CS 725/825 & IT 725

Lecture 17

Network Layer

October 28, 2024

Network Layer

Network Layer and IP

- Layer-specific function - routing
- Common functions:
  - Addressing: IP address
  - Error control: rerouting, ICMP
  - Flow control: ICMP
  - QoS: TOS field and Differentiated Services

ICMP - Internet Control Message Protocol, TOS - Type of Service

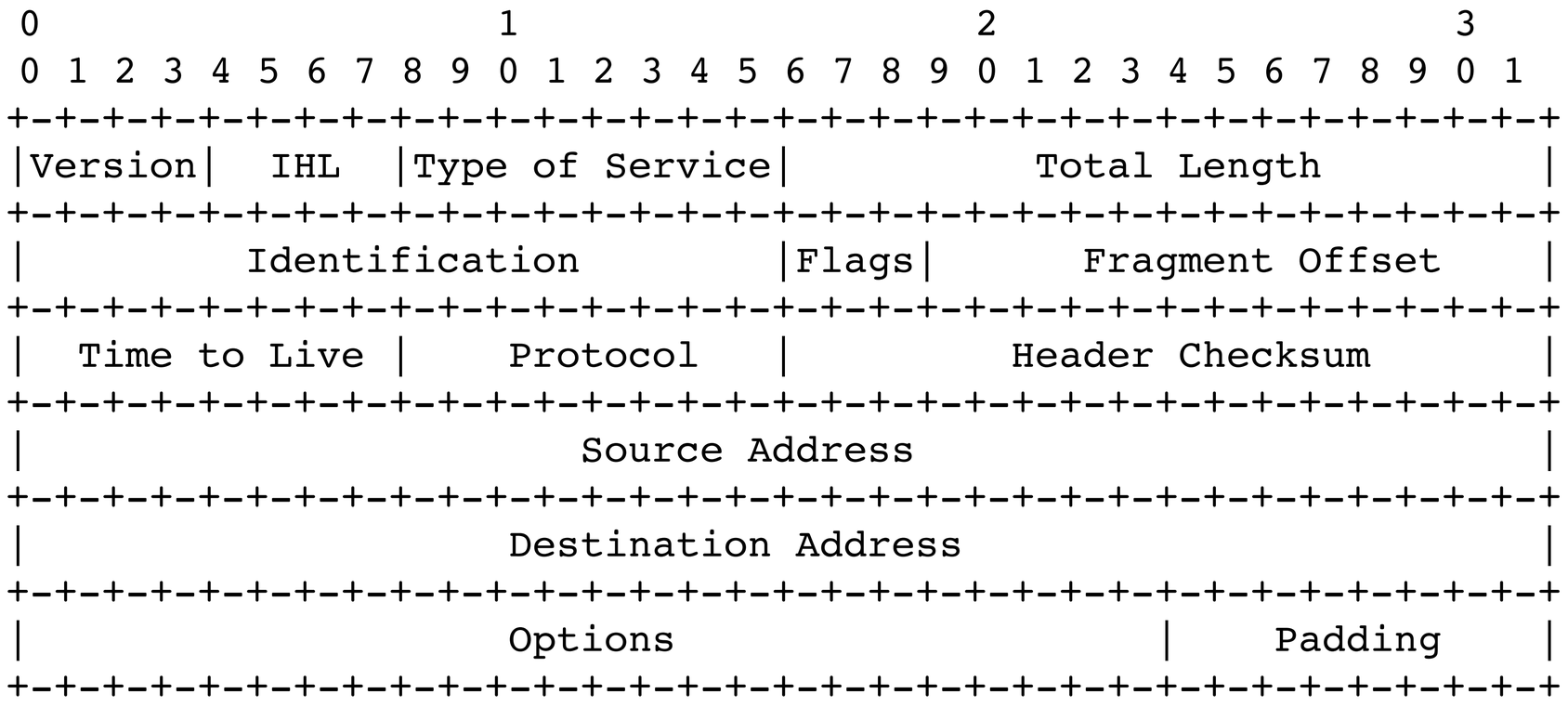
Internet Protocol (IP)

- Provides unreliable, connection-less service
- That is, packets may be:
  - lost
  - delivered out of order
  - duplicated
  - corrupted

IP Design Goals

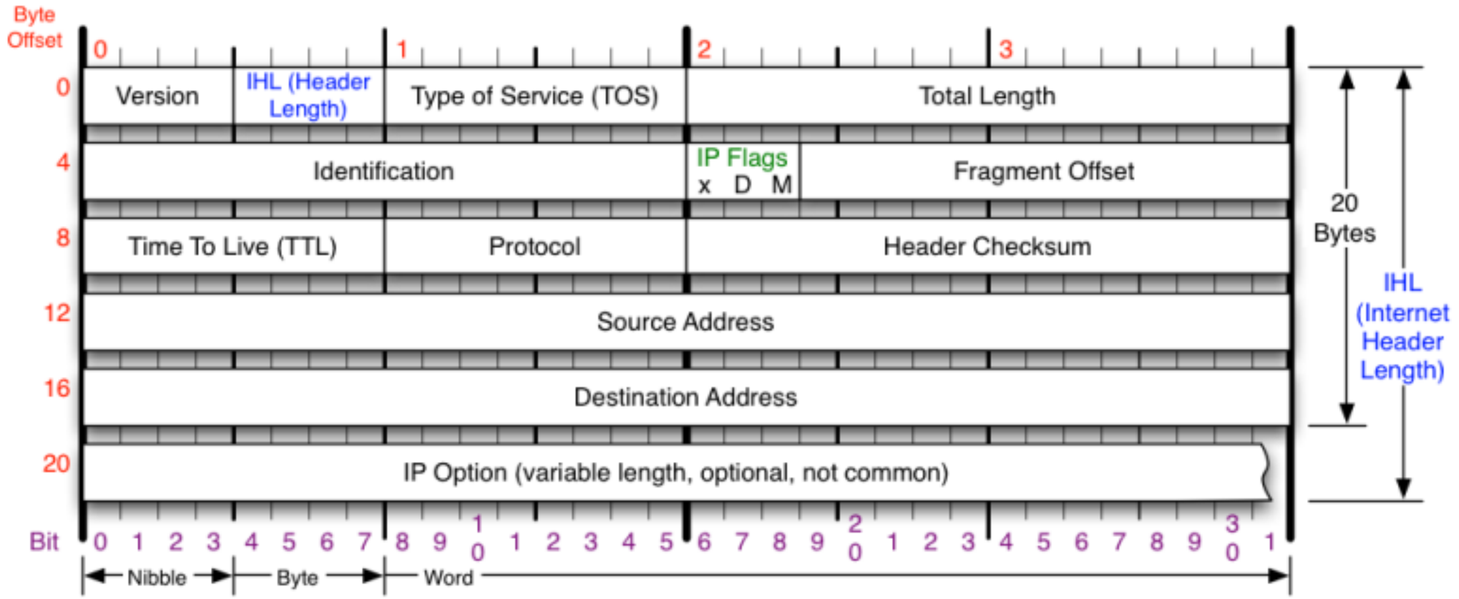
- Fields for source and destination IP addresses
- Means for error control:
  - detection of packet header corruption (L2 does the heavy lifting)
  - limiting the lifespan of a packet
- Fragmentation
  - carrying transport layer messages that are longer than what L2 can support

IPv4 Header



Source: RFC 791

IPv4 Header



Version

Version of IP Protocol. 4 and 6 are valid. This diagram represents version 4 structure only.

Header Length

Number of 32-bit words in TCP header, minimum value of 5. Multiply by 4 to get byte count.

Protocol

IP Protocol ID. Including (but not limited to):

1 ICMP	17 UDP	57 SKIP
2 IGMP	47 GRE	88 EIGRP
6 TCP	50 ESP	89 OSPF
9 IGRP	51 AH	115 L2TP

Total Length

Total length of IP datagram, or IP fragment if fragmented. Measured in Bytes.

Fragment Offset

Fragment offset from start of IP datagram. Measured in 8 byte (2 words, 64 bits) increments. If IP datagram is fragmented, fragment size (Total Length) must be a multiple of 8 bytes.

Header Checksum

Checksum of entire IP header

IP Flags

x D M

x 0x80 reserved (evil bit)
D 0x40 Do Not Fragment
M 0x20 More Fragments follow

RFC 791

Please refer to RFC 791 for the complete Internet Protocol (IP) Specification.

Image source: http://nmap.org/book/tcpip-ref.html

Fragmentation

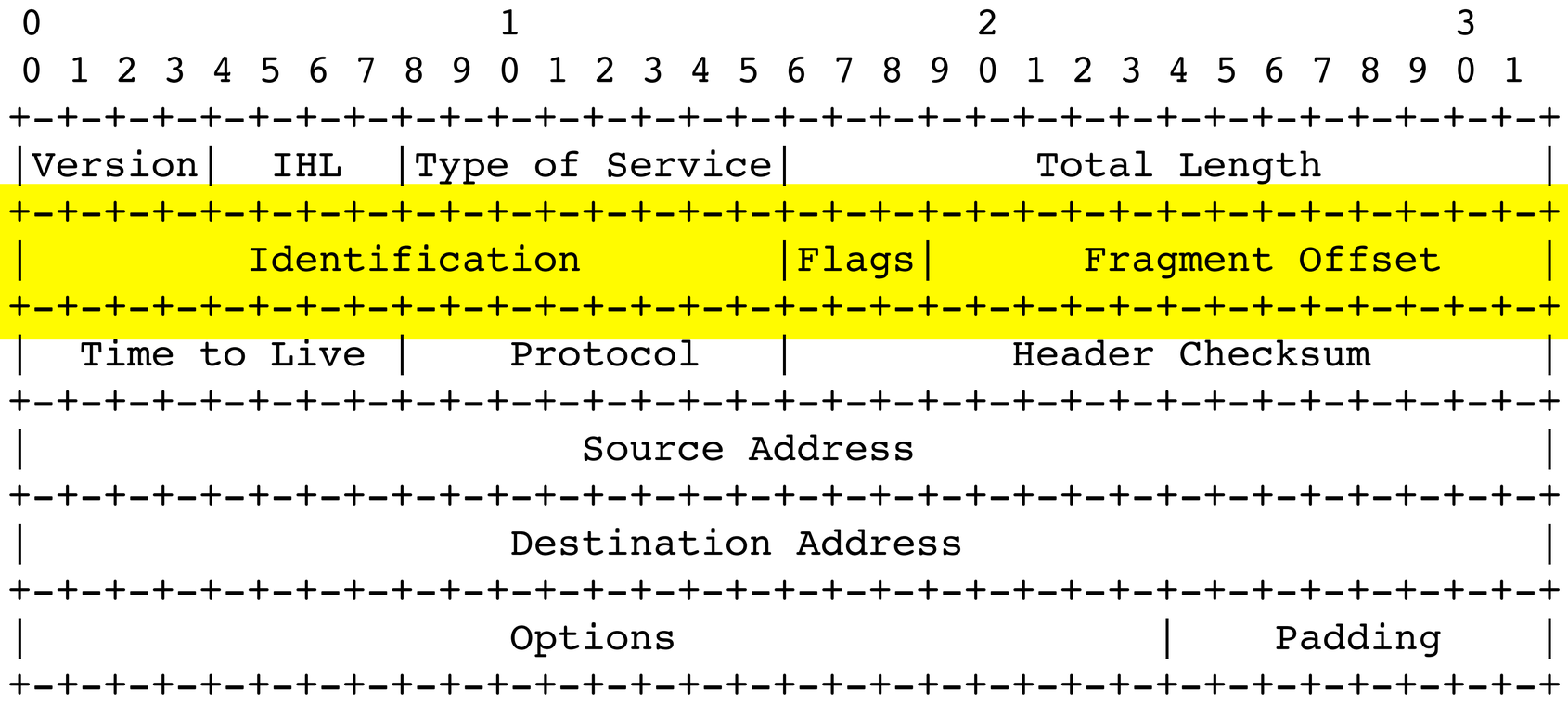
- Problem:
  - network layer needs to deliver a PDU* that is longer than what the link layer permits
- Solutions:
  - do nothing (drop the packet)
  - drop and inform (default behavior of IPv6)
  - fragment, i.e., break the PDU to smaller units (fragments) and reassemble them at the destination (default behavior of IPv4)

* PDU - protocol data unit, a fancy term for *packet*

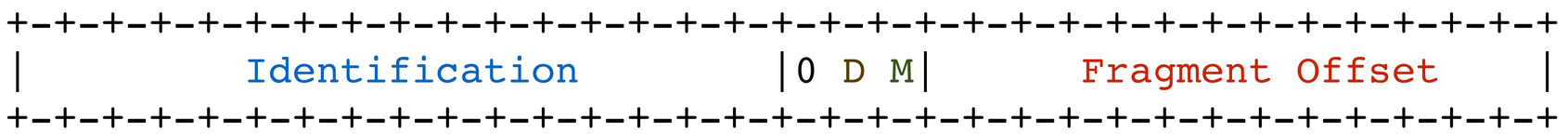
Fragmentation

- Allow packet reassembly (obviously)
  - need information on where a fragment fits the original packet
- Be able to distinguish between fragments of different packets
- Be able to further fragment a fragment…
  - and be able to reassemble the original packet in “one go” (not reassemble fragments that are then reassembled, etc.)

IPv4 Fragmentation



IPv4 Fragmentation



- Identification (16 bits)
  - identifies the original fragmented packet
- Fragment Offset (13 bits)
  - specifies the location of the fragment in the packet
  - only 13 most significant bits of the 16 bit value stored in the field, the remaining must be 0
- M (1 bit) - More Fragments
- D (1 bit) - Do Not Fragment

Fragmentation - example*

Packet length 4,500 bytes, MTU 2,500 bytes:

Fragment	Total bytes	Header bytes	Data bytes	"More fragments" flag	Fragment offset (bytes)
1	2500	20	2480	1	0
2	2040	20	2020	0	310

Fragments above reaching a link with MTU 1,500 bytes:

Fragment	Total bytes	Header bytes	Data bytes	"More fragments" flag	Fragment offset (bytes)
1	1500	20	1480	1	0
2	1020	20	1000	1	185
3	1500	20	1480	1	310
4	560	20	540	0	495

* Example shamelessly stolen from Wikipedia
(http://en.wikipedia.org/wiki/IPv4#Fragmentation_and_reassembly)

ICMP

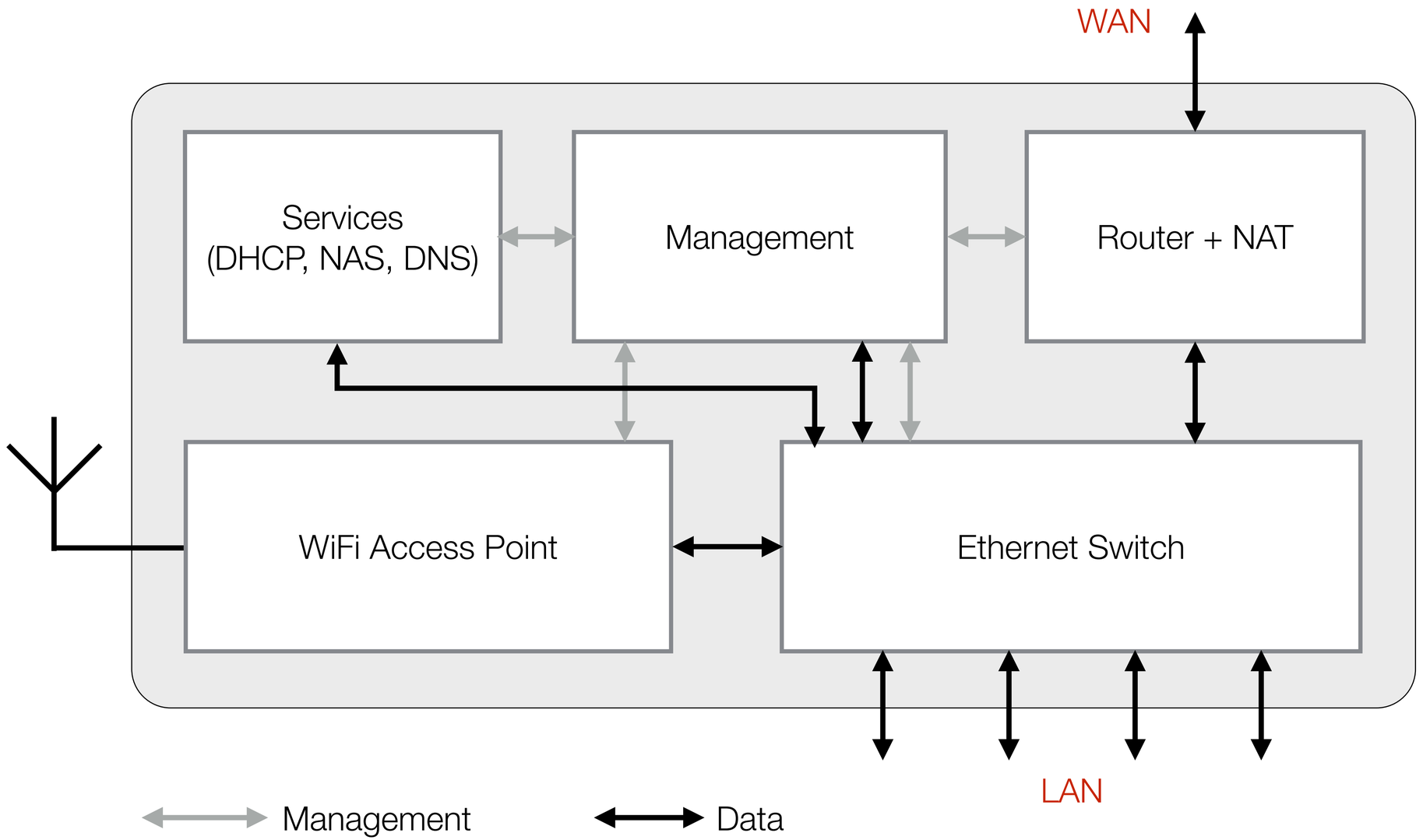
- **Internet Control Message Protocol**
- Runs on top of IP but still within the network layer
- **Examples**:
  - ping - Echo Request/Reply
  - traceroute - Time Exceeded
  - “No route to host” - Destination Unreachable
  - Source Quench (deprecated by RFCs 1812 and 6633)

Is a home router a router?

Is a home router a router?

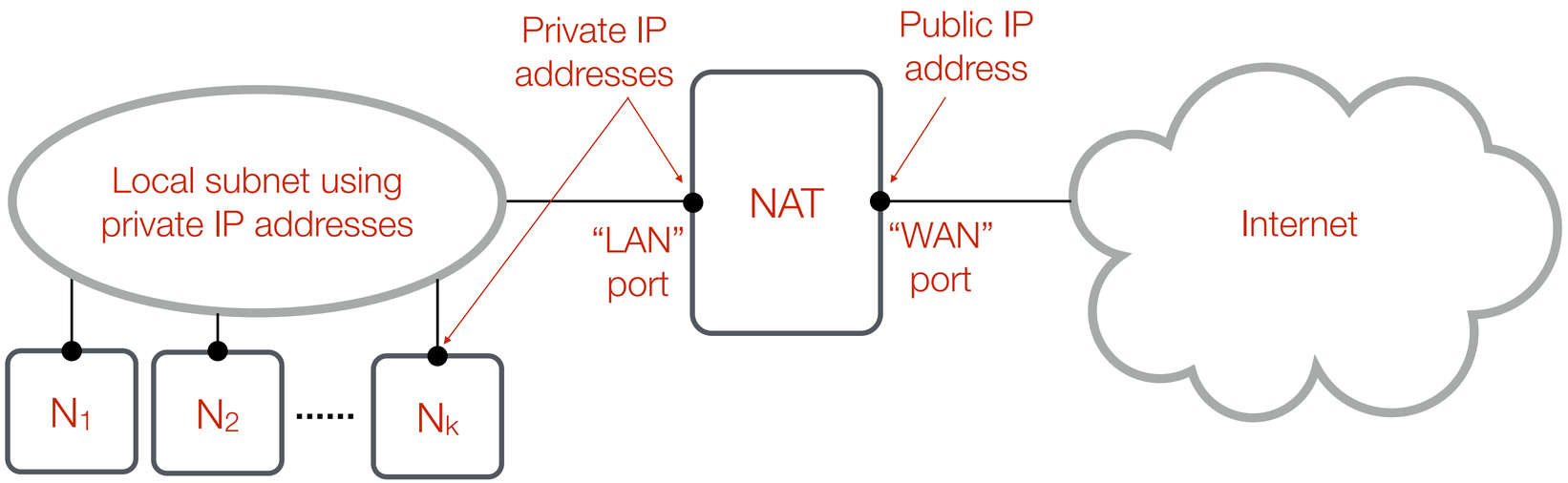
- Ethernet switch
- WiFi Access Point
- IP router
- Network Address Translation (NAT)
- DHCP server
- (NAS server)
- (print server)
- (DNS server)

Anatomy of a Home Router



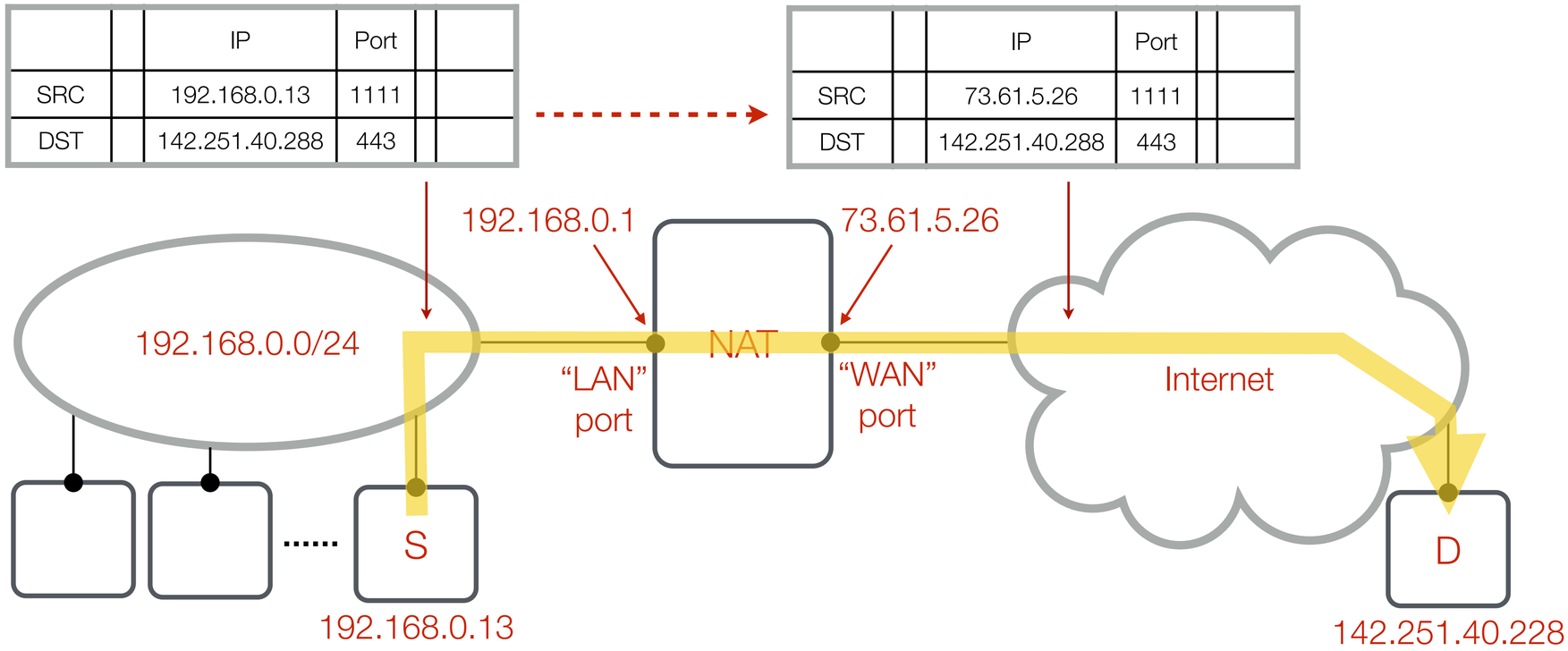
NAT

- Network Address Translation
- Motivation:
  - allow multiple nodes to share a single IP address
  - prevent external traffic from entering the local network



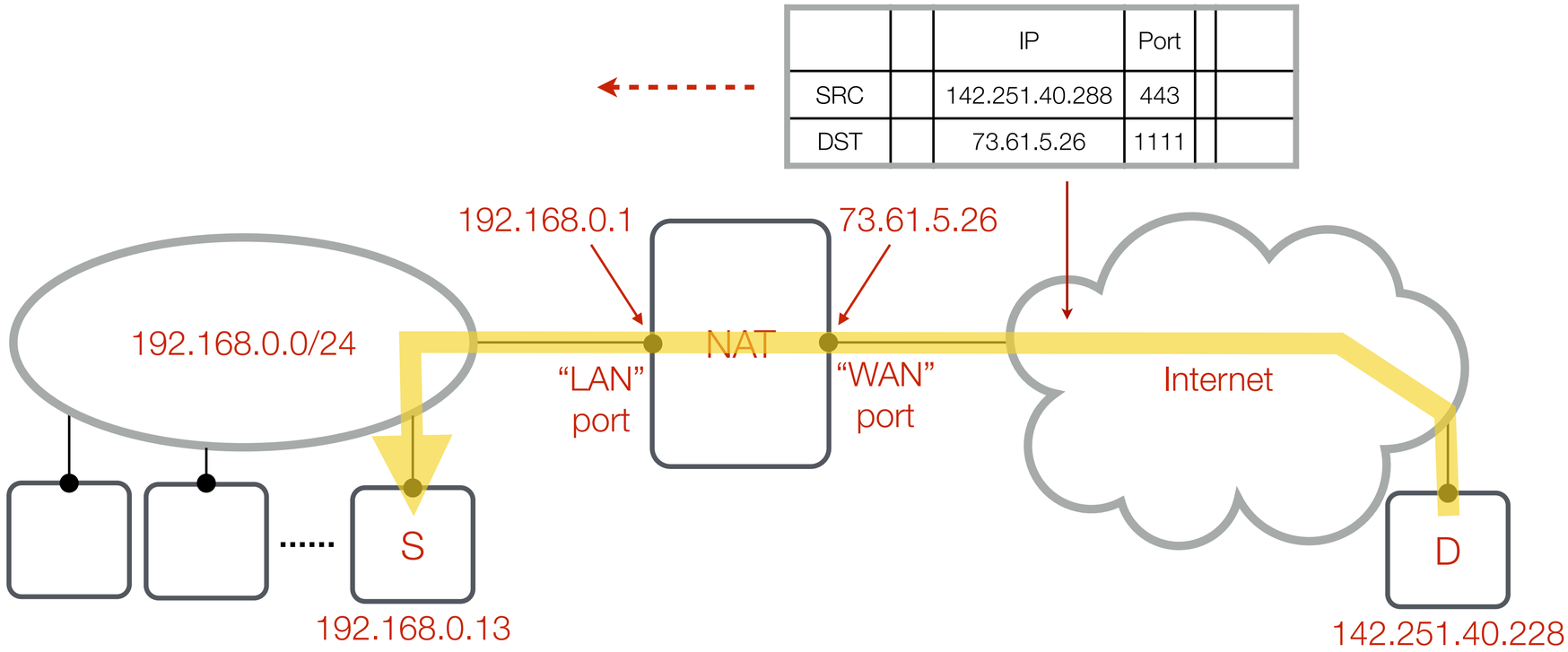
NAT

- Communication is initiated from a local node (S)
- Local (private) source IP address changed to the public IP address of the NAT box



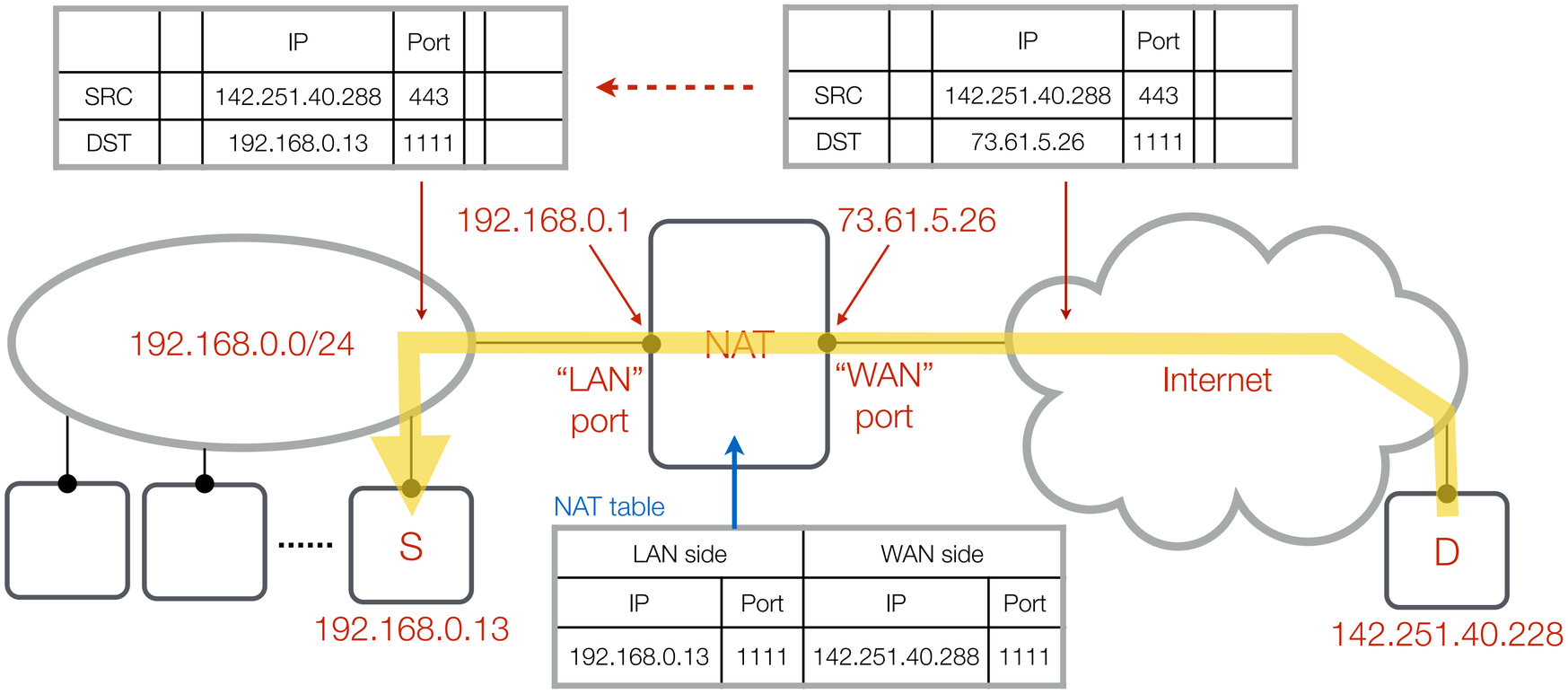
NAT

- Response is delivered to the NAT box
- The NAT box needs to know which local IP (and port) to use to deliver the packet



NAT

- NAT box observes outgoing connection requests and keeps track of a src/dst IP/port translation in NAT table
- Source IP/port is used to look up the table



NAT

- There can be a port number conflict on the WAN side, so the port numbers can and do get translated too