IPV6 TUTORIAL

Ron Bonica – Juniper Networks

Engineering Simplicity

MOTIVATION

Big

Address

Space

And a few

other things

Addressing Architecture

RFC 4291, February 2006

IPV6 ADDRESSES

- IPv6 addresses are 128-bit identifiers for interfaces and sets of interfaces
- This definition is frequently stretched by encoding other things in addresses
  - Extra DiffServ codepoints
  - Segment Routing functions and arguments

TEXTUAL REPRESENTATION

- The preferred form is x:x:x:x:x:x:x:x, where the 'x's are one to four hexadecimal digits of the eight 16-bit pieces of the address
  - ABCD:EF01:2345:6789:ABCD:EF01:2345:6789
  - 2001:DB8:0:0:8:800:200C:417A
- It is not necessary to write the leading zeros in an individual field, but there must be at least one numeral in every field
- The use of "::" indicates one or more groups of 16 bits of zeros
  - 2001:DB8:0:0:8:800:200C:417A
  - 2001:DB8::8:800:200C:417A

TEXTUAL REPRESENTATION (CONTINUED)

- An alternative form that is convenient when dealing with a mixed environment of IPv4 and IPv6 nodes is x:x:x:x:x:x:d.d.d.d, where the 'x's are the hexadecimal values of the high-order pieces of the address, and the 'd's are the decimal values of the low-order pieces of the address (standard IPv4 representation).
  - 0:0:0:0:0:0:13.1.68.3
  - ::13.1.68.3
- There are many ways to represent some IPv6 addresses
  - See RFC 5925 for preference recommendation

ADDRESS TYPES

- Unicast - An identifier for a single interface.
  - A packet sent to a unicast address is delivered to the interface identified by that address.
- Anycast - An identifier for a set of interfaces.
  - A packet sent to an anycast address is delivered to one of the interfaces identified by that address.
- Multicast - An identifier for a set of interfaces.
  - A packet sent to a multicast address is delivered to all interfaces identified by that address.
- There are no broadcast addresses in IPv6.

UNICAST ADDRESS SCOPES (I)

- Global Unicast Address (GUA)
  - Globally unique
  - Fully routable
- Link-Local Address
  - Unique to the link
  - Routers will not forward packets that have Link-Local source or destination addresses
  - Useful for on-link protocols (e.g., Neighbor Discovery)

UNICAST ADDRESS SCOPES (II)

- Unique Local Unicast Address (ULA)
  - Guaranteed to be unique within a domain
  - Probably globally unique, but no guarantees
  - Used to isolate subnetworks from the global Internet
- Unspecified
  - Indicates that the node has not been assigned an address yet
  - Used in Neighbor Discovery (ND)
  - Routers will not forward packets that have Unspecified source or destination addresses

UNICAST ADDRESS SCOPES (III)

- Loopback
  - Devices use this address to send packets to themselves
  - Should not appear on the wire
  - Routers will not forward packets that have Loopback source or destination addresses

UNICAST ADDRESS FORMATTING (I)

- Global Unicast Address (GUA)
  - Standard
    - Global Routing Prefix (N bits)
    - Subnet ID (M bits)
    - Interface ID (128 – N – M bits)
    - N + M = 64, except for point-to-point interfaces [RFC 6164]
  - IPv4 Mapped IPv6 Address
    - Zeros (80 bits)
    - FFFF (16 bits)
    - IPv4 Address (32 bits)

UNICAST ADDRESS FORMATTING (II)

- Unique Local Unicast Address (ULA)
  - FC00 (7 bits)
  - L (1 bit) Method by which Global ID was generated
  - Global ID (40 bits)
  - Interface ID (64 bits)
- Link-local Address
  - FE80 (10 bits)
  - Zeros (54 bits)
  - Interface ID (64 bits)

UNICAST ADDRESS FORMATTING (III)

- Loopback
  - Value ::1
- Unspecified
  - Value ::0

MULTICAST ADDRESS FORMAT

- Value FF (8 bits)
- Flags (4 bits)
  - Rendez-vous address embedded in address
  - Prefix (i.e., address is assigned based on prefix)
  - Transient (i.e., not well-known)
- Scope (4 bits)
- Group ID (nominally 112 bits)
  - Shorter if R or P bits are set

WELL-KNOWN MULTICAST ADDRESSES

- All Nodes
  - FF01::1
  - FF02::1
- All Routers
  - FF01::2
  - FF02::2
  - FF05::2

ADDRESSING MODEL

- IPv6 addresses of all types are assigned to interfaces, not nodes
- All interfaces are required to have at least one Link-Local unicast address
- A single interface may also have multiple IPv6 addresses of any type (unicast, anycast, and multicast) or scope
- Unicast addresses with a scope greater than link-scope are not needed for interfaces that are not used as the origin or destination of any IPv6 packets to or from non-neighbors
- Exception for Link Aggregate Groups (LAG)

Protocol Specification

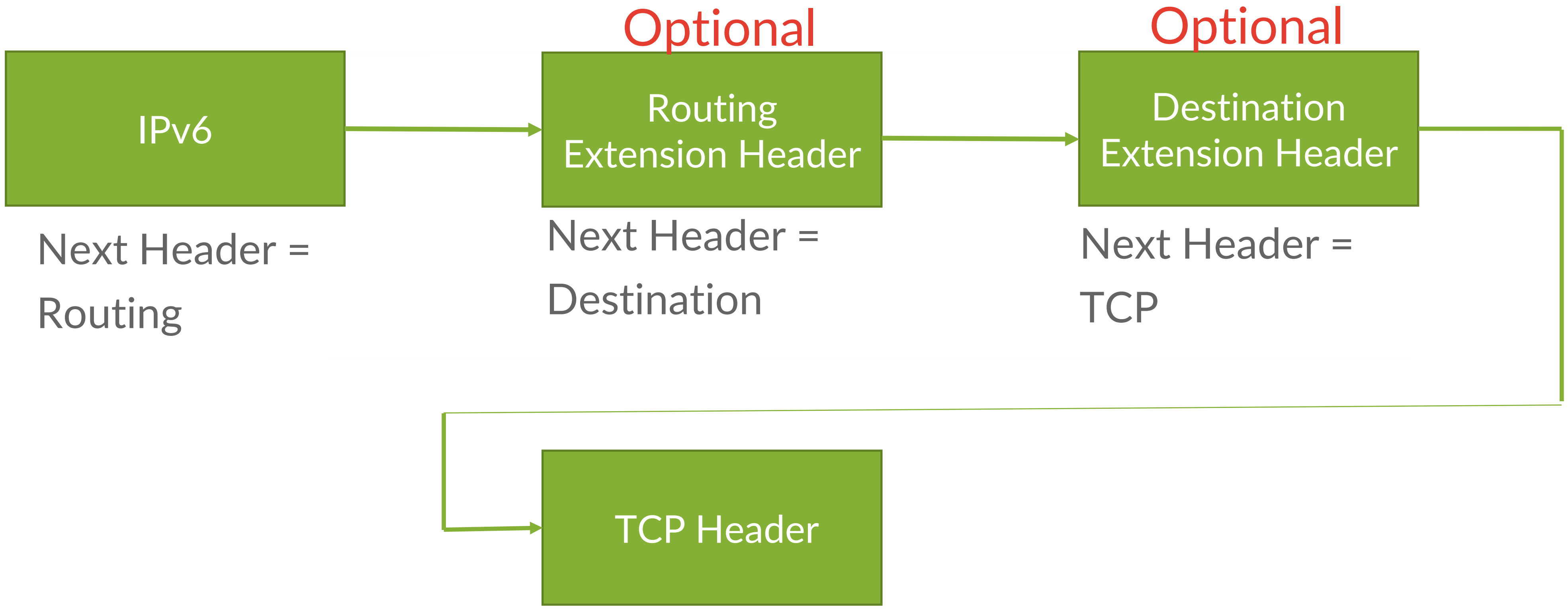
RFC 8200, July 2017

IPV6 SERVICES

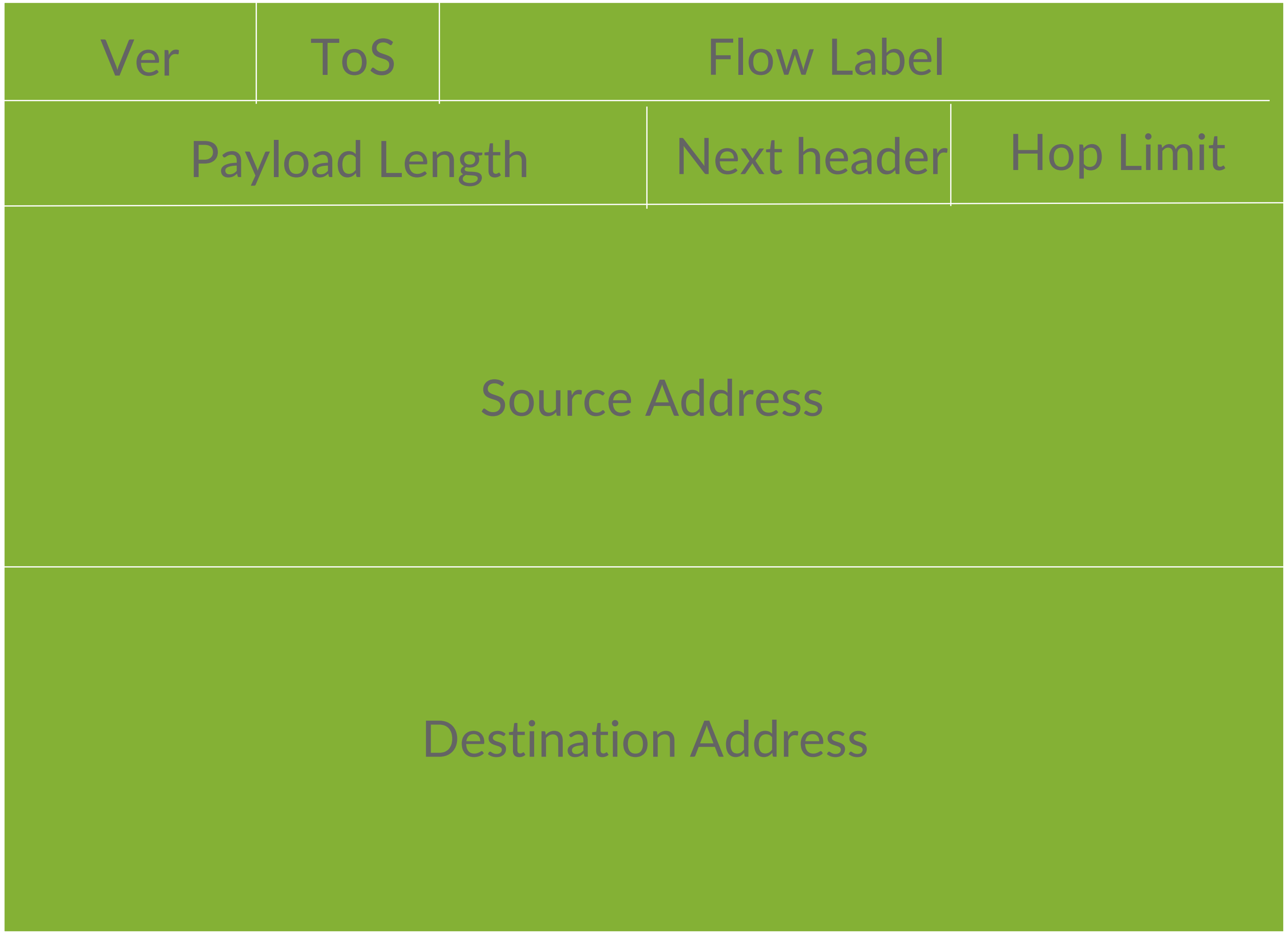
Basic	Extended
• Next-hop identification and forwarding • DifServ and ECN • Flow Identification (for ECMP) • Loop prevention • Delivery to upper-layer protocols	• Traffic steering • Fragmentation • Authentication • Encryption • Future extensions Optional

OPTIONAL EXTENSION HEADERS SUPPORT EXTENDED SERVICES

IPv6 Header Chain



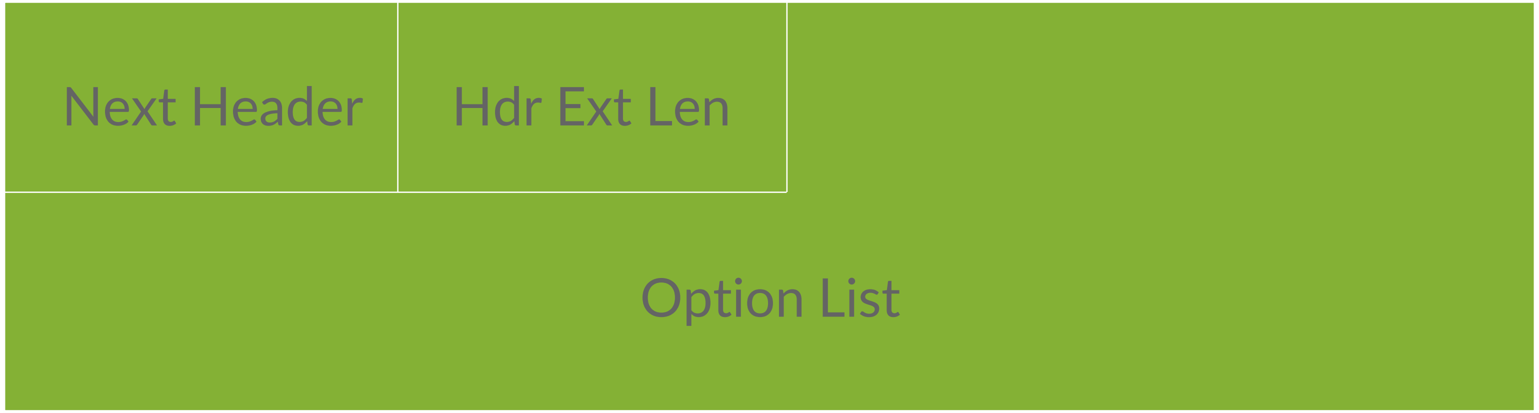
IPV6 HEADER SUPPORTS BASIC SERVICES



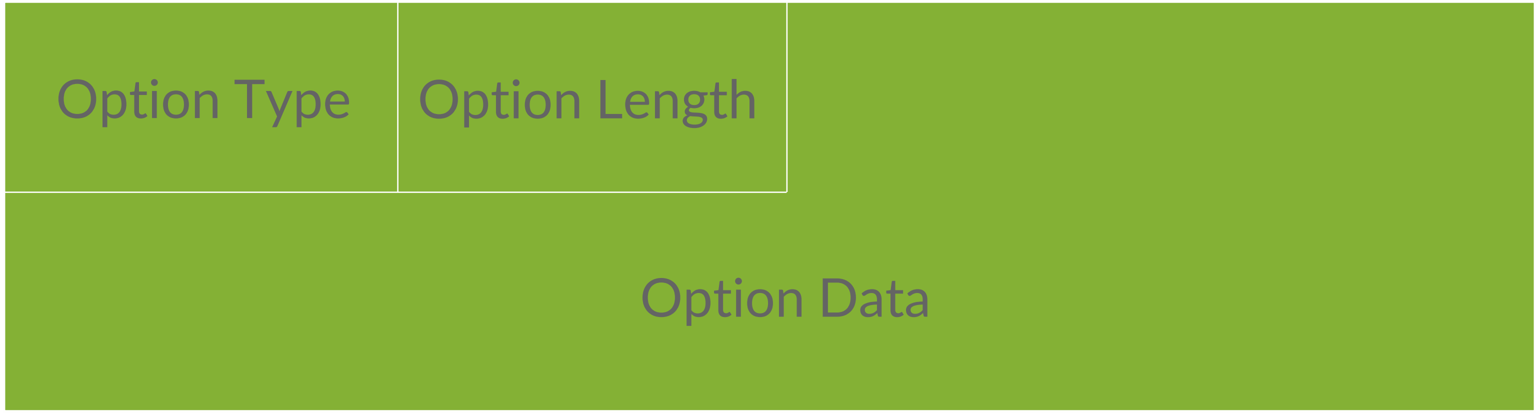
IPV6 FACILITATES ECMP LOAD BALANCING

- Three-tuple provides sufficient entropy for ECMP Load Balancing
  - Source Address
  - Destination Address
  - Flow Label
- All three can be found in fixed positions in the basic IPv6 header
- An implementation may load-balance among ECMPs using a five-tuple that includes source and destination port
  - But this requires the implementation to parse more of the header chain
  - Not always possible

THE HOP-BY-HOP AND DESTINATION OPTIONS EXTENSION HEADERS



OPTIONS LIST ENTRY (I.E., AN OPTION)



OPTION TYPE

- First two bits (ACT bits) indicate the required behavior when the processing node does not recognize the option
  - 00: Skip the option and process the next
  - 01: Discard the packet
  - 10: Discard the packet and send an ICMP message
  - 11: Discard the packet and, if its destination address was not multicast, send an ICMP message
- Third bit (CHG bit) indicates whether Option data can change on route to the packet's destination

IGNORING THE HOP-BY-HOP OPTIONS HEADER

- Nodes may ignore the Hop-by-hop Extension header
- “While [RFC2460] required that all nodes must examine and process the Hop-by-Hop Options header, it is now expected that nodes along a packet's delivery path only examine and process the Hop-by-Hop Options header if explicitly configured to do so.” – RFC 8200

ROUTING EXTENSION HEADER

Next Header	Hdr Ext Len	Routing Type	Segments Left
Type Specific Data			

- Used for traffic steering
- Many Routing Types defined
- Segments Left indicates number of segments yet to be visited
- - Type specific data includes a list of addresses to be copied to the basic IPv6 header destination address

FRAGMENT HEADER

Next Header	Reserved	Fragment Offset	RES	M
Identification				

- Contains information needed for packet reassembly

IPV6 MAKES FRAGMENTATION FRAGILE

- An IPv6 packet can only be fragmented at its source node
- The source node relies on PMTUD to determine whether fragmentation is required
- PMTUD is itself fragile
  - Filtered or lost ICMPv6 Packet Too Big messages
  - So, the source node can send packets into PMTU black holes
- Workaround
  - IPv6 links are required to support 1280-byte MTUs
  - Hosts generally refrain from sending packets larger than 1280-byte
  - Fragmentation is rarely required

EXTENSION HEADER PLACEMENT

Extension headers are arranged in an order that minimizes resources required to process them

IPV6 EXTENSION HEADERS

Listed in the order that they appear in a packet

Name	Contents	Processing node
HBH	Any optional information	Every node along a packet's delivery path
Destination Options	Any optional information	Every destination node
Routing	A list of waypoints (destination nodes) along the packet's delivery path	Every destination node
Fragment	Fragmentation / reassembly information	Ultimate destination node
Authentication	Authentication information	Ultimate destination node
Encrypted Security Payload (ESP)	Encrypted payload and security parameters	Ultimate destination node
Destination Options	Any optional information	Ultimate destination node

IPV6 EXTENSIBILITY: THE BRIGHT SIDE

- IPv6 can be extended to accommodate any future requirement
  - Add information to existing extension headers
- No need for more extension headers
  - Just add new option to the hop-by-hop or destination options header
- There may never be a "next version" of the Internet Protocol
  - Because IPv6 is extensible

IPV6 EXTENSIBILITY: THE DARK SIDE

- Most router implementations cannot process extension headers on the fast path
  - Punt to the slow path
  - Denial of Service Vulnerability
- Many operators block packets with extension headers [RFC 7872]
- IETF currently redefining the Hop-by-hop options extension header to overcome vulnerability
  - Stay tuned

ICMPV6

- Many holdovers from IPv4
  - Destination Unreachable
  - Packet Too Big
  - Time Exceeded
- Many new messages to support
  - Neighbor discovery
  - IPv6 mobility
  - Other stuff

Neighbor Discovery

RFC 4861, September 2007

NEIGHBOR DISCOVERY REPLACES ARP

Neighbor Discovery Replaces ARP in IPv6

IPv6 Neighbor Table replaces ARP Table

Five new ICMPv6 messages

ICMPV6 MESSAGES

- Router Solicitation
  - Solicits one or more Router Advertisement messages
- Router Advertisement
  - Advertises an on-link router and its attributes
- Neighbor Solicitation
  - Solicits one or more Neighbor Advertisement messages
- Neighbor Advertisement
  - Advertises an on-link neighbor and its attributes
- Redirect
  - Redirects a destination address to a better next-hop

ROUTER SOLICITATION MESSAGE

- IPv6 fields
  - Source Address: An IP address assigned to the sending interface, or the unspecified address if no address is assigned to the sending interface
  - Destination Address: Typically, the all-routers multicast address
  - Hop count: 255
- ICMP fields
  - Type: 133; Code: 0; Checksum
  - Source link-layer address: The link-layer address of the sender, if known. Must not be included if the Source Address is the unspecified address. Otherwise, it SHOULD be included on link layers that have addresses.

ROUTER ADVERTISEMENT MESSAGE (I)

- IPv6 fields
  - Source Address: The link-local address assigned to the interface from which this message is sent
  - Destination Address: The Source Address of an invoking Router Solicitation or the all-nodes multicast address
  - Hop count: 255

ROUTER ADVERTISEMENT MESSAGE (II)

- ICMP fields
  - Type: 134; Code: 0; Checksum
  - Cur Hop Limit: The default value that should be placed in the Hop Count field of the IP header for outgoing IP packets
  - M-bit: Indicates that addresses are available via Dynamic Host Configuration Protocol
  - O-bit: Indicates that other configuration information is available via DHCPv6
  - Router Lifetime: The lifetime associated with a default router. Zero if not a default router.
  - Reachable Time: The time that a node assumes a neighbor is reachable after having received a reachability confirmation
  - Retransmit Timer: The time between retransmitted Neighbor Solicitation messages
  - Source Link-layer Address: The link-layer address of the interface from which the Router Advertisement is sent
  - MTU - Link MTU
  - Prefix Information - Prefixes that are on-link and/or are used for SLAAC

NEIGHBOR SOLICITATION MESSAGE

- IPv6 fields
  - Source Address: Either an address assigned to the interface from which this message is sent or the unspecified address
  - Destination Address: Either the solicited-node multicast address corresponding to the target address, or the target address
  - Hop count: 255
- ICMP fields
  - Type: 135; Code: 0; Checksum
  - Target - The IP address of the target of the solicitation. Must not be a multicast address.
  - Source link-layer address - The link-layer address for the sender. Must not be included when the source IP address is the unspecified address.

NEIGHBOR ADVERTISEMENT MESSAGE (I)

- IPv6 fields
  - Source Address - An address assigned to the interface from which the advertisement is sent
  - Destination Address - For solicited advertisements, the Source Address of an invoking Neighbor Solicitation or, if the solicitation's Source Address is the unspecified address, the all-nodes multicast address. For unsolicited advertisements typically the all-nodes multicast address.
  - Hop count -255

NEIGHBOR ADVERTISEMENT MESSAGE (II)

- ICMP fields
  - Type: 136; Code: 0; Checksum
  - Router flag: When set, the R-bit indicates that the sender is a router
  - Solicited flag: When set, the S-bit indicates that the advertisement was sent in response to a Neighbor Solicitation from the destination address
  - Target Address: For solicited advertisements, the Target Address field in the Neighbor Solicitation message that prompted this advertisement. For an unsolicited advertisement, the address whose link-layer address has changed.
  - Target Link-Layer Address: The link-layer address for the target, i.e., the sender of the advertisement

REDIRECT MESSAGE

- IPv6 fields
  - Source Address: The link-local address assigned to the interface from which this message is sent
  - Destination Address: The Source Address of the packet that triggered the redirect
  - Hop count: 255
- ICMP fields
  - Type: 137; Code: 0; Checksum
  - Target Address: An IP address that is a better first hop to use for the ICMP Destination Address
  - Destination Address: The IP address of the destination that is redirected to the target
  - Target Link: Layer Address: The link-layer address for the target
  - Redirected header: As much as possible of the IP packet that triggered the sending of the Redirect

Stateless Auto Address Configuration (SLAAC)

RFC 4862, September 2007

SLAAC VERSUS DHCPV6

A node can be configured by SLAAC or DHCP

Or both

CONCEPT OF OPERATION

- A node forms a link-local address by appending an interface identifier to the well-known link local prefix (i.e., fe800::/64)
- The node performs Duplicate Address Detection (DAD) procedures
  - Send an NS message with target equal to the link-local address
  - If no response, proceed
- The node sends an RS message and receives an RA message
- The node forms a GUA address by appending an interface identifier to a prefix received in the RA message
- The node performs Duplicate Address Detection (DAD) procedures
  - Send an NS message with target equal to the GUA address


THANK YOU
JUNIPER NETWORKS
Engineering
Simplicity