



# Programming and Communications III: Internet

Jordi Ricard Onrubia Palacios

Departament d'Informàtica i Enginyeria Industrial Universitat de Lleida

# Programming and Communications

## III: Internet

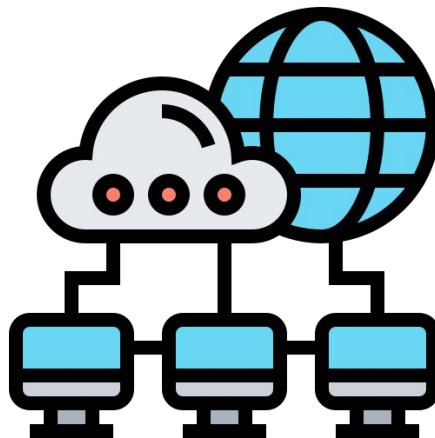


- Overview
- ❑ Architecture
  - ❑ Communication Concept ISO/OSI
  - ❑ Implementation TCP/IP
- ❑ Transmission
  - ❑ Packed data
  - ❑ Ethernet
- ❑ IP
- ❑ TCP
- ❑ HTTP



# Overview

The Internet is the global system of interconnected computer networks (a set of computers sharing resources located on or provided by a communication endpoint) that use the Internet protocol suite (TCP/IP) to link devices worldwide.



# Programming and Communications

## III: Internet



- ❑ Overview
- Architecture
  - ❑ Communication
    - Concept ISO/OSI
  - ❑ Implementation
    - TCP/IP
- ❑ Transmission
  - ❑ Packed data
  - ❑ Ethernet
- ❑ IP
- ❑ TCP
- ❑ HTTP



# Architecture: Communication Concept ISO/OSI

- Conceptual model that characterizes and standardizes the communication functions of a telecommunication or computing system without regard to its underlying internal structure and technology.
- Its goal is the interoperability of diverse communication systems with standard protocols.
- The model partitions a communication system into abstraction layers.
- A layer serves the layer above it and is served by the layer below it.
- Protocols enable an entity in one host to interact with a corresponding entity at the same layer in another host.



<b>Application</b>	High-level APIs, including resource sharing, remote file access
<b>Presentation</b>	Translation of data between a networking service and an application; including character encoding, data compression and encryption/decryption.
<b>Session</b>	Managing communication sessions, i.e. continuous exchange of information in the form of multiple back-and-forth transmissions between two nodes.
<b>Transport</b>	Reliable transmission of data segments between points on a network, including segmentation, acknowledgement and multiplexing.
<b>Network</b>	Structuring and managing a multi-node network, including addressing, routing and traffic control.
<b>Link</b>	Reliable transmission of data frames between two nodes connected by a physical layer.
<b>Physical</b>	Transmission and reception of raw bit streams over a physical medium.





# Implementation TCP/IP

The TCP/IP model is a concise version of the OSI model. It contains four layers, unlike seven layers in the OSI model.

Application	Application (FTP, SMTP, HTTP, etc.)
Presentation	
Session	
Transport	TCP (host to host)
Network	IP
Link	Network access (usually Ethernet)
Physical	





## Implementation TCP/IP

1. Network Access: This layer corresponds to the combination of Data Link Layer and Physical Layer of the OSI model. It looks out for hardware addressing and the protocols present in this layer allows for the physical transmission of data.
2. Internet: This layer parallels the functions of OSI's Network layer. It defines the protocols which are responsible for logical transmission of data over the entire network.
3. Transport: It is responsible for end-to-end communication and error-free delivery of data. It shields the upper-layer applications from the complexities of data. The two main protocols present in this layer are :
  - a. Transmission Control Protocol (TCP) – Provides reliable and error-free communication between end systems. It performs sequencing and segmentation of data. It also has acknowledgment feature and controls the flow of the data through flow control mechanism. It is a very effective protocol but has a lot of overhead due to such features.
  - b. User Datagram Protocol (UDP) – Does not provide any such features. It is the go-to protocol if your application does not require reliable transport as it is very cost-effective. UDP is connectionless.





## Implementation TCP/IP

4. Application Layer: This layer performs the functions of top three layers of the OSI model: Application, Presentation and Session Layer. It is responsible for node-to-node communication and controls user-interface specifications. Protocols other than those present in the linked article are :
  - a. HTTP and HTTPS: HTTP stands for Hypertext transfer protocol. It is used by the World Wide Web to manage communications between web browsers and servers. HTTPS stands for HTTP-Secure. It is a combination of HTTP with SSL(Secure Socket Layer). It is efficient in cases where the browser need to fill out forms, sign in, authenticate and carry out bank transactions.
  - b. SSH: SSH stands for Secure Shell. It is a terminal emulations software similar to Telnet. The reason SSH is more preferred is because of its ability to maintain the encrypted connection. It sets up a secure session over a TCP/IP connection.

# Programming and Communications

## III: Internet

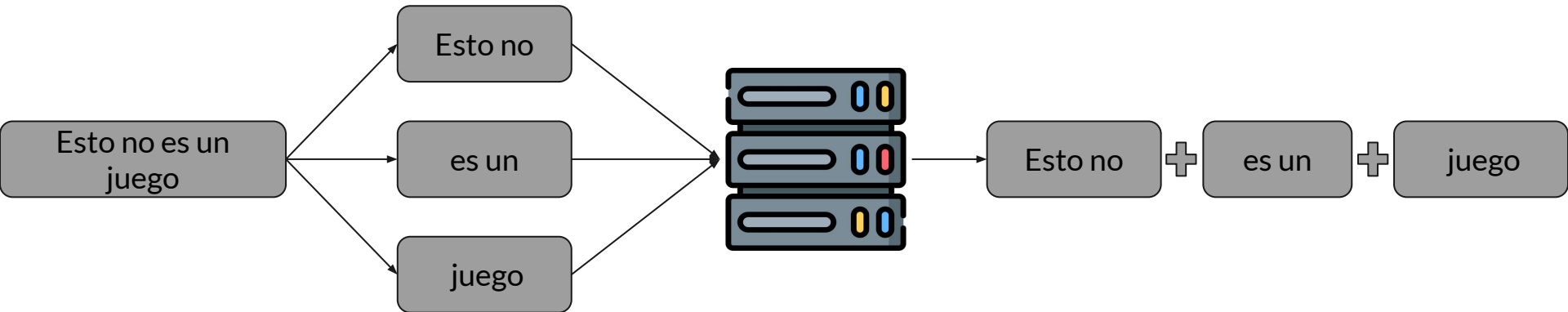


- ❑ Overview
- ❑ Architecture
  - ❑ Communication Concept ISO/OSI
  - ❑ Implementation TCP/IP
- Transmission
  - ❑ Packet data
  - ❑ Ethernet
- ❑ IP
- ❑ TCP
- ❑ HTTP



## Transmission: Packed Data (simplified)

2. Send them



1. Divide data into small blocks (packets)

3. Reassemble them in the proper sequence



# Transmission: Packet Data

Issues:

- Where do we send them?
- In which order do we need to reassemble them?
- Which is the maximum and the minimum size for a packet?
- Who is sending them?
- The message got corrupted?
- ...





# Transmission: Ethernet

Ethernet is a family of computer networking technologies commonly used in local area networks (LAN), metropolitan area networks (MAN) and wide area networks (WAN)

Application	Application (FTP, SMTP, HTTP, etc.)
Presentation	
Session	
Transport	TCP (host to host)
Network	IP
Link	Network access (usually Ethernet)
Physical	



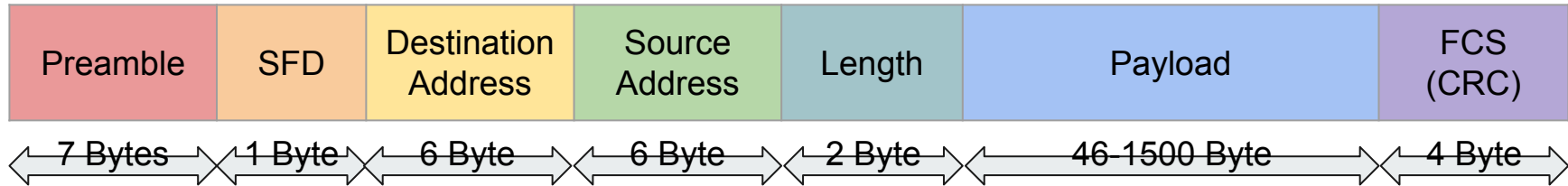
# Transmission: Ethernet

- Link layer (Ethernet):
  - In charge of the communication physical connected elements
  - Access control
  - The concept of address is born
  - Shared medium → necessity of identifying elements



## Transmission: Ethernet

Ethernet divides files to be transmitted into packets, which provide addressing, sequencing and other information, which allows the receiving network device to properly reassemble the file.





## Transmission: Ethernet

- Preamble: The preamble of an Ethernet packet consists of a 56bit (7bytes) pattern of alternating 1 and 0 bits, allowing devices on the network to easily synchronize their receiver clocks
- Source Address: MAC address of the device which sends the packet
- Destination Address: MAC address of the device that should receive the packet







## Transmission: Ethernet -> MAC Address Bonus

- MAC address:
  - The Media Access Control (MAC) address provides a product serial number, or physical address, which is used to identify a particular device on a network
  - It is a 48bit address space and contains potentially  $2^{48}$  or 281,474,976,710,656 possible MAC addresses
  - An Ethernet packet can be sent to:
    - Unicast : to one single device
    - Broadcast (ff:ff:ff:ff:ff:ff): to all devices
    - Multicast: a group of devices (that have to do a process to be part of the group)





## Transmission: Ethernet -> MAC Address Bonus

- Type: Type of the payload
- Data: Minimum of 42 octets and maximum of 1500 octets
- Length of the Payload:
  - Frames must be at least 64 bytes long, not including the preamble, so, if the data field is shorter than 46 bytes, it must be compensated by the Pad field.
  - The reason for specifying a minimum length lies with the collision-detect mechanism.



## Transmission: Ethernet -> MAC Address Bonus

- FCS: The frame check sequence (FCS) is a four-octet cyclic redundancy check (CRC) that allows detection of corrupted data within the entire frame as received on the receiver sides

<https://quickbirdstudios.com/blog/validate-data-with-crc/>





# Transmission: Ethernet -> MAC Address Bonus

## CSMA/CD

- CSMA/CD is a Media Access Control (MAC) protocol
- It defines how network devices respond when two devices attempt to use a data channel simultaneously and encounter a data collision
- The CSMA/CD rules define how long the device should wait if a collision occurs





# Transmission: Ethernet -> MAC Address Bonus

## CSMA/CD

- CSMA/CD is a Media Access Control (MAC) protocol
- It defines how network devices respond when two devices attempt to use a data channel simultaneously and encounter a data collision
- The CSMA/CD rules define how long the device should wait if a collision occurs





## Transmission: Ethernet -> MAC Address Bonus

### CSMA/CD: Carrier Sense

- Carrier sense is the ability of a network interface card (NIC) to check the network for any communication
- Obviously if there is data being transmitted over the network, the NIC should not attempt to transmit data
- If there is no traffic on the network, the NIC will then attempt to transmit the data



# Transmission: Ethernet -> MAC Address Bonus

## CSMA/CD: Medium Access

- The MA (multiple access) part of CSMA/CD tells us that there will be multiple devices using the same network

## CSMA/CD: Collision Detect

- The CD (collision detect) part of CSMA/CD states that we need a method for detecting a collision
- To detect collisions, the device must transmit two times the transmission period time



# Transmission: Ethernet -> MAC Address Bonus

## The Collision Detection and Solution Process

1. A collision is detected.
2. Devices involved in the collision keep transmitting for a short period of time, to make sure all devices on the network see the collision (also referred to as the jamming signal)
3. Each device sees the jamming signal, and invokes the back-off algorithm. Each device will have a random timer that determines when it can transmit again.
4. When the back-off timer expires, devices are free to transmit data again. Devices involved in the collision earlier do not have priority to transmit data.



# Programming and Communications

## III: Internet



- ❑ Overview
- ❑ Architecture
  - ❑ Communication Concept ISO/OSI
  - ❑ Implementation TCP/IP
- ❑ Transmission
  - ❑ Packet data
  - ❑ Ethernet
- IP
- ❑ TCP
- ❑ HTTP



# Internet Protocol (IP)

The Internet Protocol (IP) is the network layer communications protocol in charge of routing the datagrams.

Application	Application (FTP, SMTP, HTTP, etc.)
Presentation	
Session	
Transport	TCP (host to host)
Network	IP
Link	Network access (usually Ethernet)
Physical	





# Internet Protocol (IP)

What IP is not:

- A connection-oriented protocol. The receiver node can have duplicate packets, packets outside the sequence, etc...
- A reliable protocol as it does not provide any control mechanism for the proper reception of the packets. The reliability has to be provided by upper (in the stack) protocols



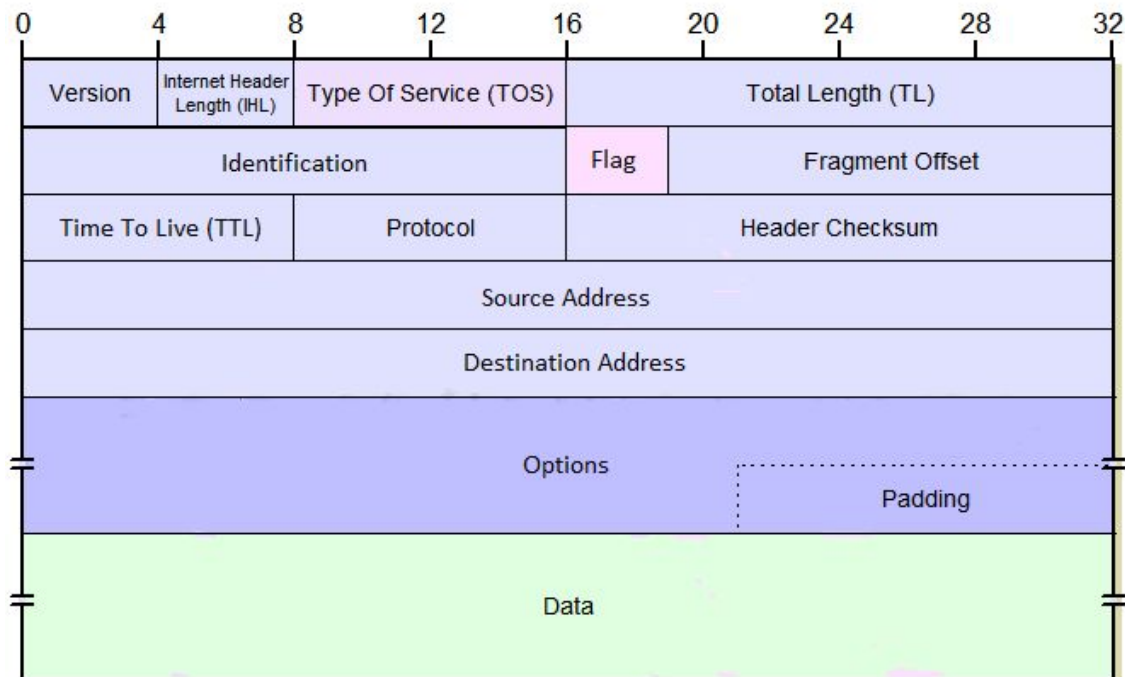
# Internet Protocol (IP)

What IP does:

- To define the format of the packet in the transmission
- To fragment and reassemble the data so it can be transmitted
- To define an addressing schema to identify each node in the network
- To route the packets in the network



# Internet Protocol (IP)





# Internet Protocol (IP)

- Version: Identifies the version of IP used to generate the datagram.
- Internet Header Length (IHL): Specifies the length of the IP header, in 32-bit words.
- Type Of Service (TOS): A field designed to carry information to provide quality of service features.
- Total Length (TL): Specifies the total length of the IP datagram
- Identification: This field contains a 16-bit value that is common to each of the fragments belonging to a particular message.
- Flags: Reserved, not in use; DF, do not fragment; MF, more fragments.
- Fragment Offset: When fragmentation of a message occurs, this field specifies the offset, or position, in the overall message where the data in this fragment goes.



# Internet Protocol (IP)

- Time to live (TTL): Specifies how long the datagram is allowed to “live” on the network, in terms of router hops. Each router decrements the value of the TTL field (reduces it by one) prior to transmitting it. If the TTL field drops to zero, the datagram is assumed to have taken too long a route and is discarded.
- Protocol: Identifies the upper layer protocol to which it is to be delivered.
- Header Checksum: A checksum computed over the header to provide basic protection against corruption in transmission.
- Source Address: The 32-bit IP address of the originator of the datagram.
- Destination Address: The 32-bit IP address of the intended recipient of the datagram.



# Internet Protocol (IP)

- Options: One or more of several types of options may be included after the standard headers in certain IP datagrams.
- Padding: If one or more options are included, and the number of bits used for them is not a multiple of 32, enough zero bits are added to “pad out” the header to a multiple of 32 bits (4 bytes).
- Data: The data to be transmitted in the datagram, either an entire higher-layer message or a fragment of one.





# Internet Protocol (IP)

What IP does:

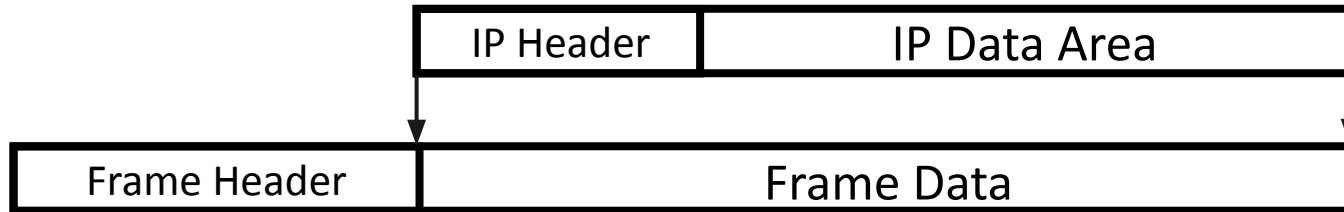
- To define the format of the packet in the transmission
- To fragment and reassemble the data so it can be transmitted
- To define an addressing schema to identify each node in the network
- To route the packets in the network





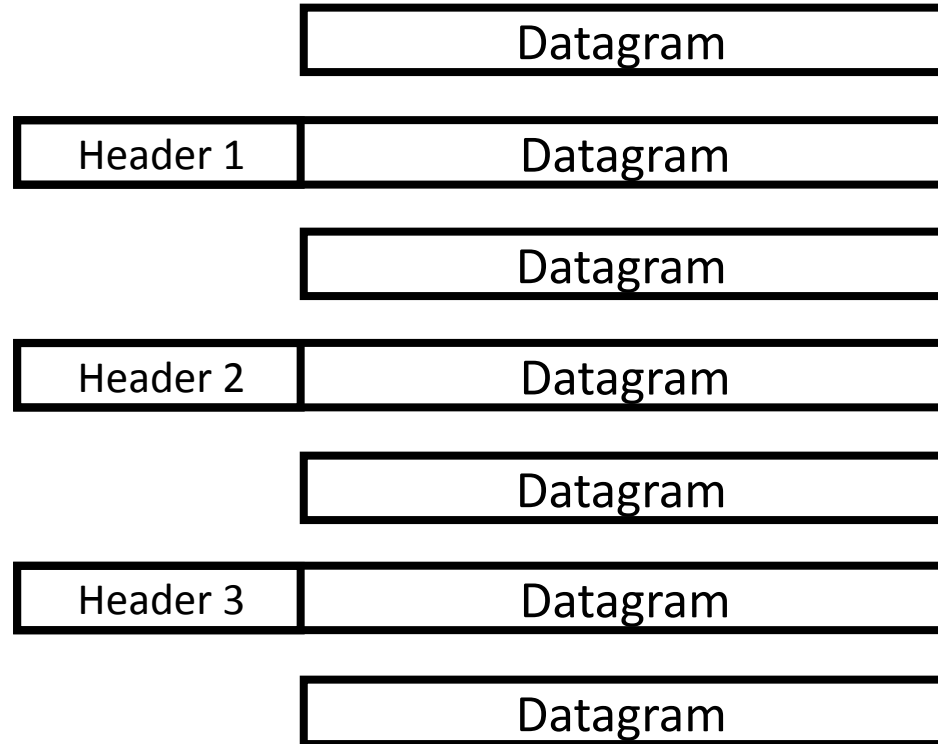
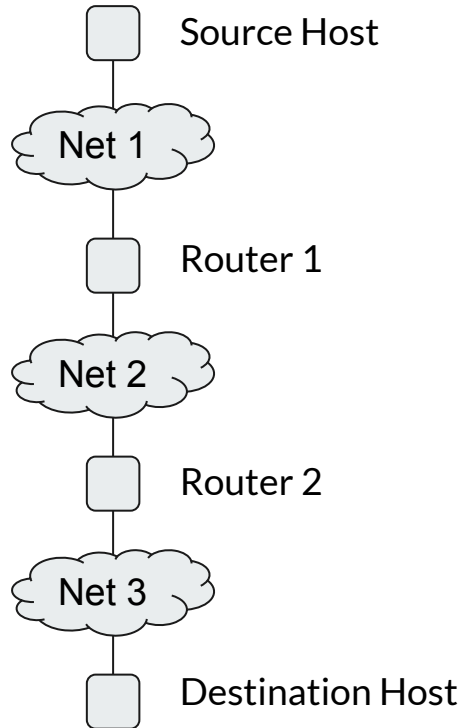
## Internet Protocol (IP)

When an IP datagram is sent across a physical network it is placed in the data area of a frame and the frame type is set to IP





# Internet Protocol (IP)





# Internet Protocol (IP)

- **MTU & Datagram Size**
  - Maximum transmission unit - max of data that a frame can carry on a given network
  - A packet may have to cope with different MTU sizes as it passes over an internet
- **Fragmentation**
  - A datagram that is larger than MTU is fragmented into smaller datagrams





# Internet Protocol (IP)

- Reassembly
  - Is done at the final host
    - routers require less state information
    - fragments can take different routes
  - Header fields indicate when the data is a fragment and also where it belongs
  - Whole datagram is lost if any fragment is lost



# Internet Protocol (IP)

What IP does:

- To define the format of the packet in the transmission
- To fragment and reassemble the data so it can be transmitted
- To define an addressing schema to identify each node in the network
- To route the packets in the network





# Internet Protocol (IP)

## IP addressing

- In an IP network each host has a unique identification given by its IP
- An IP address is a block of 32bits and generally it is written in decimal notation divided by points
- This notation divides the address in blocks of 8bits, assigns decimal values to each block and separates them with points

11000001 10010000 00001000 00001110 → 193.144.8.14





# Internet Protocol (IP)

Each IP address has two parts:

- Network: initial part of the address. It is the same for a group of hosts
- Host: final part of the address that identifies each host
- Broadcast: Network with host part all to 1

Two options:

- Classful: In each IP address the first 5 bits show to which class the IP belongs :
  - Depending on the number of bits devoted to each of the parts of an IP address it is defined 5 kind of classes
- Classless: Network mask determines network part (classless)
  - Classless Inter-Domain Routing (CIDR) notation
  - Bits devoted to network







# Internet Protocol (IP)

## Five Different Classes of IPv4 Addresses

Class	First Octet decimal (range)	First Octet binary (range)	IP range	Subnet Mask	Hosts per Network ID	# of networks
Class A	0 — 127	0XXXXXXXX	0.0.0.0-127.255.255.255	255.0.0.0	$2^{24} - 2$	$2^7$
Class B	128 — 191	10XXXXXX	128.0.0.0-191.255.255.255	255.255.0.0	$2^{16} - 2$	$2^{14}$
Class C	192 — 223	110XXXXX	192.0.0.0-223.255.255.255	255.255.255.0	$2^8 - 2$	$2^{21}$
Class D (Multicast)	224 — 239	1110XXXX	224.0.0.0-239.255.255.255			
Class E (Experimental)	240 — 255	1111XXXX	240.0.0.0-255.255.255.255			





## Internet Protocol (IP)

Network all 0	Indicates a host on own network
Host all 0	Identifies a network
Host all 1	Broadcast address in an specific network
0.0.0.0	Used for hosts in configuration process
127.X.X.X	Loopback address
10.X.X.X, 172.16.X.X, 172.31.X.X, 192.168.X.X	Address range used for private networks, they cannot get out from an institution ambit
255.255.255.255	Broadcast address in an own network



# Internet Protocol (IP)

What IP does:

- To define the format of the packet in the transmission
- To fragment and reassemble the data so it can be transmitted
- To define an addressing schema to identify each node in the network
- To route the packets in the network



# Network Layer (IP)



- IP Routing:

- Final devices: in charge of generating or receiving data
- Intermediate devices: in charge of routing data from source to destination
- Routing Table:
  - Network, subnetwork or destination host
  - Network mask
  - Next routing device (explicit or default)
- **Each entry in a routing table is a route**
- **Special entry named default route**

# Network Layer (IP)



Network, Subnetwork or Destination Host	Mask	Next Hop
192.168.1.0	255.255.255.0	0.0.0.0
192.168.2.0	255.255.255.0	192.168.1.1
default	0.0.0.0	192.168.10.100



# Network Layer (IP)



- IP Routing algorithm:

Given a destination IP address D, and network prefix N:

```
if ( N matches a directly connected network address )
```

```
    Deliver datagram to D over that network link;
```

```
else if ( The routing table contains a route for N )
```

```
    Send datagram to the next-hop address listed in the routing table;
```

```
else if ( There exists a default route )
```

```
    Send datagram to the default route;
```

```
else
```

```
    Send a forwarding error message to the originator;
```

# Network Layer (IP)



- Own IP: 192.168.1.10
- Own Mask: 255.255.255.0
- Dest IP: 192.168.1.11

Network, Subnetwork or Destination Host	Mask	Next Hop
192.168.1.0	255.255.255.0	0.0.0.0
192.168.2.0	255.255.255.0	192.168.1.1
0.0.0.0	0.0.0.0	192.168.10.100

# Network Layer (IP)



- Own IP: 192.168.1.10
- Own Mask: 255.255.255.0
- Dest IP: 192.168.2.11

Network, Subnetwork or Destination Host	Mask	Next Hop
192.168.1.0	255.255.255.0	0.0.0.0
192.168.2.0	255.255.255.0	192.168.1.1
0.0.0.0	0.0.0.0	192.168.10.100



# Network Layer (IP)



- Own IP: 192.168.1.10
- Own Mask: 255.255.255.0
- Dest IP: 200.34.78.23

Network, Subnetwork or Destination Host	Mask	Next Hop
192.168.1.0	255.255.255.0	0.0.0.0
192.168.2.0	255.255.255.0	192.168.1.1
0.0.0.0	0.0.0.0	192.168.10.100

# ARP: Address Resolution Protocol



- The address resolution protocol is responsible for converting the higher-level protocol addresses (IP addresses) to physical network addresses.
  - ARP request (ff:ff:ff:ff:ff) → asks who has IP x.x.x.x
  - ARP reply (unicast) ← I have x.x.x.x and my MAC (can learn MAC through Ethernet headers)

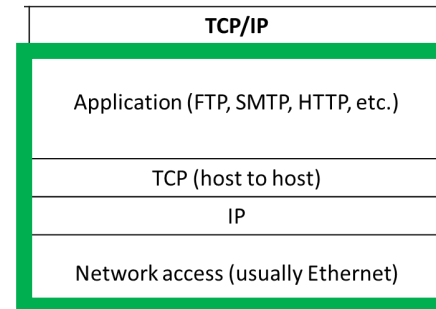
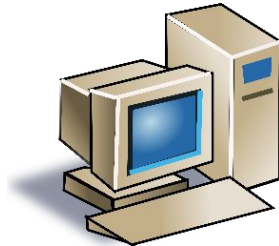


# Network Layer (IP)



- What does a final device does?

- Generates data
- Establishes to whom it wants to send data
- Sends its own data to next hop

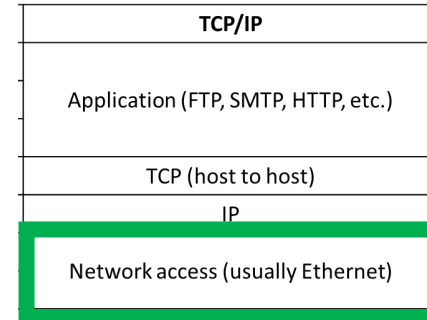


# Network Layer (IP)



- What does a Data Layer device does?

- Network access layer
- Switching device

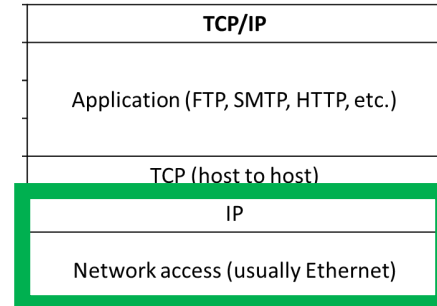


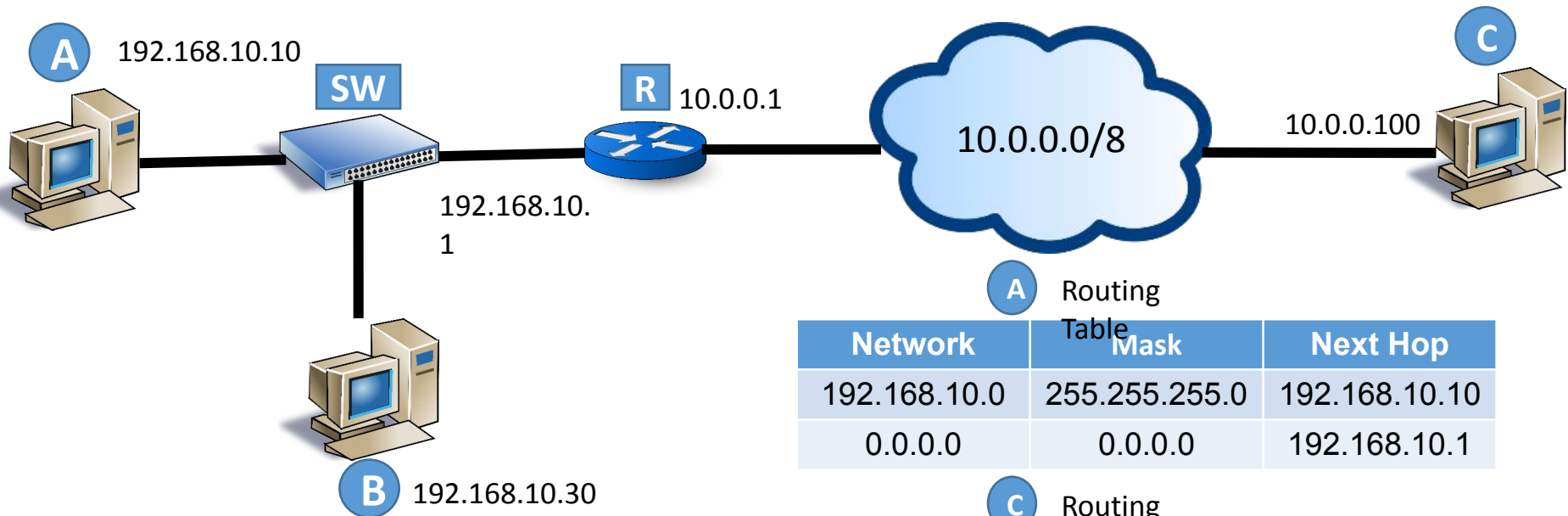
# Network Layer (IP)



- What does a Network Layer device does?

- IP Layer
- Routes packets across its interfaces
- Receives a packet and executes routing algorithm





A Routing Table

Network	Mask	Next Hop
192.168.10.0	255.255.255.0	192.168.10.10
0.0.0.0	0.0.0.0	192.168.10.1

C Routing Table

Network	Mask	Next Hop
10.0.0.0	255.0.0.0	10.0.0.100
0.0.0.0	0.0.0.0	10.0.0.1

R Routing Table

Network	Mask	Next Hop
10.0.0.0	255.0.0.0	10.0.0.1
192.168.10.0	255.255.255.0	192.168.10.1

# Programming and Communications

## III: Internet



- ❑ Overview
- ❑ Architecture
  - ❑ Communication Concept ISO/OSI
  - ❑ Implementation TCP/IP
- ❑ Transmission
  - ❑ Packet data
  - ❑ Ethernet
- ❑ IP
- TCP
- ❑ HTTP



# TCP

Application	Application (FTP, SMTP, HTTP, etc.)
Presentation	
Session	
Transport	TCP (host to host)
Network	IP
Link	Network access (usually Ethernet)
Physical	







# TCP

Layer 4 services may include:

- **Connection-oriented communication:** semi-permanent connection is established before any useful data can be transferred, and where a stream of data is delivered in the same order as it was sent.
- **Same order delivery:** The network layer doesn't generally guarantee that packets of data will arrive in the same order
- **Reliability:** Recovery of packets that may be lost during transport due to network congestion and errors



# TCP

Layer 4 services may include:

- **Flow control:** Rate of data transmission control between two nodes. This can also be used to improve efficiency by reducing buffer underrun
- **Congestion avoidance:** Congestion control can control traffic entry into a telecommunications network to avoid congestive collapse
- **Multiplexing:** Ports can provide multiple endpoints on a single node

Two main protocols used in this layer, TCP and UDP



# TCP: Procols

## TCP:

Reliable connection:

- Acknowledges packets
- Waits for remote acknowledgements (if timeout expires sends again the packet)
- Data can't be lost

**For example, while downloading a file, it is not desired to loose any information(bytes) as it may lead to corruption of downloaded content.**

## UDP:

- UDP provides a comparatively simpler but unreliable service by sending packets from one host to another.
- UDP does not take any extra measures to ensure that the data sent is received by the target host or not.
- If data is lost across the transmission, it is not resent.

**For example while streaming a video, loss of few bytes of information due to some reason is acceptable as this does not harm the user experience much.**



# Programming and Communications

## III: Internet



- ❑ Overview
- ❑ Architecture
  - ❑ Communication Concept ISO/OSI
  - ❑ Implementation TCP/IP
- ❑ Transmission
  - ❑ Packet data
  - ❑ Ethernet
- ❑ IP
- ❑ TCP
- HTTP

A short horizontal bar with a teal left half and an orange right half.

# HTTP

The Hypertext Transfer Protocol (HTTP), model for distributed, collaborative, hypermedia information systems. In a stateless way.

Application	Application (FTP, SMTP, HTTP, etc.)
Presentation	
Session	
Transport	TCP (host to host)
Network	IP
Link	Network access (usually Ethernet)
Physical	



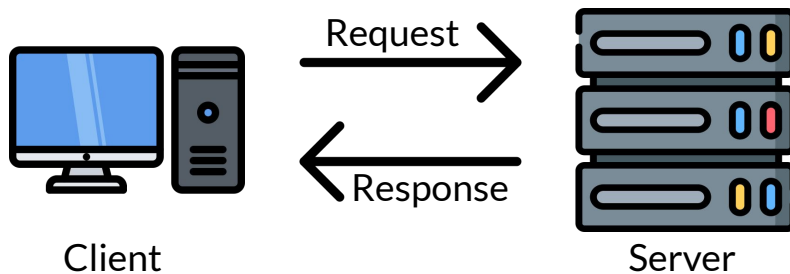


# HTTP

HTTP allows for communication between a variety of hosts and clients, supporting a mixture of network configurations.

Although the communication usually is done over TCP/IP (port 80 by default), any reliable transport can be used.

The communication is done via request/response. The client sends an HTTP request message, and the server returns the message with an HTTP response.





# HTTP

Requests are done by means of URLs( Uniform Resource Locators)

- The host points the where we want to request
- The resource path what we want to request
- The query parameters (optional) can help us to filter what we want
- We will see more about HTTP requests in future lessons ...

http://ip:port/path/to/resource?a=b&x=y  
Protocol      Host      Resource Path      Query Params

http://www.domain.com/path/to/resource?a=b&x=y  
Protocol      Host      Resource Path      Query Params





# HTTP

Responses have two important things (apart from the headers):

- A status message: An informational message about the result of the requests
- A payload (if needed): Data recovered from the server
- We will see more about HTTP requests in future lessons ...







# HTTP

```
POST / HTTP/1.1
```

```
Host: localhost:8000
```

```
User-Agent: Mozilla/5.0 (Macintosh;... )... Firefox/51.0
```

```
Accept: text/html,application/xhtml+xml,..., */*;q=0.8
```

```
Accept-Language: en-US,en;q=0.5
```

```
Accept-Encoding: gzip, deflate
```

```
Connection: keep-alive
```

```
Upgrade-Insecure-Requests: 1
```

```
Content-Type: multipart/form-data; boundary=-12656974
```

```
Content-Length: 345
```

Request headers

General headers

Entity headers

```
-12656974
```

```
(more data)
```





# HTTP

HTTP/1.1 200 OK

Access-Control-Allow-Origin: \*

Connection: Keep-Alive

Content-Encoding: gzip

Content-Type: text/html; charset=utf-8

Date: Wed, 10 Aug 2016 13:17:18 GMT

Etag: "d9b3b803e9a0dc6f22e2f20a3e90f69c41f6b71b"

Keep-Alive: timeout=5, max=999

Last-Modified: Wed, 10 Aug 2016 05:38:31 GMT

Server: Apache

Set-Cookie: csrftoken=.....

Transfer-Encoding: chunked

Vary: Cookie, Accept-Encoding

X-Frame-Options: DENY

Response headers

Entity headers

General headers

(body)





## More info at:

<https://www.geeksforgeeks.org/layers-of-osi-model/>

[http://www.tcpipguide.com/free/t\\_IPDatagramGeneralFormat.htm](http://www.tcpipguide.com/free/t_IPDatagramGeneralFormat.htm)

<https://www.geeksforgeeks.org/differences-between-tcp-and-udp/>

<https://www.geeksforgeeks.org/modulo-2-binary-division/>

<https://developer.mozilla.org/en-US/docs/Web/HTTP/Overview>

[https://www.uobabylon.edu.iq/eprints/publication\\_3\\_19816\\_1548.pdf](https://www.uobabylon.edu.iq/eprints/publication_3_19816_1548.pdf)

