Internet, Principes et Protocoles (IPP)

2. Routing and Network Layer

Table of Contents

- ~~Routing~~
  - ~~Static Routing~~
  - ~~Dynamic Routing~~
    - ~~Distance Vector~~
    - ~~LSPs~~
- IP protocol
- ARP Protocol

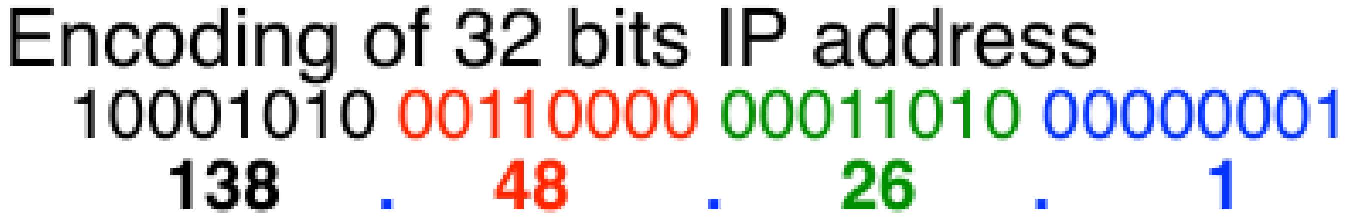
Network Layer

NETWORK LAYER: Forwarding information from source to destination, through multiple routers if necessary (Routing and logical addressing).

- provides unreliable connectionless service
    - some packets can be lost
    - packets can suffer from transmission errors
    - packets can be misordered

IP Addressing

- One IP address identifies one interface on one endhost or router
- Made of 4 bytes (32 bits)

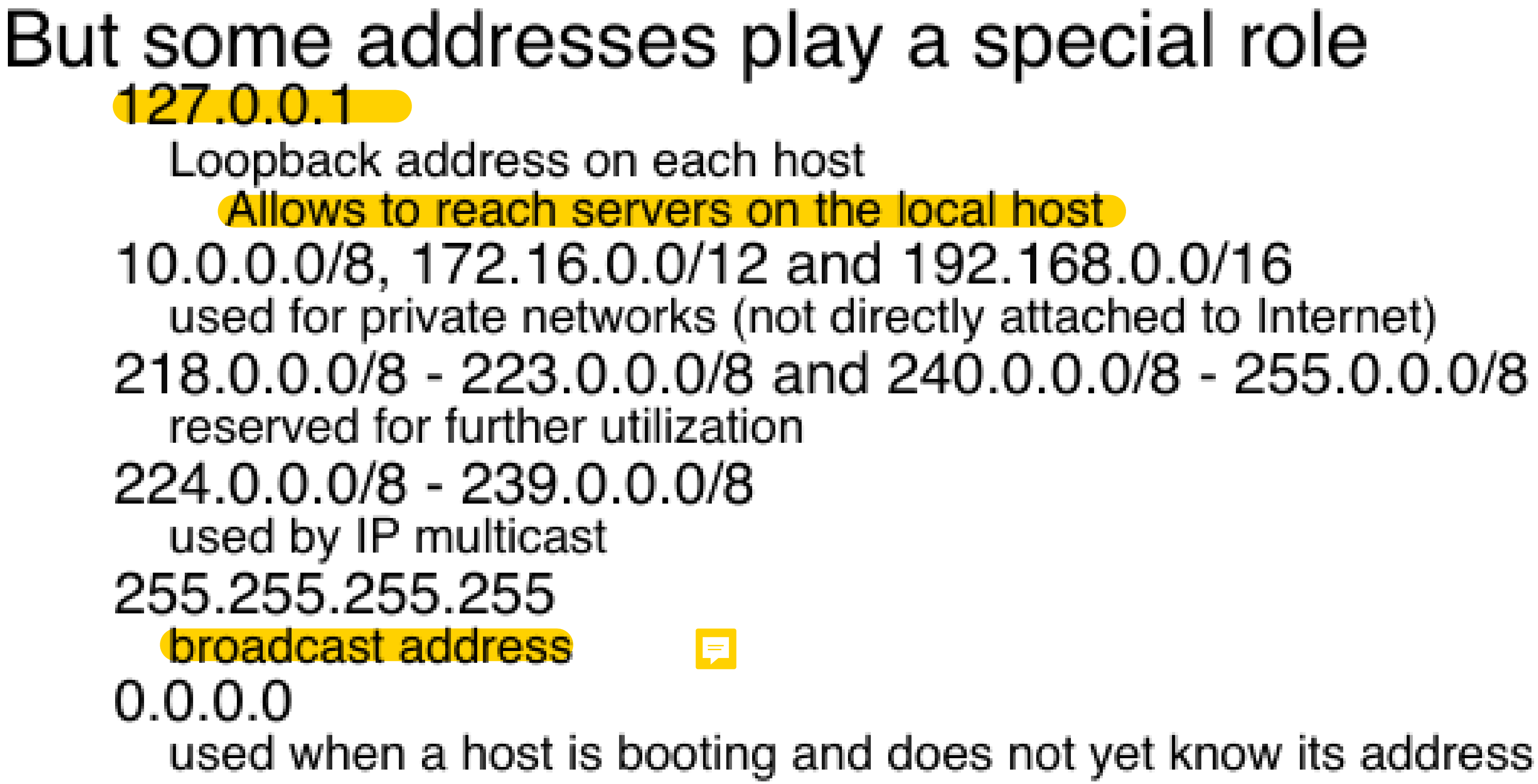


Encoding of 32 bits IP address

10001010 00110000 00011010 00000001

138 . 48 . 26 . 1

IP Addressing

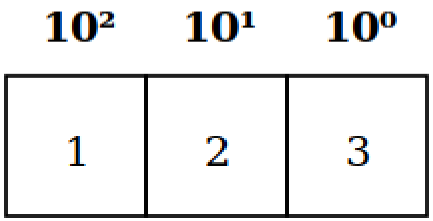


But some addresses play a special role

- 127.0.0.1
  - Loopback address on each host
    - Allows to reach servers on the local host
- 10.0.0.0/8, 172.16.0.0/12 and 192.168.0.0/16
  - used for private networks (not directly attached to Internet)
- 218.0.0.0/8 - 223.0.0.0/8 and 240.0.0.0/8 - 255.0.0.0/8
  - reserved for further utilization
- 224.0.0.0/8 - 239.0.0.0/8
  - used by IP multicast
- 255.255.255.255
  - broadcast address
- 0.0.0.0
  - used when a host is booting and does not yet know its address

Binary 101

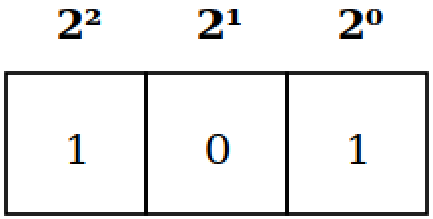
- Base 10 - 123



10^2	10^1	10^0
1	2	3

= 1*100 + 2*10 + 3*1 = 123

- Base 2 - 101



2^2	2^1	2^0
1	0	1

= 1*4 + 0*2 + 1*1 = 5

Binary 101

- What is 1111 1111 in decimal?

Binary 101

- What is 1111 1111 in decimal?

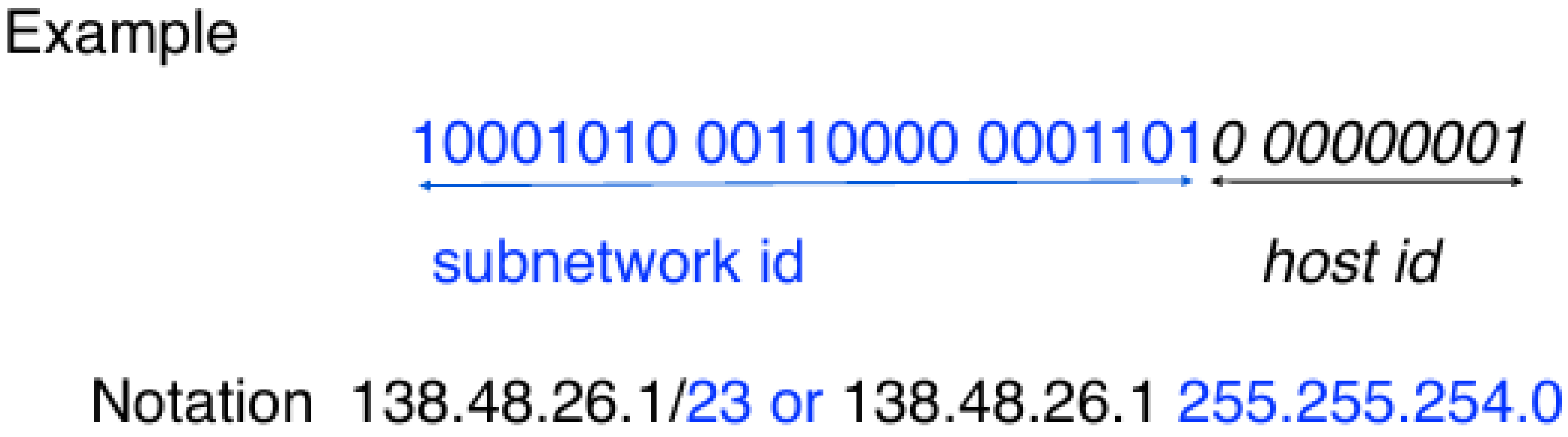
2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
1	1	1	1	1	1	1	1

= 1*128 + 1*64 + 1*32 + 1*16 + 1*8 + 1*4 + 1*2 + 1*1

= 255

IP Address mask

- There are only 3,706,452,992 public IP addresses available
- We use (private) subnets, and masks to identify them.



Example

10001010 00110000 00011010 00000001

subnetwork id

host id

Notation 138.48.26.1/23 or 138.48.26.1 255.255.254.0

IP Address mask

At home

- Public IP is the IP of the router (ex : 13.41.1.5).
- Private IP is the IP of the computers (often 192.168.0.0/24)

IP Address Mask

The network is identified by the IP range it defines. The bits “masked” don’t change.

Exercise: What is the range of IP s that can exist on this network: 194.23.21.0/19?

IP Address Mask

Exercise: What is the range of IP s that can exist on this network: 194.23.21.0/19?

1100 0010 . 0001 0111 . 0001 0101 . 0000 0000

Mask: 1100 0010 . 0001 0111 . 0001 0101 . 0000 0000

Smallest possible address:

1100 0010 . 0001 0111 . 0000 0000 . 0000 0000

194.23.0.0

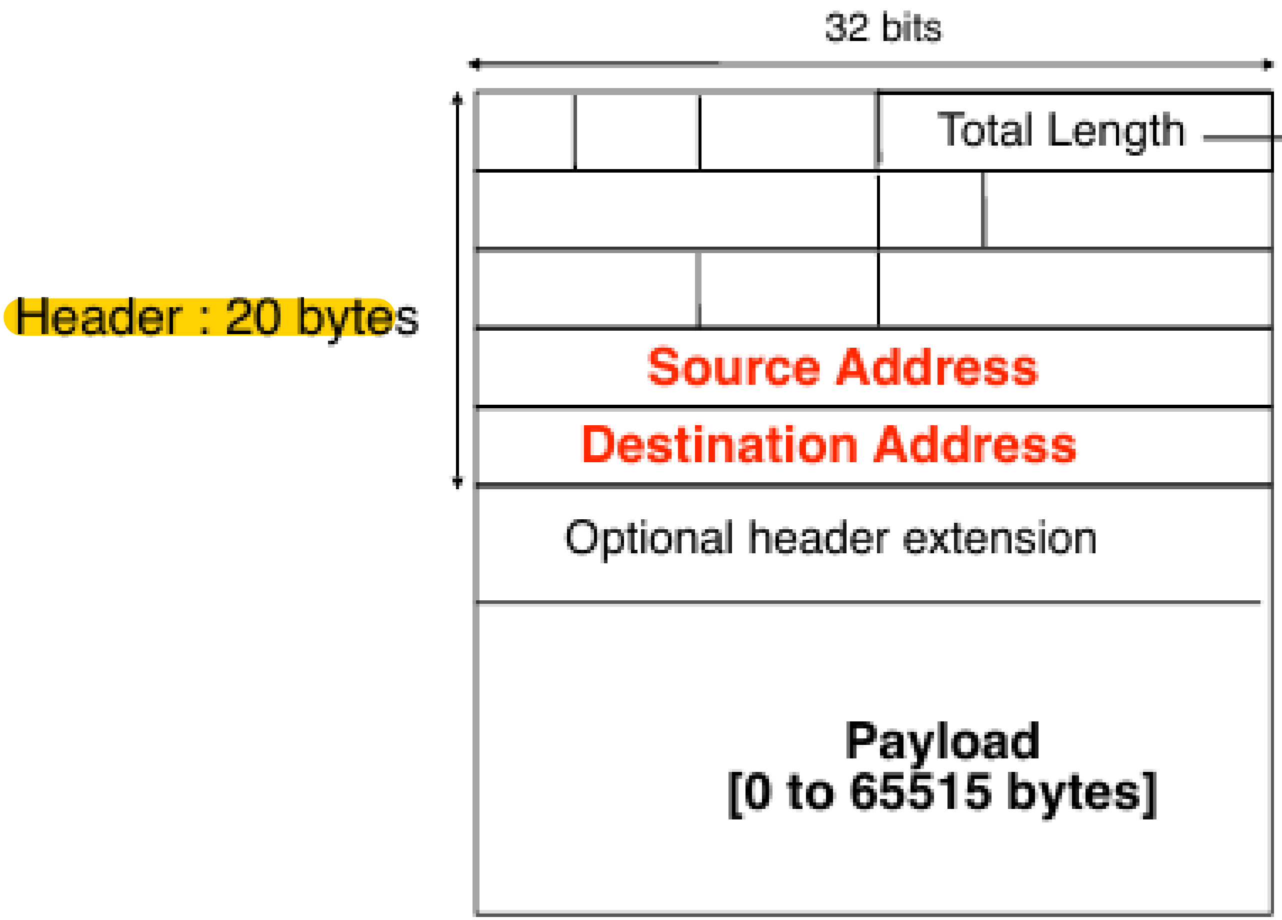
Biggest possible address:

1100 0010 . 0001 0111 . 0001 1111 . 1111 1111

194.23.31.254

IP Packet Format

IP packet format

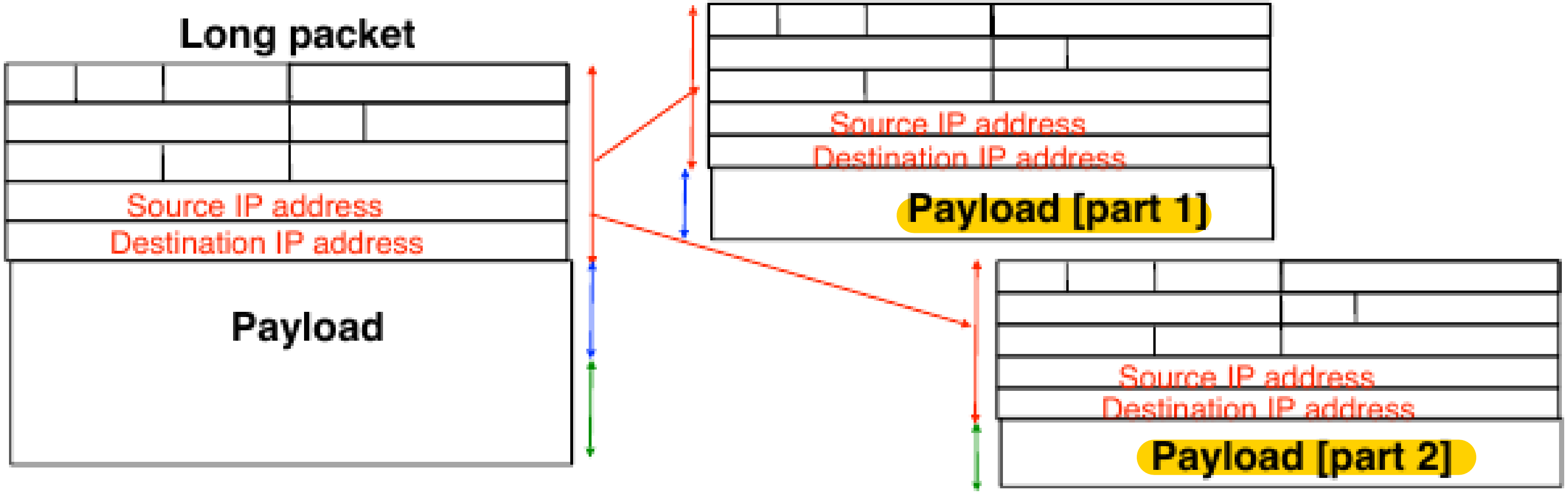


Total length of the packet in bytes
(16 bits – maximum)

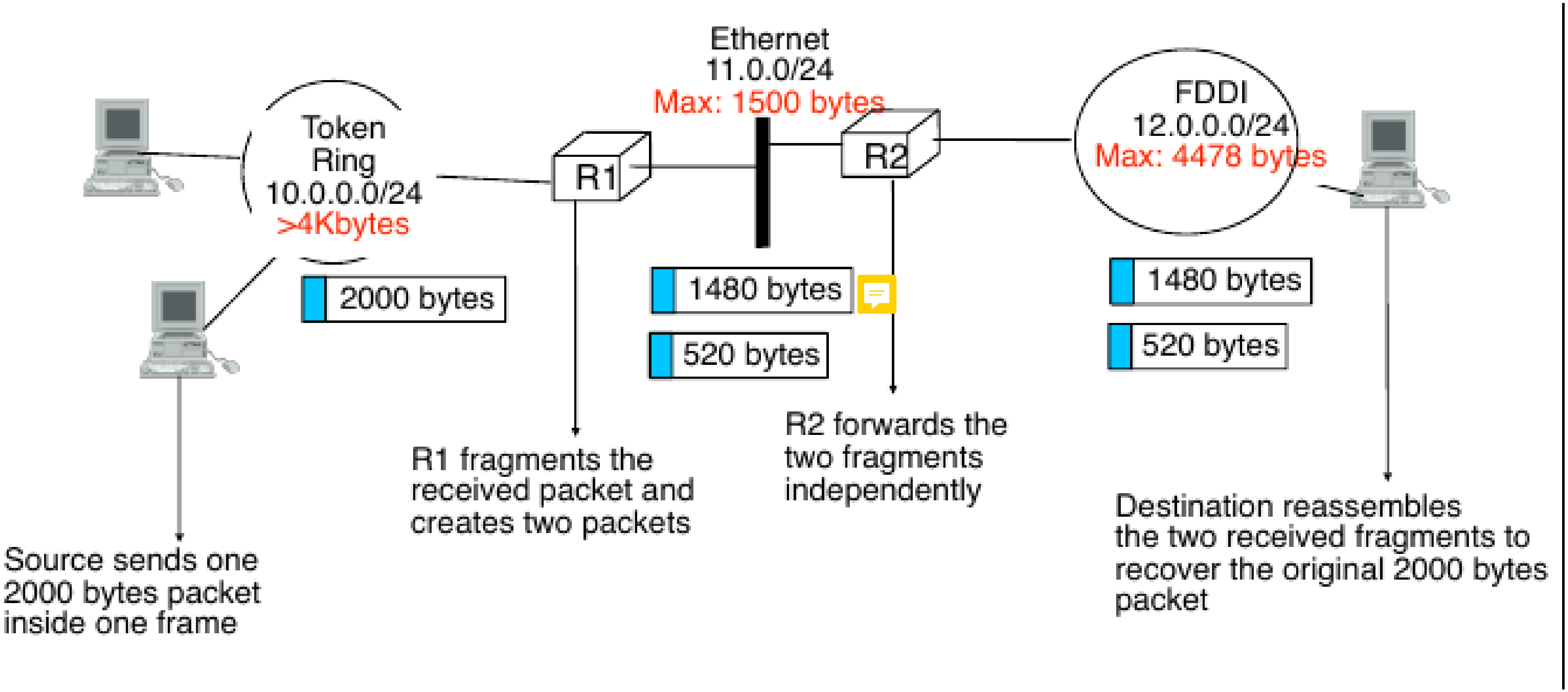
IP Fragmentation

Principe

- Each host and each router can fragment packets
  - Each fragment is a complete IP packet that contains source and destination IP addresses
- Only the destination host performs reassembly



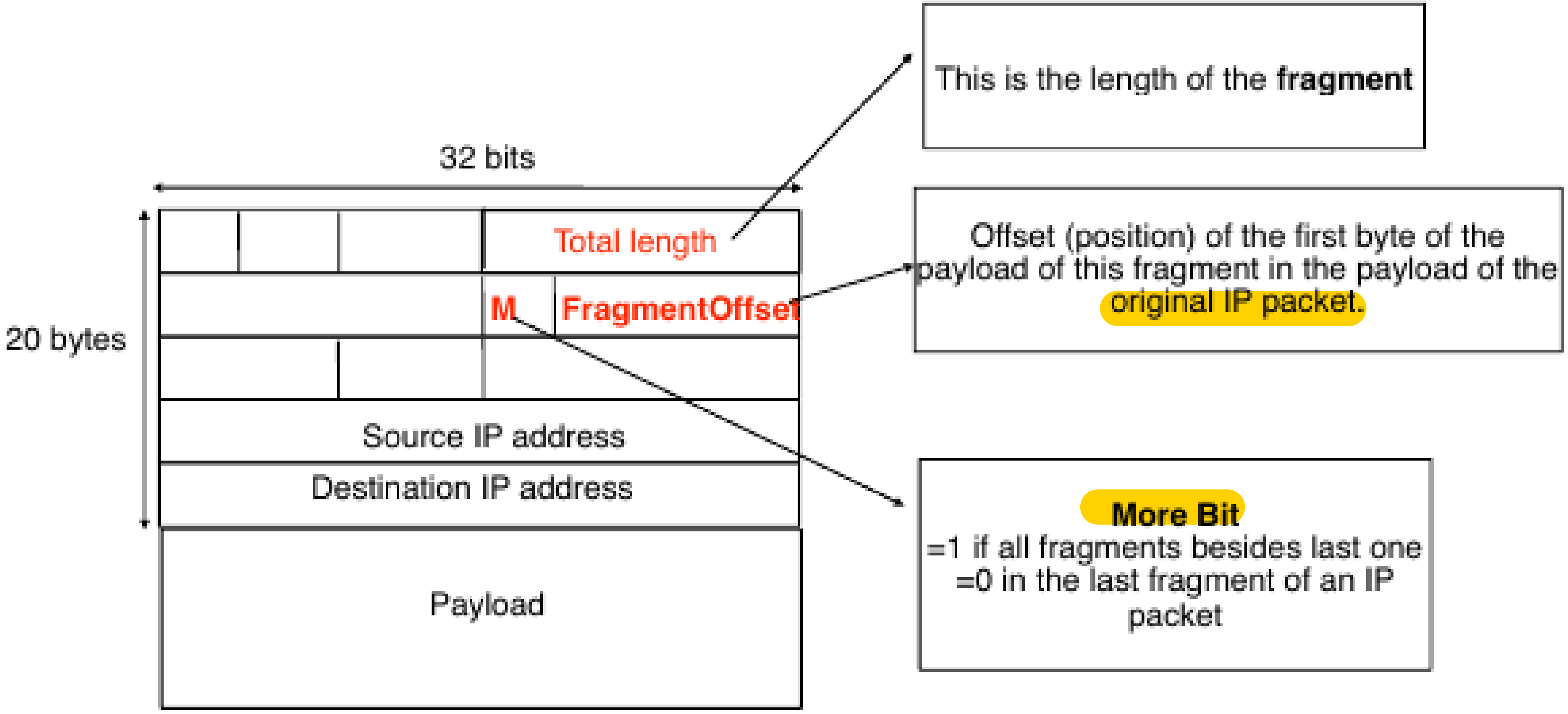
Fragmentation


Ethernet
11.0.0/24
Max: 1500 bytes
FDDI
12.0.0.0/24
Max: 4478 bytes
Token
Ring
10.0.0.0/24
>4Kbytes
R1
R2
2000 bytes
1480 bytes
520 bytes
1480 bytes
520 bytes
R1 fragments the received packet and creates two packets
R2 forwards the two fragments independently
Source sends one 2000 bytes packet inside one frame
Destination reassembles the two received fragments to recover the original 2000 bytes packet

IP Fragmentation

IP fragmentation

- Fragment the payload of IP packet
- Each fragment must be number to recover from misordering

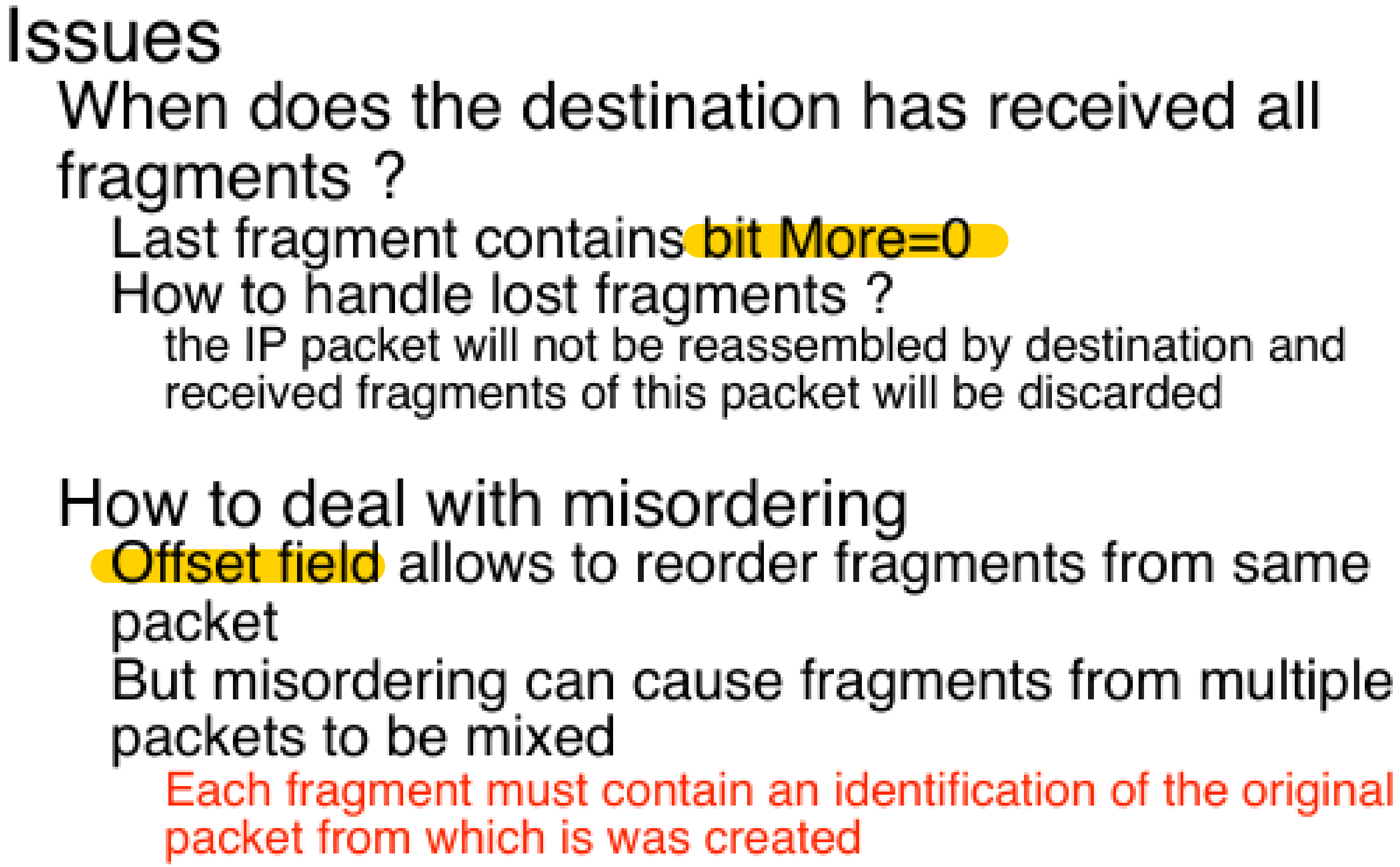


This is the length of the **fragment**

Offset (position) of the first byte of the payload of this fragment in the payload of the original IP packet.

More Bit
=1 if all fragments besides last one
=0 in the last fragment of an IP packet

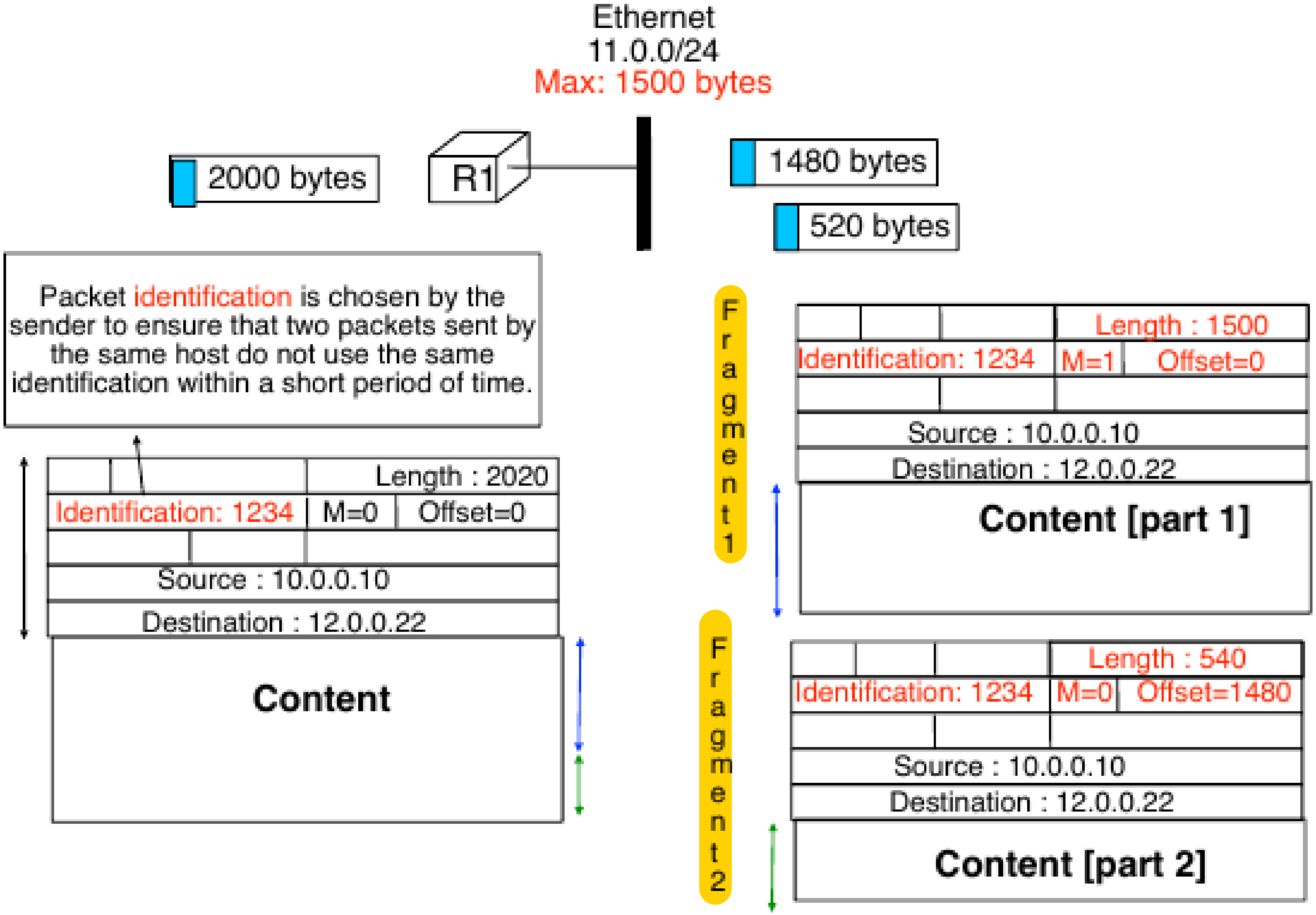
Reassembly



Issues

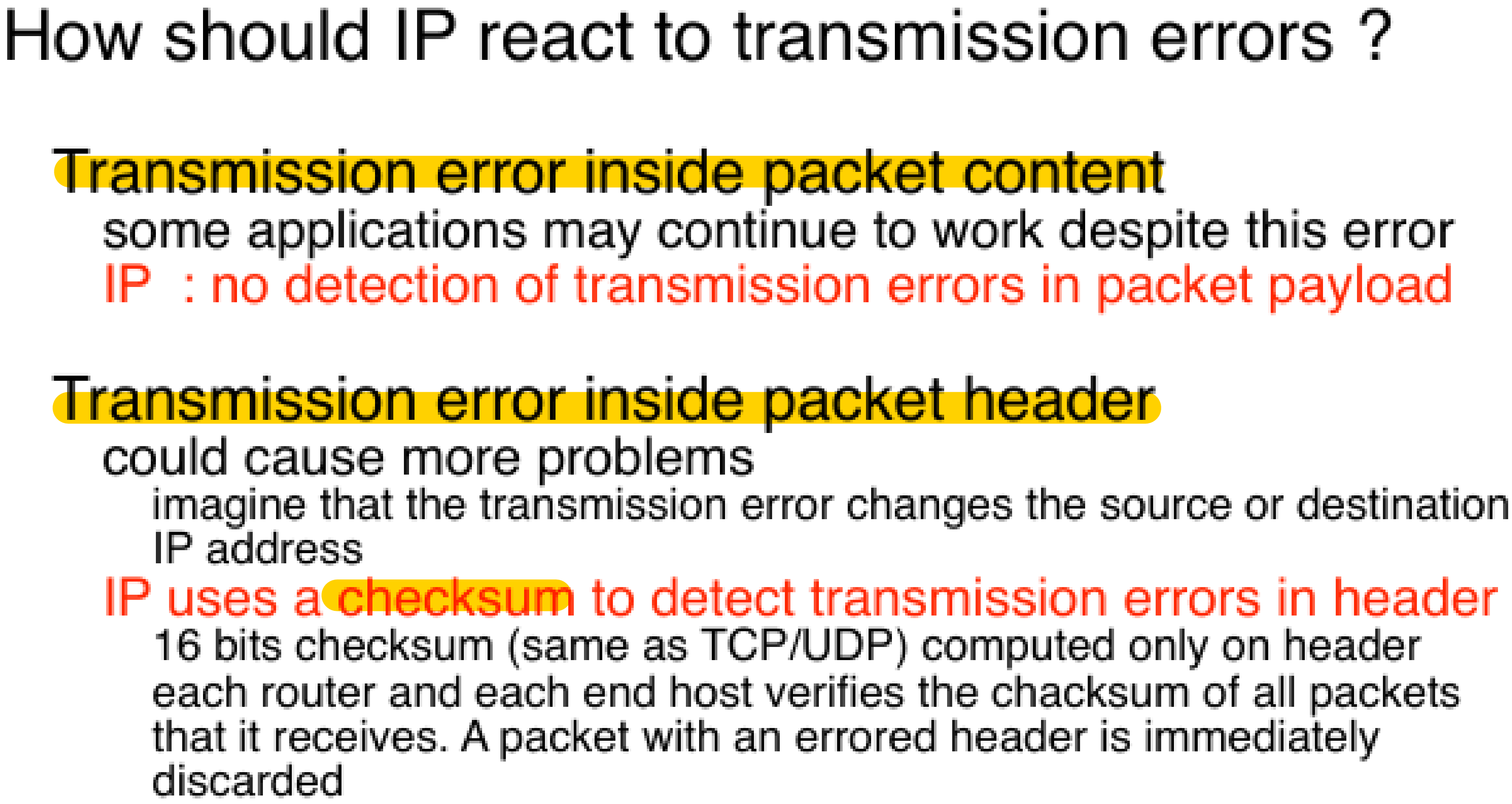
- When does the destination has received all fragments ?
  - Last fragment contains bit More=0
  - How to handle lost fragments ?
    - the IP packet will not be reassembled by destination and received fragments of this packet will be discarded
- How to deal with misordering
  - Offset field allows to reorder fragments from same packet
  - But misordering can cause fragments from multiple packets to be mixed
    - Each fragment must contain an identification of the original packet from which is was created

Fragment Identification



Packet identification is chosen by the sender to ensure that two packets sent by the same host do not use the same identification within a short period of time.

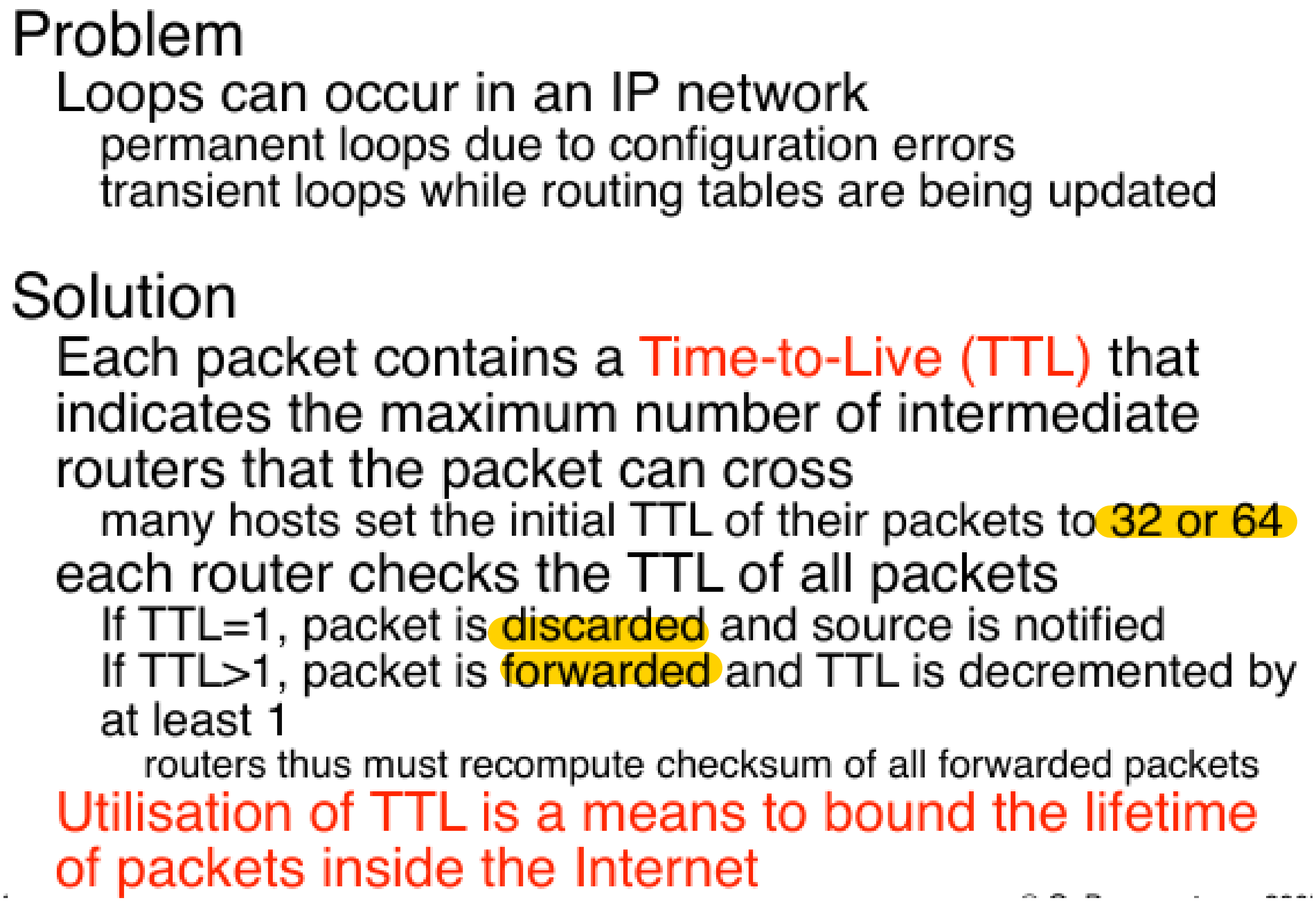
Transmision errors



How should IP react to transmission errors ?

- Transmission error inside packet content
  - some applications may continue to work despite this error
  - IP : no detection of transmission errors in packet payload
- Transmission error inside packet header
  - could cause more problems
    - imagine that the transmission error changes the source or destination IP address
  - IP uses a checksum to detect transmission errors in header
    - 16 bits checksum (same as TCP/UDP) computed only on header
    - each router and each end host verifies the chacksum of all packets that it receives. A packet with an errored header is immediately discarded

Time To Live



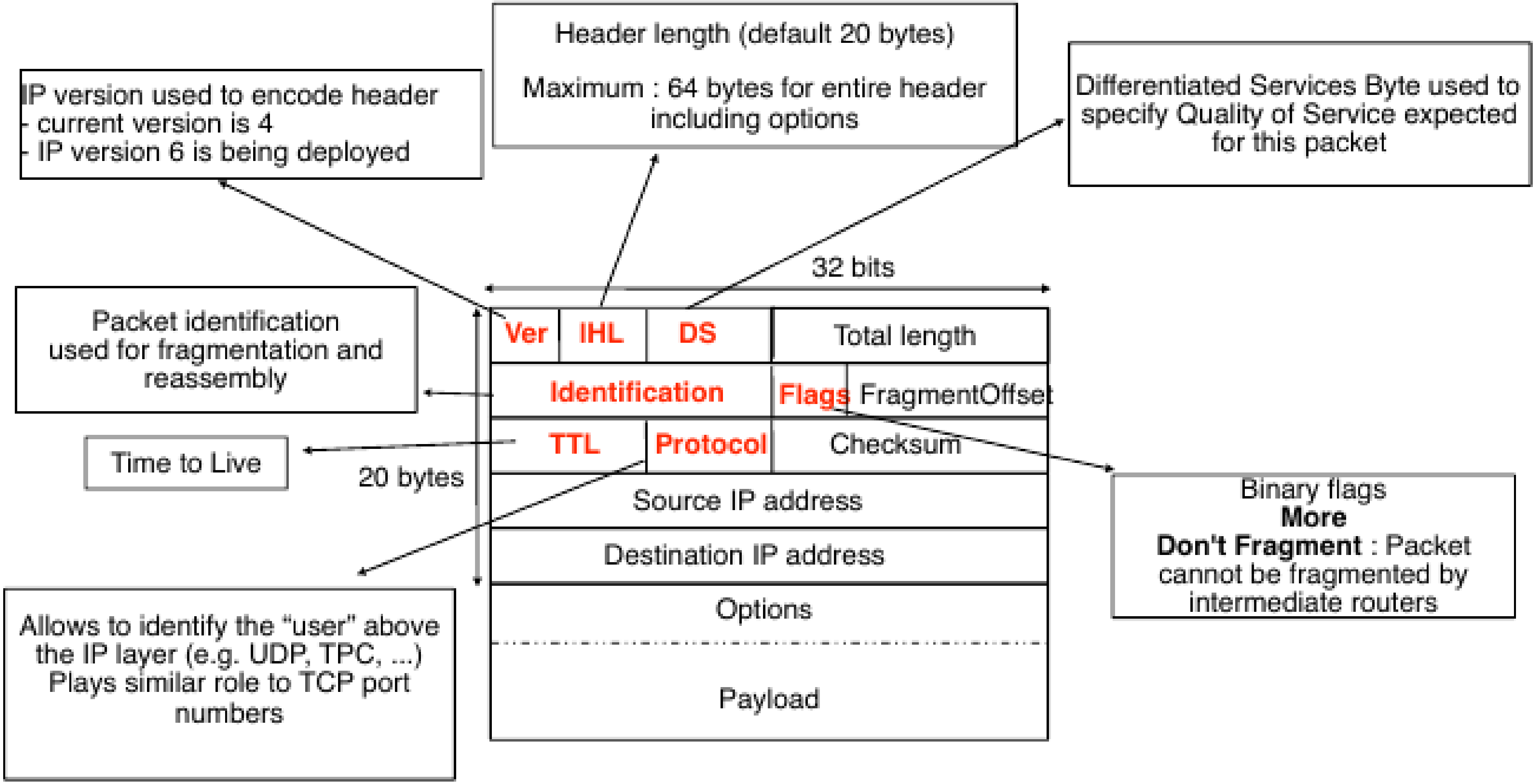
Problem

- Loops can occur in an IP network
  - permanent loops due to configuration errors
  - transient loops while routing tables are being updated

Solution

- Each packet contains a Time-to-Live (TTL) that indicates the maximum number of intermediate routers that the packet can cross
  - many hosts set the initial TTL of their packets to 32 or 64
- each router checks the TTL of all packets
  - If TTL=1, packet is discarded and source is notified
  - If TTL>1, packet is forwarded and TTL is decremented by at least 1
    - routers thus must recompute checksum of all forwarded packets
- Utilisation of TTL is a means to bound the lifetime of packets inside the Internet

IP header



IP version used to encode header
- current version is 4
- IP version 6 is being deployed

Header length (default 20 bytes)

Maximum : 64 bytes for entire header including options

Differentiated Services Byte used to specify Quality of Service expected for this packet

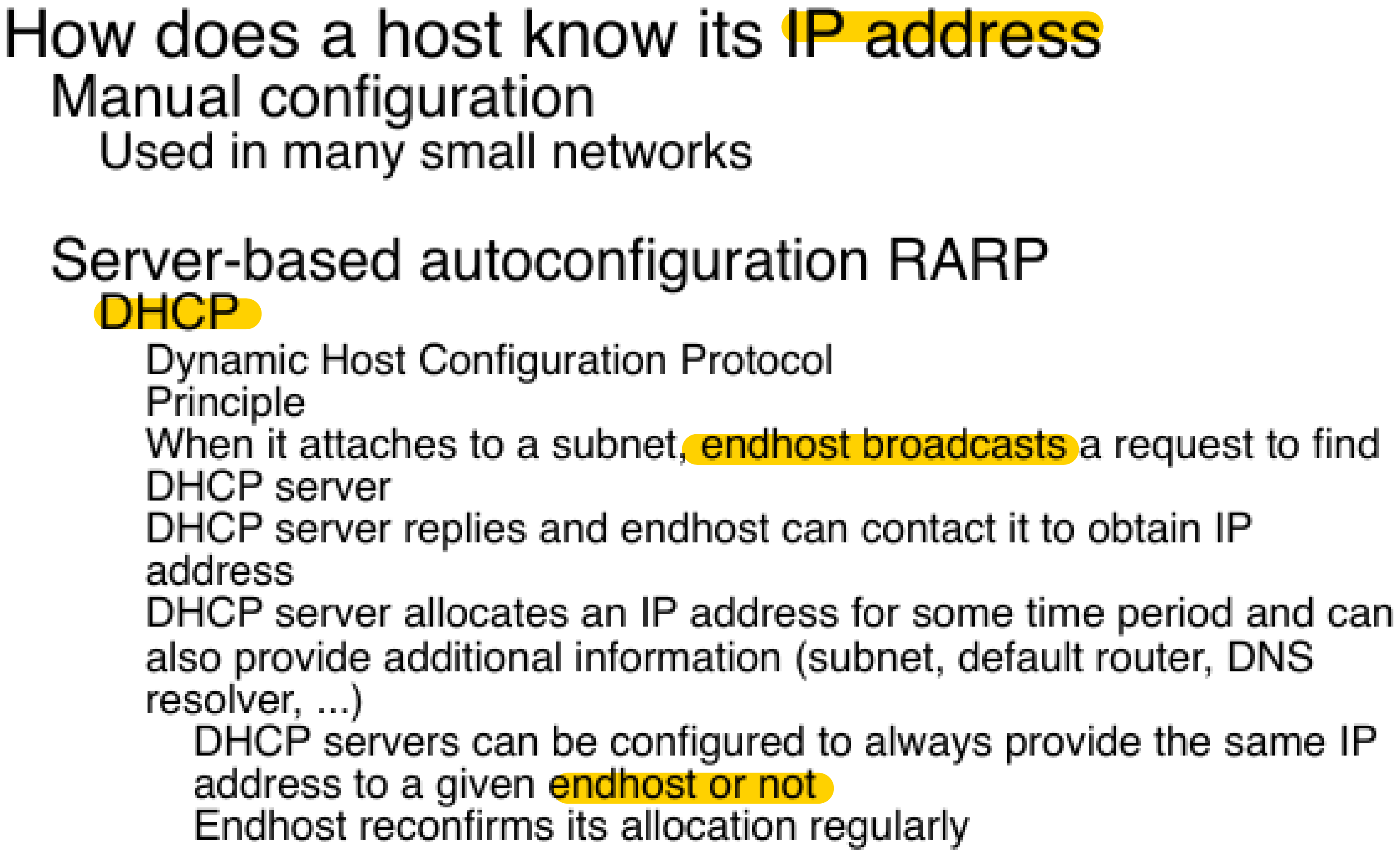
Packet identification used for fragmentation and reassembly

Time to Live

Allows to identify the “user” above the IP layer (e.g. UDP, TPC, ...) Plays similar role to TCP port numbers

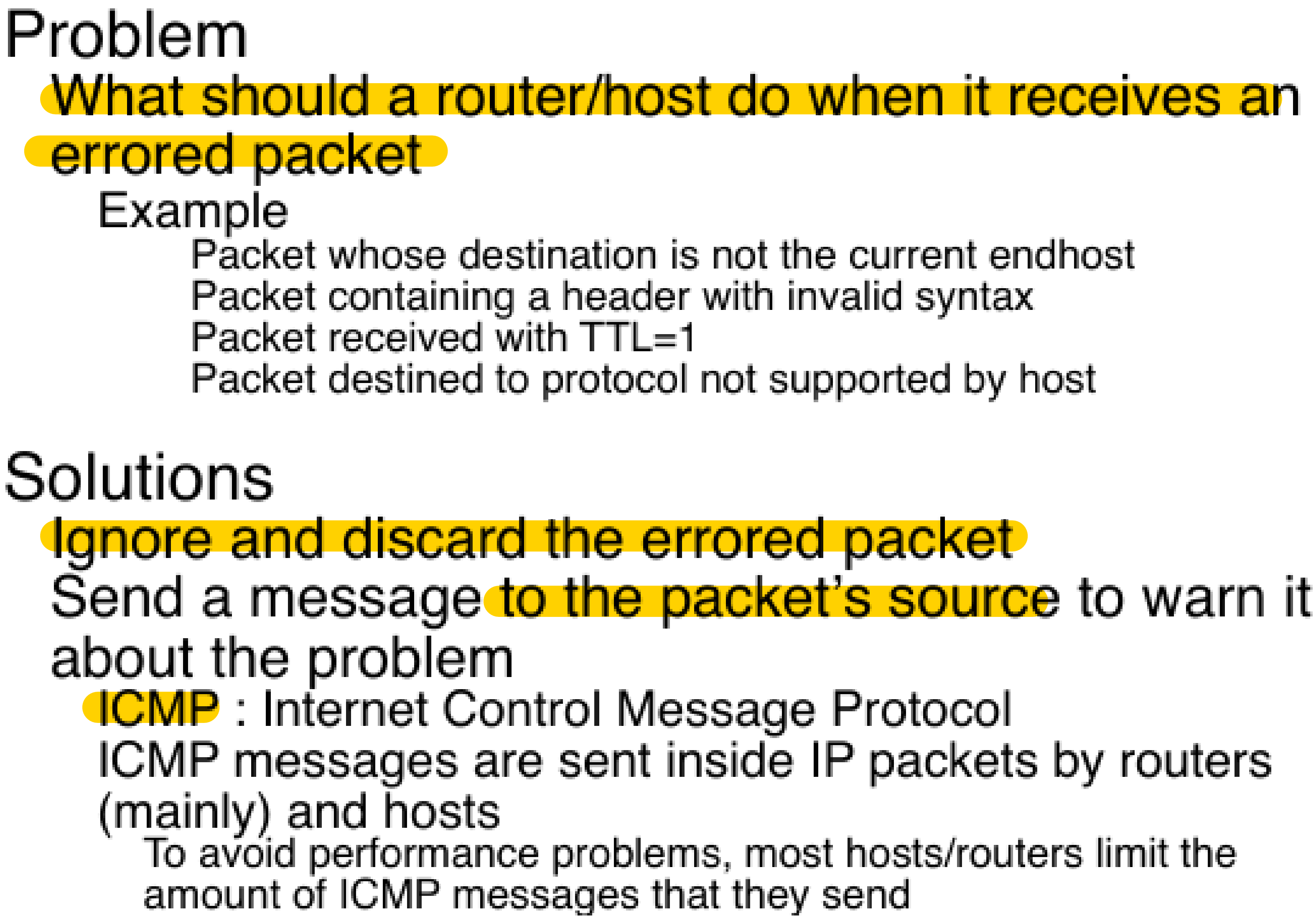
Binary flags
More
Don't Fragment : Packet cannot be fragmented by intermediate routers

IP address configuration



- How does a host know its IP address
  - Manual configuration
    - Used in many small networks
  - Server-based autoconfiguration RARP
    - DHCP
      - Dynamic Host Configuration Protocol
      - Principle
      - When it attaches to a subnet, endhost broadcasts a request to find DHCP server
      - DHCP server replies and endhost can contact it to obtain IP address
      - DHCP server allocates an IP address for some time period and can also provide additional information (subnet, default router, DNS resolver, ...)
        - DHCP servers can be configured to always provide the same IP address to a given endhost or not
        - Endhost reconfirms its allocation regularly

ICMP



Problem

- What should a router/host do when it receives an errored packet
  - Example
    - Packet whose destination is not the current endhost
    - Packet containing a header with invalid syntax
    - Packet received with TTL=1
    - Packet destined to protocol not supported by host

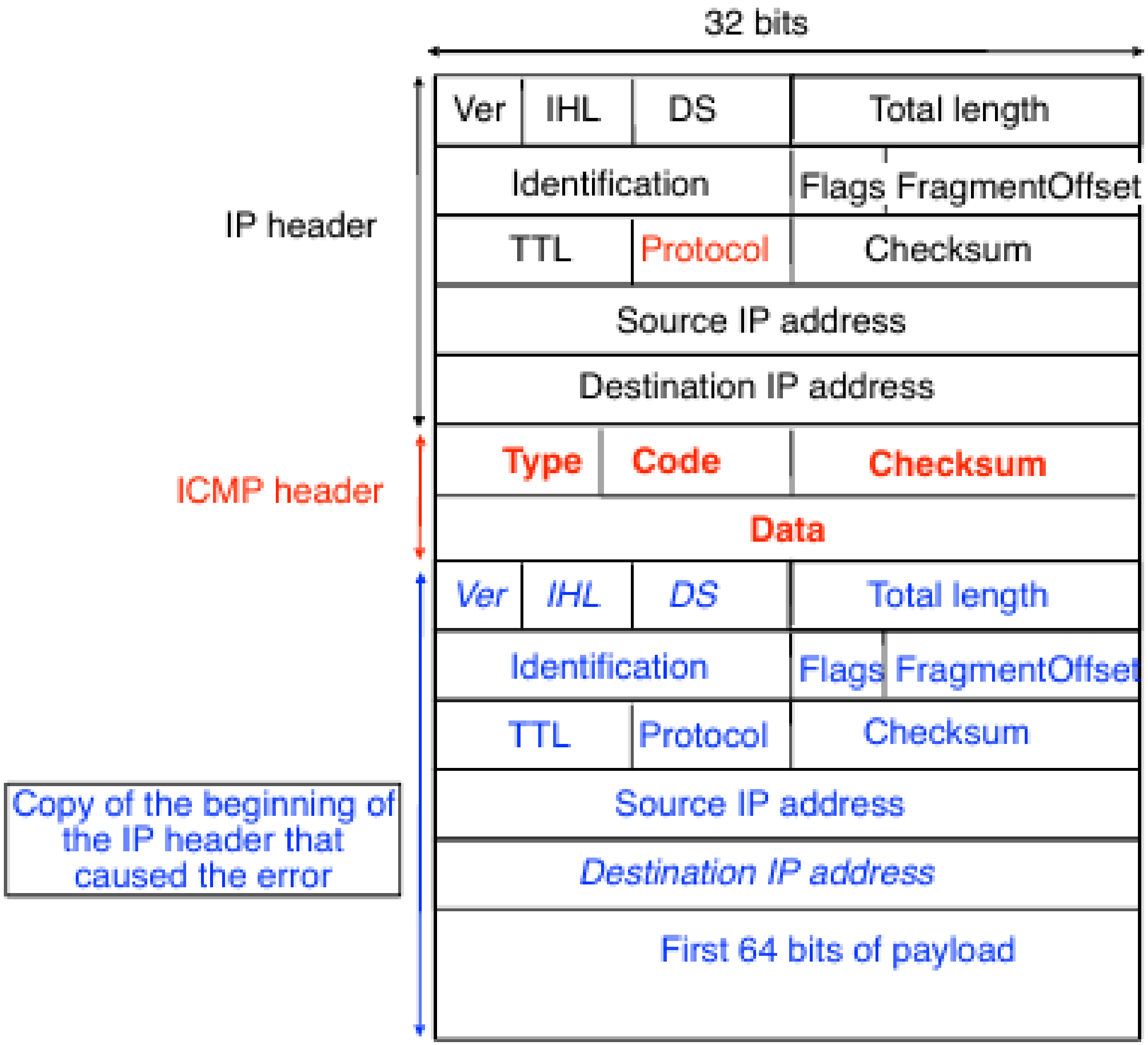
Solutions

- Ignore and discard the errored packet
- Send a message to the packet's source to warn it about the problem
  - ICMP : Internet Control Message Protocol
  - ICMP messages are sent inside IP packets by routers (mainly) and hosts
    - To avoid performance problems, most hosts/routers limit the amount of ICMP messages that they send

ICMP

- Routing error
  - Destination unreachable
    - Final destination of packet cannot be reached
      - Network unreachable for entire subnet
      - Host unreachable for an individual host
      - Protocol/Port unreachable for protocol/port on a reachable host
    - Redirect
      - The packet was sent to an incorrect first-hop router and should have been instead sent to another first-hop router
- Error in the IP header
  - Parameter Problem
    - Incorrect format of IP packet
  - TTL Exceeded
    - Router received packet with TTL=1
  - Fragmentation
    - the packet should have been fragmented, but its DF flag was true

ICMP message



32 bits

IP header

Ver | IHL | DS | Total length

Identification | Flags | FragmentOffset

TTL | Protocol | Checksum

Source IP address

Destination IP address

ICMP header

Type | Code | Checksum

Data

Copy of the beginning of the IP header that caused the error

Ver | IHL | DS | Total length

Identification | Flags | FragmentOffset

TTL | Protocol | Checksum

Source IP address

Destination IP address

First 64 bits of payload

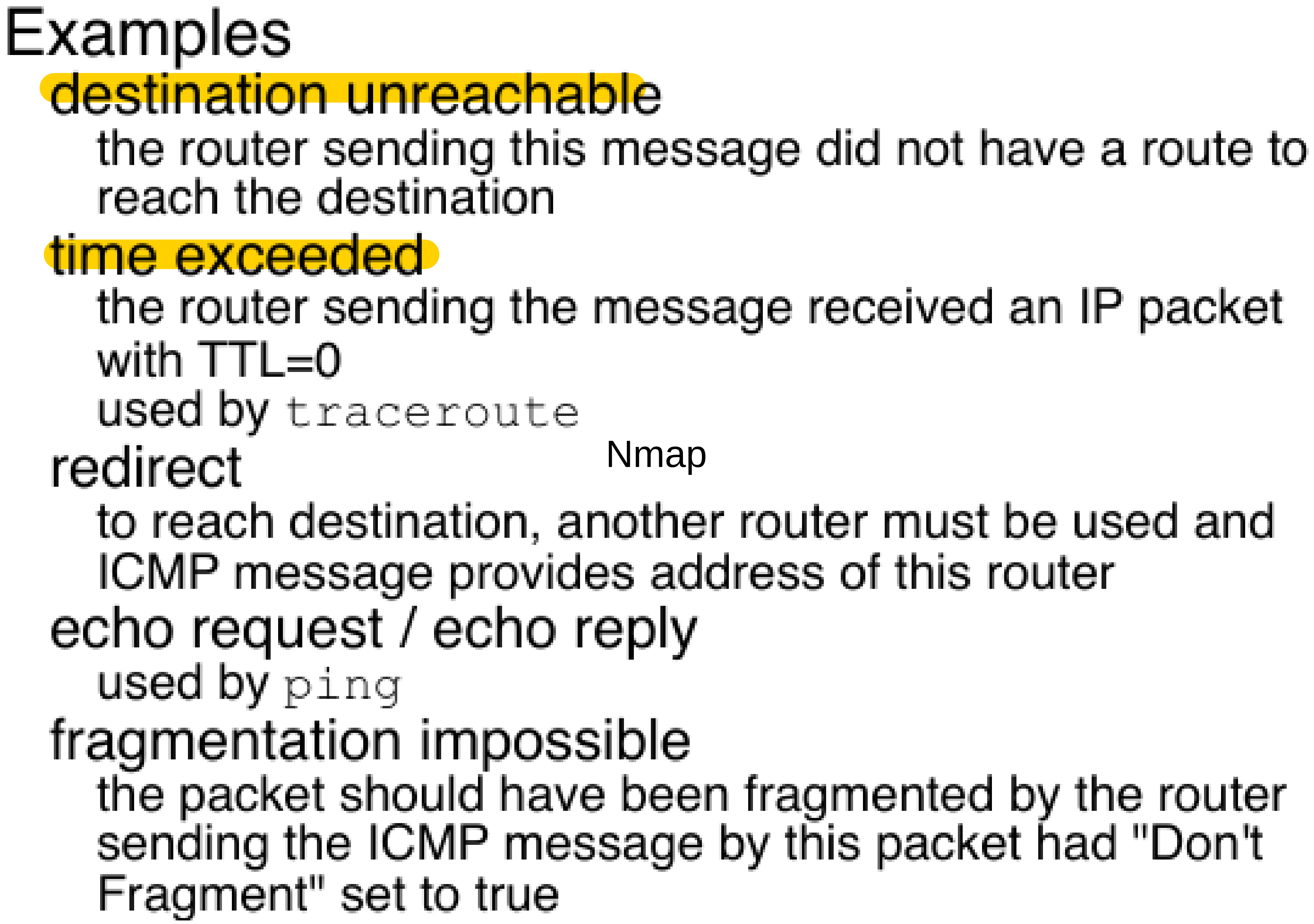
Protocol = 1 (for ICMP)

Type = the type of error

Code = the specific error code

Data = additional information

ICMP Usage



Examples

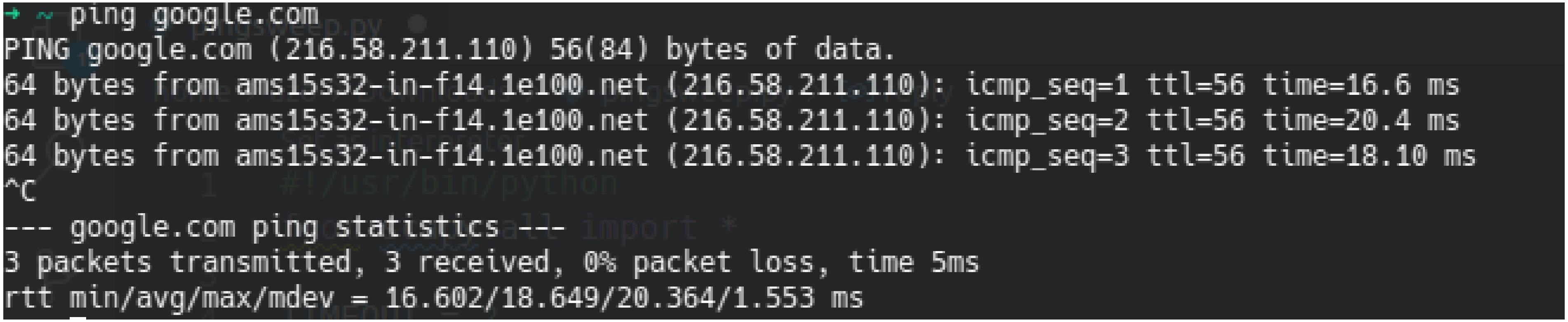
- destination unreachable
  - the router sending this message did not have a route to reach the destination
- time exceeded
  - the router sending the message received an IP packet with TTL=0
  - used by `traceroute`

Nmap

- redirect
  - to reach destination, another router must be used and ICMP message provides address of this router
- echo request / echo reply
  - used by `ping`
- fragmentation impossible
  - the packet should have been fragmented by the router sending the ICMP message by this packet had "Don't Fragment" set to true

ICMP Usage

- A machine can send an ICMP ECHO request, to see if a target is reachable.
  (Also called “pinging” a machine)

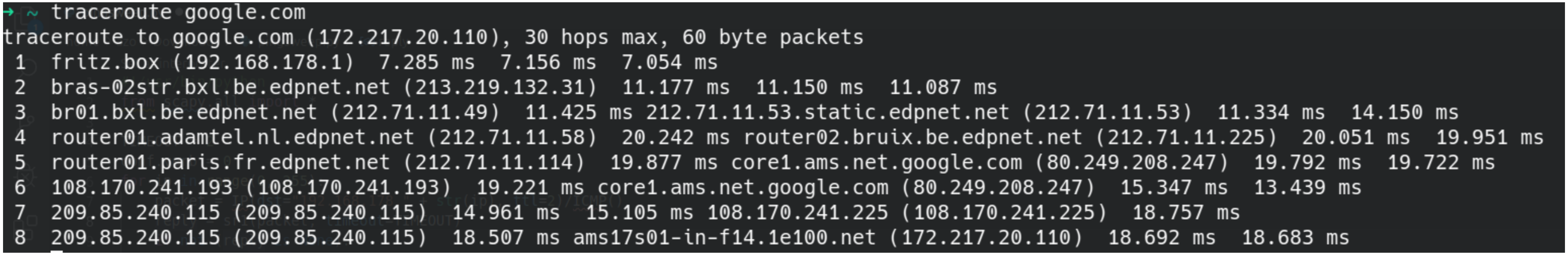


```
➜ ~ ping google.com
PING google.com (216.58.211.110) 56(84) bytes of data.
64 bytes from ams15s32-in-f14.1e100.net (216.58.211.110): icmp_seq=1 ttl=56 time=16.6 ms
64 bytes from ams15s32-in-f14.1e100.net (216.58.211.110): icmp_seq=2 ttl=56 time=20.4 ms
64 bytes from ams15s32-in-f14.1e100.net (216.58.211.110): icmp_seq=3 ttl=56 time=18.10 ms
^C
--- google.com ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 5ms
rtt min/avg/max/mdev = 16.602/18.649/20.364/1.553 ms
```

- Problem?
  - Attackers use this to find hosts/servers on the network.
  - Example

ICMP Usage

- We can find the route a packet takes, to see where the network fails.



```
→ ~ traceroute google.com
traceroute to google.com (172.217.20.110), 30 hops max, 60 byte packets
 1  fritz.box (192.168.178.1)  7.285 ms  7.156 ms  7.054 ms
 2  bras-02str.bxl.be.edpnet.net (213.219.132.31)  11.177 ms  11.150 ms  11.087 ms
 3  br01.bxl.be.edpnet.net (212.71.11.49)  11.425 ms 212.71.11.53.static.edpnet.net (212.71.11.53)  11.334 ms  14.150 ms
 4  router01.adamtel.nl.edpnet.net (212.71.11.58)  20.242 ms router02.bruix.be.edpnet.net (212.71.11.225)  20.051 ms  19.951 ms
 5  router01.paris.fr.edpnet.net (212.71.11.114)  19.877 ms core1.ams.net.google.com (80.249.208.247)  19.792 ms  19.722 ms
 6  108.170.241.193 (108.170.241.193)  19.221 ms core1.ams.net.google.com (80.249.208.247)  15.347 ms  13.439 ms
 7  209.85.240.115 (209.85.240.115)  14.961 ms  15.105 ms 108.170.241.225 (108.170.241.225)  18.757 ms
 8  209.85.240.115 (209.85.240.115)  18.507 ms ams17s01-in-f14.1e100.net (172.217.20.110)  18.692 ms  18.683 ms
```

- Traceroute manipulates the TTL in order to get an answer from the machines.
  - Starts with TTL = one, first hop answer with an unreachable ICMP paquet.
  - Increment TTL by one until reaching final destination. Every router/host will answer with an ICMP message until reached
- Example