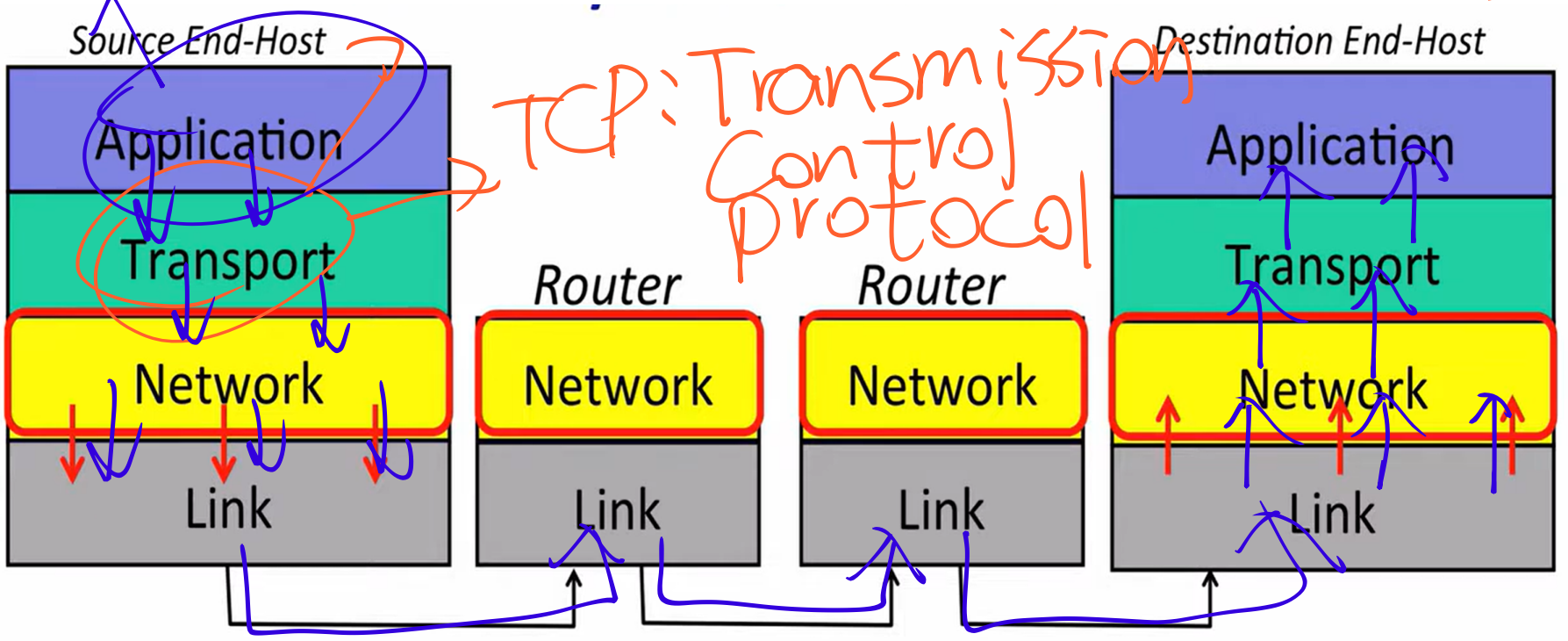
What the Internet is

The 4 Layer Internet Model

API

UDP: User Datagram Protocol

TCP: Transmission Control Protocol



The Network Layer is special

We must use the Internet Protocol (IP)

① IP makes a best-effort attempt to deliver our datagrams to the other end. But it makes no promises.

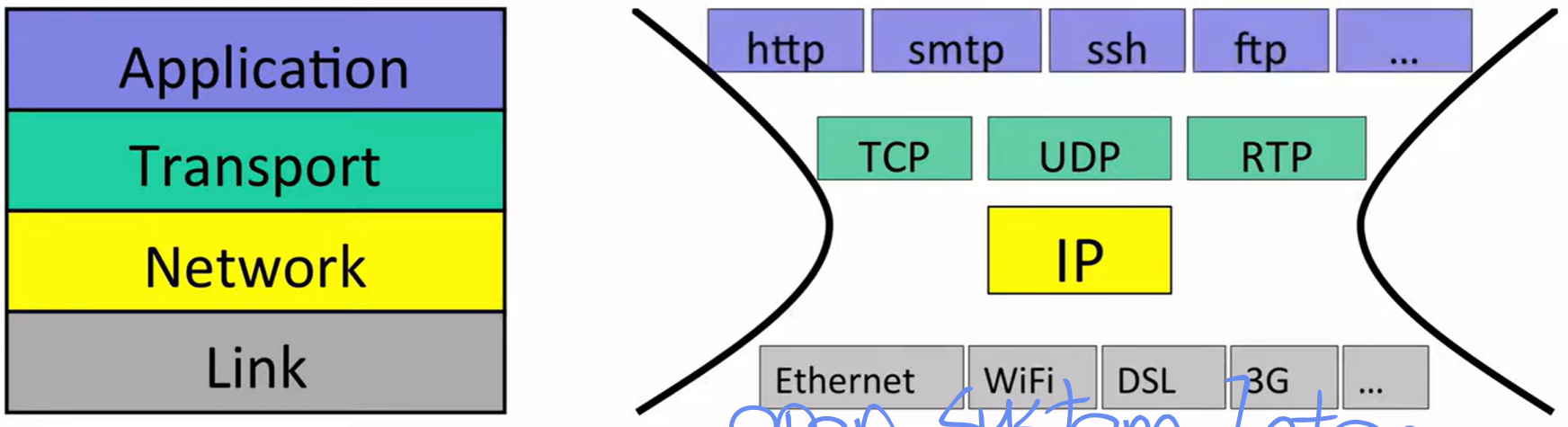
② IP datagrams can get lost, can be delivered out of order, and can be corrupted. There are no guarantees.

Summary of 4 Layer Model

Application	Bi-directional reliable byte stream between two applications, using application-specific semantics (e.g. http, bit-torrent).
Transport	Guarantees correct, in-order delivery of data end-to-end. Controls congestion.
Network	Delivers datagrams end-to-end. Best-effort delivery – no guarantees. Must use the Internet Protocol (IP).
Link	Delivers data over a single link between an end host and router, or between routers

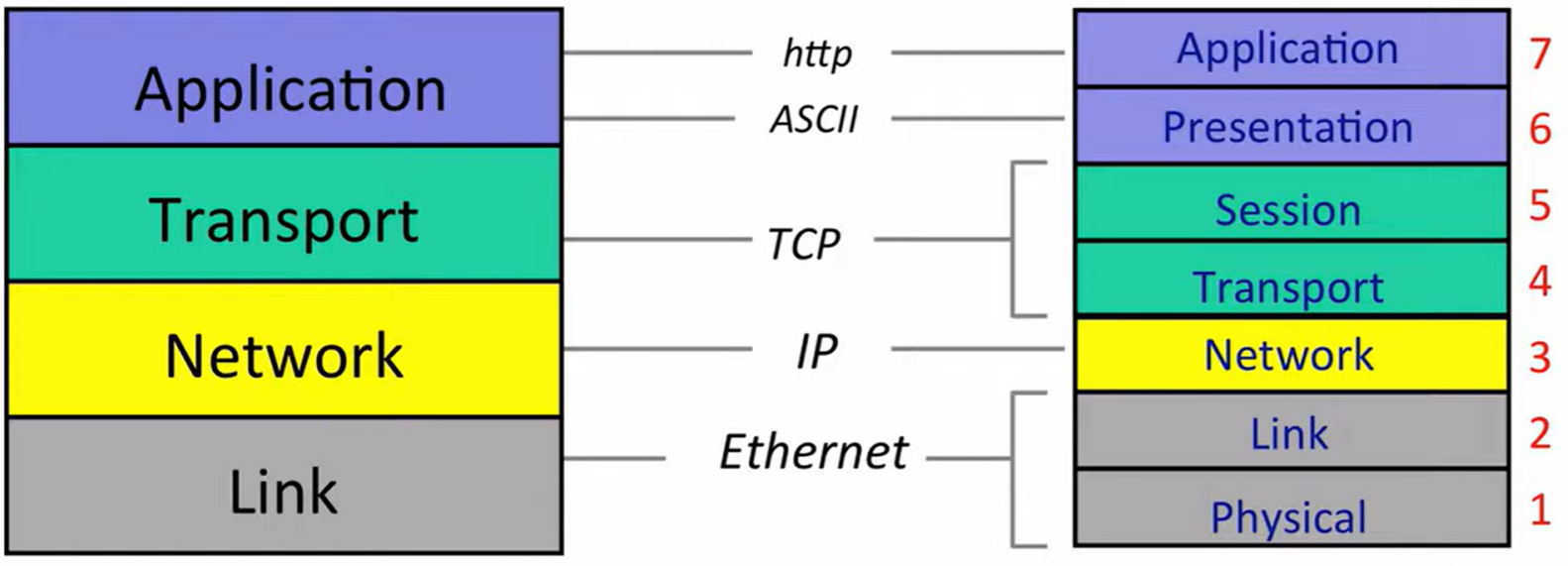
The two more things you should know

IP is the "thin waist"



open system Interconnect

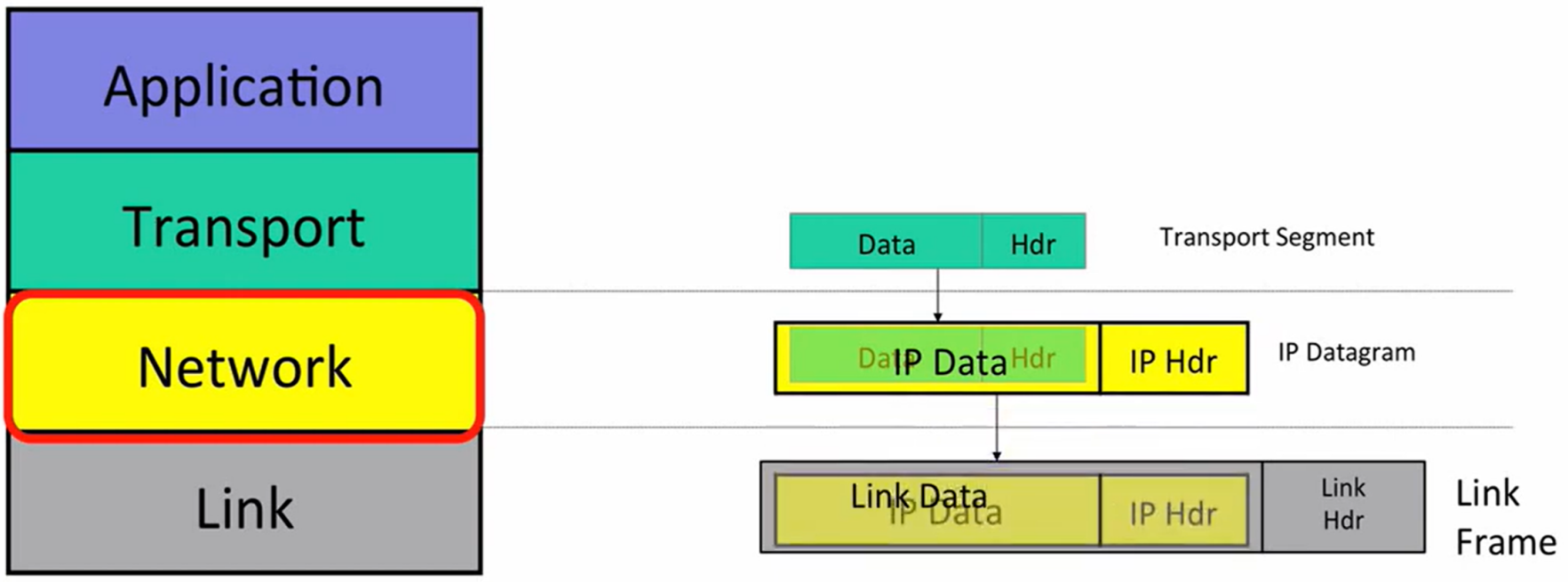
The 7-layer OSI Model



The 7-layer OSI Model

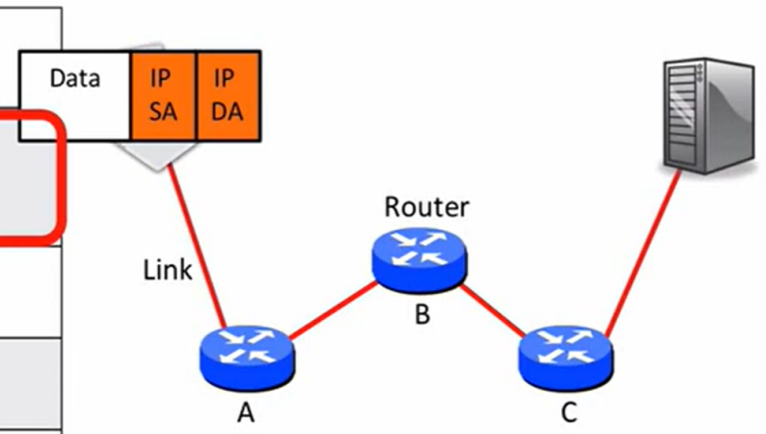
The IP Service

The Internet Protocol (IP)



The IP Service Model

Property	Behavior
Datagram	Individually routed packets. Hop-by-hop routing.
Unreliable	Packets might be dropped.
Best effort	…but only if necessary.
Connectionless	No per-flow state. Packets might be mis-sequenced.



Why the IP service so simple?

① simple, dumb, minimal: Faster, more streamlined and lower cost to build and maintain.

② the end-to-end principle: Where possible, implement features in the end hosts.

③ Allows a variety of reliable (or unreliable) services to be built on top.

④ Works over any link layer: IP makes very few assumptions about the link layer below.

The IP Service Model (Features)

① Tries to prevent packets looping forever.

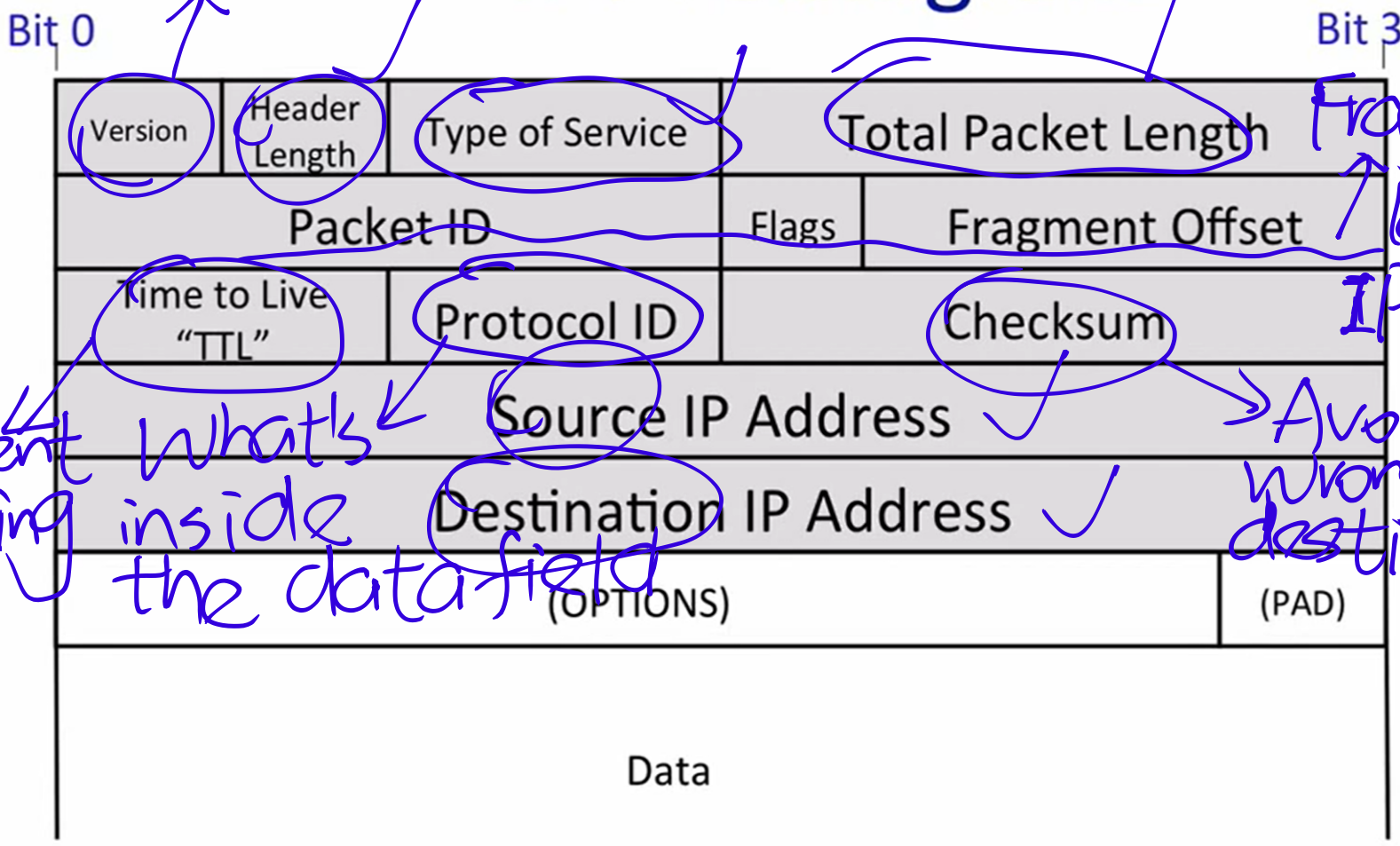
② Will fragment packets if they are too long.

③ Uses a header checksum to reduce chances of delivering datagram to wrong destination.

④ Allows for new versions of IP
- IPv4 with 32 bit address
- IPv6 with 128 bit address.

⑤ Allows for new options to be added to header.

IPv4 Datagram



IPv4/IPv6

up to 64KB (header + data)

Fragment large IP packet

prevent looping

What's inside the data field

Avoid wrong destination

Life of a Packet

Application : stream of data

↓ ↓

Transport: segments of data

↓ ↓

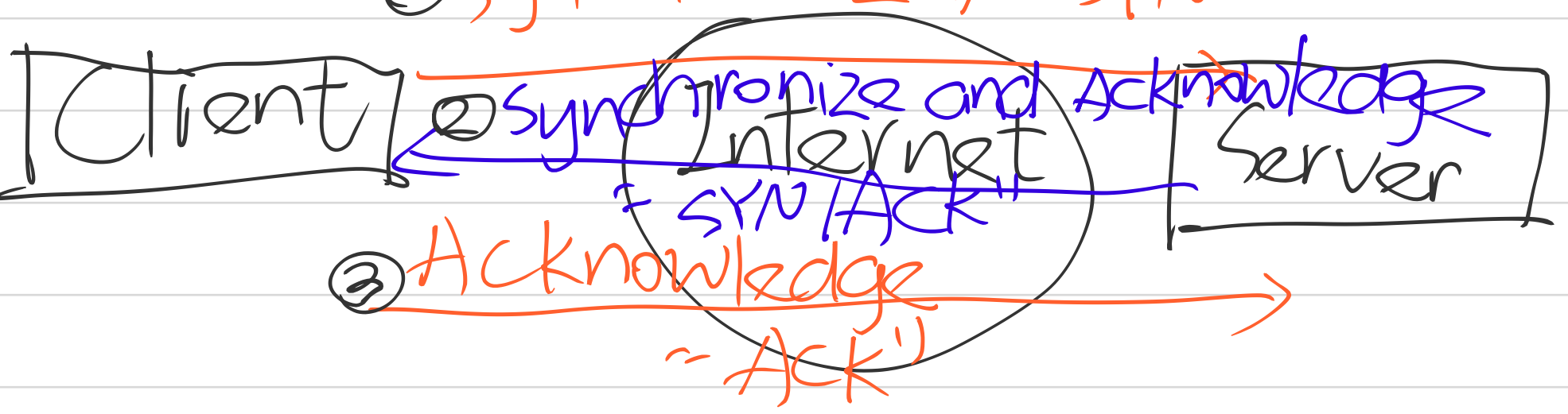
Network: Packets of data

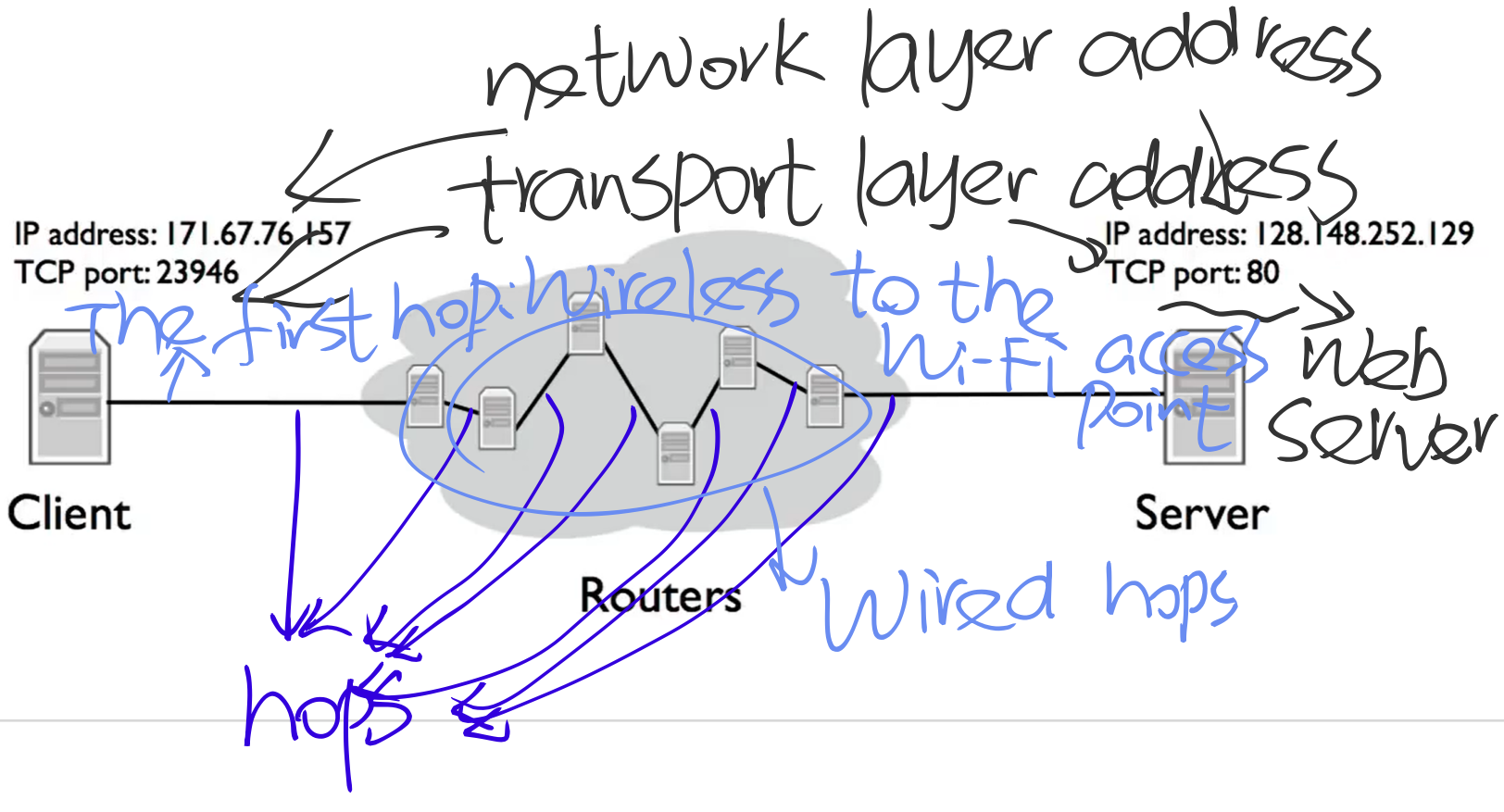
TCP Byte Stream

"3-way handshake"

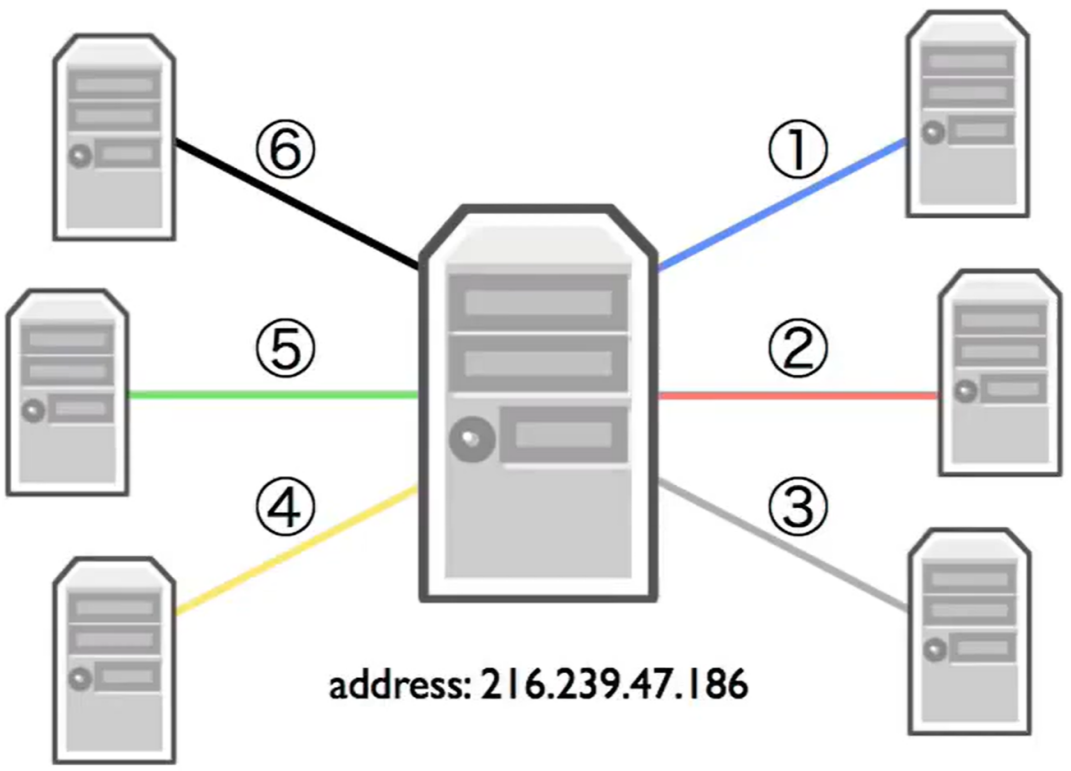
("SYN" + "SYN/ACK" + "ACK")

① synchronize, "SYN"





Inside Each Hop



forwarding table

dest	link
default	1
pattern A	5
pattern B	2
pattern C	4
pattern D	2
pattern E	6
pattern F	1

Packet Switching (Principle)

Packet: A self-contained unit of data that carries information necessary for it to reach its destination.

Packet switching: Independently for each arriving packet, pick its outgoing link. If the link is free, send it. Else hold the packet for later.

No per-flow state required (simple)

Flow: A collection of datagrams belonging to the same end-to-end communication e.g. a TCP connection.

Packet switches don't need state for each flow - each packet is self-contained

No per-flow state to be added/removed
No per-flow state to be stored
No per-flow state to be changed upon failure

Efficient sharing of links (Efficient)

Data traffic is bursty:

- Packet switching allows flows to use all available link capacity.
- Packet switching allows flows to share link capacity

This is call "statistical Multiplexing"

Layering (Principle)

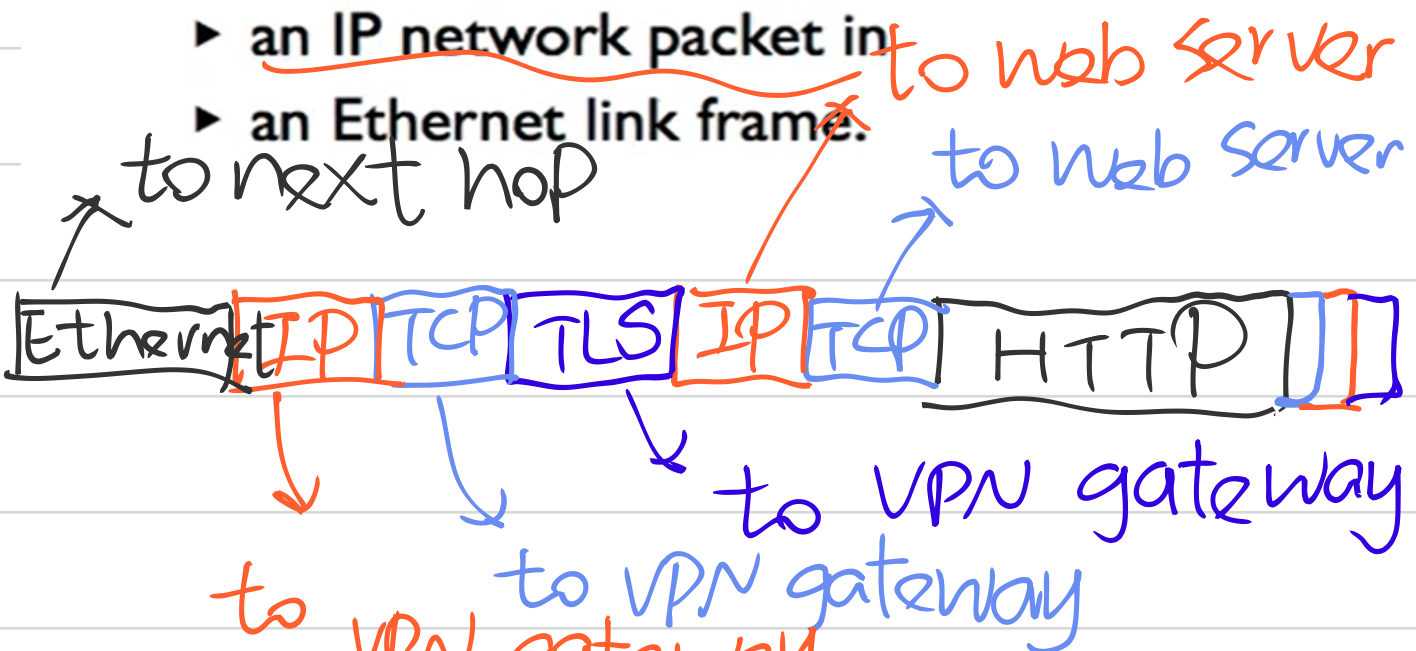
Six reasons:

1. Modularity
2. Well defined service
3. Reuse
4. Separation of concerns
5. Continuous improvement
6. Peer-to-peer communication

Encapsulation (Principle)

(the result of combining layering and Packet switching)

- Encapsulation allows you to layer recursively
- Example: Virtual Private Network (VPN):
  - HTTP (web) application payload in
  - a TCP transport segment in
  - an IP network packet in
  - a secured TLS presentation message in
  - a TCP transport segment in
  - an IP network packet in
  - an Ethernet link frame.



already

Memory, Byte Order, and Packet Formats

Memory

Endianness:

① Little endian: LSB is at lowest address

Makes most sense from an addressing/computational standpoint

② Big endian: MSB is at lowest address

Makes most sense to a human reader.

Network Byte Order

Different processors have different endianness

eg. Little endian: x86; Big endian: ARM

To interoperate, they need to agree how to represent multi-byte fields

Network Byte Order is big endian

Packet Formats

We have already talked before.

Name and Addresses: IPv4

Goal of IP address

① Stitch many different networks together
② Need network-independent unique address

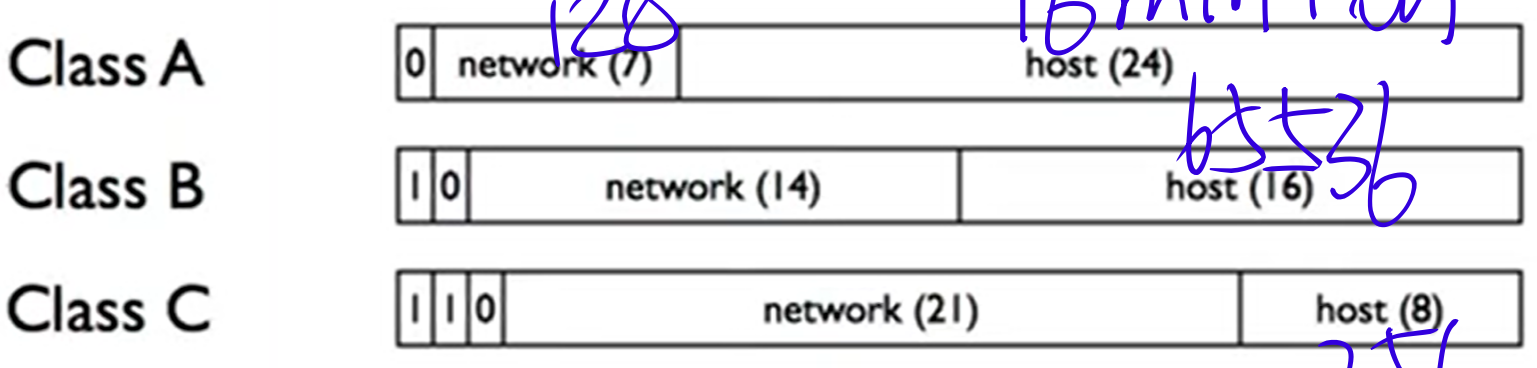
IPv4: Internet Protocal, version 4

- An IPv4 address identifies a device on the Internet
  - Layer 3 (network) address
- 32 bits long (4 octets): a.b.c.d 4×8bits
  - Example: 171.64.64.64
  - Example: 128.30.76.82
  - Example: 12.22.58.30
- Netmask: apply this mask, if it matches, in the same network
  - Netmask of 255.255.255.0 means if the first 24 bits match
  - Netmask of 255.255.252.0 means if the first 22 bits match
  - Netmask of 255.128.0.0 means if the first 9 bits match
  - Smaller netmask (fewer 1s) means larger network

Address Structure

historical

- Originally hierarchical: network + host
  - Network to get to correct network (administrative domain)
  - Host to get to correct device in network (within administrative domain)
- Originally 3 classes of addresses: class A, class B, class C



today

- Still assign contiguous ranges of addresses to nearby networks
  - Class A, B, C is too coarse grained (e.g., MIT dorms!)
  - http://news.stanford.edu/news/1999/january27/itss127.html
- Classless Inter-Domain Routing (CIDR)
  - Address block is a pair: *address,count*
  - Counts are powers of 2, specify netmask length
  - 171.64.0.0/16 means any address in the range 171.64.0.0 to 171.64.255.255
  - A /24 describes 256 addresses, a /20 describes 4,096 addresses
- Stanford today has 5 /16 blocks -- 325,000 addresses

IPv4 Address Assignment

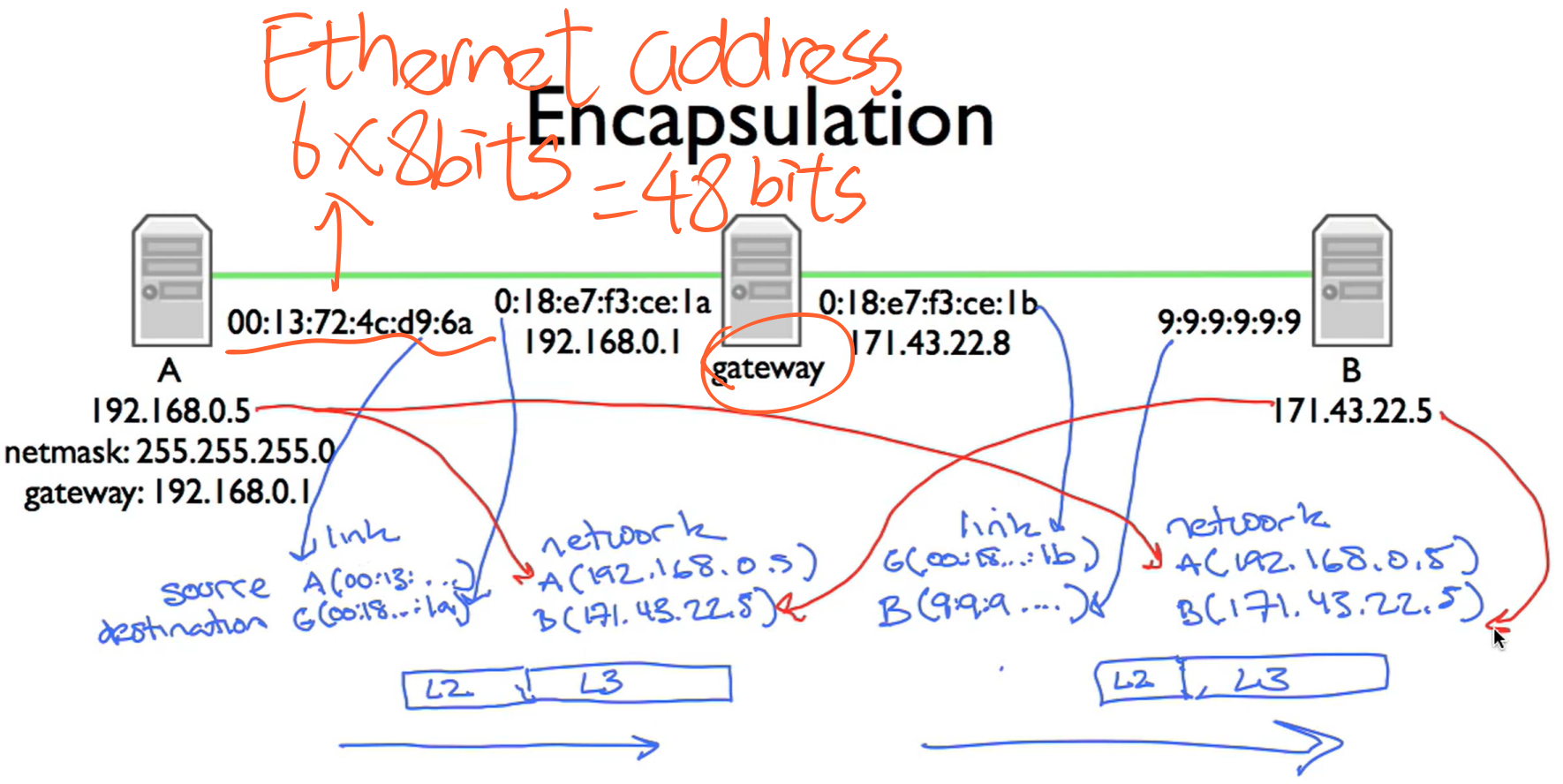
- IANA: Internet Assigned Numbers Authority
  - Internet Corporation for Assignment of Names and Numbers (ICANN)'s job
- IANA gives out /8s to Regional Internet Registries (RIRs)
  - Ran out in February 2011, in special end case of giving 1 to each RIR
- RIRs responsible for geographic regions, each has own policy
  - AfriNIC: Africa
  - ARIN: U.S.A., Canada, Carribean, Antarctica
  - APNIC: Asia, Australia, New Zealand
  - LACNIC: Latin America, Carribean
  - RIPE NCC: Europe, Russia, Middle East, Central Asia

Longest Prefix Match

- Algorithm IP routers use to chose matching entry from forwarding table
- Forwarding table is a set of CIDR entries
  - An address might match multiple entries
  - E.g., 171.33.0.1 matches both entries on right
- Algorithm: use forwarding entry with the longest matching prefix
  - Longest prefix match will chose link 5 for 171.33.0.1

dest	link
0.0.0.0/0	1
171.33.0.0/16	5

Address Resolution Protocol (ARP)



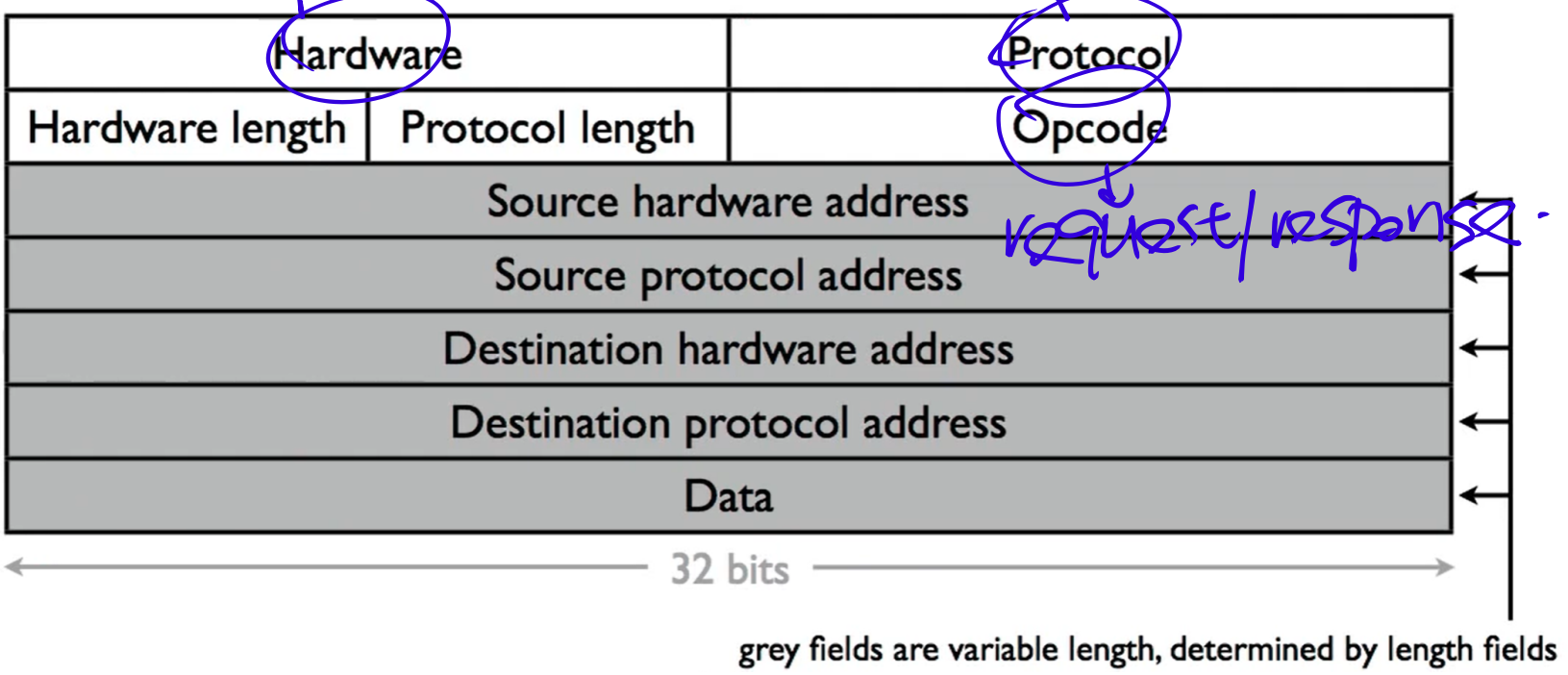
ARP

- Generates mappings between layer 2 and layer 3 addresses
  - Nodes cache mappings, cache entries expire
- Simple request-reply protocol
  - "Who has network address X?"
  - "I have network address X."
- Request sent to link layer broadcast address
- Reply sent to requesting address (not broadcast)
- Packet format includes redundant data
  - Request has sufficient information to generate a mapping
  - Makes debugging much simpler
- No "sharing" of state: bad state will die eventually

ARP Packet Format (RFC826)

what link layer

what network Protocol



request/response.

Hardware ⇔ Ethernet

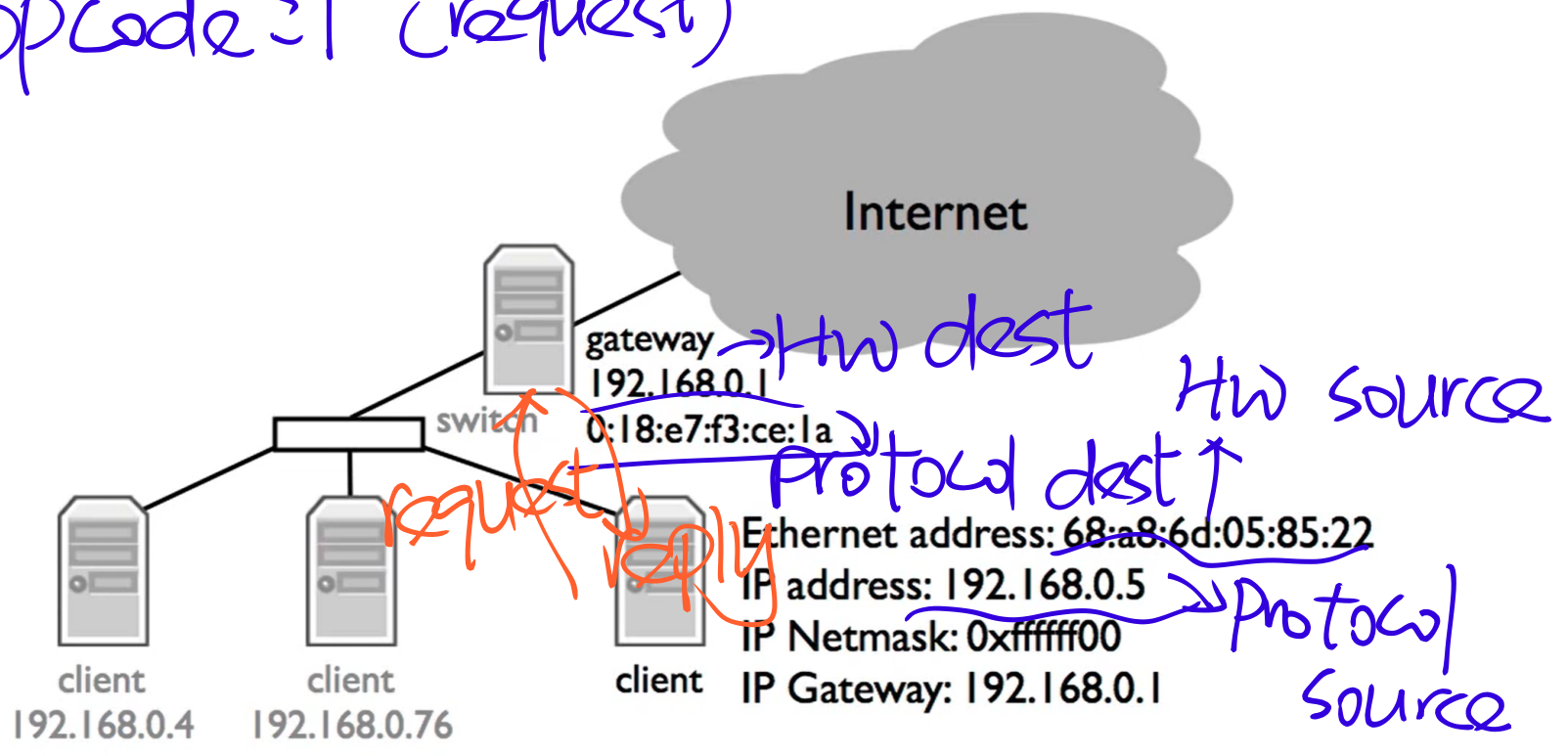
Protocol ⇔ IP

Hw length: 6 (48 bits)

Protocol length: 4 (32 bits)

opcode: 1 (request)

ARP Request



What you learned

How an application uses the Internet

The structure of the Internet: The 4 layer model

The Internet protocol (IP): What it is

Basic architectural ideas and principles

- Packet switching
- Layering
- Encapsulation