

# Computer Networks

Mohammed El-Hajj

Jacobs University Bremen

October 25, 2021



JACOBS  
UNIVERSITY



# Course Content

- 1.Introduction
- 2.Fundamental Networking Concepts
- 3.Local Area Networks (IEEE 802)
- 4.Internet Network Layer (IPv4, IPv6)**
- 5.Internet Routing (RIP, OSPF, BGP)
- 6.Internet Transport Layer (UDP, TCP)
- 7.Firewalls and Network Address Translators
- 8.Domain Name System (DNS)
- 9.Abstract Syntax Notation 1 (ASN.1)
- 10.External Data Representation (XDR)
- 11.Augmented Backus Naur Form (ABNF)
- 12.Electronic Mail (SMTP, IMAP)
- 13.Document Access and Transfer (HTTP, FTP)

# Part 4: Internet Network Layer

18 Concepts and Terminology

19 Internet Protocol Version 6

20 Internet Protocol Version 4

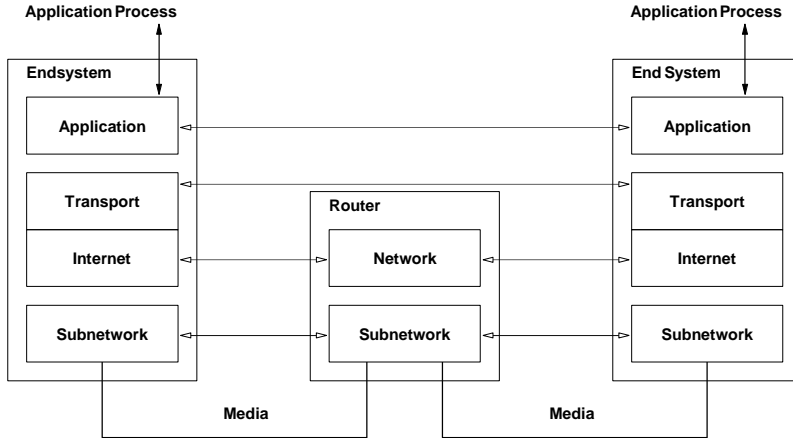
# Section 18: Concepts and Terminology

18 Concepts and Terminology

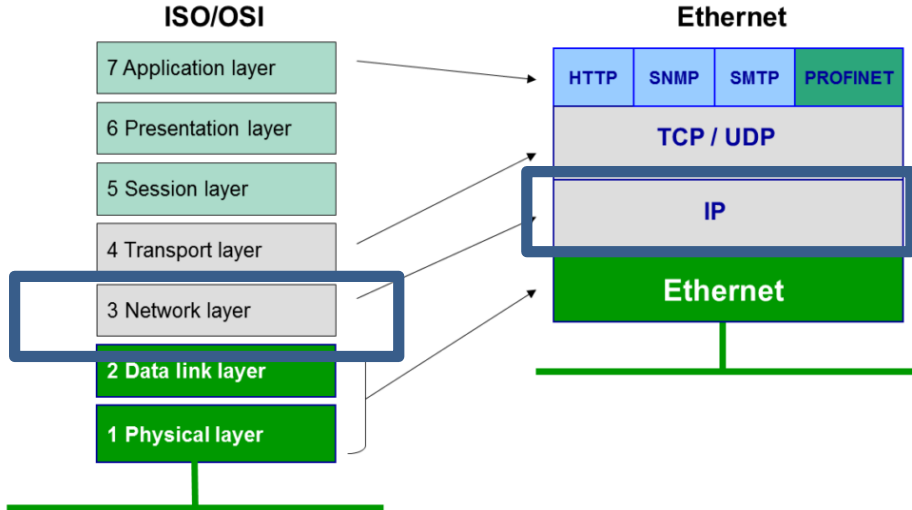
19 Internet Protocol Version 6

20 Internet Protocol Version 4

# Internet Reference Model



# Internet Reference Model – Our Focus



# Terminology (1/2)

- A *node* is a device which implements an Internet Protocol (such as IPv4 or IPv6).
- A *router* is a node that forwards IP packets not addressed to itself.
- A *host* is any node which is not a router.
- A *link* is a communication channel below the IP layer which allows nodes to communicate with each other (e.g., an Ethernet).
- The *neighbors* is the set of all nodes attached to the same link.
- An *interface* is a node's attachment to a link.
- An *IP address* identifies an interface or a set of interfaces.

# Terminology (2/2)

- An *IP prefix* is the initial part of an IP address identifying an IP network. The IP prefix is commonly defined by the number of the initial bits of an IP address that are identifying an IP network, the so called *prefix length*.
- An *interface identifier* is the portion of an IP address that identifies an interface in a certain IP network.
- An *IP packet* is a bit sequence consisting of an IP header and the payload.
- The *link MTU* is the maximum transmission unit, i.e., maximum packet size in octets, that can be conveyed over a link.
- The *path MTU* is the the minimum link MTU of all the links in a path between a source node and a destination node.



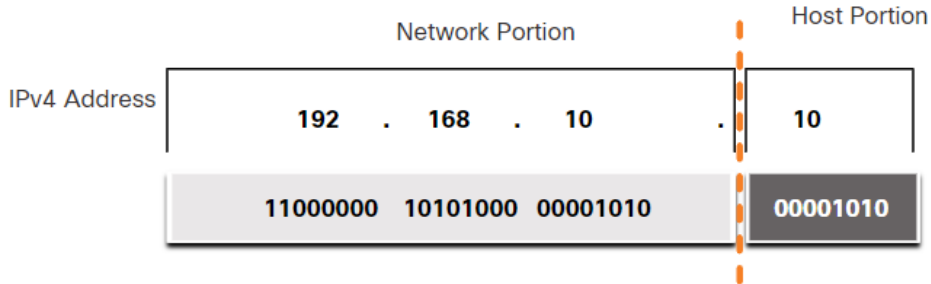
# Internet Address / Prefix Assignment

- Manual: A network administrator assigns an IP prefix manually to an interface.
- System: A networking stack automatically assigns a prefix to an interface (e.g., 127.0.0.1/8 or ::1/128 for a loopback interface).
- Stateless automatic configuration: A networking stack automatically calculates and assigns an IP prefix (e.g., deriving an interface identifier from a lower-layer address and combining it with a learned prefix).
- Stateful automatic configuration: A networking stack obtains an prefix from a service providing IP addresses on request (e.g., DHCP).
- Temporary addresses: A networking stack generates temporary addresses from a known prefix in order to enhance privacy.

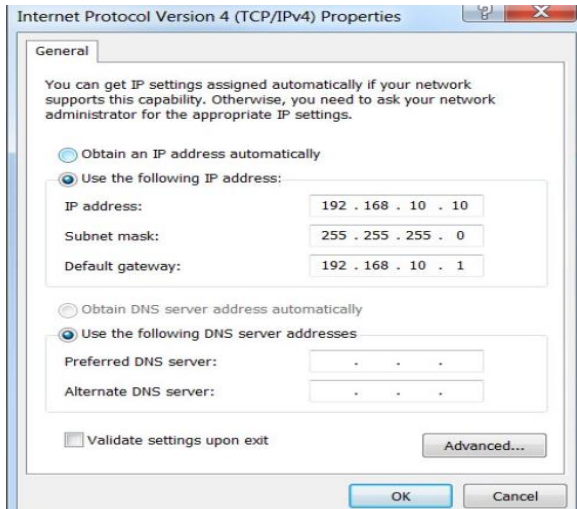
# IPv4 Address Structure

## Network and Host Portions

- An IPv4 address is a 32-bit hierarchical address that is made up of a network portion and a host portion.
- When determining the network portion versus the host portion, you must look at the 32-bit stream.

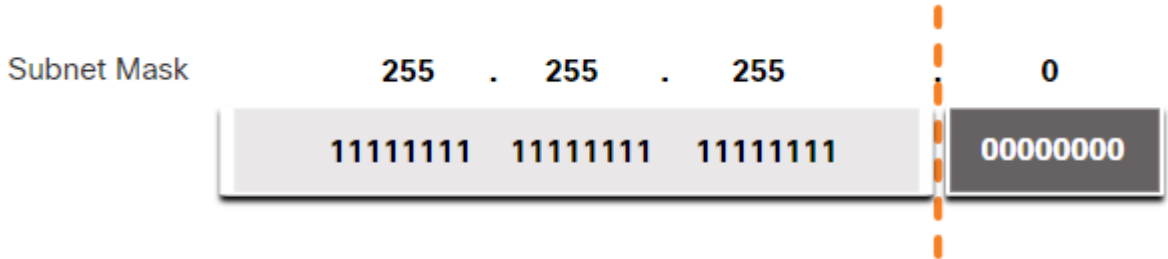


# The Subnet Mask (1-2)



# The Subnet Mask (2-2)

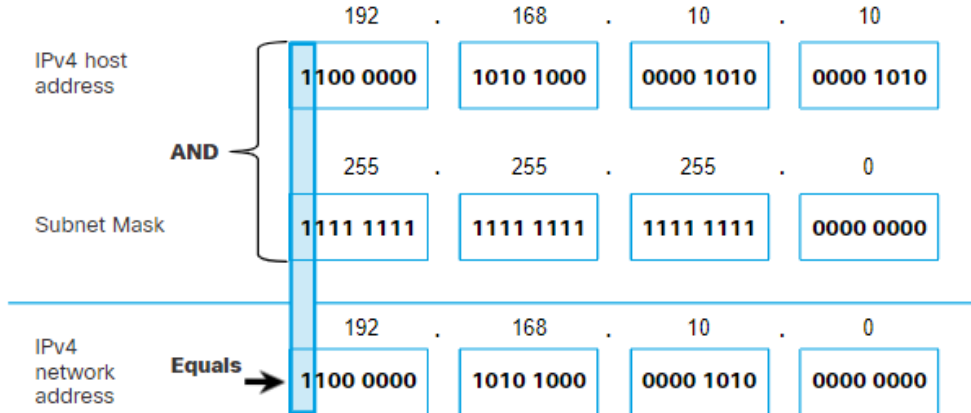
The 32-bit subnet mask in dotted decimal and binary formats.



# Prefix Length

Subnet Mask	32-bit Address	Prefix Length
255.0.0.0	11111111.00000000.00000000.00000000	/8
255.255.0.0	11111111.11111111.00000000.00000000	/16
255.255.255.0	11111111.11111111.11111111.00000000	/24
255.255.255.128	11111111.11111111.11111111.10000000	/25
255.255.255.192	11111111.11111111.11111111.11000000	/26
255.255.255.224	11111111.11111111.11111111.11100000	/27
255.255.255.240	11111111.11111111.11111111.11110000	/28
255.255.255.248	11111111.11111111.11111111.11111000	/29
255.255.255.252	11111111.11111111.11111111.11111100	/30

# Determining the Network: Logical AND



# Jacobs University's IP Networks

- Jacobs University currently uses the global IPv4 address blocks 212.201.44.0/22 and 212.201.48.0/23. How many IPv4 addresses can be used in these two address spaces?
- 212.201.44.0/22:  $2^{32-22} - 2 = 2^{10} - 2 = 1022$   
212.201.48.0/23:  $2^{32-23} - 2 = 2^9 - 2 = 510$
- Jacobs University currently uses the global IPv6 address block 2001:638:709::/48. How many IPv6 addresses can be used?
- 2001:638:709::/48:  $2^{128-48} - 2 = 2^{80} - 2$  which is 1208925819614629174706174.
- If you equally distribute the addresses over the campus area ( $30 \cdot 10^4 m^2$ ), what is the space covered per address?

# Internet Network Layer Protocols

IPv4

VS

IPv6

Example: 127.255.255.255

Example:

2001:0db8:85a3:0000:0000:8a2e:0370:7334

- IPv6:
  - The *Internet Protocol version 6* (IPv6) provides for transmitting datagrams from sources to destinations using 16 byte IPv6 addresses
  - The *Internet Control Message Protocol version 6* (ICMPv6) is used for IPv6 error reporting, testing, auto-configuration and address resolution
- IPv4:
  - The *Internet Protocol version 4* (IPv4) provides for transmitting datagrams from sources to destinations using 4 byte IPv4 addresses
  - The *Internet Control Message Protocol version 4* (ICMPv4) is used for IPv4 error reporting and testing
  - The *Address Resolution Protocol* (ARP) maps IPv4 addresses to IEEE 802 addresses



# IP Forwarding

- IP addresses can be divided into a part which identifies a network (the network prefix) and a part which identifies an interface of a node within that network (the interface identifier).
- The *forwarding table* realizes a mapping of the network prefix to the next node (next hop) closer to the destination and the local interface used to reach the next node.
- For every IP packet, the entry in the forwarding table with the longest matching prefix for the destination address has to be found (longest prefix match).
- A default forwarding table entry (if it exists) uses a zero-length prefix, that is either 0.0.0.0/0 (IPv4) or ::/0 (IPv6).

# Displaying the Routing Table

```
Administrator: Command Prompt

IPv4 Route Table
=====
Active Routes:
Network Destination        Netmask          Gateway          Interface        Metric
0.0.0.0                    0.0.0.0          192.168.0.1      192.168.0.101    25
127.0.0.0                  255.0.0.0        On-link          127.0.0.1        306
127.0.0.1                  255.255.255.255  On-link          127.0.0.1        306
127.255.255.255            255.255.255.255  On-link          127.0.0.1        306
192.168.0.0                255.255.255.0    On-link          192.168.0.101    281
192.168.0.101              255.255.255.255  On-link          192.168.0.101    281
192.168.0.255              255.255.255.255  On-link          192.168.0.101    281
192.168.56.0               255.255.255.0    On-link          192.168.56.1     266
192.168.56.1               255.255.255.255  On-link          192.168.56.1     266
192.168.56.255             255.255.255.255  On-link          192.168.56.1     266
224.0.0.0                  240.0.0.0        On-link          127.0.0.1        306
224.0.0.0                  240.0.0.0        On-link          192.168.56.1     266
224.0.0.0                  240.0.0.0        On-link          192.168.0.101    281
255.255.255.255            255.255.255.255  On-link          127.0.0.1        306
255.255.255.255            255.255.255.255  On-link          192.168.56.1     266
255.255.255.255            255.255.255.255  On-link          192.168.0.101    281
=====
Persistent Routes:
Network Address            Netmask          Gateway Address  Metric
0.0.0.0                    0.0.0.0          192.168.80.251  Default
```

# Longest Prefix Matching in Routers

Prefix	Next Hop
--------	----------

192.24.0.0/18	D
---------------	---

192.24.12.0/22	B
----------------	---

Prefix	Next Hop
192.24.0.0/18	D
192.24.12.0/22	B

192.24.63.255

D /18

192.24.15.255

B

192.24.12.0

D

192.24.0.0

More  
specific

IP Address

# IP Forwarding Table Management

- Entries of the IP forwarding table may be created by different entities:

Destination	Subnet mask	Interface
128.75.43.0	255.255.255.0	Eth0
128.75.43.0	255.255.255.128	Eth1
192.12.17.5	255.255.255.255	Eth3
default		Eth2

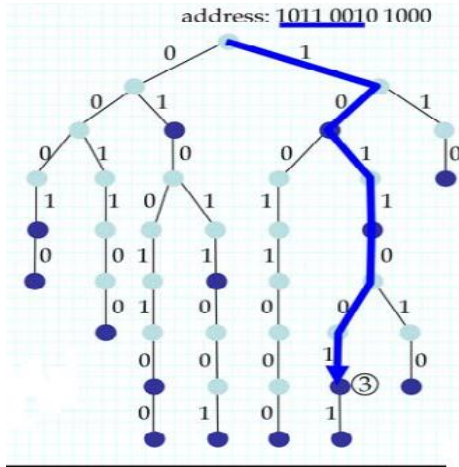
- Manual: A network administrator creates entries in the IP forwarding table manually.
- System: A networking stack automatically creates forwarding entries (e.g., when assigning a prefix to a network interface).
- Automatic Configuration Protocols: Protocols discovering valid prefixes or obtaining IP addresses and prefixes dynamically from a pool may create suitable IP forwarding table entries.
- Routing Protocols: Distributed routing protocols create and maintain one or more routing tables that these routing tables feed data into the IP forwarding table.
- Some implementations support multiple forwarding tables that can be selected by certain packet properties.

# Longest Prefix Match: Binary Trie

routing table

prefix	next hop
10*	7
01*	5
110*	3
1011*	5
0001*	0
0101 1*	7
0001 0*	1
0011 00*	2
1011 001*	3
1011 010*	5
0100 110*	6
0100 1100*	4
<del>1011 0011*</del>	<del>8</del>
1001 1000*	10
0101 1001*	9

address: 1011 0010 1000

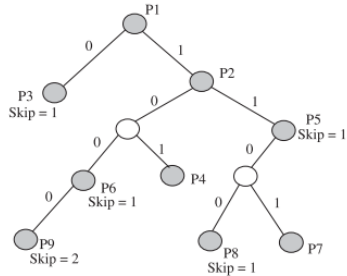
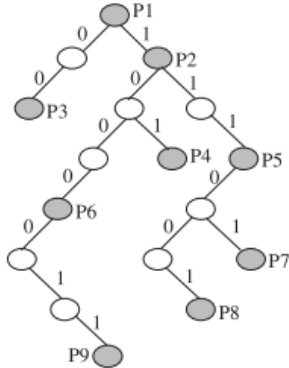


- A binary trie is the representation of the binary prefixes in a tree.

# Longest Prefix Match: Path Compressed Trie

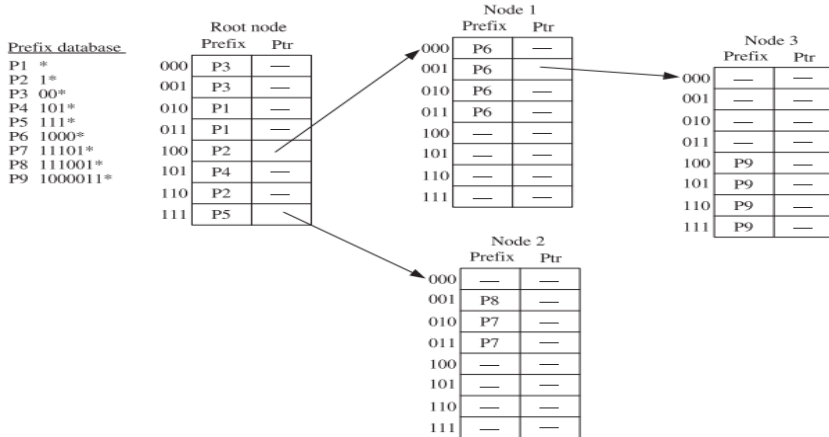
### Prefix database

P1	*
P2	1*
P3	00*
P4	101*
P5	111*
P6	1000*
P7	11101*
P8	111001*
P9	1000011*



- A path compressed trie is obtained by collapsing all one-way branch nodes.
- The number attached to nodes indicates the next (absolute) bit to inspect.
- While walking down the tree, you verify in each step that the prefix still matches the prefix stored at each node.

# Longest Prefix Match: Two-bit stride multibit Trie



- A two-bit multibit trie reduces the number of memory accesses.

# Section 19: Internet Protocol Version 6

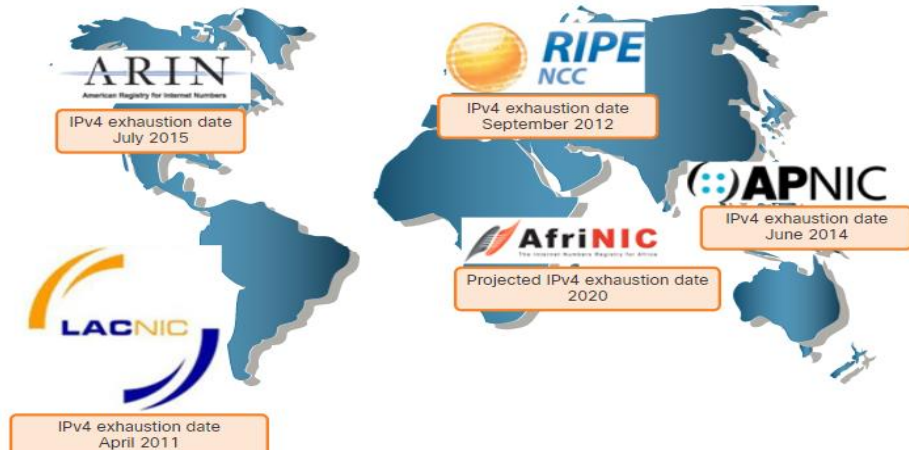
18 Concepts and Terminology

19 Internet Protocol Version 6

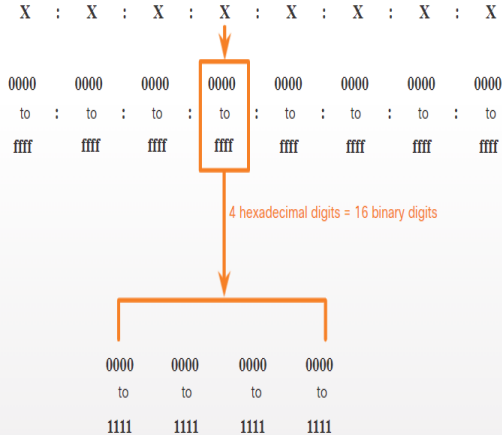
20 Internet Protocol Version 4



# Need for IPv6



# IPv6 Addressing Formats (1/3)



```
2001 : 0db8 : 0000 : 1111 : 0000 : 0000 : 0000 : 0200
2001 : 0db8 : 0000 : 00a3 : abcd : 0000 : 0000 : 1234
2001 : 0db8 : 000a : 0001 : c012 : 9aff : fe9a : 19ac
2001 : 0db8 : aaaa : 0001 : 0000 : 0000 : 0000 : 0000
fe80 : 0000 : 0000 : 0000 : 0123 : 4567 : 89ab : cdef
fe80 : 0000 : 0000 : 0000 : 0000 : 0000 : 0000 : 0001
fe80 : 0000 : 0000 : 0000 : c012 : 9aff : fe9a : 19ac
fe80 : 0000 : 0000 : 0000 : 0123 : 4567 : 89ab : cdef
0000 : 0000 : 0000 : 0000 : 0000 : 0000 : 0000 : 0001
0000 : 0000 : 0000 : 0000 : 0000 : 0000 : 0000 : 0000
```

# IPv6 Addressing Formats (2/3) -Omit Leading Zeros

## Omitting Leading 0s

Type	Format
Preferred	2001 : 0db8 : 0000 : 1111 : 0000 : 0000 : 0000 : 0200
No leading 0s	2001 : db8 : 0 : 1111 : 0 : 0 : 0 : 200
Preferred	2001 : 0db8 : 0000 : 00a3 : ab00 : 0ab0 : 00ab : 1234
No leading 0s	2001 : db8 : 0 : a3 : ab00 : ab0 : ab : 1234
Preferred	2001 : 0db8 : 000a : 0001 : c012 : 90ff : fe90 : 0001
No leading 0s	2001 : db8 : a : 1 : c012 : 90ff : fe90 : 1

# IPv6 Addressing Formats (3/3) -Double Colon

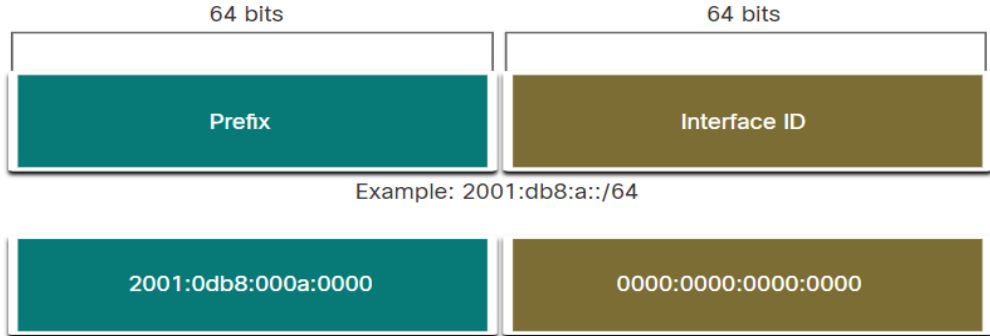
## Omitting Leading 0s and All 0 Segments

Type	Format
Preferred	2001 : 0db8 : 0000 : 1111 : 0000 : 0000 : 0000 : 0200
Compressed/spaces	2001 : db8 : 0 : 1111 : : 200
Compressed	2001:db8:0:1111::200
Preferred	2001 : 0db8 : 0000 : 0000 : ab00 : 0000 : 0000 : 0000
Compressed/spaces	2001 : db8 : 0 : 0 : ab00 ::
Compressed	2001:db8:0:0:ab00::
Preferred	2001 : 0db8 : aaaa : 0001 : 0000 : 0000 : 0000 : 0000
Compressed/spaces	2001 : db8 : aaaa : 1 ::
Compressed	2001:db8:aaaa:1::

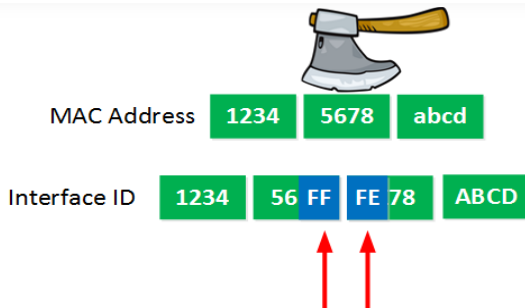
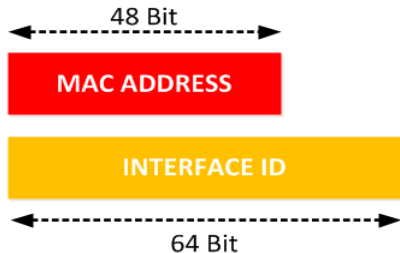
# IPv6 Interface Identifier

- Interface identifiers in IPv6 unicast addresses are used to uniquely identify interfaces on a link.
- For unicast addresses, except those that start with binary 000, interface identifiers are generally required to be 64 bits long.
- Combination of the interface identifier with a network prefix leads to an IPv6 address.
- Link local unicast addresses have the prefix fe80::/10.
- Interface identifier may be obtained from an IEEE 802 MAC address using a modified EUI-64 format, but this has privacy issues.
- Alternatively, it is possible to use temporary interface identifiers that keep changing.

# IPv6 prefix length

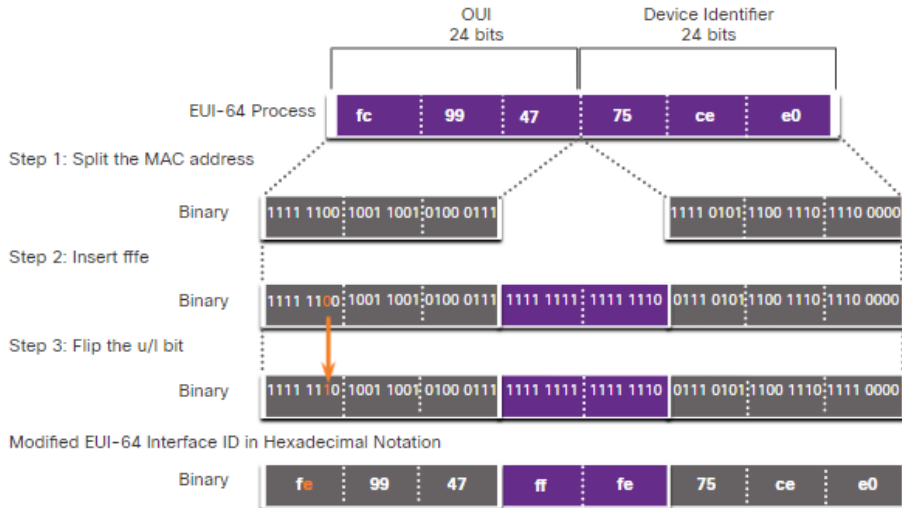


# Modified EUI-64 Format (1/2)



- Modified EUI-64 format can be obtained from IEEE 802 MAC addresses.
- Warning: Ipv6 addresses derived from MAC addresses can be used to track mobile nodes used in different networks.
- Solution: Temporary addresses with interface identifiers based on time-varying random bit strings and relatively short lifetimes.

# Modified EUI-64 Format (2/2)





# IPv6 Multicast Addresses

8 bits	4 bits	4 bits	112 bits
FF	Flags	Scope	Group ID

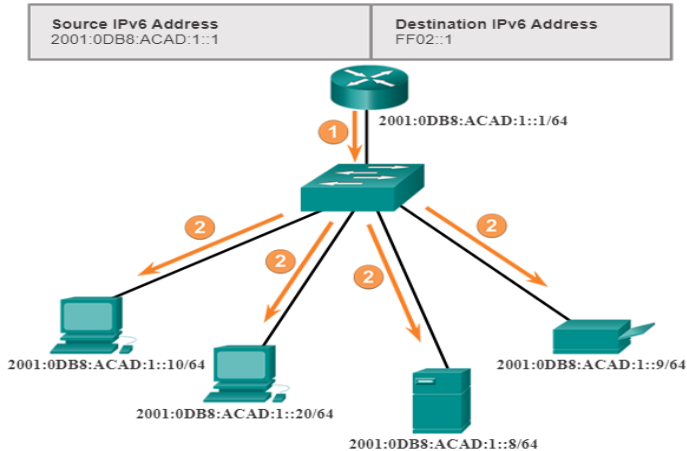
Address	Description
---------	-------------

ff02::1	All nodes on the local link.
ff02::2	All routers on the local link.
ff02::3	All hosts on the local link.
ff02::1:2	All DHCP servers and relay agents on a local link.
ff02::fb	All multicast DNS servers on a local link.

- IPv6 multicast addresses use the prefix ff00::/8.
- The addresses listed above are some of the well-known multicast addresses.
- Applications can, of course, allocate additional multicast addresses.

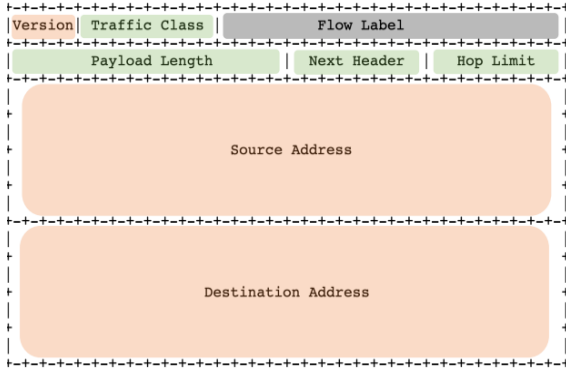
# IPv6 Multicast Addresses

## IPv6 All-nodes Multicast Communications



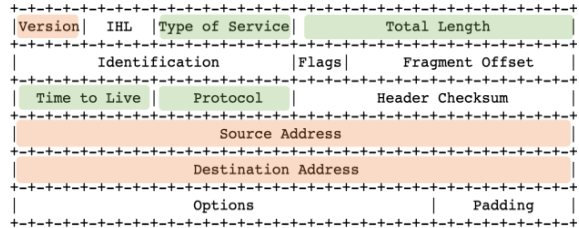
# IPv6 Packet Format (RFC 8200)

## IPv6 Header

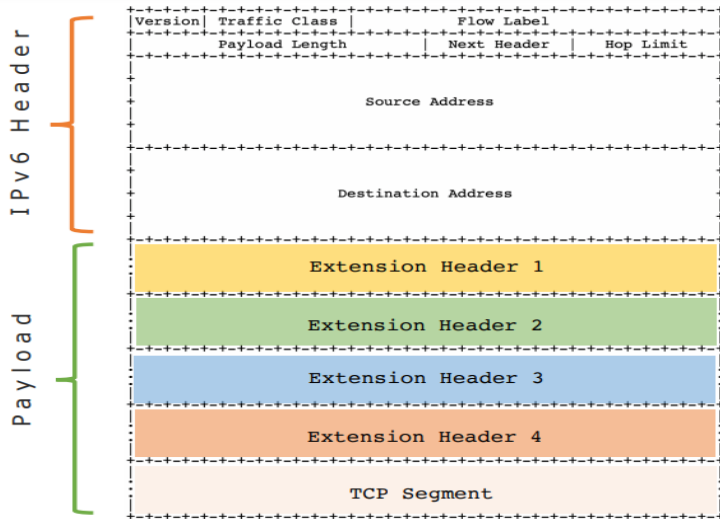


- Keep the same name -- 3 parts
- Different names with similar functions -- 4 parts
- New function -- 1 part

## IPv4 Header

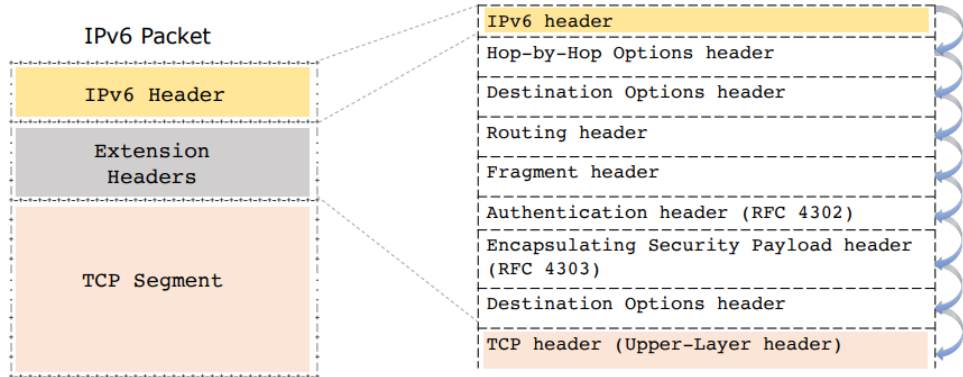


# IPv6 Extension Header



Maybe the payload is composed of several extension headers and the TCP segment.

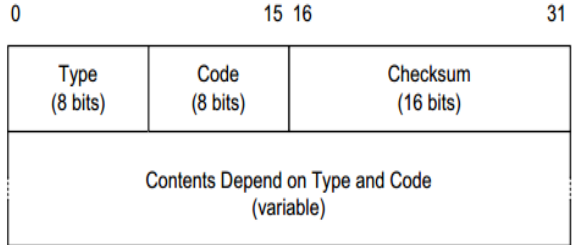
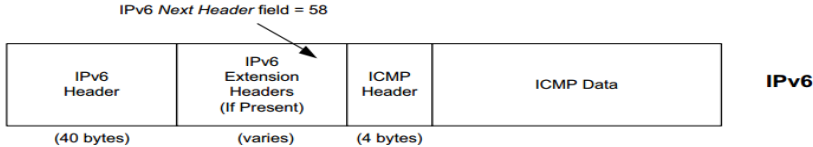
# IPv6 Extension Header



# IPv6 Forwarding

- IPv6 packets are forwarded using the longest prefix match algorithm.
- IPv6 addresses have relatively long prefixes, which allows network operators to achieve better address aggregation, which reduces the number of forwarding table entries needed in the backbone infrastructure.
- Due to the length of the prefixes, it is crucial to use a longest prefix match algorithm whose complexity does not depend on the number of entries in the forwarding table or the average prefix length.
- Due to better aggregation possibilities, IPv6 forwarding tables can be expected to be shorter than IPv4 forwarding tables

# IPv6 Error Handling (ICMPv6) (RFC 4443)



- The Internet Control Message Protocol Version 6 (ICMPv6) is used
  - to report error situations,
  - to run diagnostic tests,
  - to auto-configure IPv6 nodes, and
  - to supports the resolution of IPv6 addresses to link-layer addresses.

# IPv6 Neighbor Discovery (RFC 4861)

- Discovery of the routers attached to a link.
- Discovery of the prefixes used on a link.
- Discovery of parameters such as the link MTU or the hop limit for outgoing packets.
- Automatic configuration of IPv6 addresses.
- Resolution of IPv6 addresses to link-layer addresses.
- Determination of next-hop addresses for IPv6 destination addresses.
- Detection of unreachable nodes which are attached to the same link.
- Detection of conflicts during address generation.
- Discovery of better alternatives to forward packets.



# IPv6 over IEEE 802.3 (RFC 2464)

- Frames containing IPv6 packets are identified by the value 0x86dd in the IEEE 802.3 type field.
- The link MTU is 1500 bytes, which corresponds to the IEEE 802.3 maximum frame size of 1500 byte.
- The mapping of IPv6 addresses to IEEE 802.3 addresses is table driven. Entries in so called address translation tables can be either statically configured or dynamically learned using the neighbor discovery protocol.
- IPv6 over IEEE 802.3 does not use IEEE LLC encapsulation.

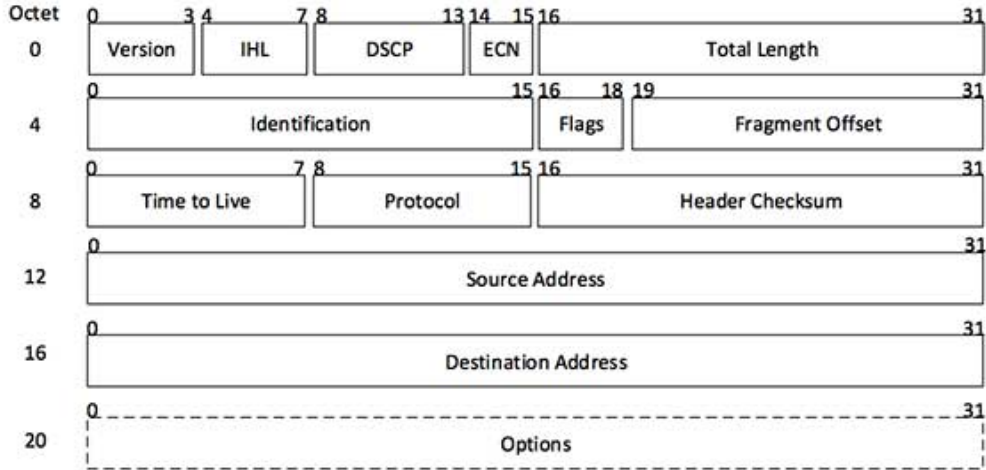
# Section 20: Internet Protocol Version 4

18 Concepts and Terminology

19 Internet Protocol Version 6

20 Internet Protocol Version 4

# IPv4 Packet Format (RFC 791)



# IPv4 Error Handling (ICMPv4)

- The Internet Control Message Protocol (ICMP) is used to inform nodes about problems encountered while forwarding IP packets.
  - Echo Request/Reply messages are used to test connectivity.
  - Unreachable Destination messages are used to report why a destination is not reachable.
  - Redirect messages are used to inform the sender of a better (shorter) path.
- Can be used to trace routes to hosts:
  - Send messages with increasing TTL starting with one and interpret the ICMP response message.
  - Pack additional data into the request to measure latency.
- ICMPv4 is an integral part of IPv4 (even though it is a different protocol).

# ICMPv4 Echo Request/Reply

No.	Delta	Cum.Bytes	Time	Source	Destination	Protocol	Length	Info
21	0.000039	2296	2.946012	google-public-dns-a.google.com	192.168.2.14	DNS	128	Standard query response 0x9b6d No such name
22	0.049102	2370	2.995114	192.168.2.14	b.resolvers.level3.net	ICMP	74	Echo (ping) request id=0x0001, seq=13/3328, ttl
23	0.024354	2444	3.019468	b.resolvers.level3.net	192.168.2.14	ICMP	74	Echo (ping) reply id=0x0001, seq=13/3328, ttl
24	0.501033	2528	3.520501	192.168.2.14	google-public-dns-a.google.com	Standard query	84	PTR 1.2.168.192.in-addr

Frame 22: 74 bytes on wire (592 bits), 74 bytes captured (592 bits)	
Ethernet II, Src: h0m4l4f4r_6a:31:6c (Sc:1ac:4c:6a:31:6c), Dst: 192.168.2.1 (94:44:52:39:72:e5)	
Internet Protocol version 4, Src: 192.168.2.14 (192.168.2.14), Dst: b.resolvers.level3.net (4.2.2.2)	
version: 4	
Header length: 20 bytes	
Differentiated services field: 0x00 (psci 0x00: default; LCN: 0x00: Not-ECT (Not LCN-capable transport))	
Total Length: 60	
Identification: 0x07fd (2045)	
Flags: 0x00	
Fragment offset: 0	
Time to live: 128	
Protocol: ICMP (1)	
Header checksum: 0x6a0a [correct]	
Source: 192.168.2.14 (192.168.2.14)	
Destination: b.resolvers.level3.net (4.2.2.2)	
[Source geoip: unknown]	
[Destination geoip: unknown]	
Internet Control Message Protocol	
Type: 8 (Echo (ping) request)	
Code: 0	
Checksum: 0x4d4e [correct]	
Identifier (NL): 1 (0x0001)	
Identifier (LL): 256 (0x0100)	
Sequence number (NL): 13 (0x000d)	
Sequence number (LL): 3328 (0x0d00)	
[Response in ttl]	
Data (32 bytes)	
Data: 616263645666768696a6b6c6d6e6f707172737475767768...	

- The ICMP echo request message (type = 8, code = 0) asks the destination node to return an echo reply message (type = 0, code = 0).
- The Identifier and Sequence Number fields are used to correlate incoming replies with outstanding requests.
- The data field may contain any additional data.

# ICMPv4 Unreachable Destinations

2 0.000078	192.168.0.205	192.168.0.4	ICMP	190 0.000078	Destination unreachable (Port unreachable)
25 3.953911	192.168.0.205	192.168.0.4	ICMP	186 3.953911	Destination unreachable (Port unreachable)
60 13.160687	192.168.0.205	192.168.0.4	ICMP	74 13.160687	Echo (ping) request id=0x0001, seq=1230
61 13.175425	192.168.0.4	192.168.0.205	ICMP	74 13.175425	6.738 Echo (ping) reply id=0x0001, seq=1230
71 14.175286	192.168.0.205	192.168.0.4	ICMP	74 14.175286	Echo (ping) request id=0x0001, seq=1336
72 14.180287	192.168.0.4	192.168.0.205	ICMP	74 14.180287	13.021 Echo (ping) reply id=0x0001, seq=1336
78 15.185084	192.168.0.205	192.168.0.4	ICMP	74 15.185084	Echo (ping) request id=0x0001, seq=1792
79 15.206197	192.168.0.4	192.168.0.205	ICMP	74 15.206197	21.133 Echo (ping) reply id=0x0001, seq=1792
100 19.010443	192.168.0.205	192.168.0.4	ICMP	190 19.010443	Destination unreachable (Port unreachable)
106 20.021230	192.168.0.205	192.168.0.4	ICMP	185 20.021230	Destination unreachable (Port unreachable)
119 23.005479	192.168.0.205	192.168.0.4	ICMP	186 23.005479	Destination unreachable (Port unreachable)
181 41.015272	192.168.0.205	192.168.0.4	ICMP	190 41.015272	Destination unreachable (Port unreachable)
200 44.012440	192.168.0.205	192.168.0.4	ICMP	186 44.012440	Destination unreachable (Port unreachable)
274 59.937311	192.168.0.205	192.168.0.4	ICMP	190 59.937311	Destination unreachable (Port unreachable)
282 61.016155	192.168.0.205	192.168.0.4	ICMP	185 61.016155	Destination unreachable (Port unreachable)
288 61.958795	192.168.0.205	192.168.0.4	ICMP	185 61.958795	Destination unreachable (Port unreachable)
299 63.045492	192.168.0.205	192.168.0.4	ICMP	182 63.045492	Destination unreachable (Port unreachable)
307 65.024739	192.168.0.205	192.168.0.4	ICMP	186 65.024739	Destination unreachable (Port unreachable)
356 78.980037	192.168.0.205	192.168.0.4	ICMP	190 78.980037	Destination unreachable (Port unreachable)
385 88.940412	192.168.0.205	192.168.0.4	ICMP	185 88.940412	Destination unreachable (Port unreachable)
410 94.010018	192.168.0.205	192.168.0.4	ICMP	190 94.010018	Destination unreachable (Port unreachable)
472 119.007769	192.168.0.205	192.168.0.4	ICMP	190 119.007769	Destination unreachable (Port unreachable)

Destination: 192.168.0.4

## Internet Control Message Protocol

Type: 3 (Destination unreachable)

Code: 3 (Port unreachable)

Checksum: 0x7fb0 [correct]

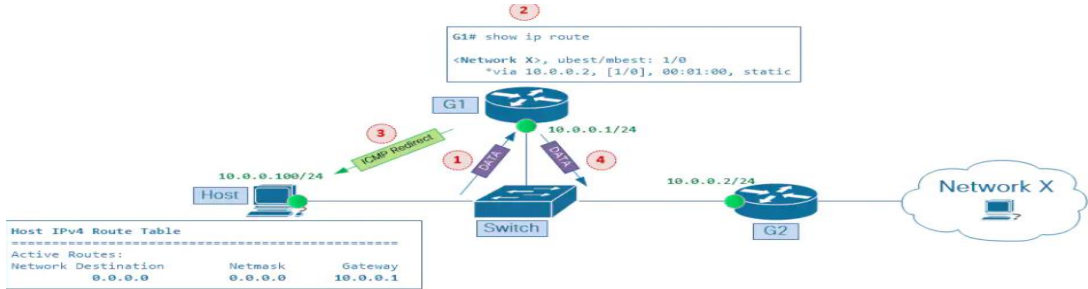
[Checksum Status: Good]

Unused: 00000000

- > Internet Protocol Version 4, Src: 192.168.0.4, Dst: 192.168.0.205
- > User Datagram Protocol, Src Port: 53, Dst Port: 59358
- > Domain Name System (response)

- The Type field has the value 3 for all unreachable destination messages.
- The Code field indicates why a certain destination is not reachable.
- The data field contains the beginning of the packet which caused the ICMP unreachable destination message.

# ICMPv4 Redirect



- The Type field has the value 5 for redirect messages.
- The Code field indicates which type of packets should be redirected.
- The Router Internet Address field identifies the IP router to which packets should be redirected.
- The data field contains the beginning of the packet which caused the ICMP redirect message.

# IPv4 Fragmentation

- IPv4 packets that do not fit the outgoing link MTU will get fragmented into smaller packets that fit the link MTU.
  - The Identification field contains the same value for all fragments of an IPv4 packet.
  - The Fragment Offset field contains the relative position of a fragment of an IPv4 packet (counted in 64-bit words).
  - The flag More Fragments (MF) is set if more fragments follow.
- The Don't Fragment (DF) flag can be set to indicate that a packet should not be fragmented.
- IPv4 allows fragments to be further fragmented without intermediate reassembly.



# Fragmentation Considered

- The receiver must buffer fragments until all fragments have been received. However, it is not useful to keep fragments in a buffer indefinitely. Hence, the TTL field of all buffered packets will be decremented once per second and fragments are dropped when the TTL field becomes zero.
- The loss of a fragment causes in most cases the sender to resend the original IP packet which in most cases gets fragmented as well. Hence, the probability of transmitting a large IP packet successfully goes quickly down if the loss rate of the network goes up.
- Since the Identification field identifies fragments that belong together and the number space is limited, one cannot fragment an arbitrary large number of packets.

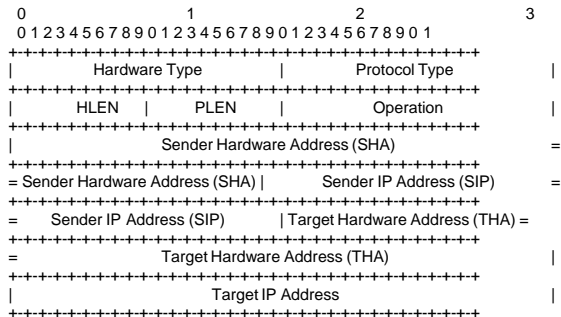
# MTU Path Discovery (RFC 1191)

- The sender sends IPv4 packets with the DF flag set.
- A router which has to fragment a packet with the DF flag turned on drops the packet and sends an ICMP message back to the sender which also includes the local maximum link MTU.
- Upon receiving the ICMP message, the sender adapts his estimate of the path MTU and retries.
- Since the path MTU can change dynamically (since the path can change), a once learned path MTU should be verified and adjusted periodically.
- Not all routers send necessarily the local link MTU. In this cases, the sender tries typical MTU values, which is usually faster than doing a binary search.

# IPv4 over IEEE 802.3 (RFC 894)

- IPv4 packets are identified by the value 0x800 in the IEEE 802.3 type field.
- According to the maximum length of IEEE 802.3 frames, the maximum link MTU is 1500 byte.
- The mapping of IPv4 addresses to IEEE 802.3 addresses is table driven. Entries in so called mapping tables (sometimes also called address translation tables) can either be statically configured or dynamically learned.
- Note that the RFC 894 approach does not provide an assurance that the mapping is actually correct...

# IPv4 Address Translation (RFC 826)



- The Address Resolution Protocol (ARP) resolved IPv4 addresses to link-layer addresses of neighboring nodes.

# ARP and RARP

- The Hardware Type field identifies the address type used on the link-layer (the value 1 is used for IEEE 802.3 MAC addresses).
- The Protocol Type field identifies the network layer address type (the value 0x800 is used for IPv4).
- ARP/RARP packets use the 802.3 type value 0x806.
- The Operation field contains the message type: ARP Request (1), ARP Response (2), RARP Request (3), RARP Response (4).
- The sender fills, depending on the request type, either the Target IP Address field (ARP) or the Target Hardware Address field (RARP).
- The responding node swaps the Sender/Target fields and fills the empty fields with the requested information.

# DHCP Version 4

- The Dynamic Host Configuration Protocol (DHCP) allows nodes to retrieve configuration parameters from a central configuration server.
- A binding is a collection of configuration parameters, including at least an IP address, associated with or bound to a DHCP client.
- Bindings are managed by DHCP servers.
- Bindings are typically valid only for a limited lifetime.
- See RFC 2131 for the details and the message formats.
- See RFC 3118 for security aspects due to lack of authentication.

# DHCPv4 Message

- The DHCPDISCOVER message is a broadcast message which is sent by DHCP clients to locate DHCP servers.
- The DHCPOFFER message is sent from a DHCP server to offer a client a set of configuration parameters.
- The DHCPREQUEST is sent from the client to a DHCP server as a response to a previous DHCPOFFER message, to verify a previously allocated binding or to extend the lease of a binding.
- The DHCPACK message is sent by a DHCP server with some additional parameters to the client as a positive acknowledgement to a DHCPREQUEST.
- The DHCPNAK message is sent by a DHCP server to indicate that the client's notion of a configuration binding is incorrect.

# DHCPv4 Message Types

- The DHCPDECLINE message is sent by a DHCP client to indicate that parameters are already in use.
- The DHCPRELEASE message is sent by a DHCP client to inform the DHCP server that configuration parameters are no longer used.
- The DHCPINFORM message is sent from the DHCP client to inform the DHCP server that only local configuration parameters are needed.



# Referenc



C. Kent and J. Mogul.  
Fragmentation Considered Harmful.  
*In Proc. SIGCOMM '87 Workshop on Frontiers in Computer Communications Technology*, August 1987.



R.P. Draves, C. King, S. Venkatachary, and B.N. Zill.  
Constructing Optimal IP Routing Tables.  
*In Proc. 18th IEEE INFOCOM 1999*, pages 88–97, New York, March 1999.



M. A. Ruiz-Sánchez, E. W. Biersack, and W. Dabbous.  
Survey and Taxonomy of IP Address Lookup Algorithms.  
*IEEE Network*, 15(2):8–23, March 2001.



D. C. Plummer.  
An Ethernet Address Resolution Protocol.  
RFC 826, MIT, November 1982.



C. Hornig.  
A Standard for the Transmission of IP Datagrams over Ethernet Networks.  
RFC 894, Symbolics Cambridge Research Center, April 1984.



J. Mogul and S. Deering.  
Path MTU Discovery.  
RFC 1191, DECWRL, Stanford University, November 1990.

# Reference



R. Droms.

Dynamic Host Configuration Protocol.  
RFC 2131, Bucknell University, March 1997.



R. Droms and W. Arbaugh.

Authentication for DHCP Messages.  
RFC 3118, Cisco Systems, University of Maryland, June 2001.



S. Deering and R. Hinden.

Internet Protocol, Version 6 (IPv6) Specification.  
RFC 2460, Cisco, Nokia, December 1998.



T. Narten, E. Nordmark, W. Simpson, and H. Soliman.

Neighbor Discovery for IP version 6 (IPv6).  
RFC 4861, IBM, Sun Microsystems, Daydreamer, Elevate Technologies, September 2007.



A. Conta, S. Deering, and M. Gupta.

Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification.  
RFC 4443, Transwitch, Cisco Systems, Tropos Networks, March 2006.



M. Crawford.

Transmission of IPv6 Packets over Ethernet Networks.  
RFC 2464, Fermilab, December 1998.

# Reference



R. Droms, J. Bound, B. Volz, T. Lemon, C. Perkins, and M. Carney.

Dynamic Host Configuration Protocol for IPv6 (DHCPv6).

RFC 3315, Cisco, Hewlett Packard, Ericsson, Nominum, Nokia Research Center, Sun Microsystems, July 2003.



IANA.

Special-Use IPv4 Addresses.

RFC 3330, Internet Assigned Numbers Authority, September 2002.



M. Blanchet.

Special-Use IPv6 Addresses.

RFC 5156, Viagenie, April 2008.



T. Narten, R. Draves, and S. Krishnan.

Privacy Extensions for Stateless Address Autoconfiguration in IPv6.

RFC 4941, IBM Corporation, Microsoft Research, Ericsson Research, September 2007.