IPv6 client sends an Information-Request message requesting configuration parameters like DNS server address.
4. DHCPv6 server sends a Reply message containing configuration parameters in response to an Information-Request towards the client.

DHCPv6 Client

DHCPv6 Message Exchange

DHCPv6 Server



Note: Steps 1 & 2 belong to Stateless Address Auto-configuration (SLAAC), using which the clients assign IPv6 address on the connecting link

Stateful DHCPv6

- Stateful DHCPv6 offers the capability of automatic allocation of reusable network addresses and additional configuration flexibility.
- Stateful DHCPv6 can be initiated by the network administrator by setting the “M” bit in RA messages sent to the client
- This mechanism is known as Stateful DHCPv6 because, the DHCPv6 server does keep track of the client address bindings



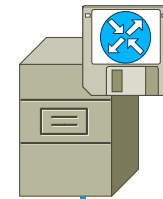
Stateful DHCPv6 – Address allocation

1. The directly connected router periodically sends out RA 'M' bit set on the link.
2. IPv6 client sends a Solicit message to locate servers.
3. DHCPv6 server sends an Advertise message to indicate that it is available for DHCPv6 service.
4. IPv6 client sends a Request message to request configuration parameters like DNS server and IPv6 address from the DHCPv6 server.
5. DHCPv6 server sends a Reply message containing assigned IPv6 addresses and configuration parameters in response to client's request.

DHCPv6 Client

DHCPv6 Message Exchange

DHCPv6 Server



Stateful
DHCPv6

Gateway Router



ICMPv6 Router Advertisement
(RA) with 'M' (Managed) bit set

Note: Steps 3 & 4 would be skipped if IPv6 client uses Rapid Commit option in Solicit message

Stateless DHCPv6

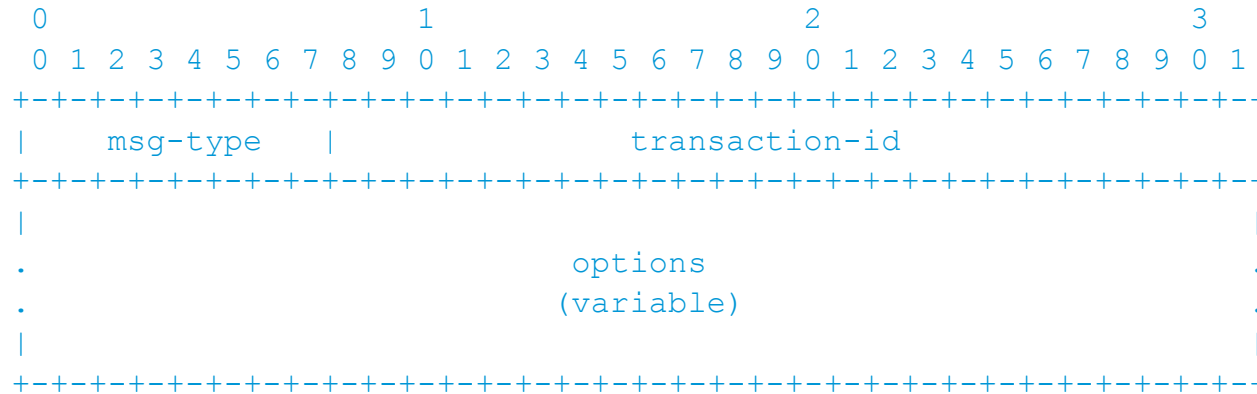
- Stateless provides no means to differentiate clients, hence used in networks which do not want to keep track or monitor IPv6 address assignment
- Using stateless DHCPv6 means that the DHCPv6 server doesn't need to keep track of any state of assigned IPv6 addresses, hence there is no need for state refreshment

Stateful DHCPv6

- Stateful DHCPv6 is used in managed networks where the DHCPv6 server needs to keep track of the client address bindings for various reasons e.g. security etc
- State maintenance/refreshment is needed for keeping track of the assigned IPv6 address
- Routers and residential gateways (RG) often need prefix delegation (DHCPv6-PD) supported ONLY via stateful DHCPv6

DHCPv6 – Client/Server Messages

- Basic message format (UDP, ports 546 and 547)



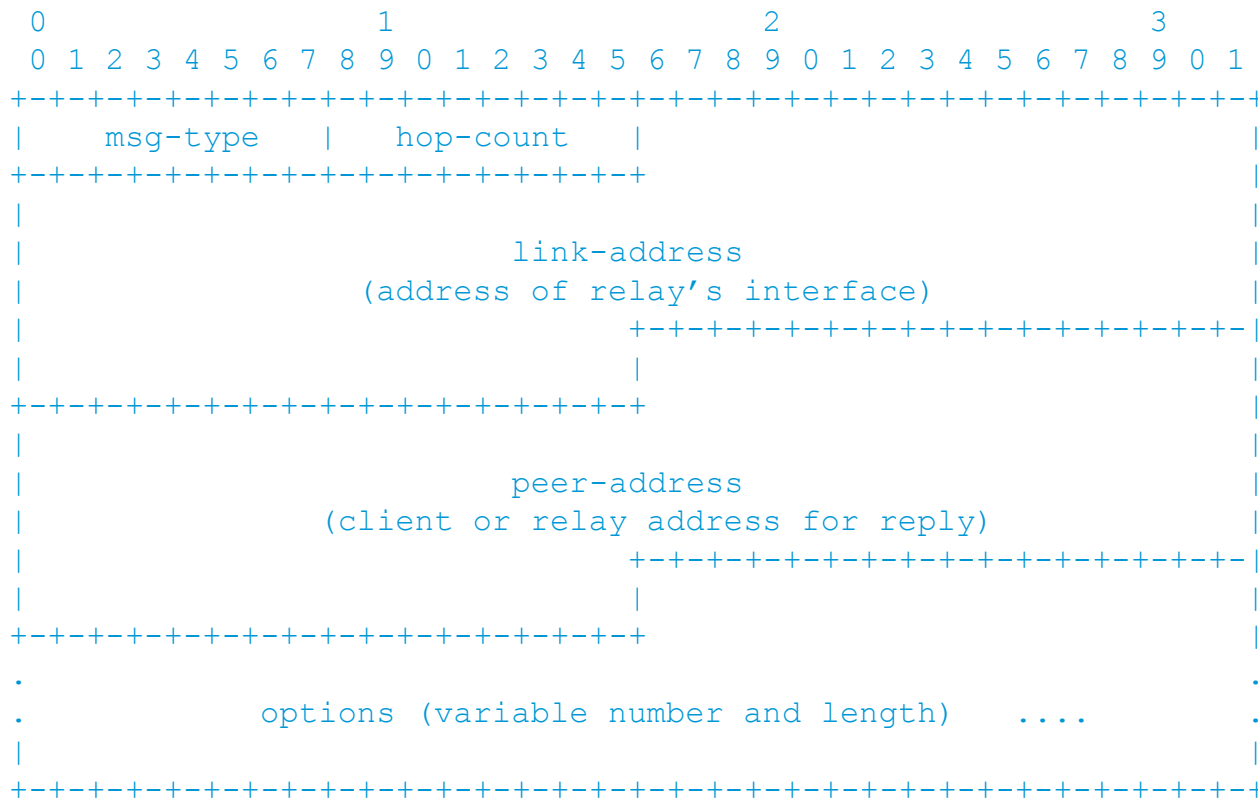
- Message Types

- Client => Server: Solicit, Request, Confirm, Renew, Rebind, Release, Decline, Information-Request
 - Server => Client: Advertise, Reply, Reconfigure
 - Relay => Relay/Server: Relay-Forward
 - Server/Relay => Relay: Relay-Reply
- Options used to carry all data (minimal fixed fields)

DHCPv6 – Relay Messages

- Relay message format

msg-type is Relay-Forward or Relay-Reply



Comparison between DHCPv6 & DHCPv4 Message Exchange

DHCPv6 Messages

- The client uses “**link-local**” address to send and receive DHCPv6 messages.
- Client, server and relay uses either reserved **multicast** or unicast address for exchanging DHCPv6 messages.

DHCPv4 Messages

- The client uses either “**0.0.0.0**” or **assigned IPv4 address** to send and receive DHCPv4 messages.
- Client, server and relay uses either **broadcast** or unicast address for exchanging DHCPv4 messages.



Comparison between DHCPv6 & DHCPv4 Message Types

DHCPv6 Message Type	DHCPv4 Message Type
Solicit (1)	DHCPDISCOVER
Advertise (2)	DHCPOFFER
Request (3), Renew (5), Rebind (6)	DHCPREQUEST
Reply (7)	DHCPACK / DHCPNAK
Release (8)	DHCPRELEASE
Information-Request (11)	DHCPINFORM
Decline (9)	DHCPDECLINE
Confirm (4)	none
Reconfigure (10)	DHCPFORCERENEW
Relay-Forw (12), Relay-Reply (13)	none

Stateful DHCPv6

Stateful is used in networks where the DHCPv6 server needs to keep track of the client address bindings for various reasons.

SLAAC

SLAAC is the simplest and most widely used way of IPv6 address assignment

Stateless DHCPv6

Stateless DHCPv6 compliments SLAAC and therefore are often used in conjunction.

Summarizing



DHCPv6 Relay Agent

- In both enterprise and broadband deployments, often DHCPv6 server serving large number of DHCPv6 clients, would reside in the server farm and hence the two would not be connected directly.
- Relay agents are used to **relay** DHCPv6 messages between client and server when they are not connected to the same link and hence is not directly reachable.
- Relay agents enhances manageability and scalability in a network.

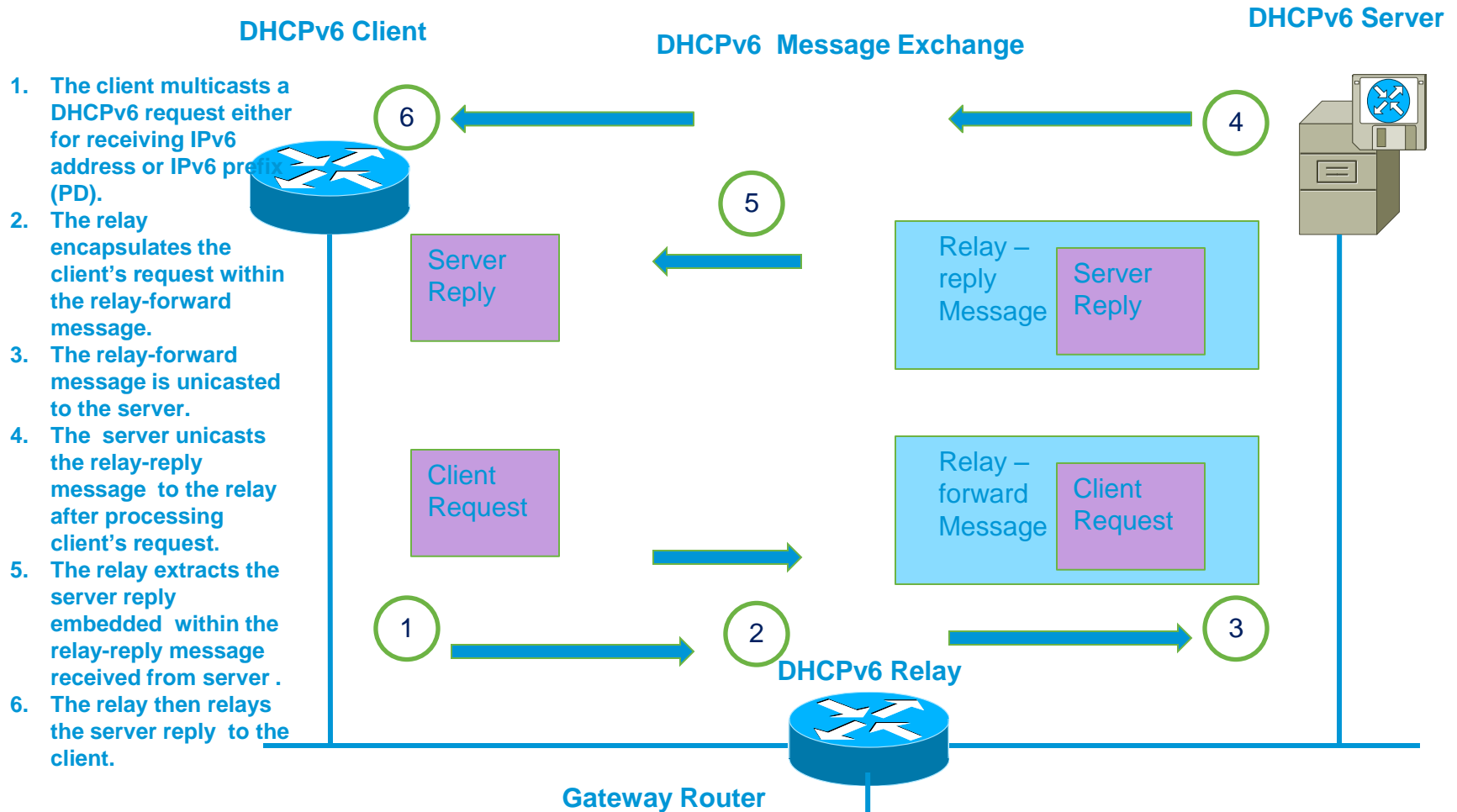


DHCPv6 Relay Agent

- The relay agents encapsulates the received messages from the directly connected DHCPv6 client, and forward these encapsulated DHCPv6 packets towards the DHCPv6 server.
- In the opposite direction, the relay agent decapsulates the packets received from the central DHCPv6 Server and forwards the same towards the client.



DHCPv6 Relay Agent



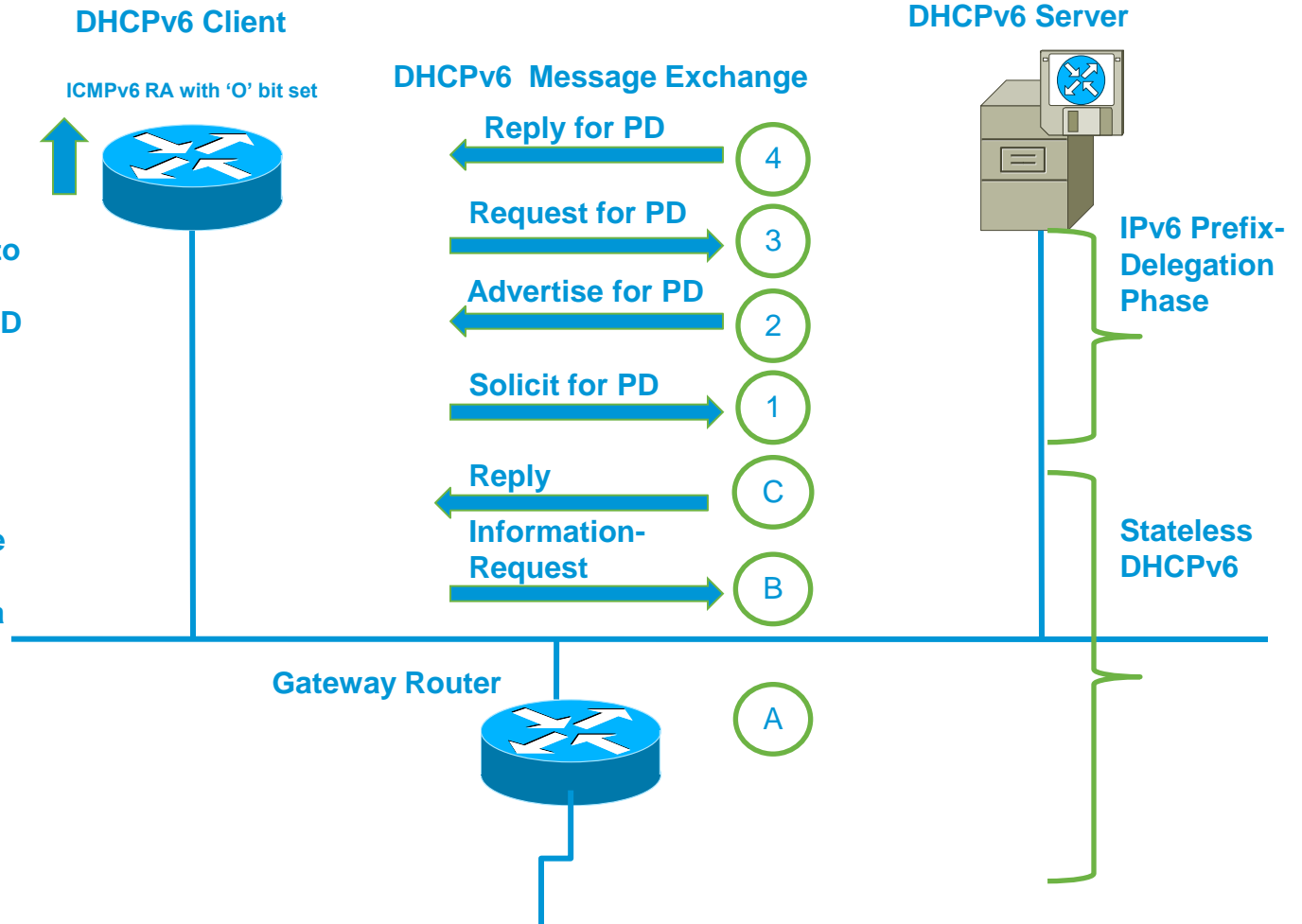
DHCPv6 Prefix Delegation(DHCPv6-PD)

- DHCPv6 enables IPv6 prefix delegation, through which an Internet Service Provider (ISP) can automate the process of assigning IPv6 prefixes to a customer for use within the customer's network.
- Once the ISP has delegated prefixes to a customer, the customer may further subnet and assign prefixes to the links in the customer's network



DHCPv6 Prefix Delegation (DHCPv6-PD)

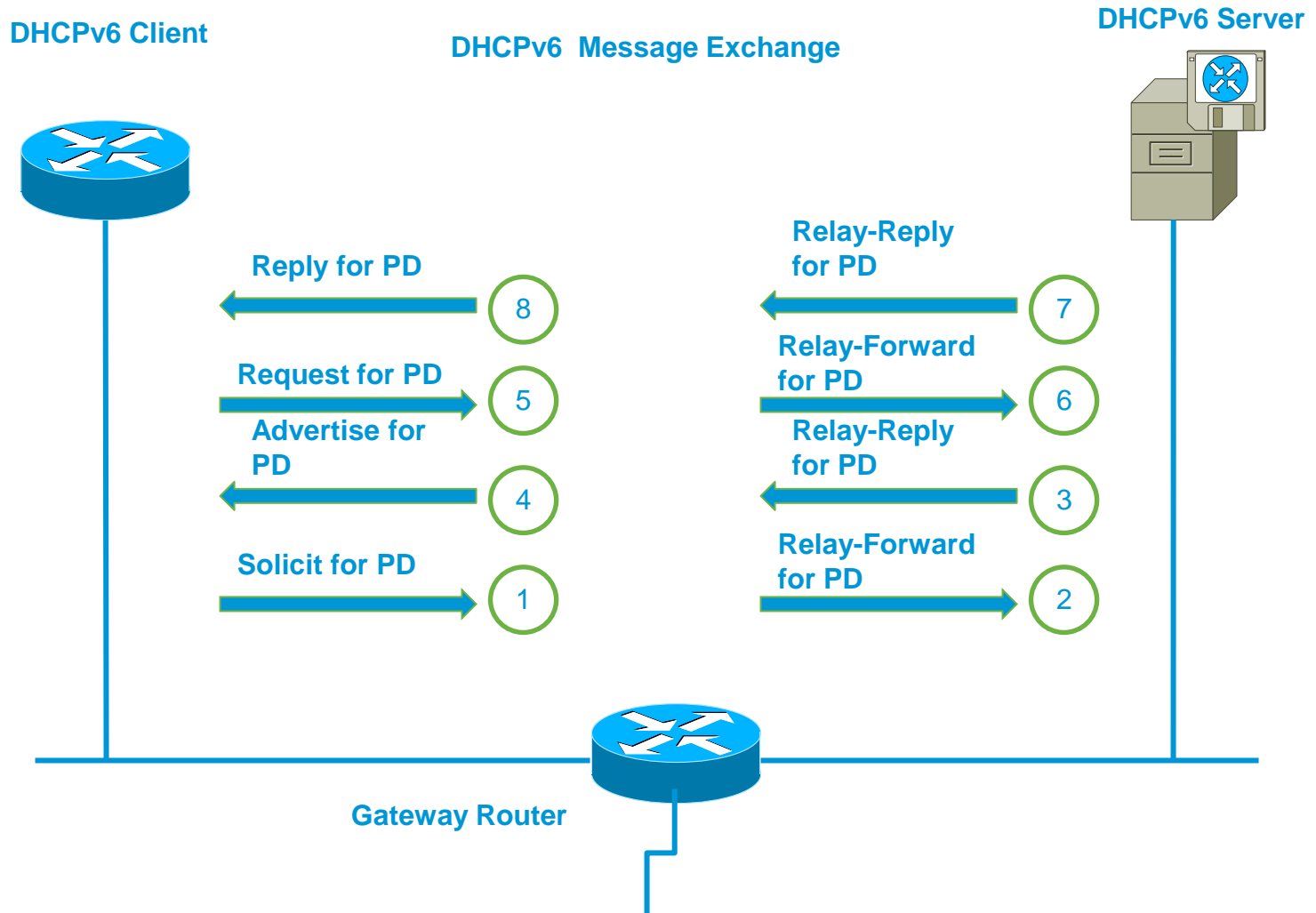
1. IPv6 client sends a Solicit message to locate servers for receiving IPv6 prefix (PD).
2. DHCPv6 server sends an Advertise message to indicate that it is available for DHCPv6-PD service.
3. IPv6 client sends a Request message for receiving PD from the DHCPv6 server which responded in the above step.
4. DHCPv6 server sends a Reply message containing assigned IPv6 prefix (PD) in response to client's request.



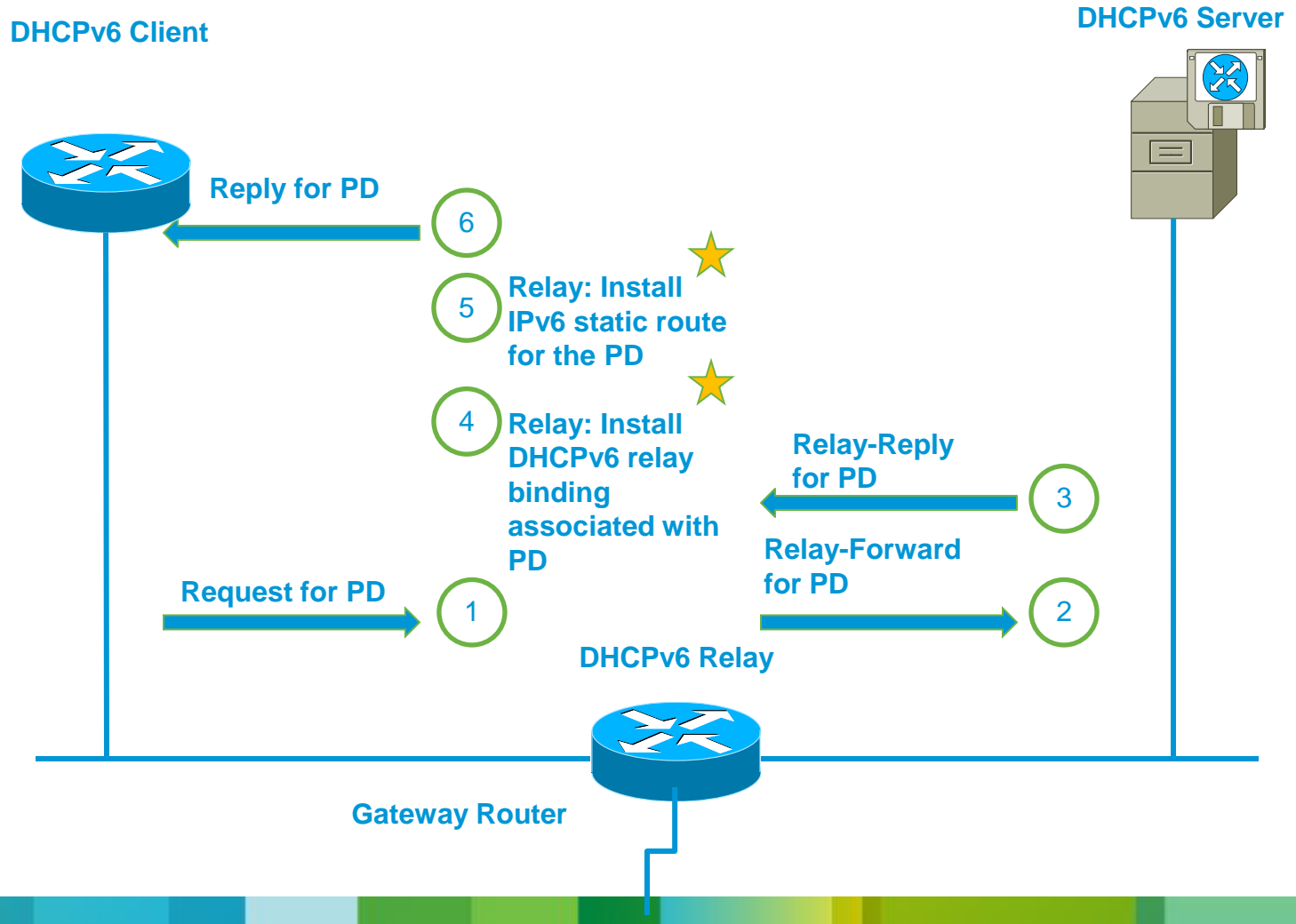
Note: Steps A, B & C belong to Stateless DHCPv6, already explained in the previous slides

Client, Server and Relay Agent Interaction for DHCPv6-PD

- ❑ Relay agents exchange DHCPv6 messages between servers and clients when the two are not connected to the same link.
- ❑ There are two relay agent messages viz. Relay-forward & Relay-reply.
- ❑ In a Relay-forward message, the received message is relayed verbatim to the next relay agent or server.
- ❑ In a Relay-reply message, the message received from server is copied and relayed to the relay agent or client.



DHCPv6 Relay Agent Notification for Prefix Delegation (PD)



Q & A

