Mid-term Summary and Feedback

ELEC-C7420 Basic principles in networking

Aalto University
School of Electrical
Engineering

Yu Xiao

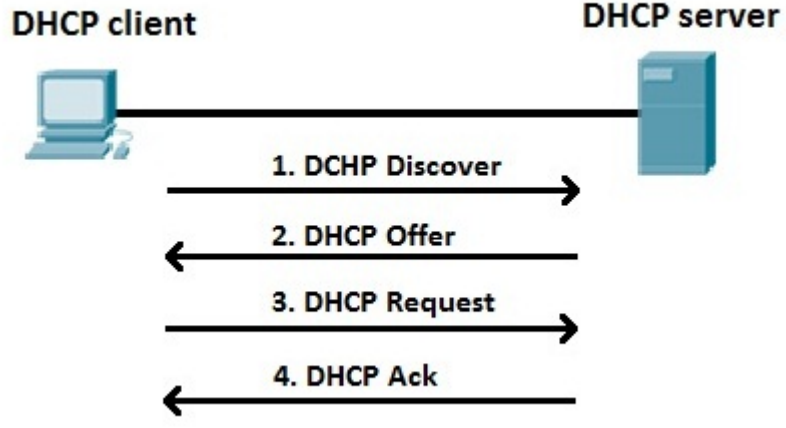
2022.02.15

Agenda

- DHCP, TCP congestion control
- What we have learnt so far
- Feedback on assignments
- Feedback from students

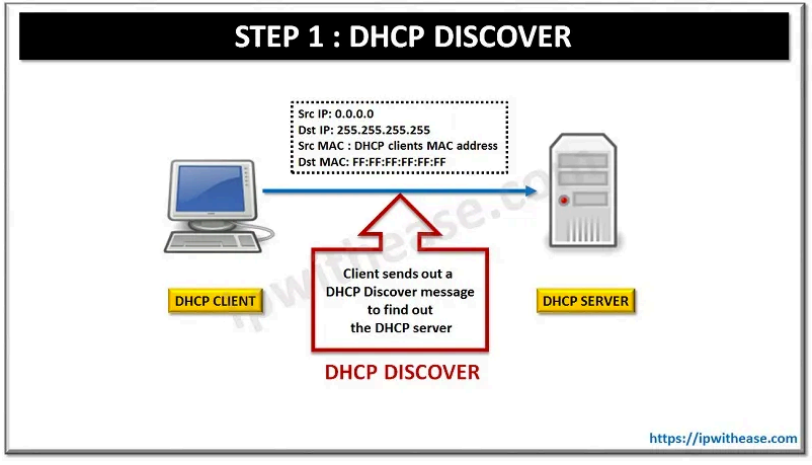
Dynamic Host Configuration Protocol (DHCP)

- DHCP is a network management protocol that is used to assign an IP address and various network parameters (e.g., default gateway, domain name, name servers, time server) to a device.
- Application layer protocol
- It runs on top of UDP/IP

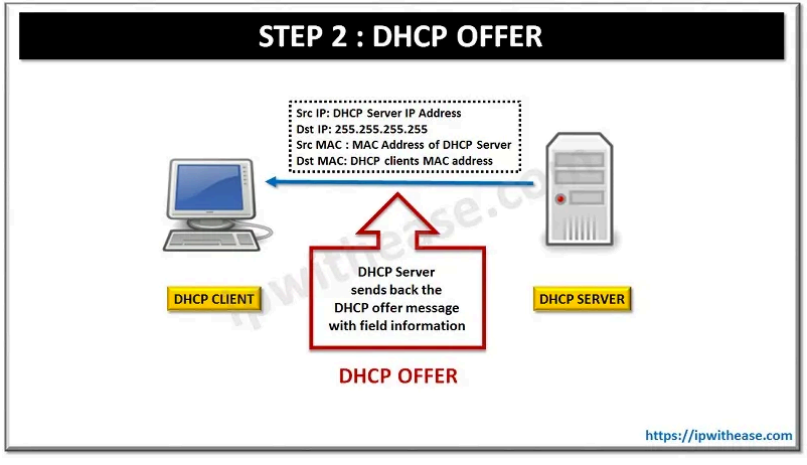


DHCP uses a well-known UDP port number 67 for the DHCP server, and the UDP port number 68 for the client.

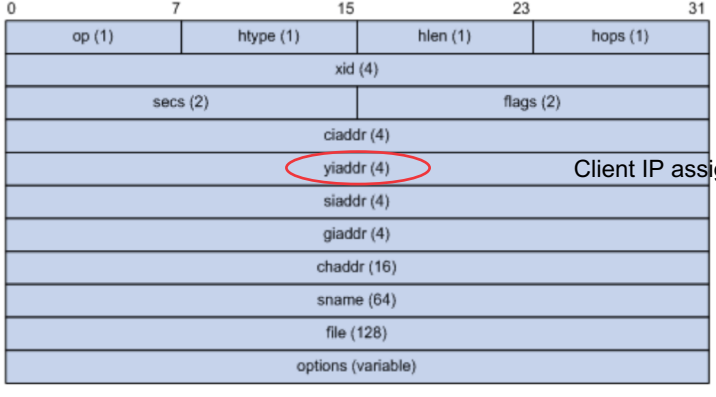
Example of the DORA process



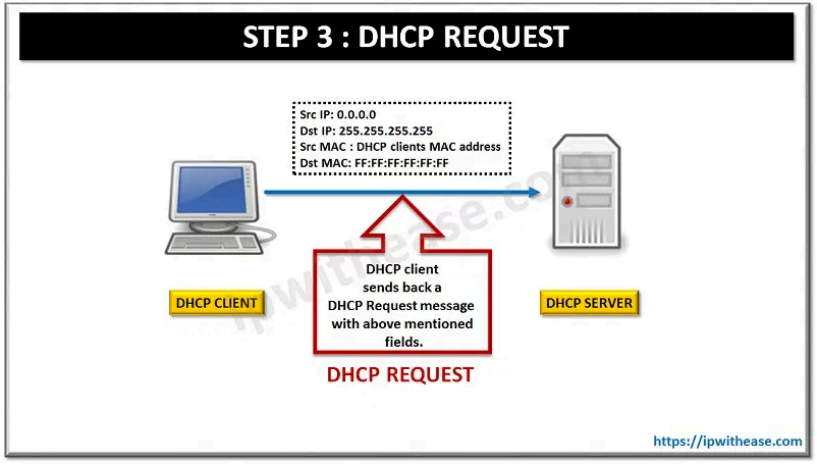
Layer 3 broadcast + Layer 2 broadcast

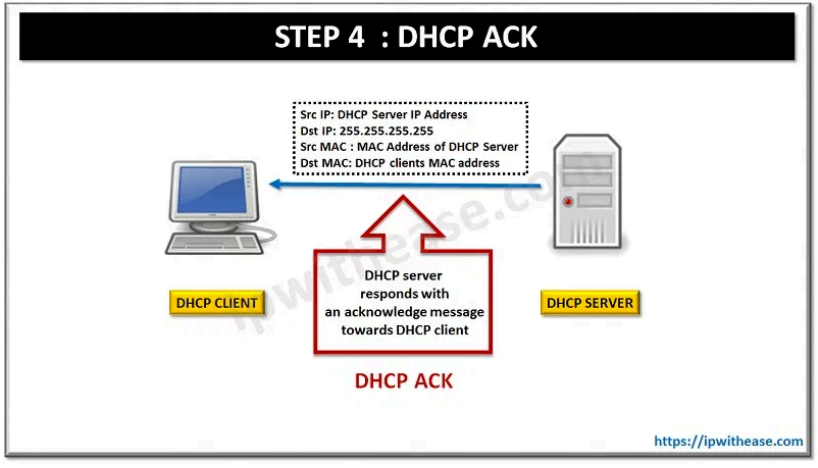


Layer 3: Still Broadcast as Client still has no IP Address
Layer 2: Unicast



Client IP assigned by the server





Layer 3: Still Broadcast as Client must have received Offer from more than one DHCP server in their domain and the DHCP client accepts the Offer that its receives the earliest and by doing a broadcast it intimates the other DHCP server to release the Offered IP address to their available pool again

DHCP

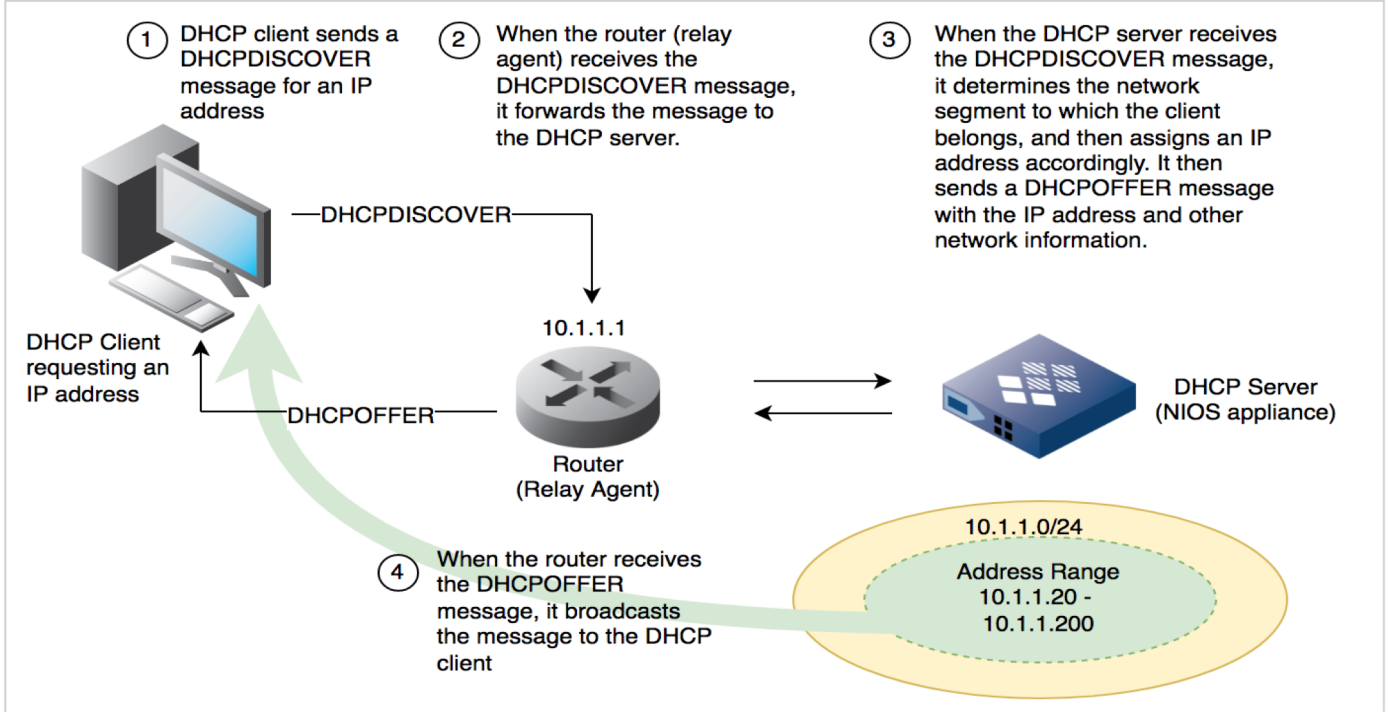
1: A DHCP client sends a broadcast packet (**DHCP Discover**) to discover DHCP servers on the LAN segment.

2: The DHCP servers receive the DHCP Discover packet and respond with **DHCP Offer** packets, offering IP addressing information.

3: If the client receives the DHCP Offer packets from multiple DHCP servers, the first DHCP Offer packet is accepted. The client responds by broadcasting a **DHCP Request** packet, requesting the network parameters from the server that responded first.

4: The DHCP server approves the lease with a **DHCP Acknowledgement** packet. The packet includes the lease duration and other configuration information.

DHCP with relay agent



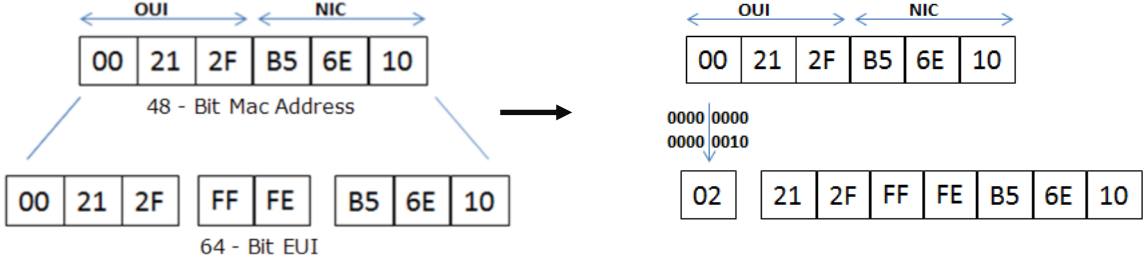
https://docs.infoblox.com/display/NAG8/IPv4+DHCP+Protocol+Overview

DHCP v6

- **SLAAC** (**S**tate**L**ess **A**ddress **A**uto **C**onfiguration): The preferred method of assigning IP addresses in an IPv6 network.
    - SLAAC devices send the router a request for the network prefix, and the device uses the prefix and its own MAC address to create an IP address. After the IP is computed, it checks to see if a duplicate IP was previously created.
- If the router does not implement SLAAC and no network prefix is received, the device sends a request to the DHCPv6 server, which responds with an IP address similar to the DHCP in IPv4.

IPv6 Interface Identifier

- Most IPv6 addresses can be divided into a **64-bit network prefix** and a 64-bit “host” portion. These host-portion bits are known officially as the **interface identifier.**
- **Modified EUI-64 Identifier (EUI: extended unique identifier)**
  - Given an Ethernet address (48 bits), insert 0xfffe between the first 3 bytes and the last 3 bytes, to get 64 bits in all.
  - Set the 7th bit of the first byte to 1



Source: Cisco

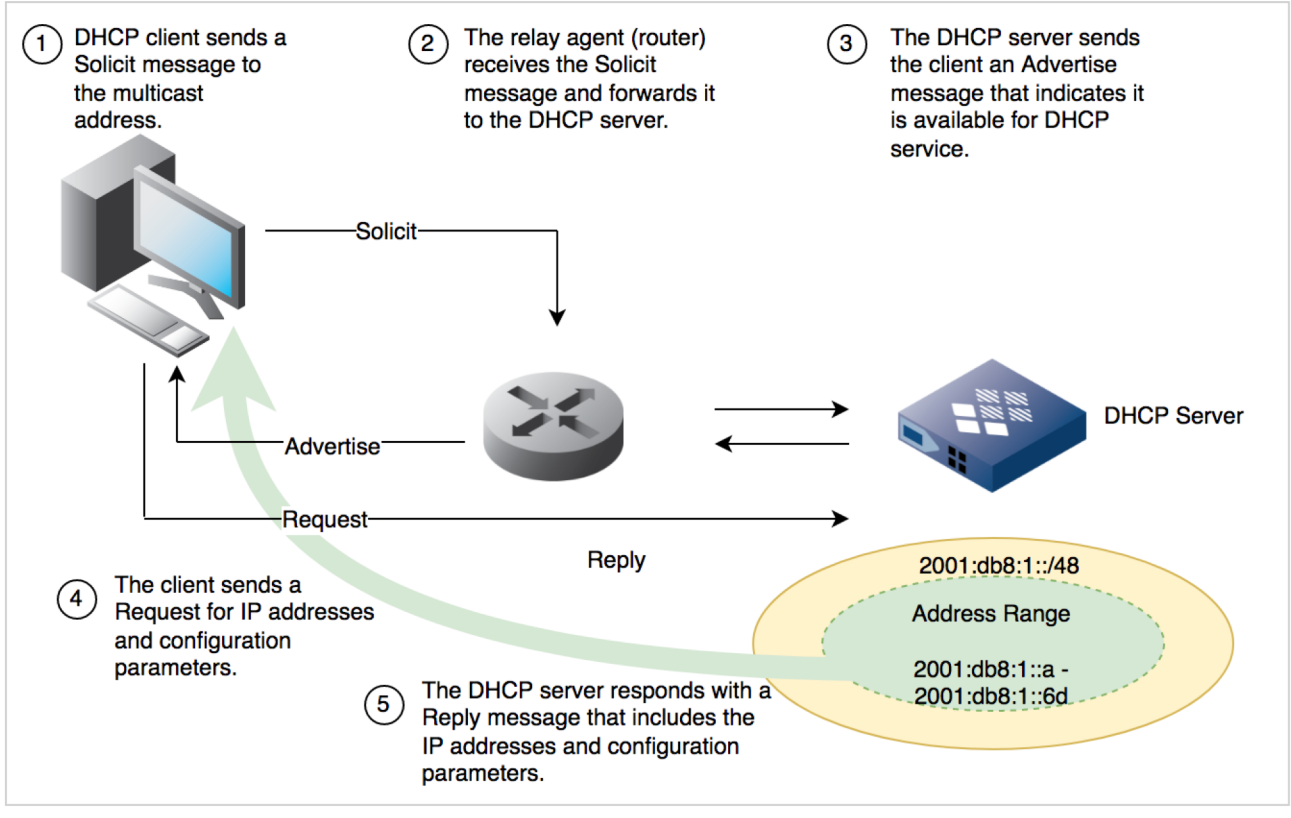
Stateful DHCPv6

Stateful DHCPv6 uses a DHCPv6 Server to centrally manage IPv6 address and prefix assignment.

- DHCPv6 Clients get IPv6 address or prefix information from the DHCPv6 Server.
- DHCPv6 Clients can obtain configuration information that is not available from other protocols, such as DNS.

Stateless DHCPv6: does not require a DHCPv6 Server to maintain any dynamic state for clients, such as DNS Server addresses.

DHCP v6



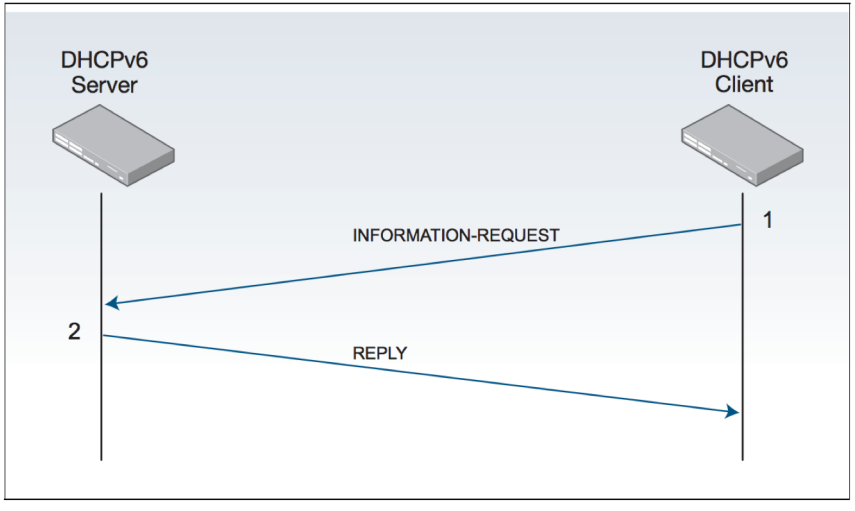
https://docs.infoblox.com/display/NAG8/IPv6+DHCP+Protocol+Overview

Normal DHCPv6 Message Exchange

1. **Solicit** - sent by a DHCPv6 Client to locate DHCPv6 Servers.
2. **Advertise** - sent by a DHCPv6 server to a DHCPv6 Client in answer to the solicit message as an affirmative message that DHCPv6 Server services are available to a DHCPv6 Client.
3. **Request** - sent by a DHCPv6 Client to a DHCPv6 Server to request configuration parameters.
4. **Reply** - sent by a DHCPv6 Server to a DHCPv6 Client with configuration information.
5. **Renew** - sent by a DHCPv6 Client to a DHCPv6 Server requesting an extension to the address lifetime.

Stateless DHCP v6

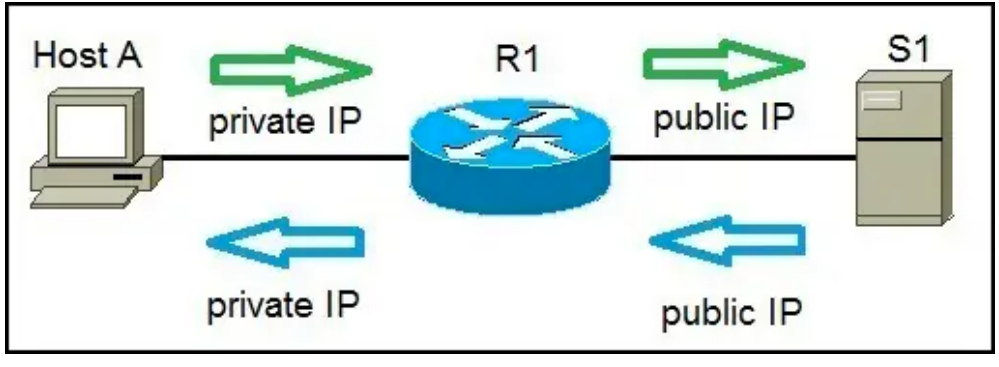
- The Client starts by sending an **INFORMATION-REQUEST** message to the Server. This request specifically excludes the assignment of any IPv6 address.
- The Server sends a **REPLY** message back to the Client to finish.



NAT

NAT (Network Address Translation) is a process of changing the source and destination IP addresses and ports.

- Address translation reduces the need for IPv4 public addresses and hides private network address ranges.
- This process is usually done by routers or firewalls.



There are three types of address translation:

Static NAT – translates one private IP address to a public one. The public IP address is always the same.

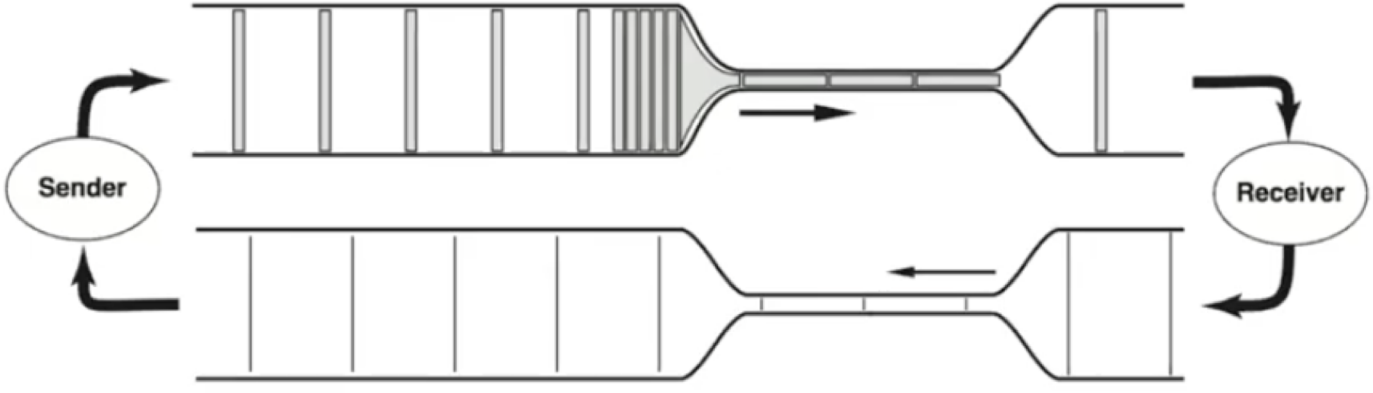
Dynamic NAT – private IP addresses are mapped to the pool of public IP addresses.

Port Address Translation (PAT) – one public IP address is used for all internal devices, but a different port is assigned to each private IP address. Also known as **NAT Overload**.

TCP Congestion Control

How to perceive congestion?

- **Bottleneck**
  - It determines the connection's maximum data-delivery rate.
  - It is where persistent queues form



How to perceive congestion?

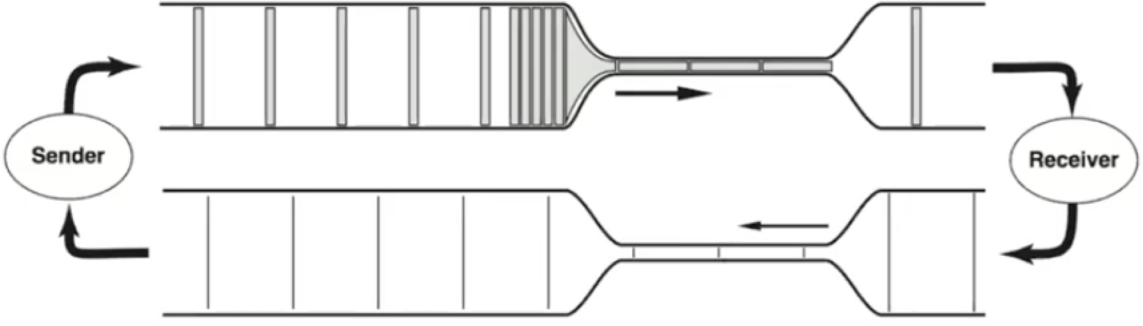
- **Implicit end-to-end feedback**
  - IP layer provides no explicit feedback to end systems regarding congestion
  - Presence of congestion inferred by the end systems based only on observed network behavior (e.g. packet loss and delay)
- **Success Event (ACK received) vs. Loss Event (timeout or duplicated ACKs)**
- **Earlier versions of TCP interprets packet loss as congestion. Does this assumption always hold?**

- **Sources of errors in wireless links**
  - Pauses due to handoff between cells
  - Mobile host out of reach of other receivers (little or no overlaps between cells)
  - Packet losses due to transmission errors in wireless links

- **In wireless lossy links, the sporadic losses are not due to congestion → unnecessary window and transmission rate reduction if loss is interpreted as congestion**

What has changed since earlier versions of TCP were invented?

- **NIC evolves from Mbps to Gbps and memory chips from KB to GB**



What would happen to loss-based congestion control if bottleneck buffers are large?

What is Bufferbloat?

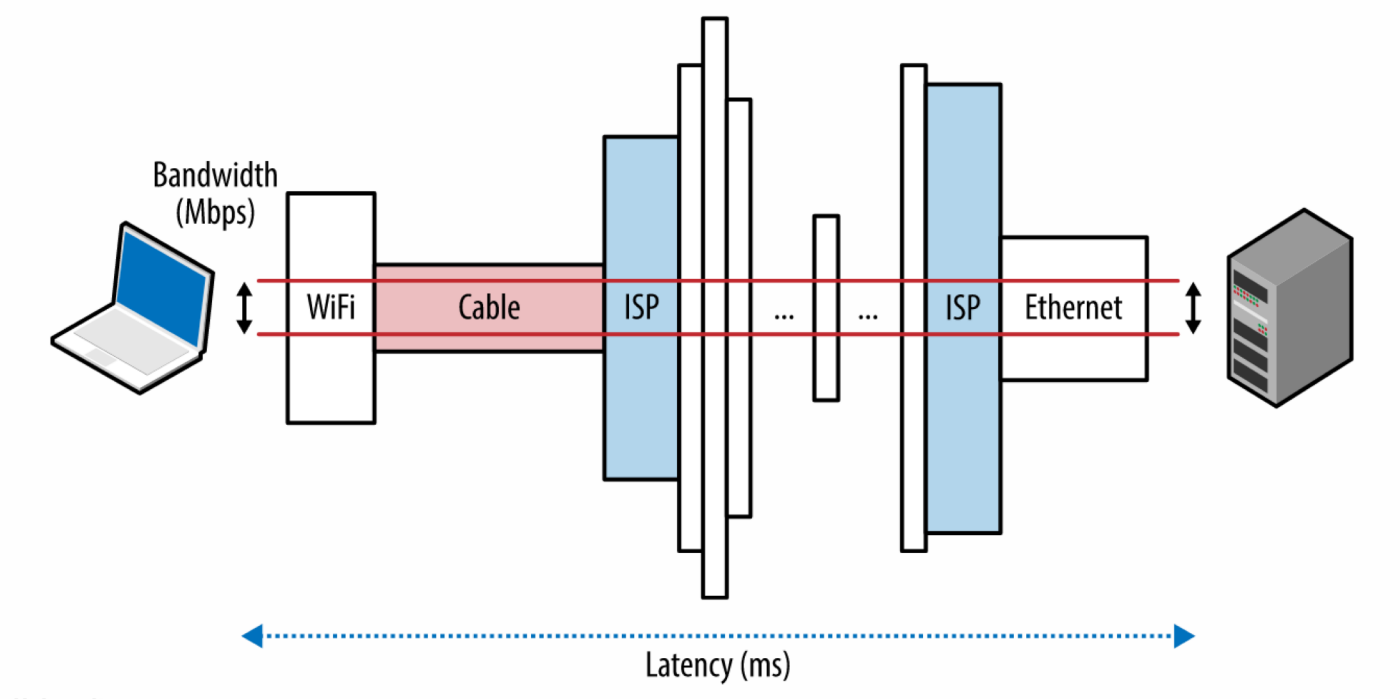
“Bufferbloat is the undesirable latency that comes from a router or other network equipment buffering too much data.”

source: bufferbloat.net

Bufferbloat is a term that was coined and popularized by Jim Gettys in 2010, and is a great example of queuing delay affecting the overall performance of the network.

Latency and Bandwidth

Latency: The time from the source sending a packet to the destination receiving it
Bandwidth: Maximum throughput of a logical or physical communication path



https://hpbn.co/primer-on-latency-and-bandwidth/#speed-is-a-feature

- **Propagation delay:** Amount of time required for a message to travel from the sender to receiver, which is a function of distance over speed with which the signal propagates.
- **Transmission delay:** Amount of time required to push all the packet's bits into the link, which is a function of the packet's length and data rate of the link.
- **Processing delay:** Amount of time required to process the packet header, check for bit-level errors, and determine the packet's destination.
- **Queuing delay:** Amount of time the packet is waiting in the queue until it can be processed.

  If the packets are arriving at a faster rate than the router is capable of processing, then the packets are queued inside an incoming buffer.

Bufferbloat

*"The underlying problem is that **many routers are now shipping with large incoming buffers under the assumption that dropping packets should be avoided at all costs**. However, this breaks TCP's congestion avoidance mechanisms, and introduces **high and variable latency delays** into the network."*

Solution: active queue management
https://queue.acm.org/detail.cfm?id=2209336

- **Congestion wasn’t reported until after a prolonged period of queueing, resulting in greater packet loss**

- **Packet loss alone is not a good proxy to detect congestion**

TCP BBR

N. Cardwell, Y. Cheng, C.S. Gunn, S.H. Yegneh and V. Jacobson. BBR: congestion based congestion control. Communications of the ACM. 60(2). February 2017.

https://www.youtube.com/watch?v=VIX45zMMZG8

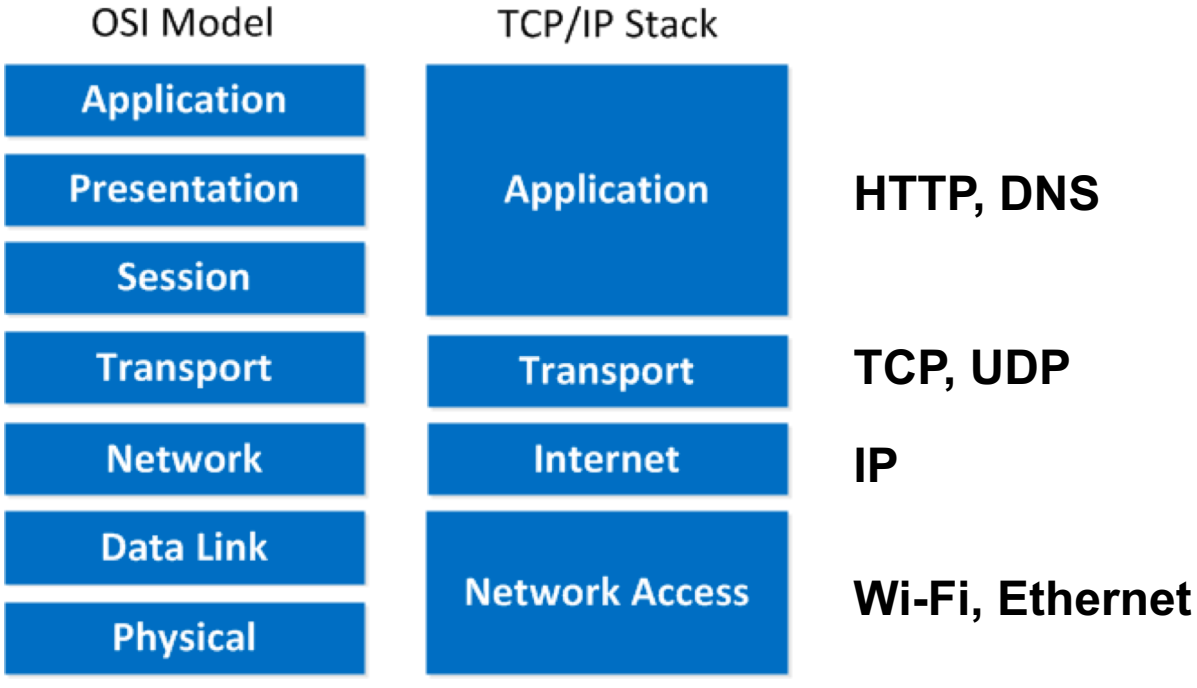
Design Criteria of BBR

- **Make network full but buffer empty → Maintain the data rate at the bottleneck bandwidth and the pipe is full**

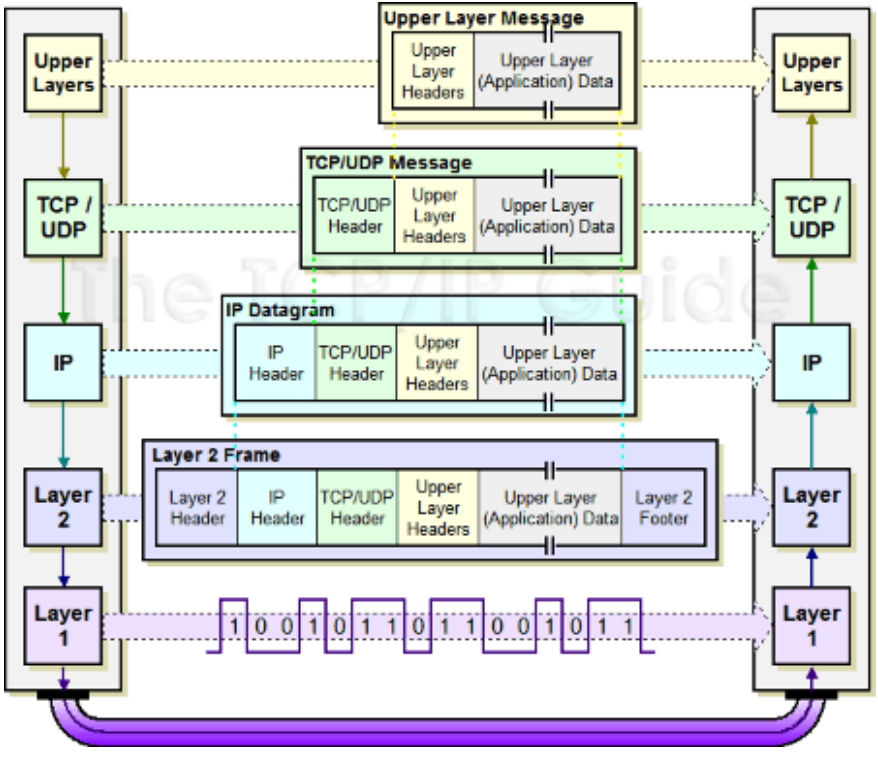
- **BBR looks at the path's bottleneck bandwidth and an estimate of the RTT to determine congestion in a network**

What we have learnt so far

- **Unicast, broadcast, multicast, anycast**
- **Personal area network, local area network, metropolitan area network, wide area network**


OSI Model
TCP/IP Stack
Application
Presentation
Session
Transport
Network
Data Link
Physical
Application
Transport
Internet
Network Access
HTTP, DNS
TCP, UDP
IP
Wi-Fi, Ethernet

Data Encapsulation



Source: https://buildingautomationmonthly.com/what-is-the-tcp-ip-stack/

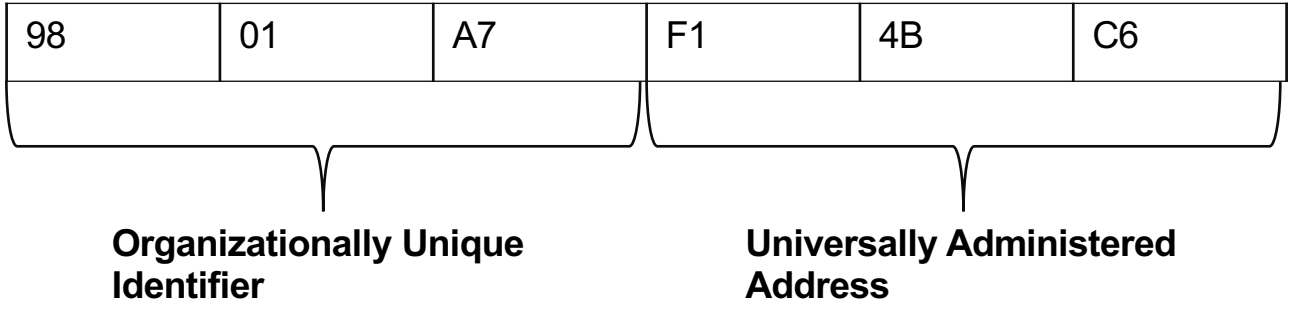
Data Link Layer (Layer 2) has two sub-layers.

- **IEEE 802.2 Logical Link Control (LLC)**
- **IEEE 802.3 Media Access Control (MAC)**

IEEE stands for Institute of Electrical and Electronics Engineers

MAC Address

- A MAC address is a **unique identifier** assigned to a network interface controller (NIC) by manufacturer
- Used for communications at the data link layer
- Ethernet addresses are 6 bytes long

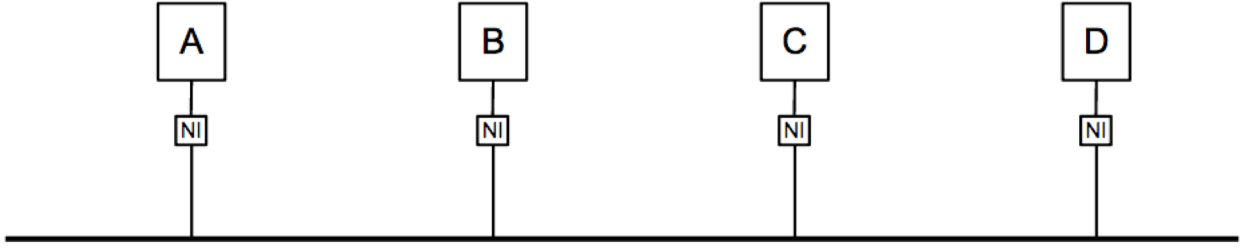


Two Operating Modes of MAC Sub-Layer

- **Half Duplex: A host can only send or receive at one time**
  - CSMA/CD
- **Full Duplex: A host can send and receive simultaneously. No collision.**
- **Duplex configuration: either manually set or auto negotiated by connected devices**
- **Duplex mismatch → poor performance**

CSMA/CD

- **Carrier Sense Multiple Access with Collision Detection (CSMA/CD)** defines how Ethernet frames get onto an Ethernet network
- CSMA/CD is designed to allow fair access by all transmission devices to shared network channels.



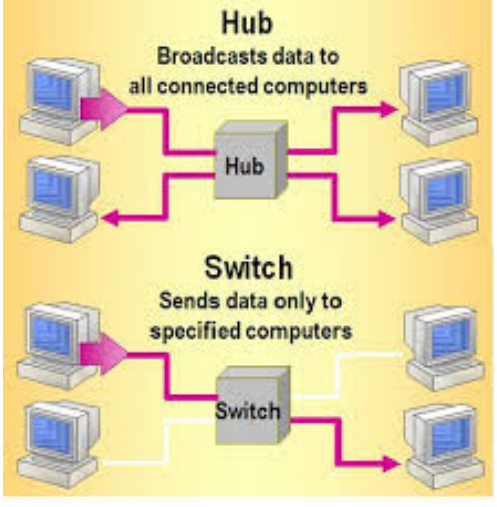
Hubs vs. Switches

Hubs

- Every station that is attached can see the traffic sent between all the other computers
- Use **CSMA/CD** to schedule transmission

Switches

- Traffic is forwarded only to the ports where it is destined.
- Multiple frames can be sent simultaneously by different stations
- Queueing: when multiple frames are sent to the same output port at the same time. Once the queue is full, packets will be dropped.



Source: hinditechy.com

- Modern Ethernet networks, built with switches and full-duplex, no longer need CSMA/CD.

5 Ports

Datagram Forwarding

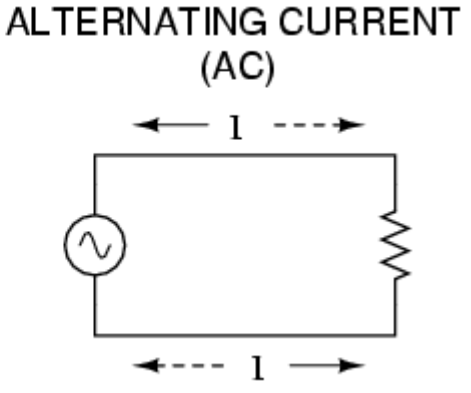
Header: destination address

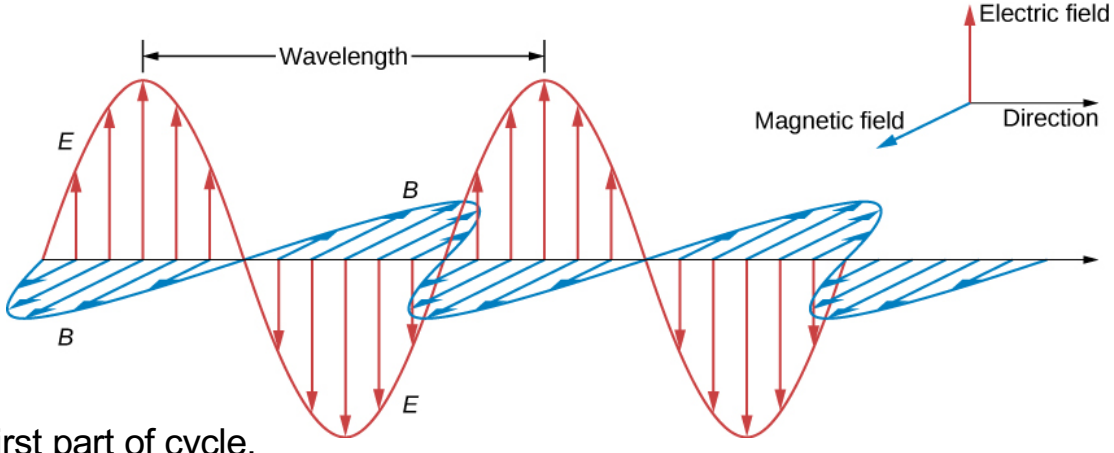
Forwarding Table

<destination, next_hop>

Radio Frequency (RF)

RF waves are electromagnetic waves generated when an alternating current goes through a conductive material.





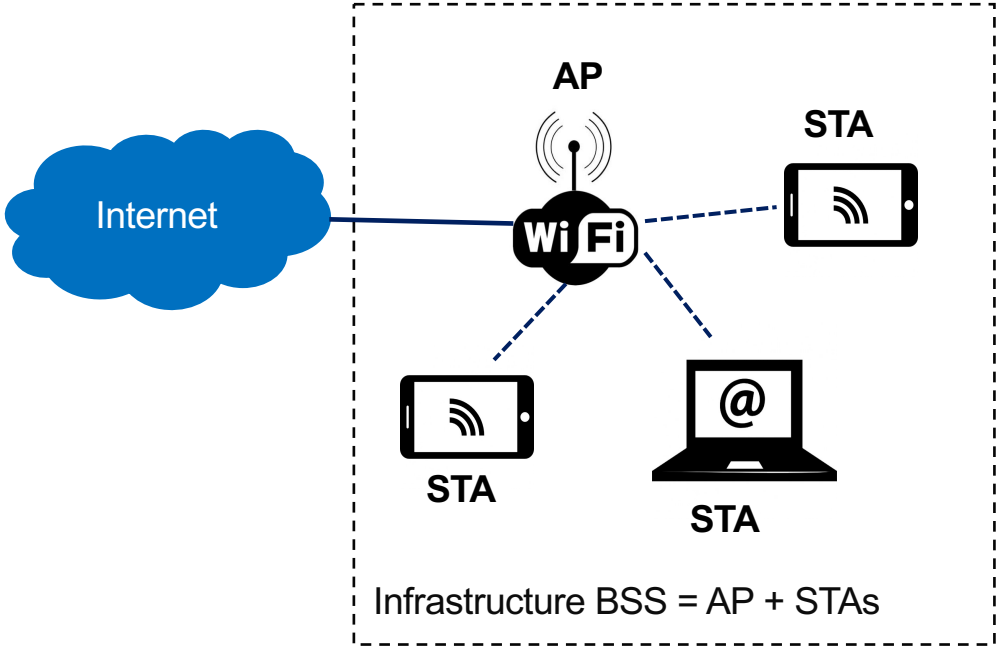
Current flows in one direction for first part of cycle, then the other direction for the second part of the cycle.

- **Frequency (Hz): how many cycles per second**
- **Wavelength**
- **Amplitude**
- RF waves travel at the speed of light in free space

  Speed of light (c) = frequency (f) x wavelength (λ)
- **Coverage area**: the area in which receiving stations can successfully receive and “understand” the signal
  - Depends on frequency and transmit power
- **RF spectrum**: the range of frequencies that are available

- **Path loss**
- **RF interference**
- **Measurement metrics**
  - Signal strength
  - Received signal strength indicator (RSSI)
  - Signal-to-noise-ratio (SNR)

Wi-Fi Architecture



- An AP has at least one antenna used for receiving and transmitting signals from and to clients
- AP converts modulated RF signals into Ethernet data, and vise versa (layer 2 translation between 802.11 and 802.3)
- An AP may have multiple MAC addresses. BSSID refers to the one of the radio interface the STA is currently connected to.

BSSID: BSS Identifier
SSID: name of the network

- **Beacons**
- **Passive scanning vs. Active scanning**
- **Association, reassociation, deassociation**

IPv4 Addresses

- **An IPv4 address consists of 4 octets (32 bits)**

  [octet] . [octet] . [octet] . [octet]

1000 1000.1110 0011.1110 1101.0110 1000 → 136.227.237.104

- Each number can be 0 to 255

#IPv4 addresses in the world?

IPv4 Subnet Mask

- An IP address is divided into two parts: network and host parts.
- Subnet mask is used for determining the network part and the host part of an IP address.
- **A subnet mask consists of 32 bits. The 1s in the subnet mask represent a network part, the 0s a host part.**
- A subnet mask must always be a series of 1s followed by a series of 0s. E.g. 255.255.0.0, 255.0.0.0.

Classless Inter-Domain Routing (CIDR)

- Introduced in 1993 by IETF (Internet Engineering Task Force)
- CIDR is based on **variable-length subnet masking (VLSM)** to allow allocation and routing based on arbitrary-length prefixes

Prefix

192.0.1.0/24 prefix length is 24.

32-24=8 bits are left for host addresses

10.0.0.0/8 32-8=24bits are left for host addresses

- **The number of usable IP addresses** can be calculated from the following formula:

2 to the power of host bits – 2

The first and the last address are the network address and the broadcast address, respectively. All other addresses inside the range could be assigned to Internet hosts.

If a company needs 12 public IP address, something like 190.5.1.1/k, what should be the value of k?

IPv6 addresses

- **IPv6 address consists of 16 octets (128 bits). IPv6 separates pairs of octets with a colon.**

  [octet] [octet] : [octet] [octet] :::::: [octet] [octet]

  For example, fedc:13:1654:310:fedc:bc37:61:3210

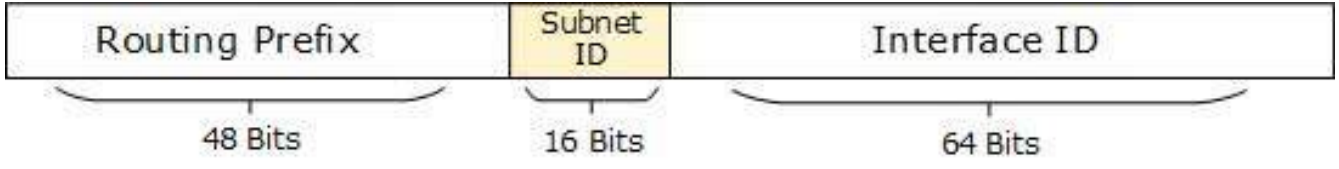
- If an address contains a long run of 0's, "::" should be used to represent many blocks of 0000
- It is possible to embed an IPv4 address in IPv6 address

  For example, ::ffff:147.126.65.141

First 80 0-bits

IPv6 Interface Identifier

- Most IPv6 addresses can be divided into a **64-bit network prefix** and a 64-bit “host” portion. These host-portion bits are known officially as the **interface identifier.**



Scope of IPv6 addresses

- The scope of a unicast address is either **global**, meaning it is intended to be globally routable, or **link-local**, meaning that it will only work with directly connected neighbors
- E.g. the loopback address **::1** **(127 0-bits followed by 1 1-bit)** is considered to have link-local scope
- **Link-local** addresses begin with the 64-bit link-local prefix consisting of the ten bits 1111 1110 10 followed by 54 more zero bits; that is, **fe80::/64**.

AnyCast

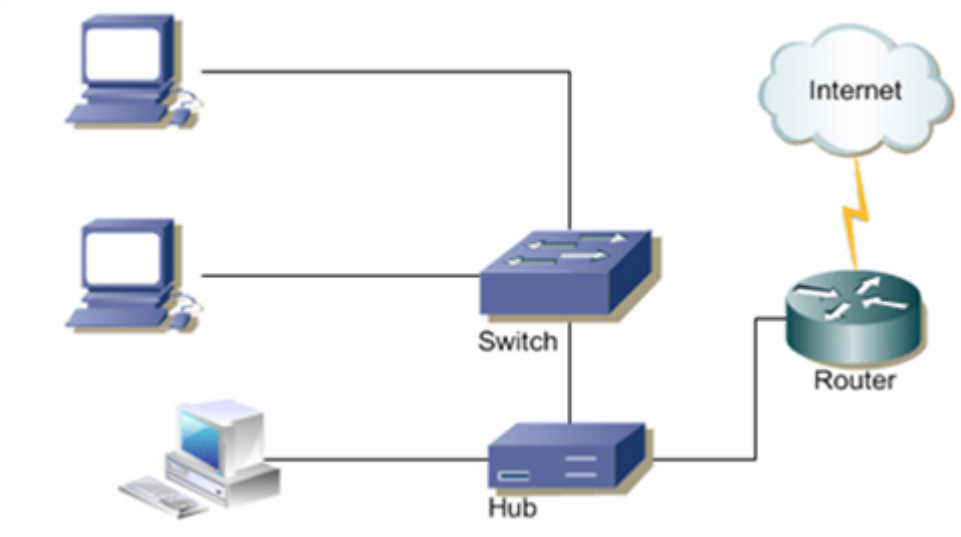
IPv6 has introduced Anycast mode of packet routing. In this mode, multiple interfaces over the Internet are assigned same Anycast IP address. Routers, while routing, send the packet to the nearest destination.

IPv6 does not support broadcast

Router vs. Switch

A router connects different networks like two LANs, two WAN's or LAN and WAN.

The main purpose of the router is to determine the smallest and best path for a packet to reach the destination.



Source: http://www.fiberopticshare.com/

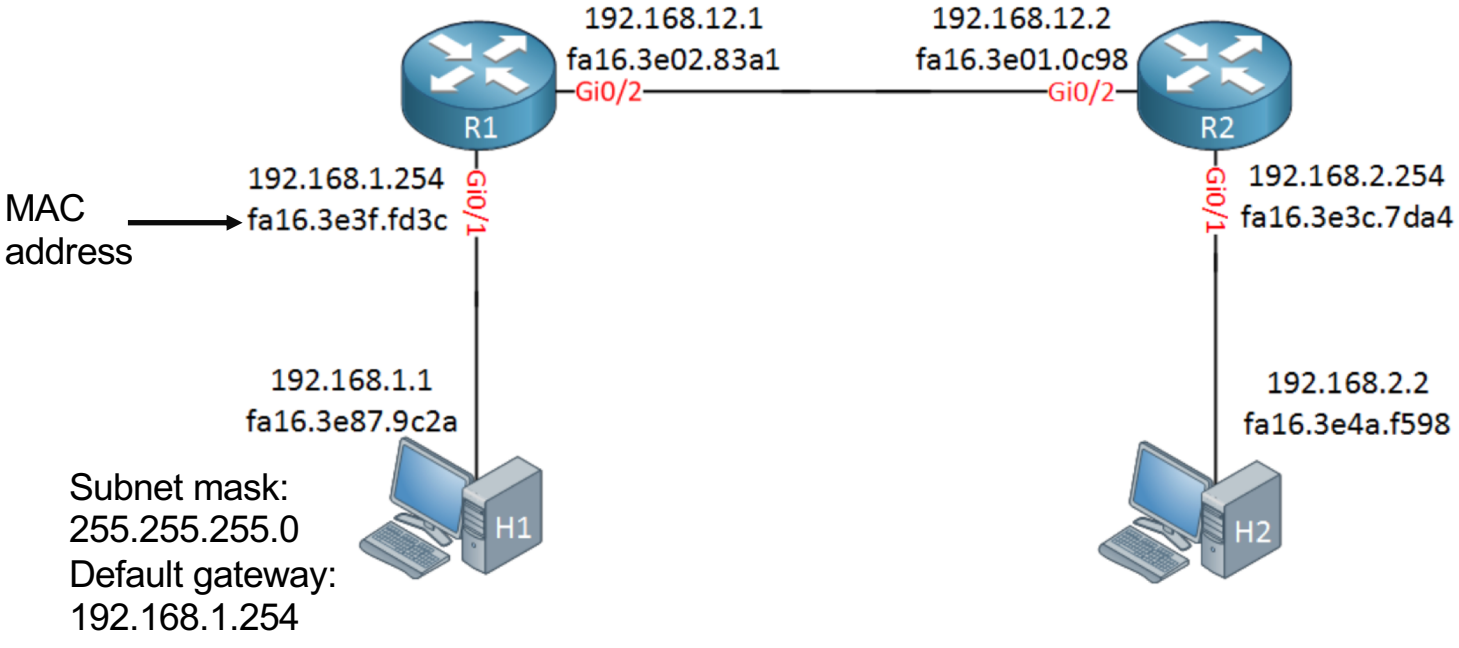
Routing Table

- Each router maintains a routing table and stores it in RAM.
- Each routing table consists of the following entries:
  - **network destination and subnet mask** – specifies a range of IP addresses.
  - **remote router** – IP address of the router used to reach that network.
  - **outgoing interface** – outgoing interface the packet should go out to reach the destination network.

All nodes on the internet have these routing tables, and this is how IP packets are routed to reach their destination.

Process of IP routing

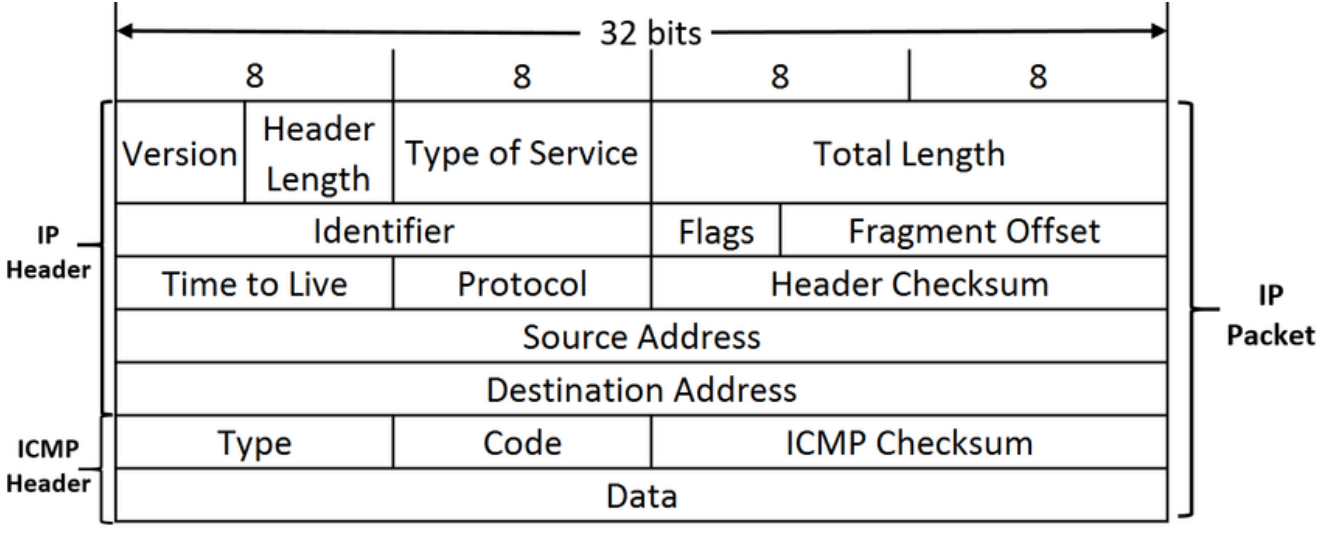
Example



Source: https://networklessons.com/cisco/ccna-routing-switching-icnd1-100-105/ip-routing-explained

Internet Control Message Protocol (ICMP)

- Used by network tools like Ping and Traceroute
- Host-to-host protocol, used for sending IP layer error and status messages
- Part of IP layer



IP

IP provides unreliable service. It makes its best effort to deliver segments between communicating hosts, but it makes no guarantees.

- It does not guarantee segment delivery
- It does not guarantee orderly delivery of segments
- It does not guarantee the integrity of the data in the segments

TCP/UDP Ports

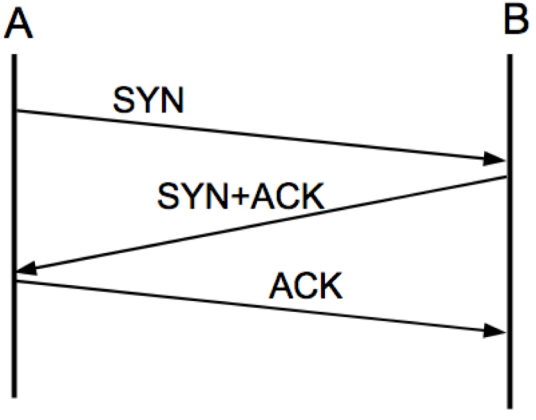
- TCP/UDP extends host-to-host delivery to process-to-process delivery
- Port <-> server process
- Ports are identified for each protocol and address combination by 16-bit unsigned numbers, known as port number.
- One IP address → 65535 TCP Ports and another 65535 UDP Ports
- Socket: <host, port>
- Socket address: <IP address, port number>, e.g. 192.168.0.10:80
- Client must know server's port

TCP	UDP
Connection-oriented	Connectionless
Reliable (Guaranteed delivery)	Best-effort delivery

Connection-oriented

- Before one application process can begin to send data to another, the two processes must "handshake" with each other
- The process that is initiating the connection is called the *client process*, while the other process is called the *server process*
- **Three-way Handshake**

  1) A sends B a packet with the SYN bit set
  2) B responds with a SYN packet of its own; the ACK bit is also set.
  3) A responds to B's SYN with its own ACK.



TCP three-way handshake

Sequence and Acknowledgement Numbers

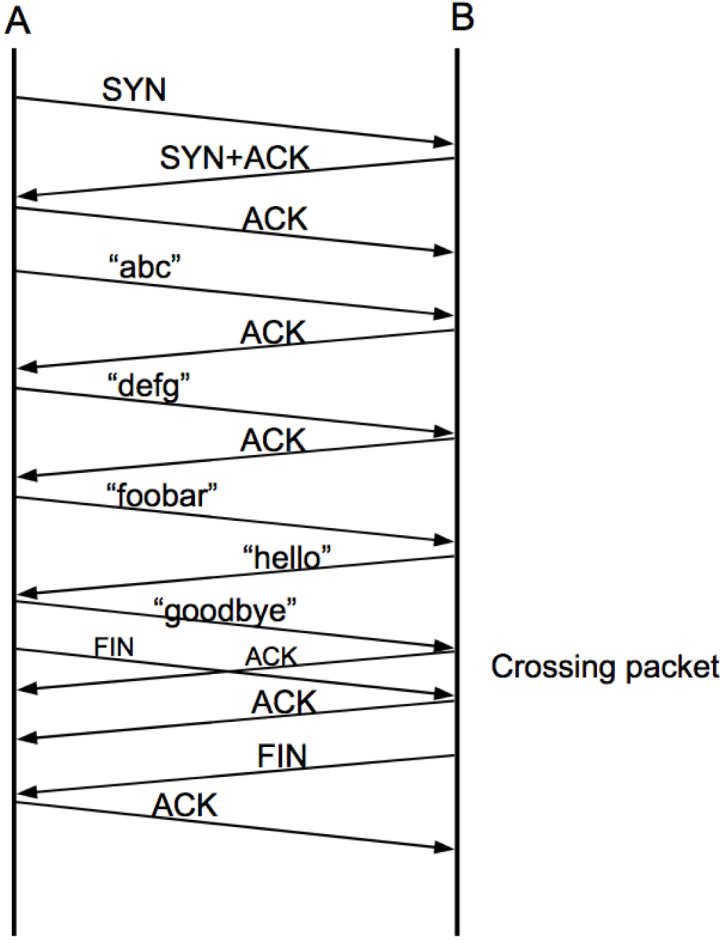
- **Numbering the data at the byte level**
- **Initial Sequence Number (ISN) is fixed for the lifetime of the connection. Each direction of a connection has its own ISN.**
- **The value of the Sequence Number, in relative terms, is the *position of the first byte of the packet in the data stream*, or the position of what would be the first byte in the case that no data was sent**
- **The value of the Acknowledgement Number, in relative terms, *represents the byte position for the next byte expected***
- **The sequence and acknowledgment numbers, as sent, represent these relative values *plus ISN***

Exercise

In terms of the sequence and acknowledgment numbers, SYNs count as 1 byte, as do FINs.

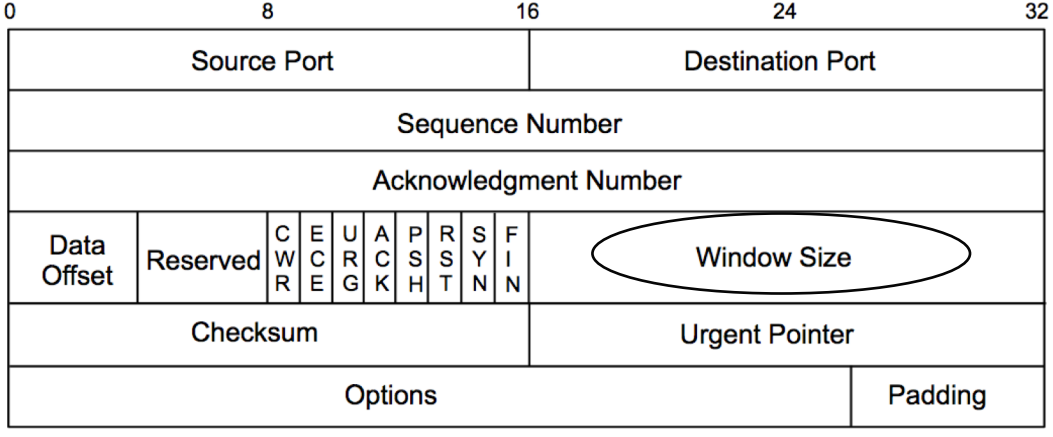
Assume that ISN = 0 on both sides

Can you calculate the seq and ack values of each segment?



TCP Flow Control

- **Window size indicates the number of bytes that a receiver is willing to accept.**
- Flow-control service is used for eliminating the possibility of the sender overflowing the receiver's buffer



TCP Congestion Control

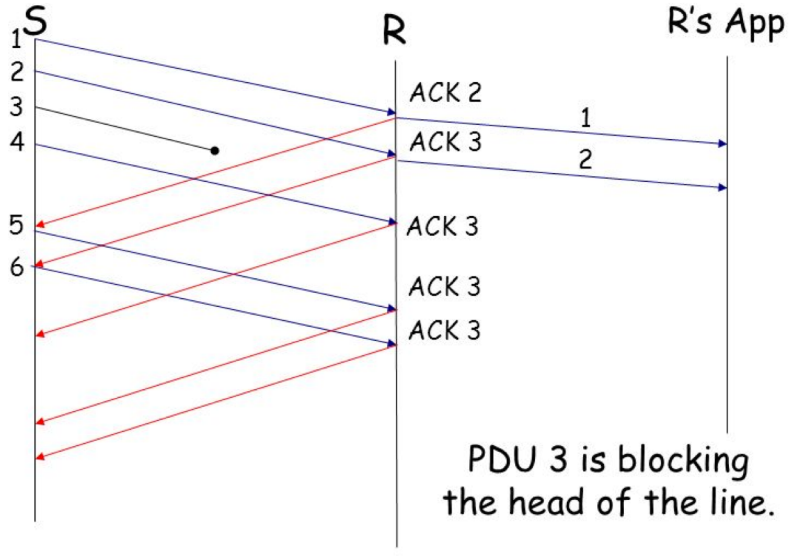
1) How does a TCP sender limit the rate at which it sends traffic into its connection?
2) How does a TCP sender perceive that there is congestion on the path between itself and the destination?
3) What algorithm should the sender use to change its send rate as a function of perceived end-to-end congestion?

TCP Congestion-Control Algorithm

Three Major Components:

- **Slow Start**
- **Congestion Avoidance**
- **Fast recovery (recommended for TCP senders, but not required)**

Head-of-Line Blocking in TCP



Source: https://www.slideshare.net/shuyaosaki/introduction-to-quic

REST (Representational State Transfer)

- **REST is an architectural style**
- **Six Constraints:**
  - Client-server
  - Uniform interface
  - Stateless
  - Layered System
  - Cacheable
  - Code on Demand (optional)

Roy Fielding's doctoral dissertation (2000)

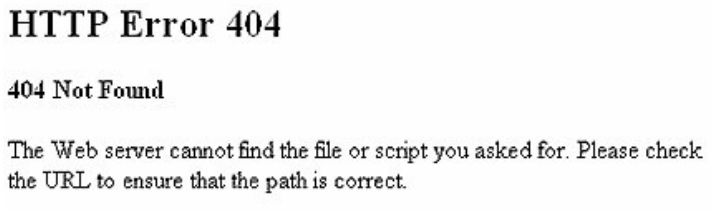
http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm

HTTP Request Methods

Method	Description
GET	Transfer a current representation of the target resource
HEAD	Same as GET, but only transfer the status line and header section
POST	Perform resource-specific processing on the request payload
PUT	Replace all current representations of the target resource with the request payload
DELETE	Remove all current representations of the target resource
CONNECT	Establish a tunnel to the server identified by the target resource
OPTIONS	Describe the communication options for the target resource
TRACE	Perform a message loop-back test along the path to the target resource

HTTP Response

- **The first line of the response is called the *status* line and has a numeric status code and a text-based reason phrase**

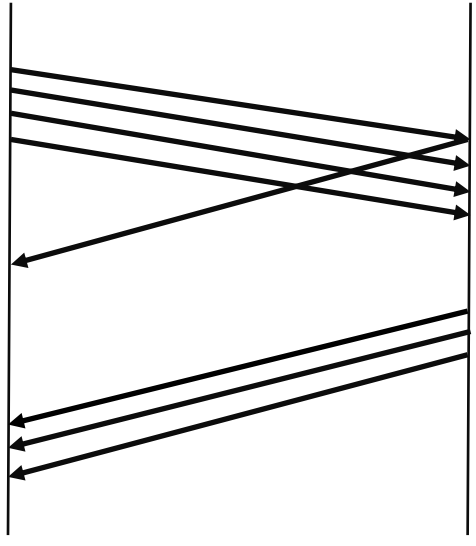


- **HTTP status codes are primarily divided into five groups:**

Informational 1XX, Successful 2XX, Redirection 3XX, Client Error 4XX, and Server Error 5XX

Head of Line Blocking

- **First-in-first-out:** The server must send its responses in the same order that the requests were received.

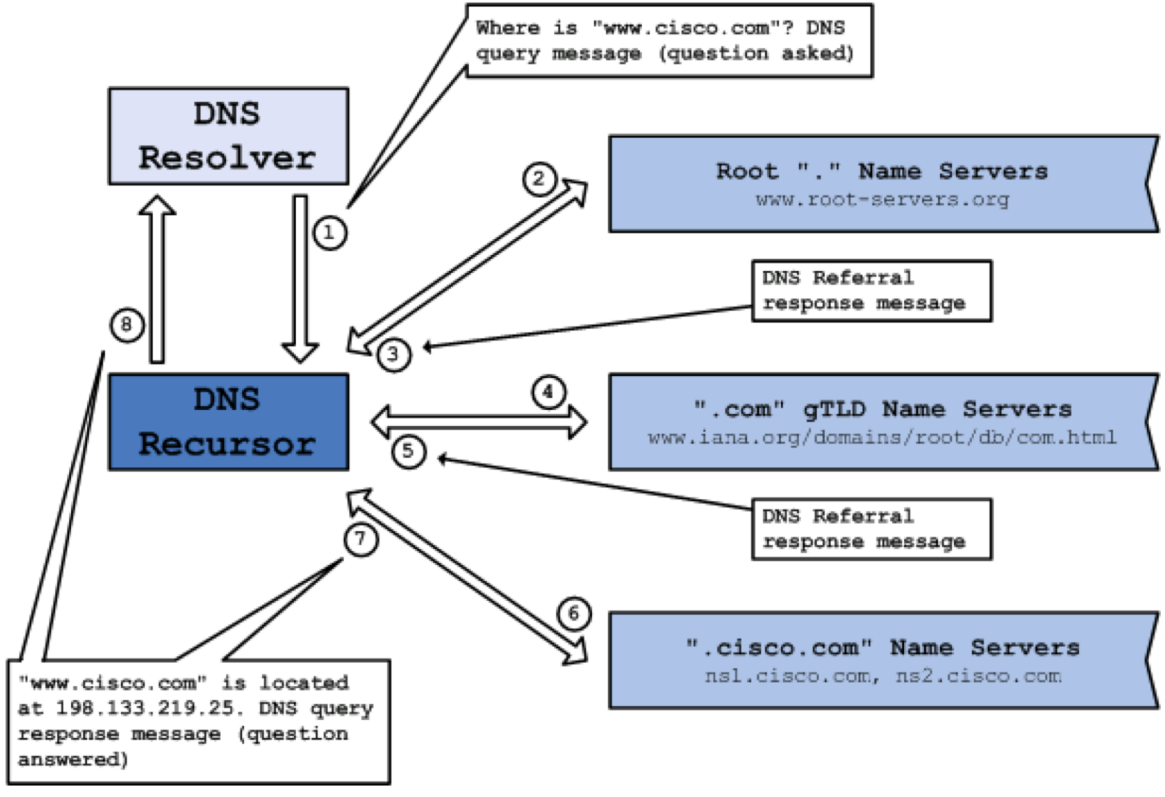


A large or slow response can block others behind it

Domain Name System (DNS)

- **Application layer protocol, responsible for mapping domain names into IP addresses**
- DNS protocol relies on UDP by default, but can also work over TCP as a fallback when firewalls block UDP.
- **DNS is a distributed database**
- **Resolver:** A DNS client that sends DNS messages to obtain information about the requested domain name space
- **A DNS name server** is a server that stores the DNS records for a domain; A DNS name server responds with answers to queries against its database.

Recursive Query



- **Authoritative Server:** A DNS server that responds to query messages with information stored in resource records for a domain name space stored on the server.
- **Recursive Resolver:** A DNS server that recursively queries for the information asked in the DNS query.

gTLD: generic top-level domain
Source: Cisco

Mid-term Exam

Aalto University
School of Electrical
Engineering

- **2 hours**
- **Max 16 points**
- **multi-choice questions. No essays.**
- **Grading of multi-choice questions:**
  - **Right minus wrong** - Users receive points equal to the number of right answers they choose minus the number of incorrect answers they choose. Users can receive a minimum of zero on a question; they cannot receive a negative mark.
    To calculate how much each answer is worth, the system takes the total number of points assigned to the question and divides it by the total number of answer choices.