



Exam Booklet

TCP/IP Networking 2022-2023

Special Addresses

0.0.0.0	absence of address
127.0.0/24 for example 127.0.0.1	this host (loopback address)
10.0.0.0/8, 172.16.0.0/12, 192.168.0.0/16	private networks (e.g. in IEW) cannot be used on the public Internet
100.64/10	private addresses for use by ISPs only (Carrier Grade NAT addresses)
192.88.99/24	6to4 relay routers
169.254.0.0/16	link local address (can be used only between systems on same LAN)
224/4	multicast
240/5	reserved
255.255.255.255/32	link local broadcast

A Few IPv6 Global Unicast Addresses

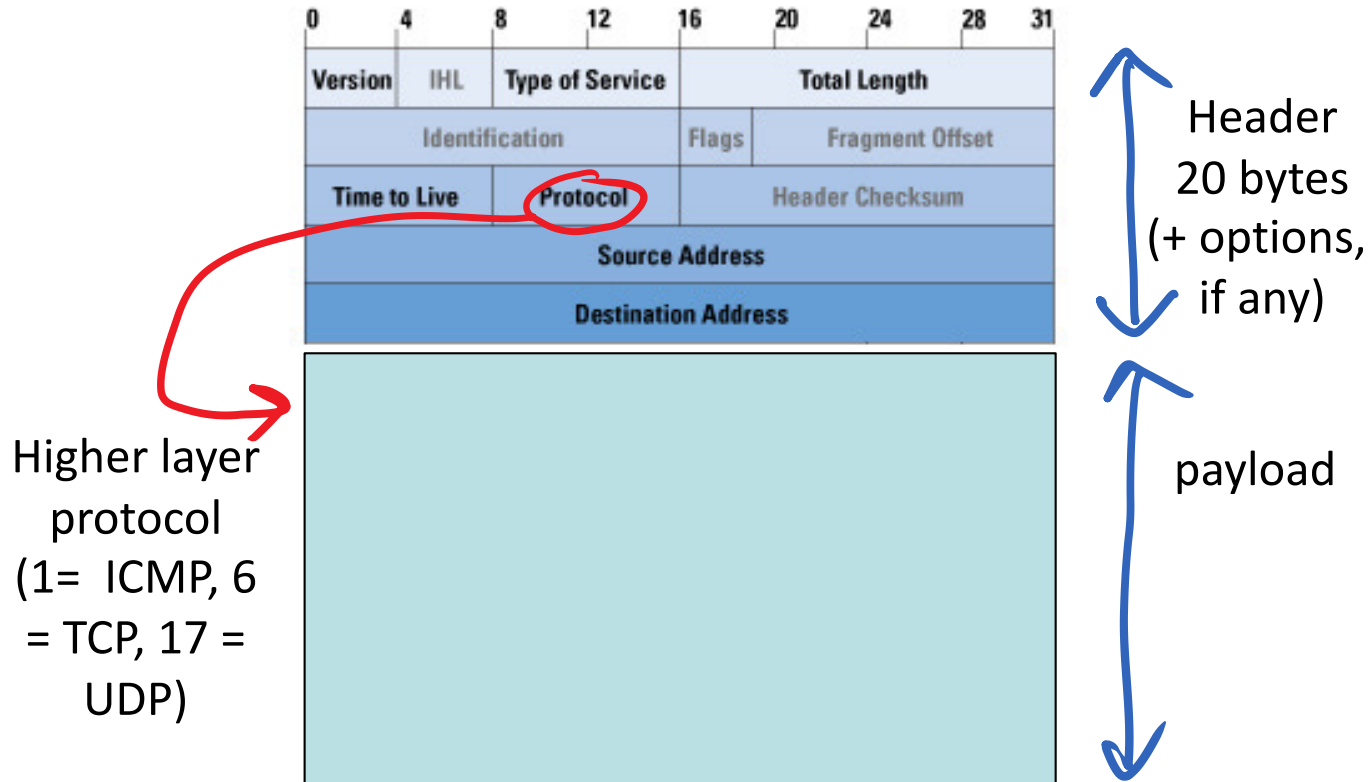
The block 2000/3 (i.e. 2xxx and 3xxx) is allocated for global unicast addresses

2001:620::/32	Switch
2001:620:618::/48	EPFL
2001:620:8::/48	ETHZ
2a02:1200::/27	Swisscom
2001:678::/29	provider independent address
2001::/32	Teredo (tunnels IPv6 in IPv4)
2002::/16	6to4 (tunnels IPv6 in IPv4)

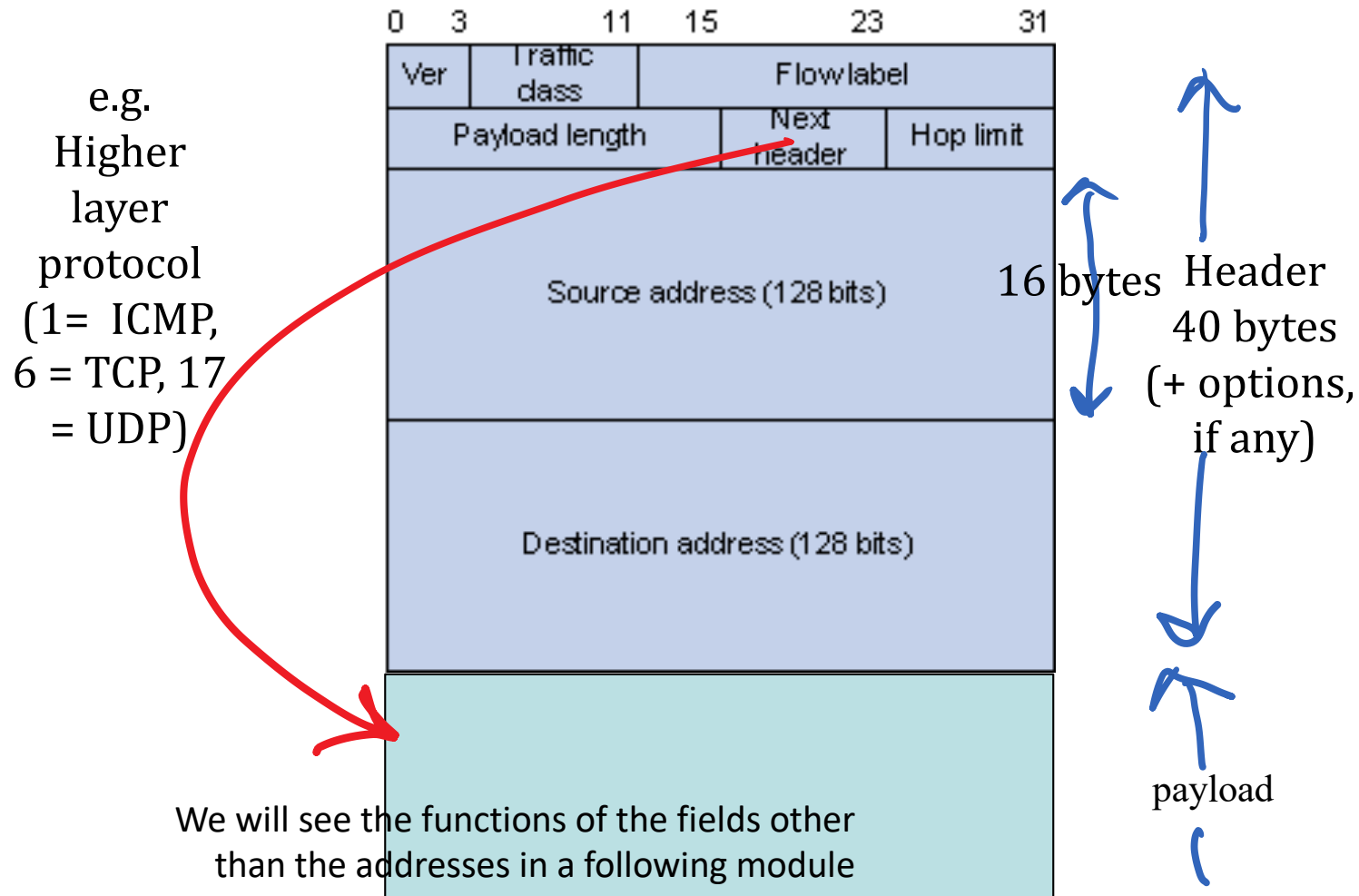
Examples of Special Addresses

	::/128	absence of address
	::1/128	this host (loopback address)
EPFL Private	fc00::/7 (i.e. fcxx: and fdxx:)	Unique local addresses = private networks (e.g. in IEW) cannot be used on the public Internet
	For example fd24:ec43:12ca:1a6:a00: 20ff:fe78:30f9	
	fe80::/10	link local address (can be used only between systems on same LAN)
	ff00::/8	multicast
	ff02::1:ff00:0/104	Solicited node multicast
	ff02::1/128	link local broadcast
	ff02::2/128	all link local routers

IPv4 Packet Format



IPv6 Packet Format



Ethernet Frame format

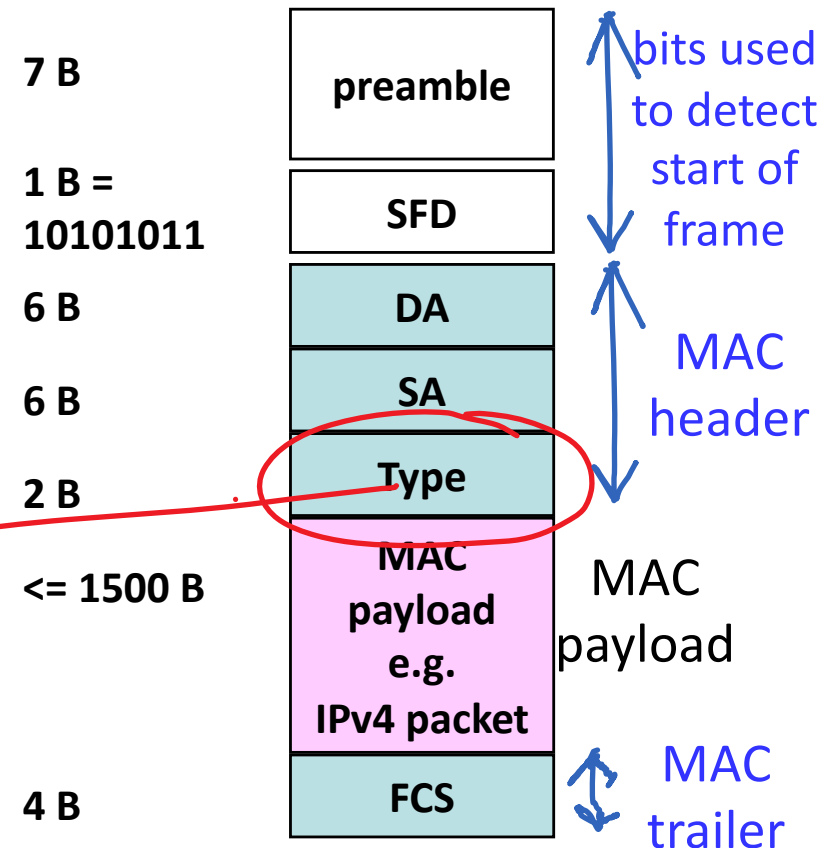
Ethernet frame = Ethernet PDU

An Ethernet frame typically transports an IP packet, sometimes also other

Type of protocol contained in the Ethernet packet (hexa):

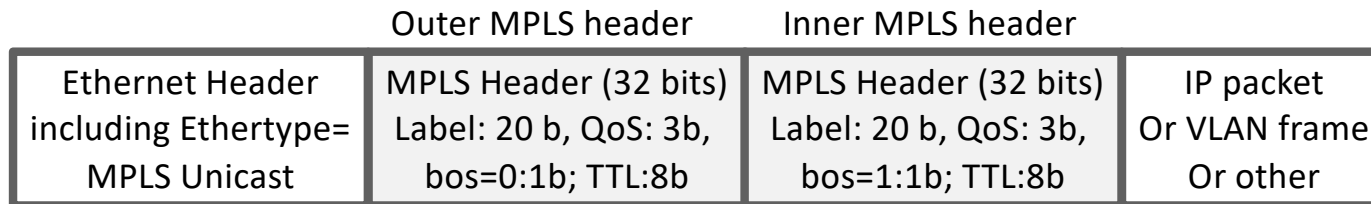
0800: IPv4
0806: ARP (used by IPv4)
86DD: IPv6
8847: MPLS unicast
88F7: Precision Time Protocol

Ethernet V.2 frame



DA = destination address
SA = source address

MPLS Header



MPLS header comes after the MAC layer header. There is one Ethertype for MPLS unicast (and one for MPLS multicast, not discussed here).

The MPLS header contains mainly the label, plus 3 bits for quality of service and a TTL field (similar to IPv4's TTL).

There can be several MPLS headers on top of one another. The innermost header is recognized by the flag "Bottom of Stack" (bos).

Multicast MAC Addresses

IP multicast address is **algorithmically** mapped to a multicast MAC address.

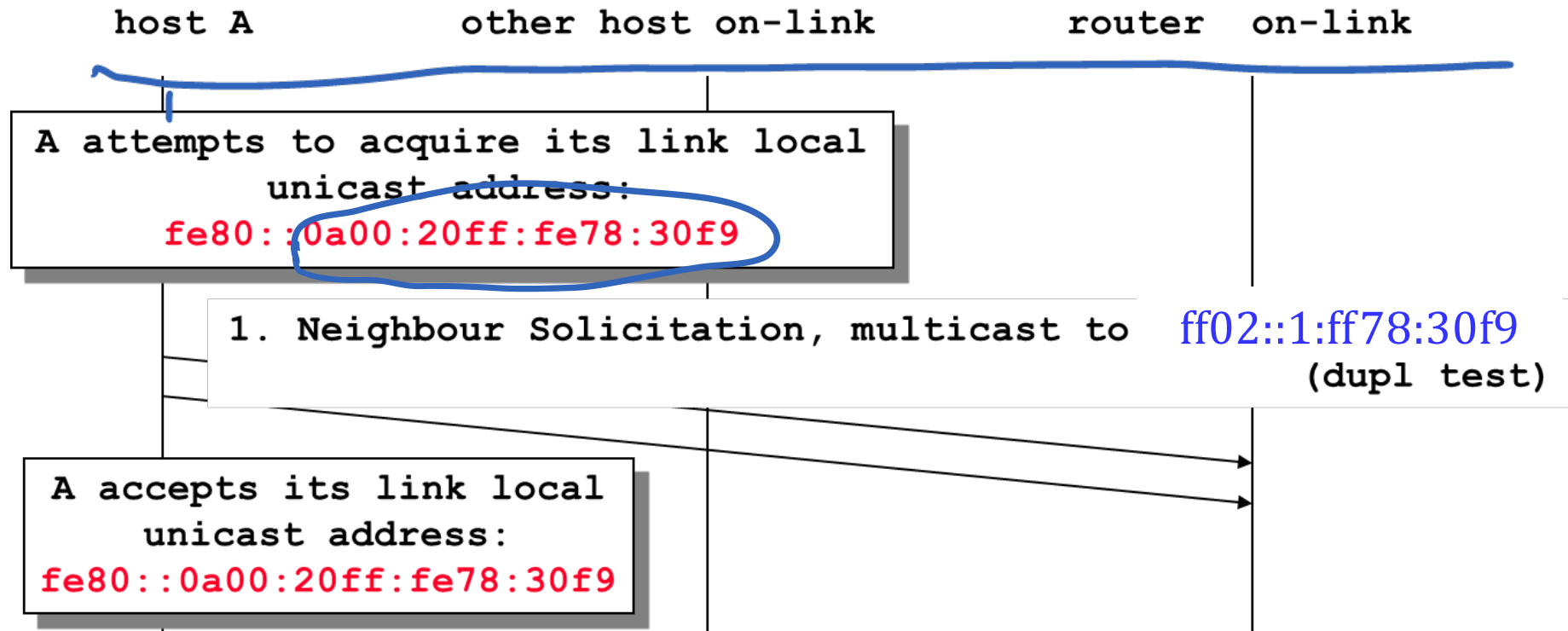
Last 23 bits of IPv4 multicast address are used in MAC address

Last 32 bits of IPv6 multicast address are used in MAC address

<i>MAC multicast addr.</i>	<i>Used for</i>
01-00-5e-XX-XX-XX	IPv4 multicast
33-33-XX-XX-XX-XX	IPv6 multicast

<i>IP dest address</i>	229.130.54.207
<i>IP dest address (hexa)</i>	e5-82-36-cf
<i>IP dest address (bin)</i>	...-10000010-...
<i>Keep last 23 bits (bin)</i>	...-00000010-...
<i>Keep last 23 bits (hexa)</i>	02-36-cf
<i>MAC address</i>	01-00-5e-02-36-cf

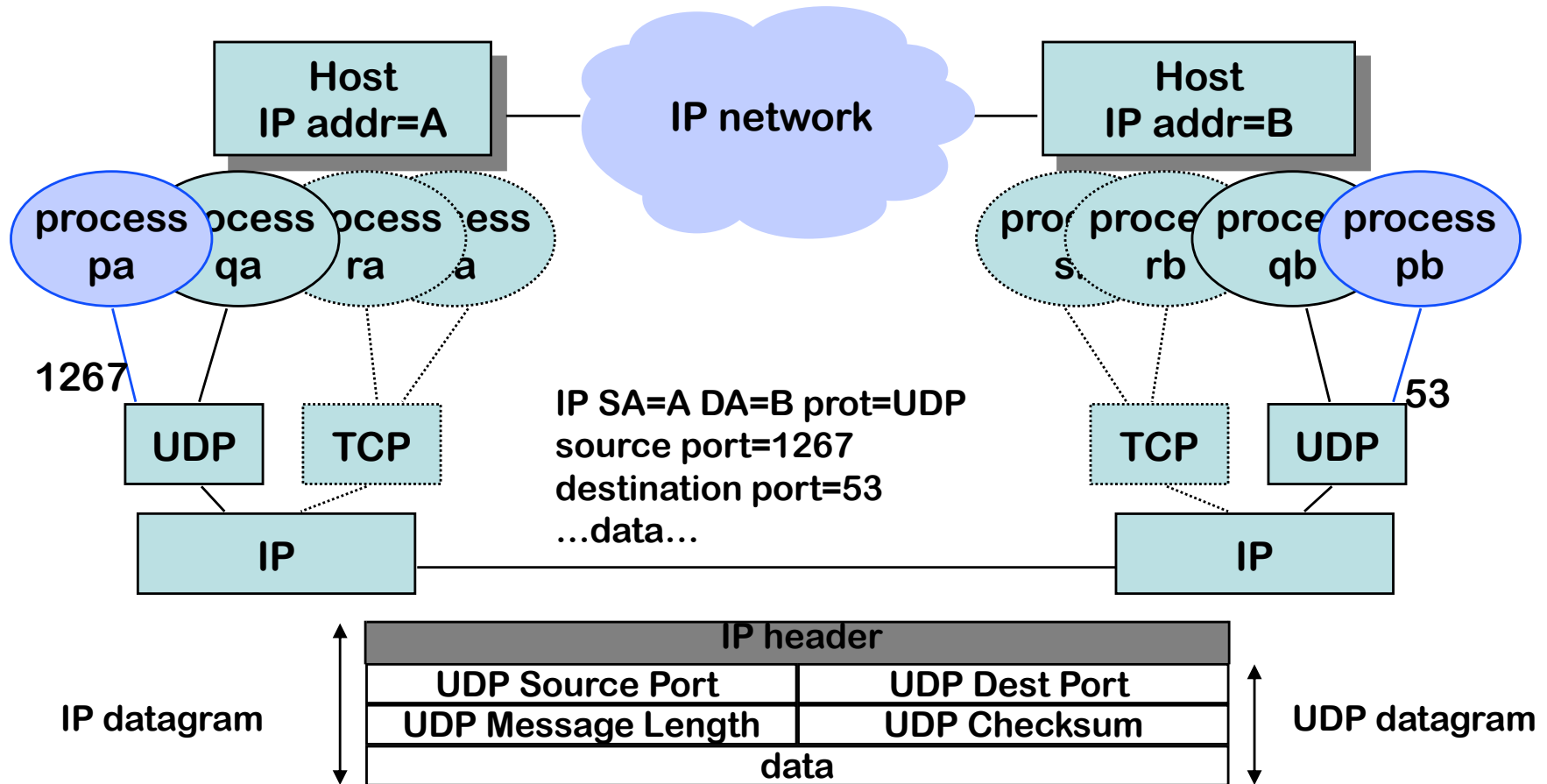
SLAAC Step 2: Duplicate Test

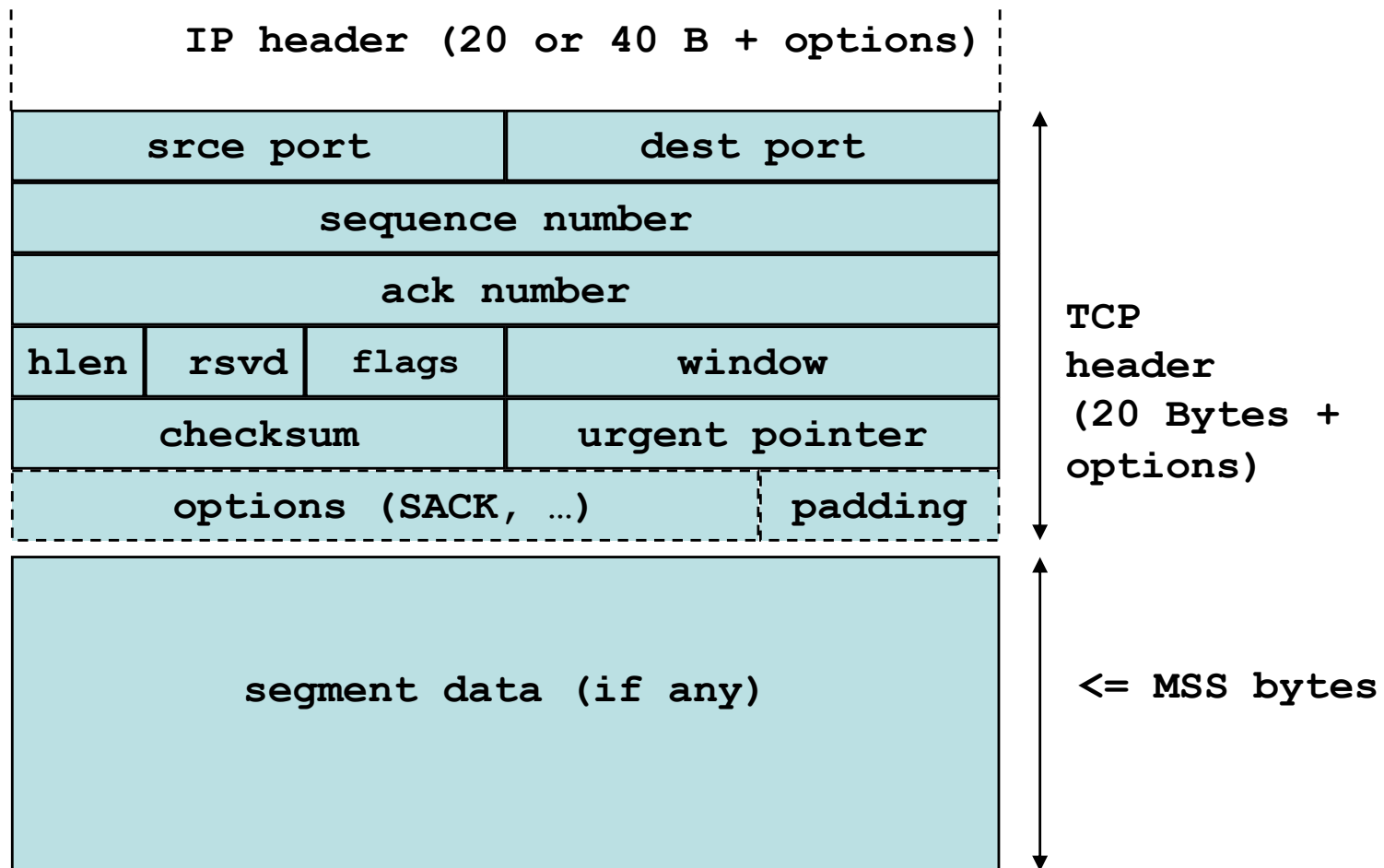


A sends a Neighbour Solicitation (NS) message to check for address duplication, sent to the **Solicited Node Multicast Address**.

Any host that would have to same link local address listens to this multicast address

UDP Uses Port Numbers





flags meaning

NS	used for explicit congestion notification
CWR	used for explicit congestion notification
ECN	used for explicit congestion notification
urg	urgent ptr is valid
ack	ack field is valid
psh	this seg requests a push
rst	reset the connection
syn	connection setup
fin	sender has reached end of byte stream

The Decision Process

The **decision process** decides which route is selected;

At most one best route to exactly the same prefix is chosen

Only one route to 2.2/16 can be chosen

But there can be different routes to 2.2.2/24 and 2.2/16

A route can be selected only if its next-hop is reachable

Routes are compared against each other using a sequence of criteria, until only one route remains. A common sequence is

0. Highest weight (Cisco proprietary)
1. Highest LOCAL-PREF
2. Shortest AS-PATH
3. Lowest MED, if taken seriously by this network
4. E-BGP > I-BGP
5. Shortest path to NEXT-HOP, according to IGP
6. Lowest BGP identifier (router-id of the BGP peer from whom route is received)

(The Cisco and FRR implementation of BGP, used in lab, have a few additional cases, not shown here)

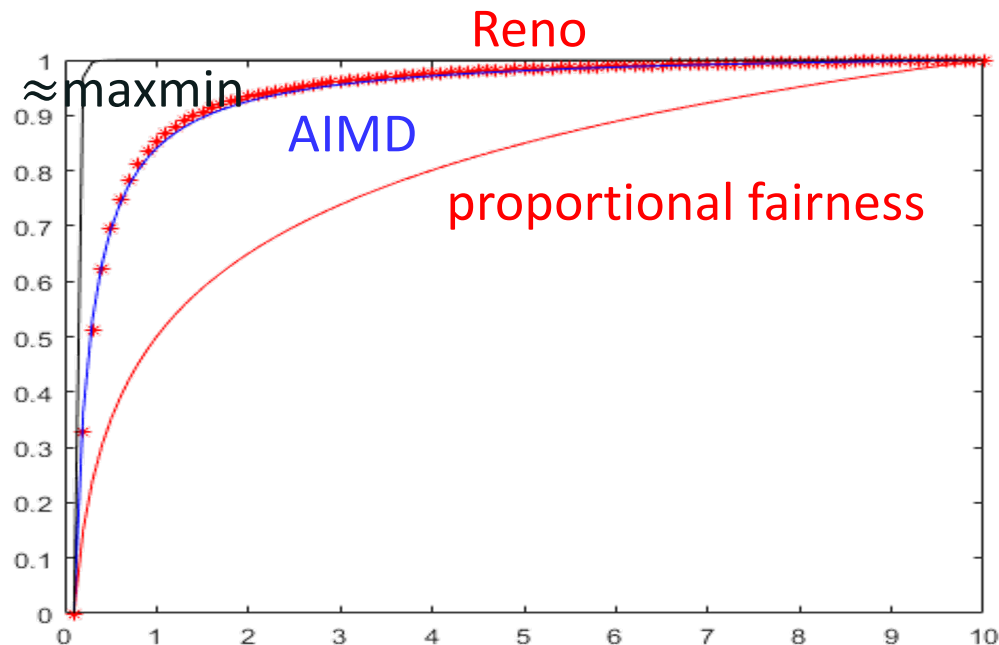
Fairness of TCP Reno

For long lived flows, the rates obtained with TCP are as if they were distributed according to utility fairness, with utility of flow i given by

$$U(x_i) = \frac{\sqrt{2}}{\tau_i} \arctan \frac{x_i \tau_i}{\sqrt{2}}$$

with $x_i = \text{rate} = W / \tau_i$, $\tau_i = \text{RTT}$

For sources that have same RTT, the fairness of TCP is between maxmin fairness and proportional fairness, closer to proportional fairness



rescaled utility

functions;

RTT = 100 ms

maxmin approx. is $U(x) = 1 - x^{-5}$

TCP Reno

Loss - Throughput Formula

Consider a *large* TCP connection (many bytes to transmit)

Assume we observe that, in average, a fraction q of packets is lost (or marked with ECN)

The throughput should be close to $\theta = \frac{MSS \cdot 1.22}{RTT \sqrt{q}}$

Formula assumes: transmission time negligible compared to RTT, losses are rare, time spent in Slow Start and Fast Recovery negligible, losses occur periodically

Cubic's Other Bells and Whistles

Cubic's Loss throughput formula

$$\theta \approx \max \left(\frac{1.054}{RTT^{0.25} q^{0.75}}, \frac{1.22}{RTT \sqrt{q}} \right)$$

in MSS per second.

Cubic's formula is same as Reno for small RTTs and small BW-delay products.

Other Cubic details

W_{max} computation uses a more complex mechanism called "fast convergence"

see Latest IETF Cubic RFC / Internet Draft

or http://elixir.free-electrons.com/linux/latest/source/net/ipv4/tcp_cubic.c

