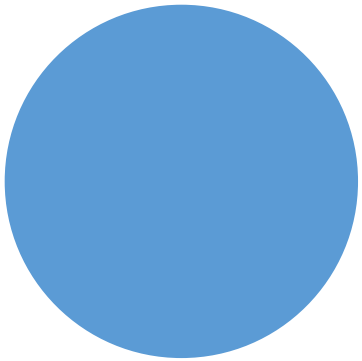



IPv6 and UDP

an IoT perspective

Transport Layer for IoT

TCP vs. UDP

	TCP	UDP
Stands for	Transmission Control Protocol	User Datagram Protocol
Protocol	Connection Oriented	Connectionless
Security	Error Checking and Reporting	Error Checking but NOT Reporting
Data Sending	Slower	Faster
Header Size	20 Bytes	8 Bytes
Implementation	Requires retransmission queues and timers	Lightweight, Stateless
Segments	Acknowledgement	No Acknowledgement
Bulks	Fragmentation and Reordering	No
Typical Applications	Email, HTTP, etc.	VoIP

TCP Applicability for IoT

Claimed issue	Discussion	Potential Solution (future)
Congestion Control	Wired=>Congested Wireless=>Lossy Channel	Experimental
Header overhead	12 bytes greater (2x) than UDP	TCP Header Compression
Long TCP connection infeasible	IoT device sleep periods may render a connection	TCP Tunning
High-latency	Three-way handshake unsuitable for short-lived (alarms, e.g.)	Experimental, TCP Fast Open
High complexity	Initially designed in 1981, when average computer run at few MHz clocks	
RTO unsuitable for IoT		CoCoA

Will it be backwards compatible?

IPv6 Basics

A Brief History of IPv4

- 1981: IP is standardized (RFC 791)
  32-bits address
  ~4.3 billion addresses (2^{32} addresses)

Decimal representation used

4 octets separated by 3 dots

192	168	1	15
1100 0000	1010 1000	0000 0001	0000 1111

1 octet = 8 bits

Classful Unicast IPv4 Addresses

(obsolete)

Address Class	First Octet Range	Mask	Number of Possible Hosts	Number of Hosts per Network	Percentage of total
Class A	0 to 127	/8	128 (2 are reserved)	16'777,214	50%
Class B	128 to 191	/16	16,384	65,534	25%
Class C	192 to 223	/24	2'097,152	254	12,5%

Examples

- 10.0.0.0/8 => Class A
- 140.70.0.0/16 => Class B
- 200.100.50.0/24 => Class C

- Organizations were able to request a pool of address from one three classes, according to their needs

A Brief History of IPv4

- 1981: IP is standardized (RFC 791)
  32-bits address
  4.3 billion addresses, 3 classes
- 1982: TCP/IP for ARPANET.
  “Internet” is defined.
- 1984: Host 1,000
- 1987: Hosts 10,000
- 1988: IANA is funded
  Internet Assigned Numbers Authority
- 1989: Hosts 100,000
- 1991: WWW
- 1992: Hosts 1’000,000
  Prediction of address depletion between 2005 and 2011

Classless Interdomain Routing (CIDR) for IPv4

RFC 1338, 1519

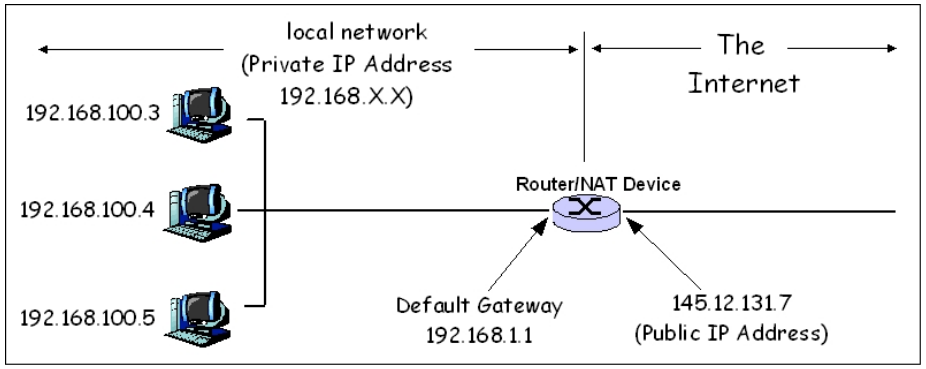
- Problem before CIDR
  - For some organizations, Class C included to few hosts (254), but Class B too many (65,534)
  - Inefficient assignment of IPv4 addresses

- New approach
  - There are no classes => Finer assignment of IPv4 addresses
  - ***Pool of addresses can be assigned with any mask***
  - *Valid assignment now (not before)*
    - *80.128.0.0/**23***

Network Address Translation (NAT) for IPv4

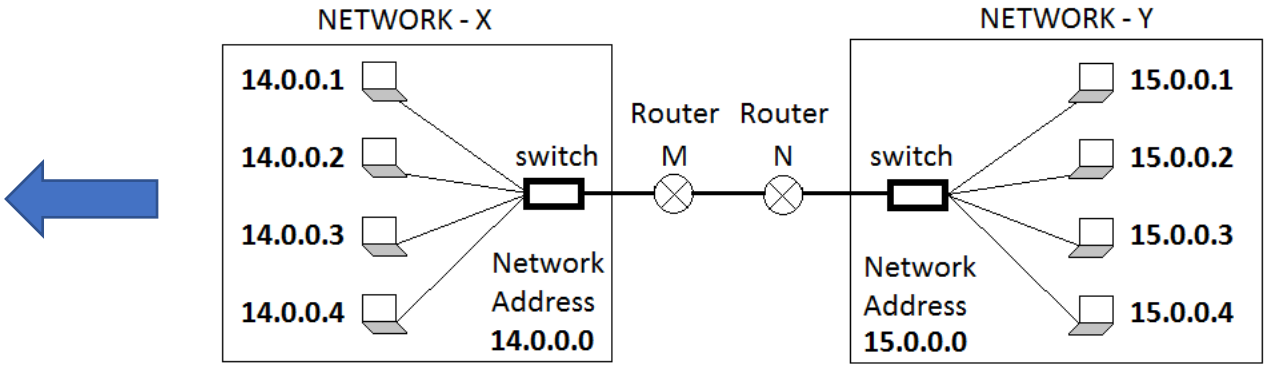
What can we do once they are depleted?

4.3 billion IPv4 addresses vs. 4.6 billion people in Asia (2019)



Creation of private networks sharing a small pool of public IPv4 addresses

Connectivity between two host from different private networks is troublesome

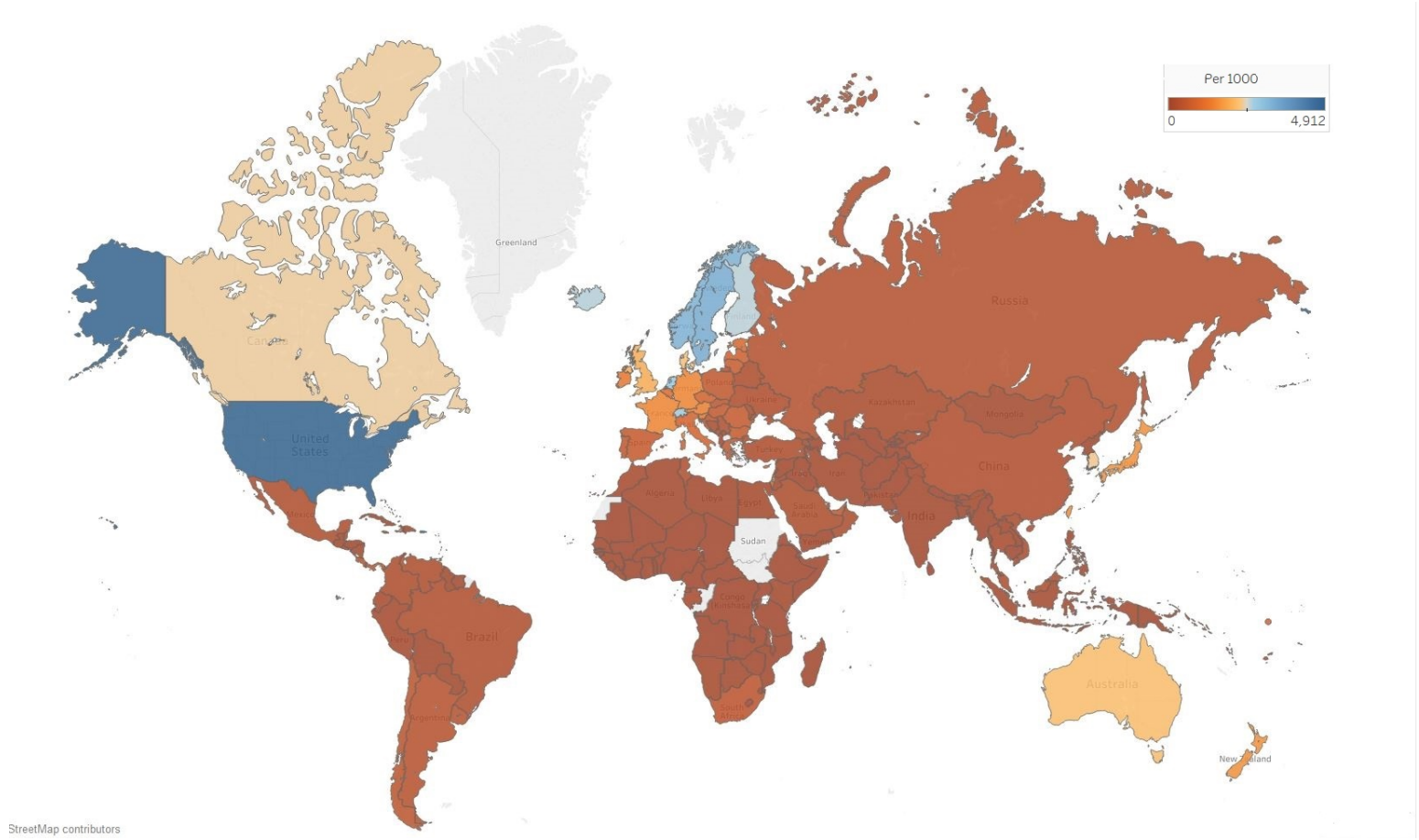


A Brief History of IPv4

- 1993: CIDR (RFC 1519)
  "As the Internet has evolved and grown over in recent years, it has become evident that it is soon to face several serious scaling problems."
- 1994: NAT is standardized (RFC 1631)
  "*The two most compelling problems facing the IP Internet are IP address depletion and scaling in routing.*"
- 1995: IPv6 (RFC 1883)
- 1997: Hosts 19'000,000
- 1998: IPv6 (RFC 2460)
- 31st of January 2011: Last two /8 blocks of IPv4 addresses allocated

- 2017: IPv6 (RFC 8200)

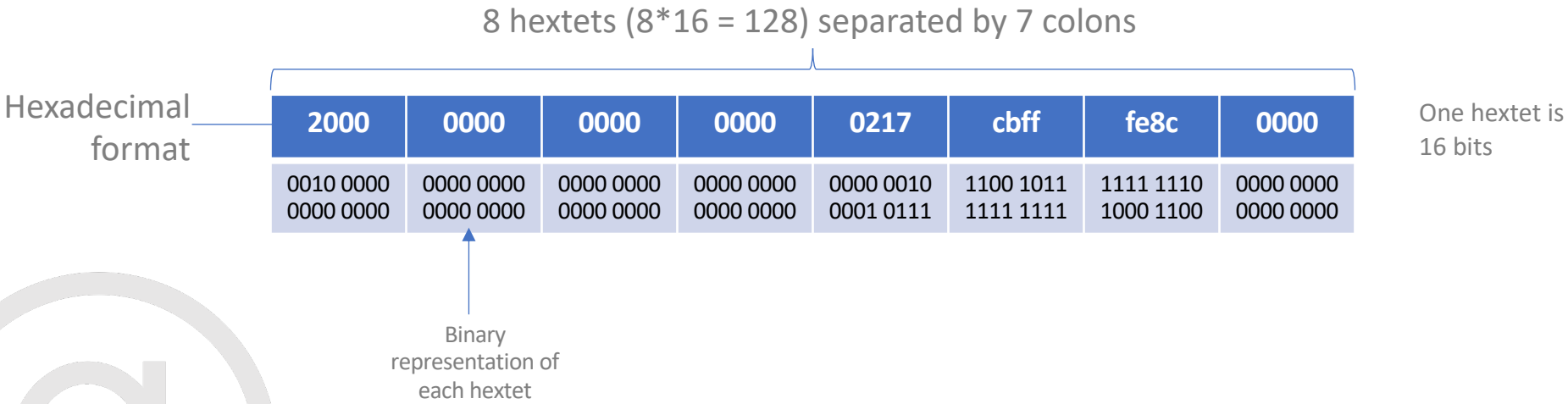
IPv4 address allocation per 1000 inhabitants



IPv6 Addresses

From 32-bits to **128-bits** addresses

2000 : 0000 : 0000 : 0000 : 0217 : cbff : fe8c : 0000



3.4x10^{38}, viz., 340 undecillion, or **340** billion billion billion billion addresses (2^{128} addresses)

Simplified IPv6 address notation rules

2000:0000:0000:0000:0217:cbff:fe8c:0000

1. Omitting leading zeros within blocks → 2000:**0:0:0:217**:cbff:fe8c:**0**

2. Double colons (::) in place of a series of zeros → 2000**::**217:cbff:fe8c:0

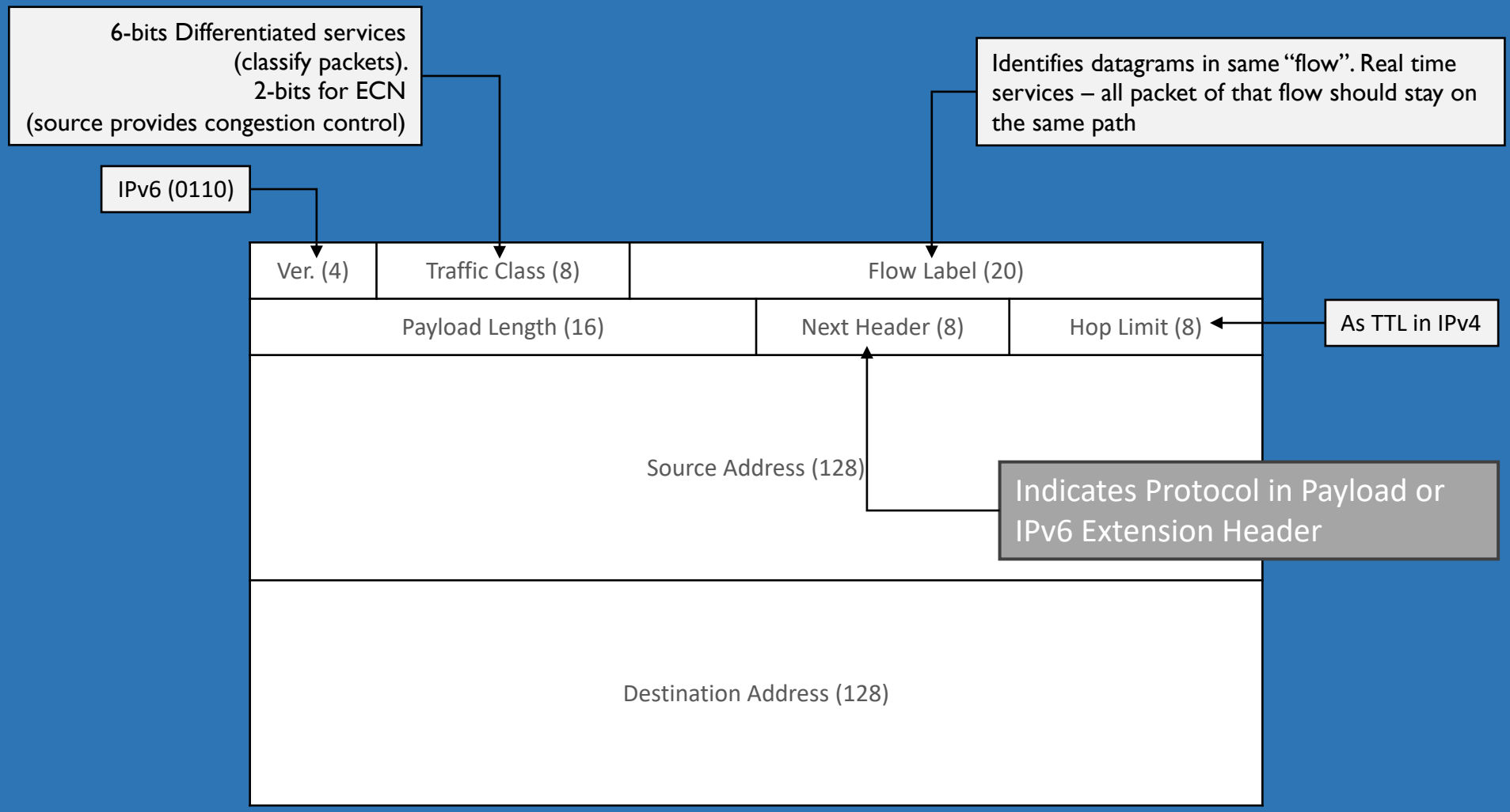
3. Only one double colons allowed: → ~~2000::0217:cbff:fe8c::~~ (invalid)

Why? The double use of :: makes it unclear how many zeros were in each 0 string originally were.

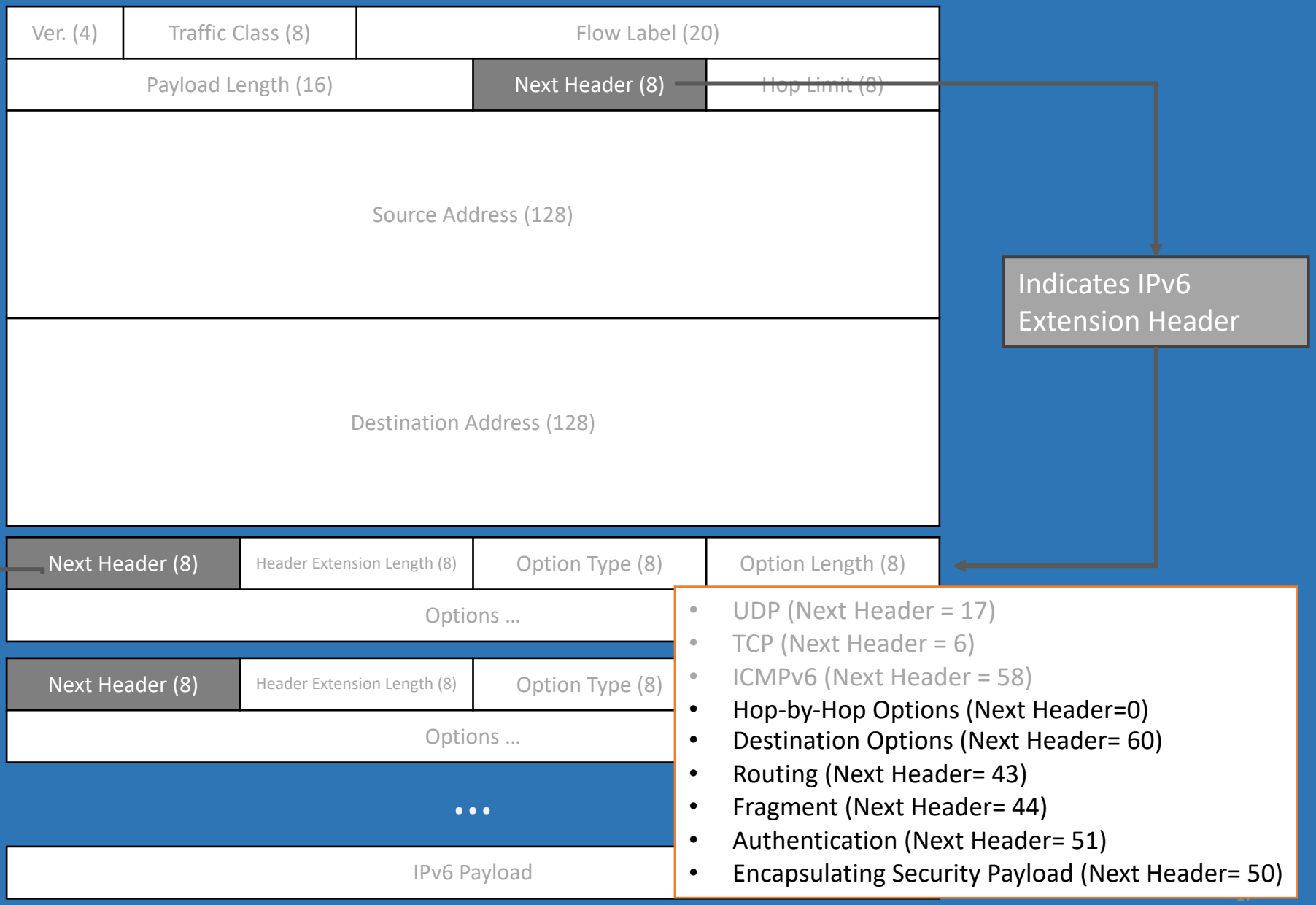
2a03:2880:f003:c07:face:b00c::2

IPv6 Datagram Format

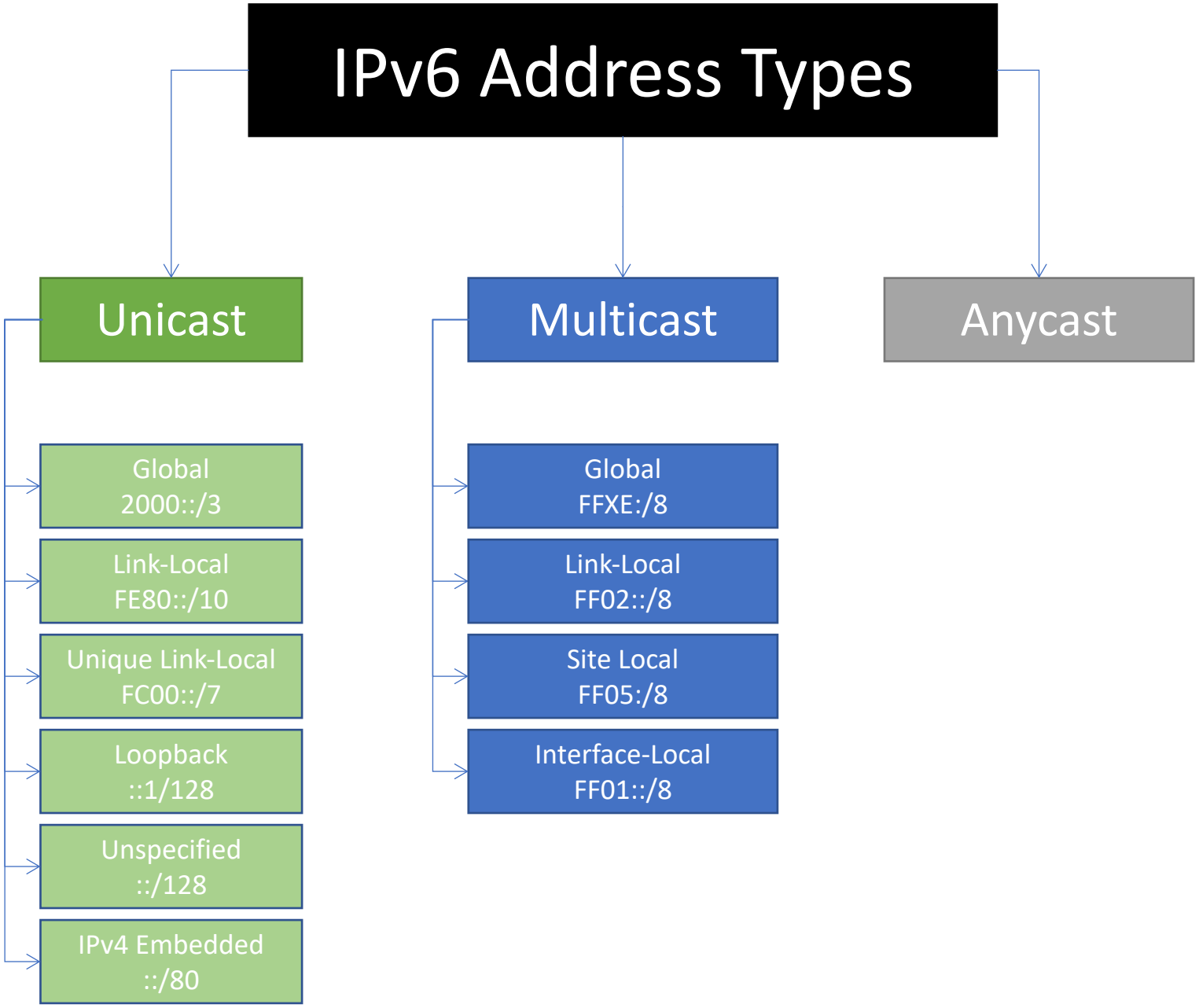
IPv6 Datagram Format



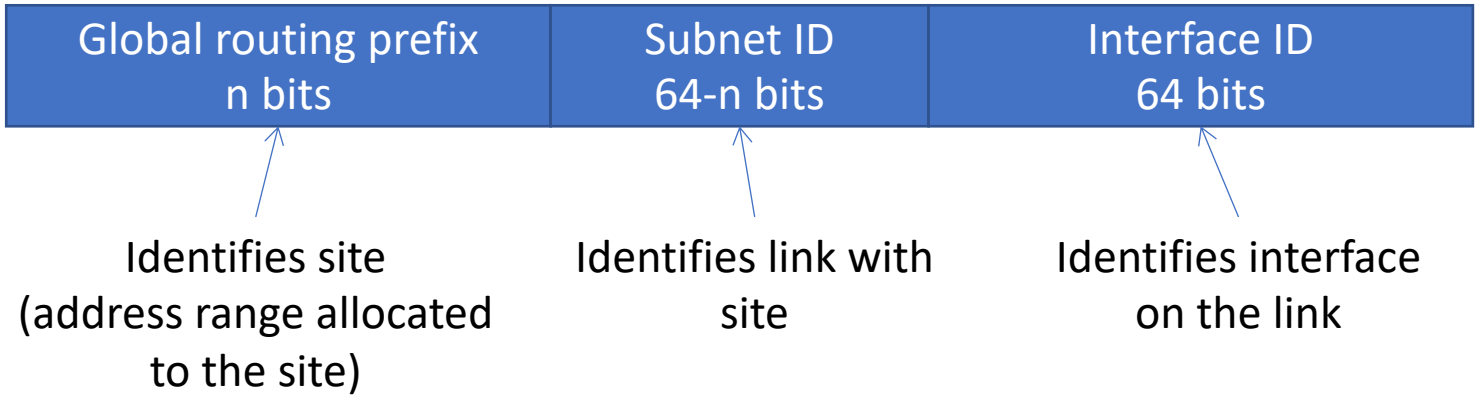
IPv6 Header Next Header



IPv6 Address Types


IPv6 Address Types
Unicast
Multicast
Anycast
Global
2000::/3
Link-Local
FE80::/10
Unique Link-Local
FC00::/7
Loopback
::1/128
Unspecified
::/128
IPv4 Embedded
::/80
Global
FFXE:/8
Link-Local
FF02::/8
Site Local
FF05:/8
Interface-Local
FF01::/8

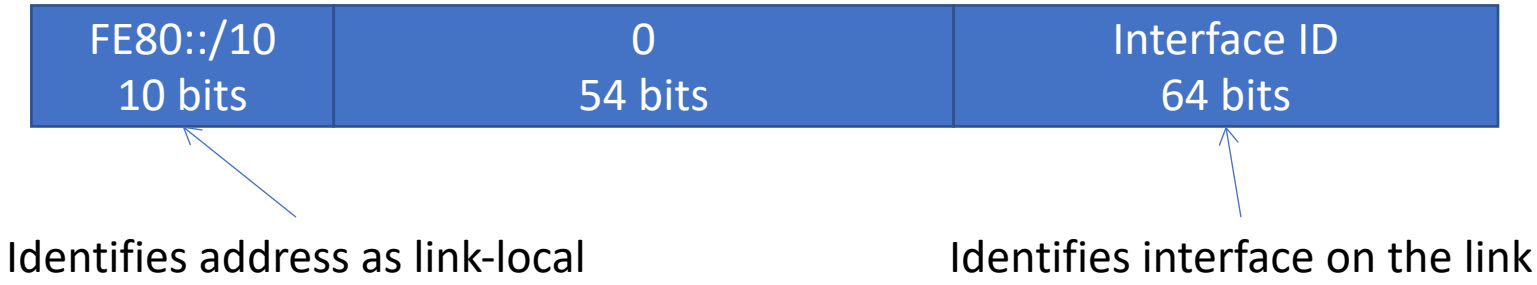
Global Unicast Address



- The global unicast address is identified by binary prefix 001 (2000::/3)
- The global routing prefix identifies the address range allocated to a site
  - The prefixes are allocated by international registry service
  - Subnet ID is allocated by site administrator for each link within site
- EUI-64 format assumes 64 bit for Interface ID
- Recommended prefix length is 48 bits
  - … leaving 16 bit Subnet ID – over 65k subnets per site
  - The Interface ID must be unique on each link (can be determined by SLAAC)

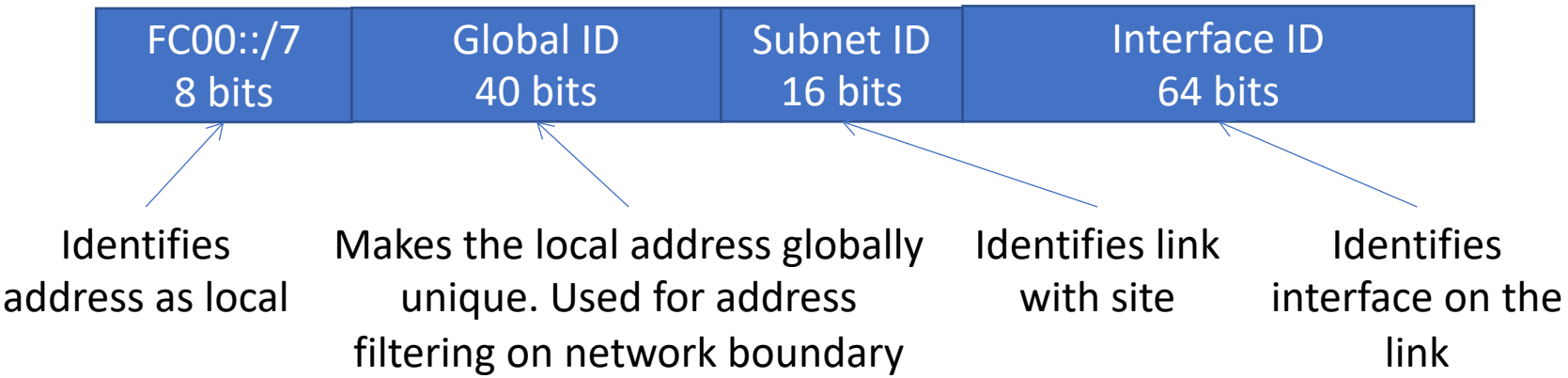
001	0 0000 0000 0000	2000::/3
001	0 XXXX XXXX XXXX	2XXX::/3
001	1 XXXX XXXX XXXX	3XXX::/3

Link-Local Unicast Address



- The link-local address is identified by binary prefix 1111 1110 10 (FE80::/10)
- The link-local addresses are used within single link and are not routable
  - The link-local addresses are generated by auto-configuration
  - The link-local addresses means „the host on this link”
  - The link-local addresses are used for auto-configuration of global address, neighbour/router discovery, for communication over LAN networks

Unique Local Unicast IPv6 Address



- The local address is identified by binary prefix 1111 110 (FC00::/7)
  - 1111 1101 (FD00::/8) – address administrated locally
  - 1111 1100 (FC00::/8) – reserved for future use
- The local addresses are *globally* unique but they should ***never be routed in the Internet*** (uniqueness assures routing security in case of misconfiguration)
- The local addresses plays the role of private address to be used by the site administrator for communication over private networks

Special IPv6 Addresses

- *Unspecified address* (allzero address) - ::/128
  - Indicates the lack of IP address, can be used as source address during the initial host configuration
  - In IPv4 it is 0.0.0.0

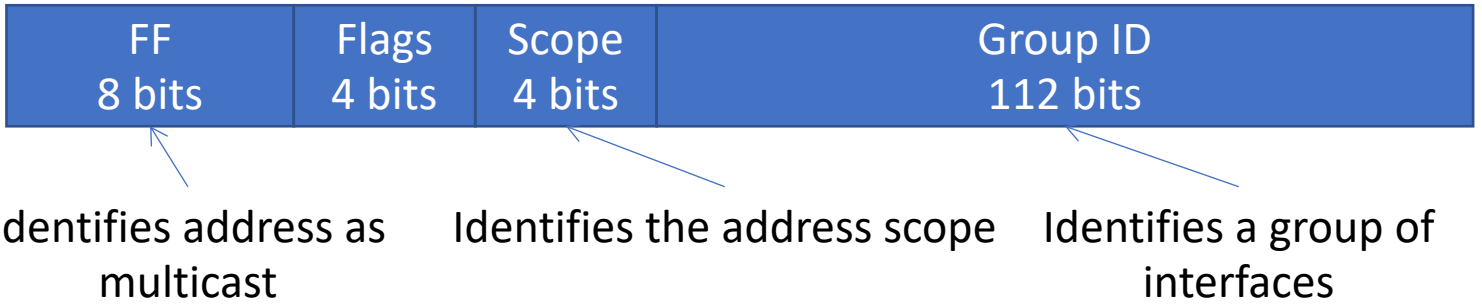
- *Loopback address* - ::1/128
  - Used as destination address in order to send packet internally within the host (from one process to the other). Its meaning is „this host”.
  - In IPv4 it is 127.0.0.1

Special IPv6 Addresses

- IPv4 Embedded Addresses
  - IPv4 compatible IPv6 address (obsolete): for tunneling IPv6 packets over IPv4 infrastructure

  - IPv4 mapped IPv6 address (in force): represents IPv4 node as IPv6 address

Multicast Address

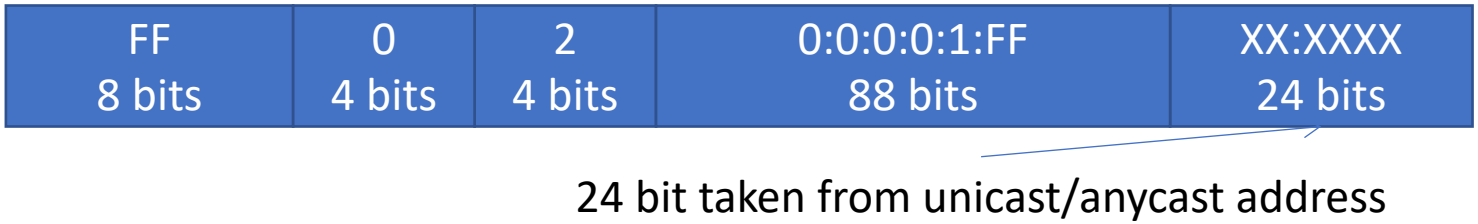


- The multicast address is identified by binary prefix 1111 1111 (FF00::/8)
- The multicast address can be assigned to a group of interfaces, all group members will receive the packets send to the multicast address
- Flags: 00PT
  - P=1: Multicast address based on network prefix (RFC3306)
  - T=0: Well-known multicast address (permanently assigned), T=1: temporary multicast address
- Scope: used to limit the range of the multicast transmission
  - 1: interface local scope (similar as loopback in unicast transmission)
  - 2: link local scope
  - 5: site local scope
  - E: global scope

Well-known Multicast addresses

- Interface-local scope
  - FF01::1 - All-nodes address
  - FF01::2 - All-routers address
- Link-local scope
  - **FF02::1 - All-nodes address**
  - **FF02::2 - All-routers address**
  - FF02::5 - All-OSPF routers address
  - FF02::6 - All-OSPF DR routers address
  - FF02::1:2 - All DHCP servers address
  - FF02::1:FFXX:XXX - Solicited-node address
- Site-local scope
  - FF05::2 - All-routers address
  - FF05::1:3 - All DHCP servers address

Solicited-Node Multicast Address



- The solicited-node multicast address is identified by prefix FF02:0:0:0:0:1:FF00::/104)
  - The solicited-node multicast address is formed by appending last 24 bits of IPv6 unicast/anycast address to the solicited-node multicast prefix
- The node must join every solicited-node multicast address generated for every unicast /anycast address assigned to the node
  - The packets send to the solicited-node multicast address are received only by this node (not all nodes on the link)
- It is used in Neighbour Discovery (ARP procedure)

ICMPv6

Internet Control Message Protocol for IPv6

IPv4 vs. IPv6

Main differences

Feature	ICMPv4	ICMPv6
Echo (ping)	YES	
Router Advertisement	YES	
Time Exceeded	YES	
Traceroute	YES	
Destination Unreachable	YES	
…		
Address Resolution	*No, through ARP*	*Built-in*
Auto-configuration	*No, partially by DHCPv4*	*Built-in*
Multicast Listener Discovery	*No, through IGMP*	*Built-in*

ICMPv6 Message Format

Type 1 Byte	Code 1 Byte	Checksum 2 Bytes	Message Body Variable

- Type – determines the message class and format
  - 0-127 – error messages
  - 128-255 – information messages
- Code – depends on the Type and provides additional information about the message
- Checksum – calculated for ICMP header and parts of the IPv6 header
- Message body – content depends on the message Type and Code

ICMPv6 Error Messages

Type	Usage
1=Destination unreachable	Packet cannot be delivered because the destination does not exists or cannot be reached due to some administrative rules
2=Packet too big	The packet length exceedes the link MTU
3=Time exceeded	The hop limit was eceeded or the fragmented packet was not feasembled in assumed time
4=Parameter problem	The packet header contains unknown fields e.g. in packet or extension header

Destination Unreachable Message

The packet due to some reason cannot be delivered to the destination

The ICMP message is send to the source address of the invoking packet

Code

0 - no route to destination

- router has no entry in the routing table for the destination address

1 - communication with dest. administratively prohibited

- this type of message can for example be sent by firewall that is configured to filter out the packet

2 – beyond scope of source address

- the scope of source and destination address is not the same e.g. the destination address is global while the source address is link-local

3 - address unreachable

- Cannot resolve layer 2 address

4 - port unreachable

- No listener for given port number at target host

5 - Source address failed ingress/egress policy

- the packet cannot be delivered due to the ingress and egress policy, packet was filtered out by the source or destination host (subset of code 1)

6 - reject route to destination

- Packets with specific prefix are blocked by an access control list or other packet filtering (subset of code 1)

Packet Too Big Message

- The packet cannot be forwarded because it exceeds the link MTU
- The ICMP message is send to the source address of the invoking packet
- The ICMP packet carries back the value of link MTU in the message body
- The code field is set to 0

Time Exceeded Message

The packet was discarded because the hop limit reached 0 or the packet cannot be reassembled

The message is send to the source address of the invoking packet

Code

0 - hop limit exceeded in transit

- The hop limit value was to low to deliver the packet or routing loop exists

1 - Fragment reassembly time exceeded

- The packet was send using fragment extension header and could not be reassembled within certain time

Parameter Problem Message

- The packet was discarded because it contains field that could not be recognized (in the IPv6 or extension header)
- The ICMP message is send to the source address of the invoking packet
- The ICMP message body contains pointer to the first byte of the unrecognized field

Code

- 0 - Erroneous header field
- 1- Unrecognised next header type
- 2- UnrecognisedIPv6 option

ICMPv6 Information Messages

Type	
128=Echo request	RFC4443 Ping command
129=Echo replay	
130=Multicast listener query	RFC2710 MLD – Multicast Listener Discovery
131=Multicast listener reoport	
132=Multicast listener done	
133=Router solicitation	**RFC2461** **NDP – Neighbor Discovery Protocol**
134=Router advertisement	
135=Neighbour solicitation	
136=Neighbour advertisement	
137=Redirect	
138=Router renumbering	RFC2894
…	

Neighbour Discovery Protocol (NDP)

NEIGHBOUR DISCOVERY

ROUTER DISCOVERY

DUPLICATE IP ADDRESS DETECTION (DAD)

NEIGHBOUR UNREACHABILITY DETECTION (NUD)

ADDRESS RESOLUTION PROTOCOL (ARP)

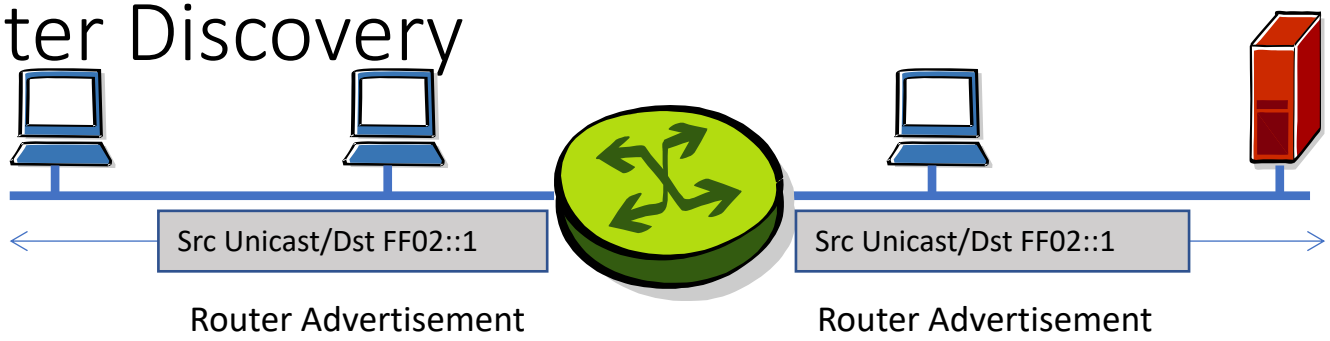
AUTO-CONFIGURATION OF IPV6 ADDRESSES

REDIRECTION

ICMPv6 Messages for NDP

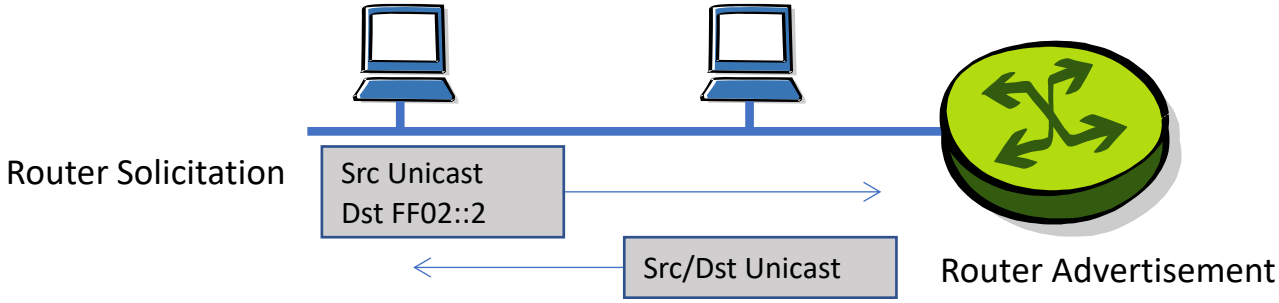
Name	Meaning	Usage
Neighbour Solicitation	*Who has IP address X?*	For *Duplicated Address Detection* (DAD), *Address Resolution,* and *Neighbour Unreachability Detection* (NUD)
Neighbour Advertisement	*I have it!* *(+ MAC address)*	Response to Neighbour Solicitation
Router Solicitation	*What's my prefix?*	For global-address auto-configuration
Router Advertisement	*The prefix is Y*	Response to Router Solicitation
Redirect Message	*There is a better route to Z*	For finding router that can forward the packet

Router Discovery



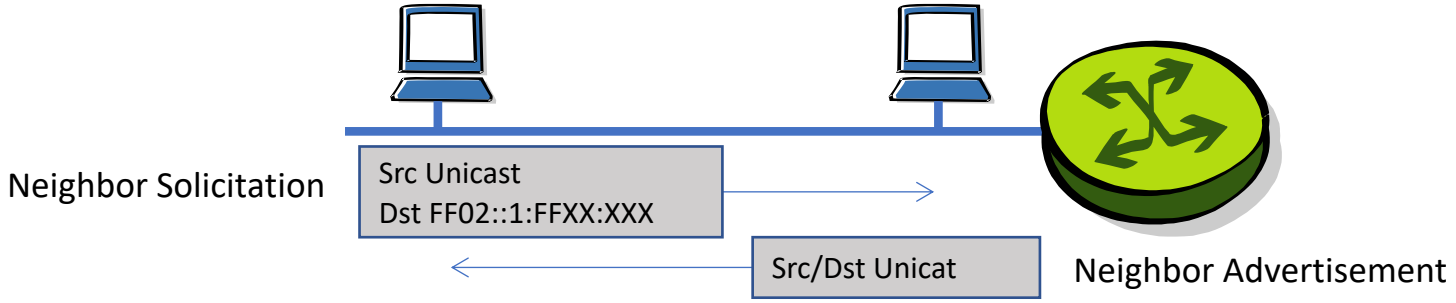
- Router send out **Router Advertisement** messages at regular intervals
  - The message is send to all nodes multicast address: FF02::1
  - Hop limit is always set to 255 (packets with lower hop limits are ignored)
  - The message contains configuration parameters for link
    - Default hop limit – used to configure the default hop count for all nodes on the link
    - Flags
      - M - statefull configuration for IP prefix (DHCP)
      - O – statefull configuration for other parameters then IP prefix
      - H = home address flag
    - Router lifetime – specifies the amount of time the router is used as default router (zero otherwise)
    - Reachable time - specifies the amount of time the router is reachable
    - Retrans time - specifies time between neighbour solicitation messages, used in NUD and ARP
    - Options: link-layer address, MTU size, prefix information

Router Discovery



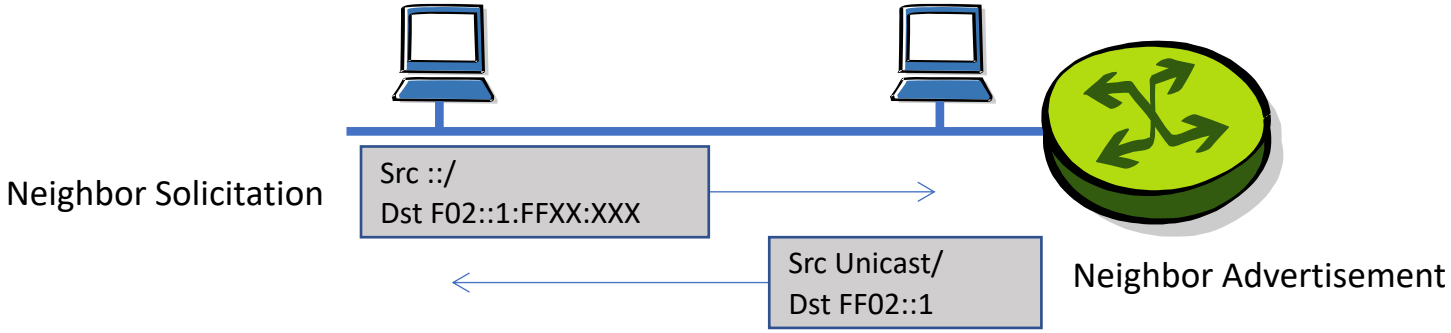
- Host can request Router Advertisement message (outside regular intervals) by sending **Router Solicitation** message
  - The message is send to all routers multicast address: FF02::2
  - Hop limit is always set to 255 (packets with lower hop limits are ignored)
  - The message contains the link-layer address of the host (only in case the IP address is known to the host)
- The Router Advertisement message is send back to the host (to source address of the Solicitation message)

Link Layer Address Resolution (ARP)



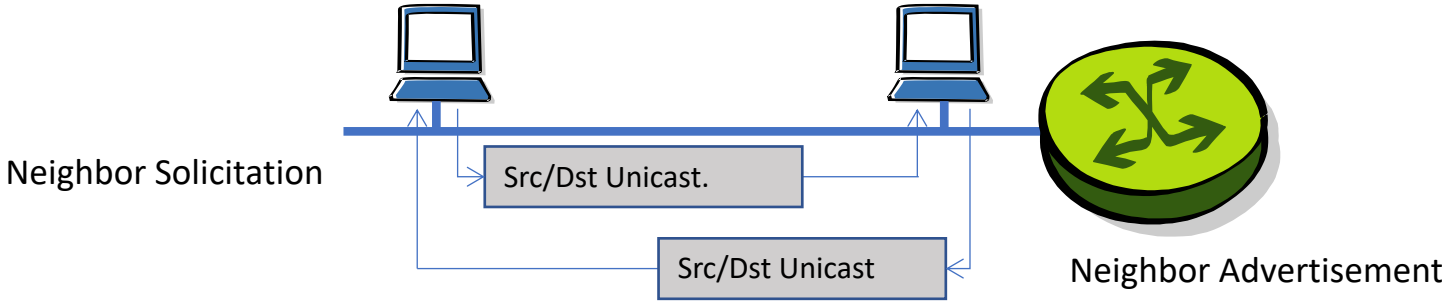
- The host that wants to resolve the link layer address sends out **Neighbor Solicitation** message
  - The message is send to solicited-node multicast address FF02::1:FFXX:XXX created from the queried address
  - The message contains link-layer address of sending host
- The target host (identified by the node solicitation multicast address) responds with **Neighbor Advertisement** message
  - The message contains link-layer address of responding host

Duplicated Address Detection (DAD)



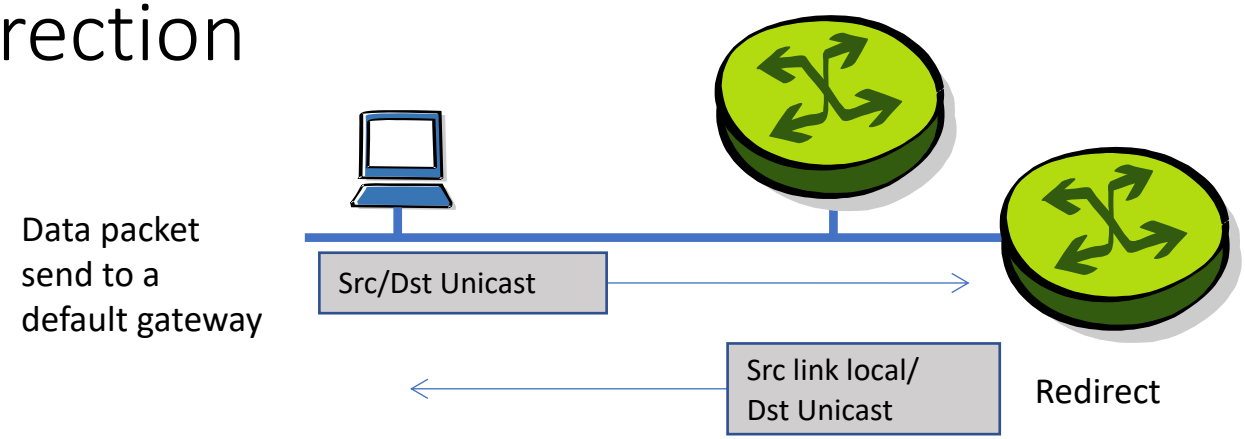
- DAD is usually used during stateless auto-configuration to verify if the IP address is in use
- The host sends out Neighbor Solicitation message
  - The message is send to solicited-node multicast address with the unspecified source address
  - The address being verified is send in target address field in the solicitation message
- If the address is in use the host identified by the node solicitation multicast address responds with Neighbor Advertisement message
  - The response is send to allnode multicast address

Neighbor Unreachability Detection (NUD)



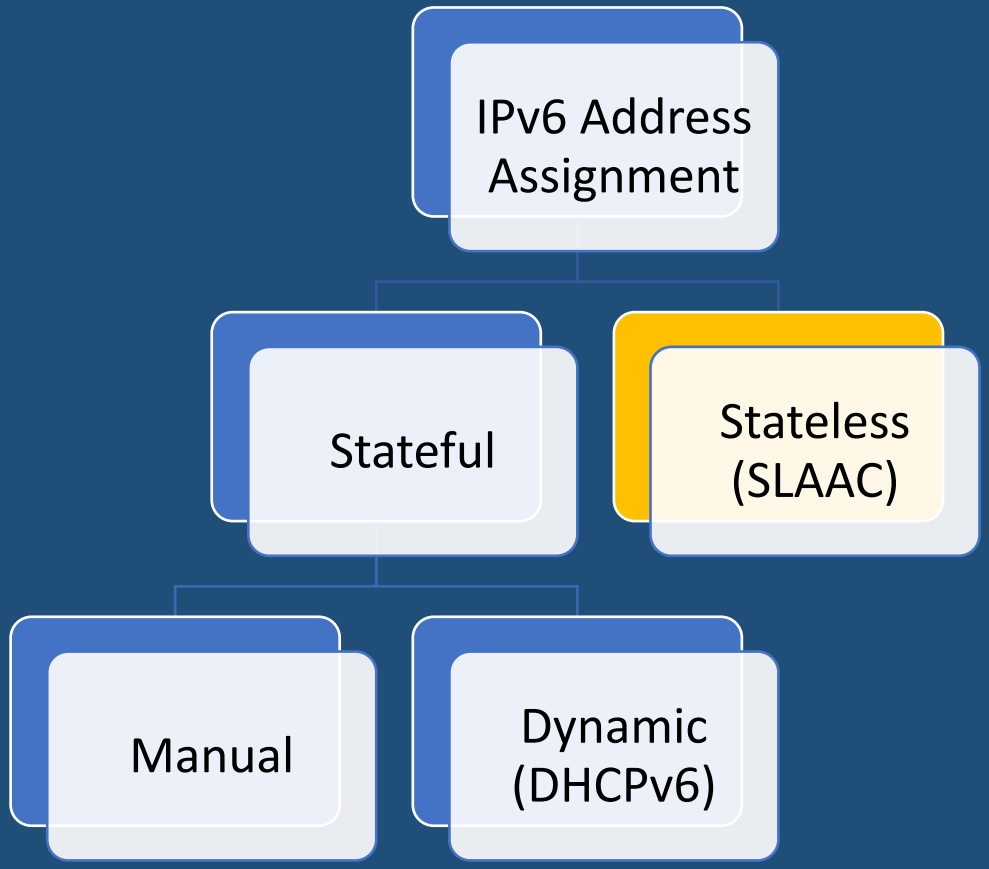
- The host that wants to verify the nieghbor reachability sends out Neighbor Solicitation message
  - The message is send to unicast address of the neighbor
  - The message contains link-layer address of sending host
- The Neighgbor responds with Neighbor Advertisement message
  - The message contains link-layer address of responding host
- It the Neighbor Advertisement is not received with some time the neighbor is declared unreachable

Redirection

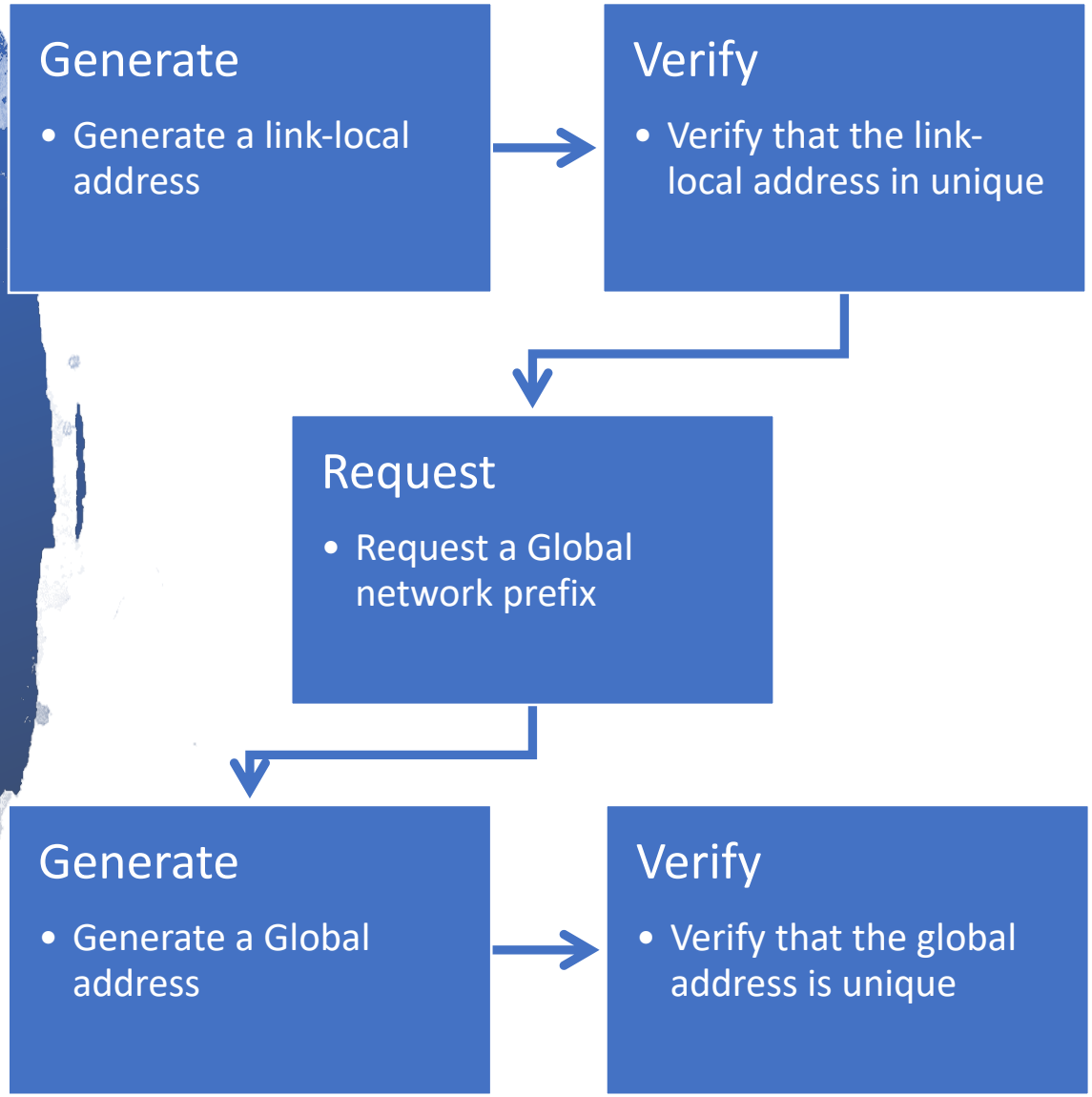


- The redirect message informs a node of a better first-hop router to reach destination
  - The message contains
    - the link-local address of the next-hop router to use for given destination
    - The redirected destination address
- If the redirected address and the next-hop address are the same it means that the destination is on the same subnet
  - In IPv6 many prefixes can exists in layer 2 network, not all of them have to be advertised by routers

IPv6 Address Assignment



StateLess Address Auto-Configuration (SLAAC)



SLAAC (Link-local Unicast)

We use the link-local 64-bits network prefix FE80::/10 and add 54 zeros (network prefix)

FE80	0000	0000	0000				

and a 48-bits Ethernet physical address - say 01-23-45-67-89-AB

(MAC address)

01	23	45	67	89	AB

We insert FF-FE in the middle

(EUI-64)
Extended Unique Identifier

01	23	45	FF	FE	67	89	AB

Set to 1 (or flip) the 7th bit of the first byte

(Interface Id)

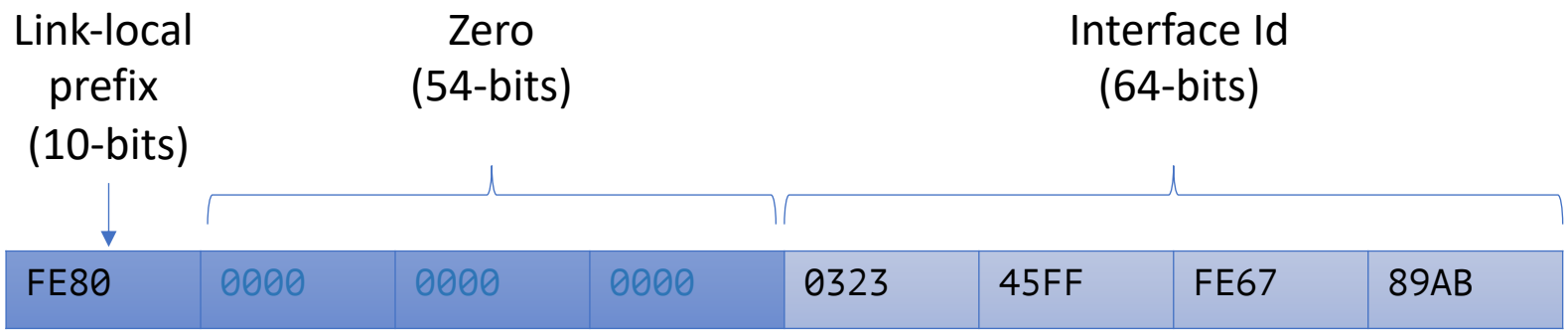
03	23	45	FF	FE	67	89	AB

Join the 64-bits prefix with the 64-bits interface id

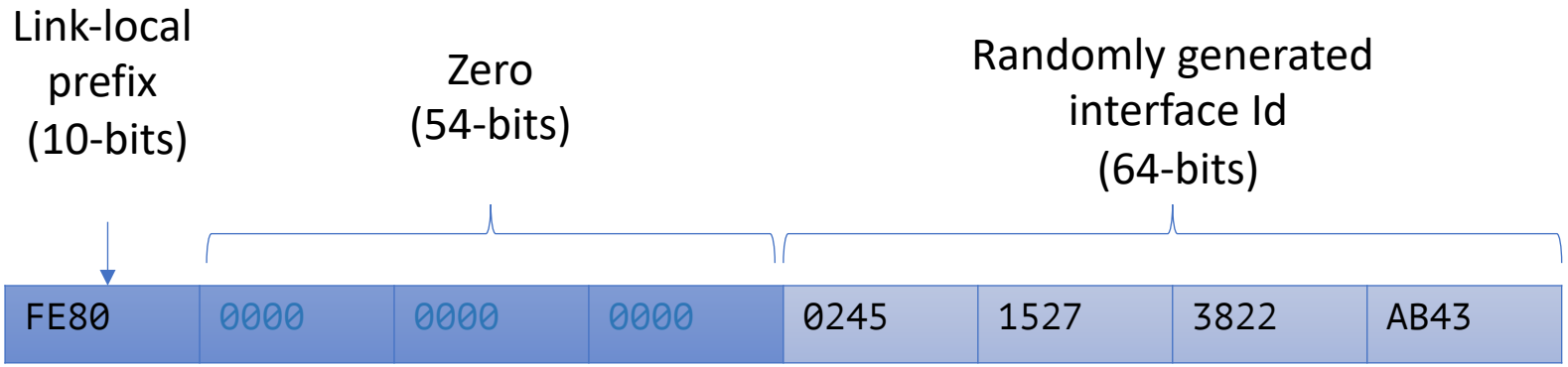
FE80	0000	0000	0000	0323	45FF	FE67	89AB

Which can be rewritten as FE80::0323:45FF:FE67:89AB

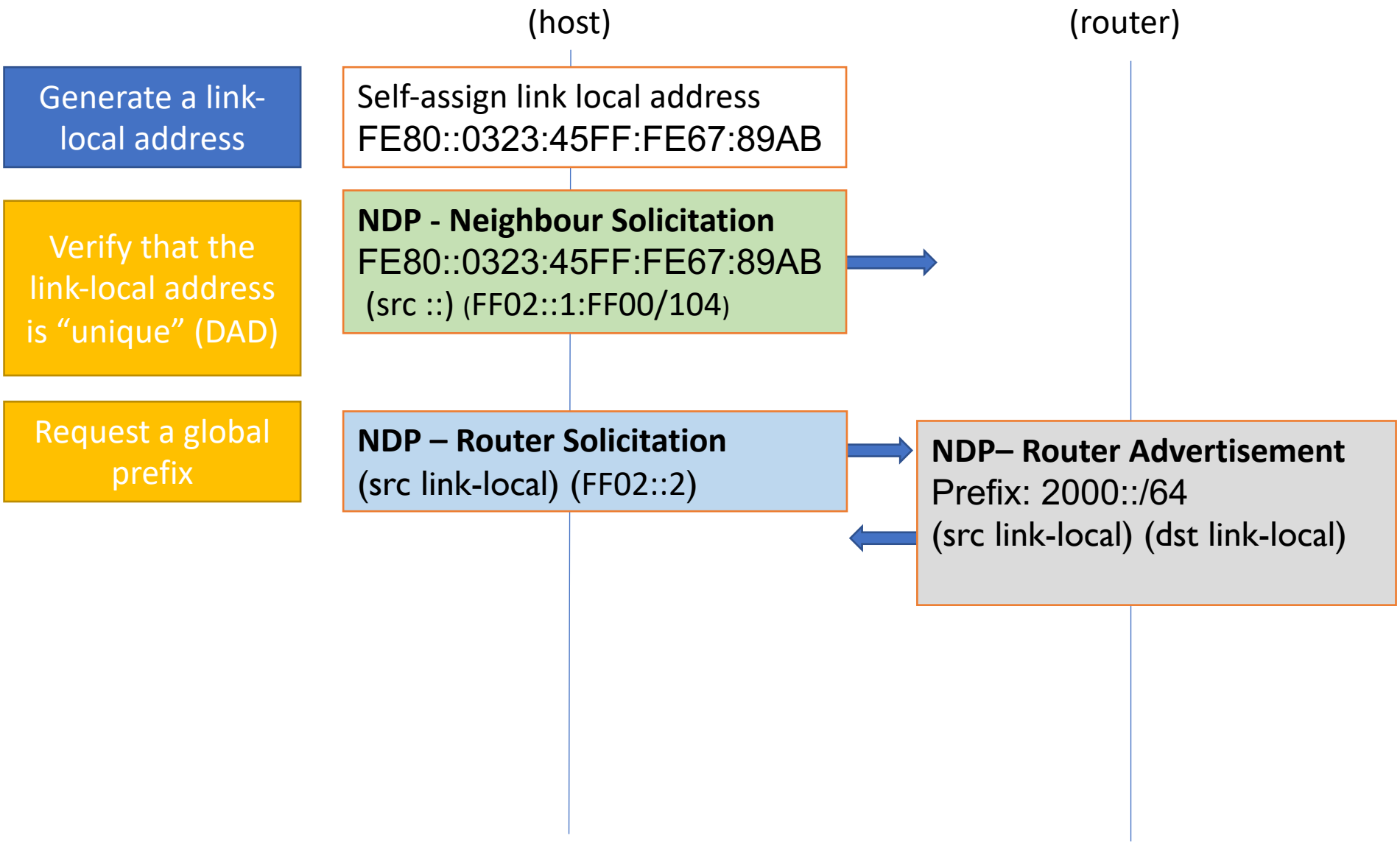
SLAAC (Link Local)



OR



SLAAC Procedure



SLAAC (GLOBAL Unicast)

We are given a 64-bits network prefix - say 2000::/64 (network prefix)

and a 48-bits Ethernet physical address - say 01-23-45-67-89-AB

(MAC address)

We insert FF-FE in the middle

(EUI-64)
Extended Unique Identifier

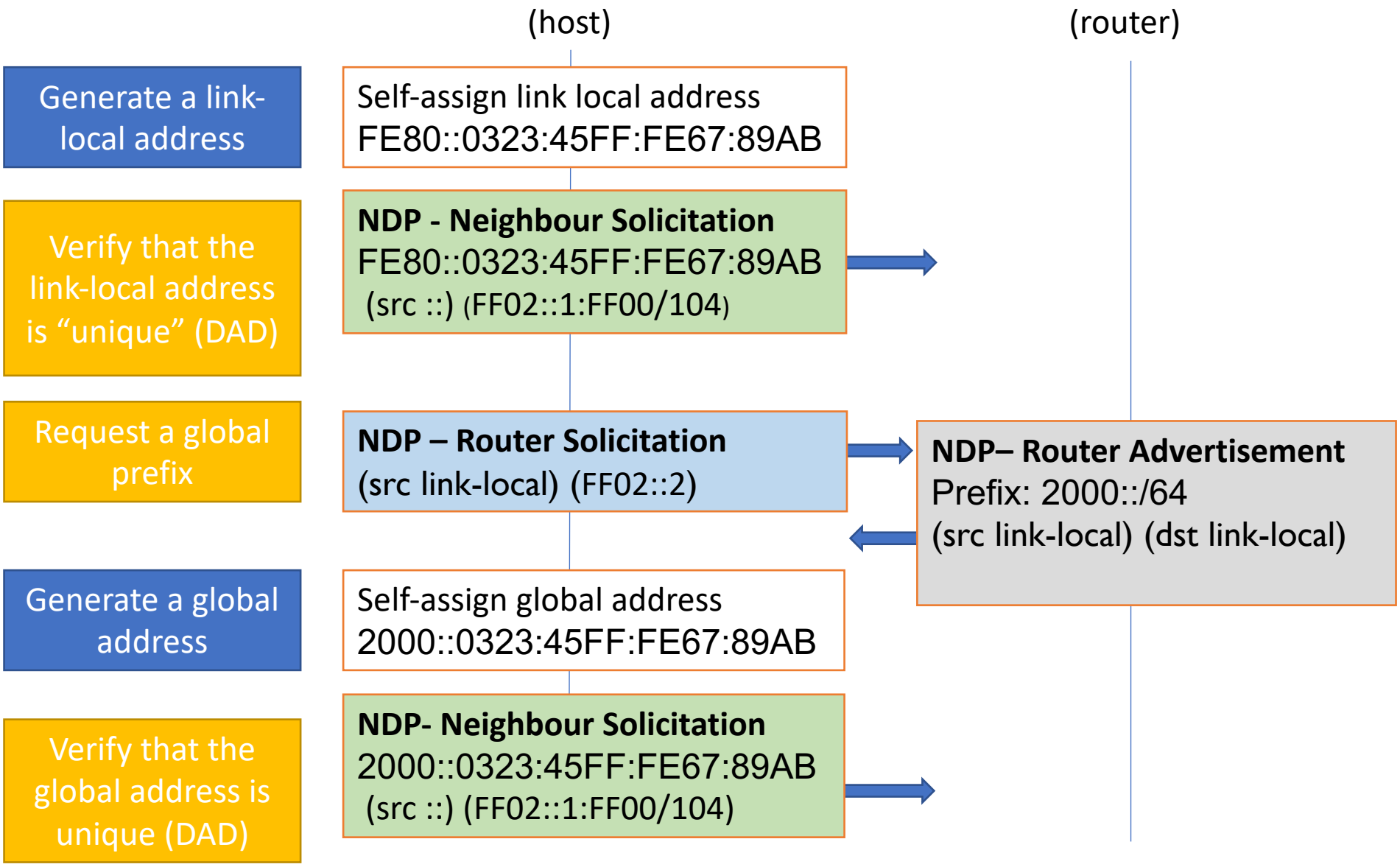
Set to 1 (or flip) the 7th bit of the first byte

(Interface Id)

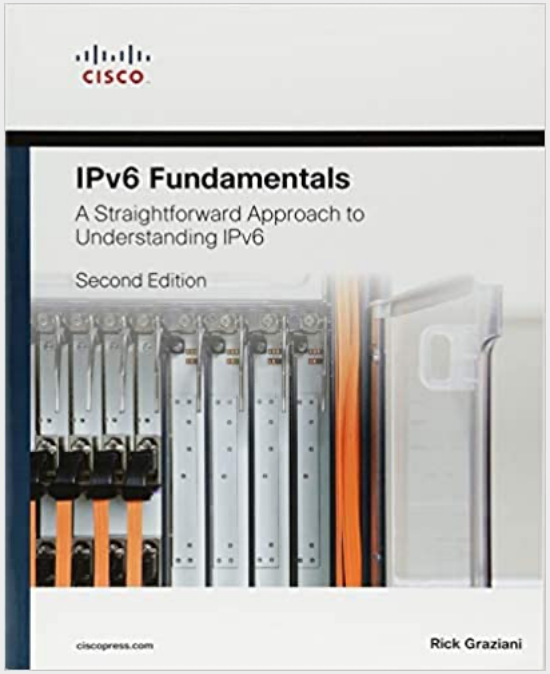
Join the 64-bits prefix with the 64-bits interface id

Which can be rewritten as 2000::0323:45FF:FE67:89AB

SLAAC Procedure



Bibliography



- **IPv6 Fundamentals: A Straightforward Approach to Understanding IPv6**. *Rick Graziani*. Cisco Press. 2013
- **TCP in the Internet of Things: from ostracism to prominence.** *Carles Gomez , Andrés Arcia-Moret , Jon Crowcroft.* IEEE Internet Computing. Jan./Feb. 2018, pp. 29-41, vol. 22.