南京大學
NANJING UNIVERSITY

Computer Networks

Wenzhong Li, Chen Tian

Nanjing University

Material with thanks to James F. Kurose, Mosharaf Chowdhury, and other colleagues.

Chapter 2. Link Layer

- Link Layer Service
  - Framing
  - Link access
  - Reliable delivery
  - Error detection and correction
- Local Area Network (LAN)
  - Token Ring
  - Ethernet
- Medium access control (Cont.)
- Bridges and Layer-2 switch
- Wireless Networks

Performance of MAC

Performance Metric

- Media Utilization
  - Time used for frame transmission vs. time the shared media is occupied

$$U = \frac{\text{Time for frame transmission}}{\text{total time for a frame}}$$

- Relative Propagation Time

$$a = \frac{\text{propagation time}}{\text{transmission time}} \quad or$$

$$a = \frac{\text{length of the data path (in bits)}}{\text{length of a standard frame (in bits)}}$$

Different Networks

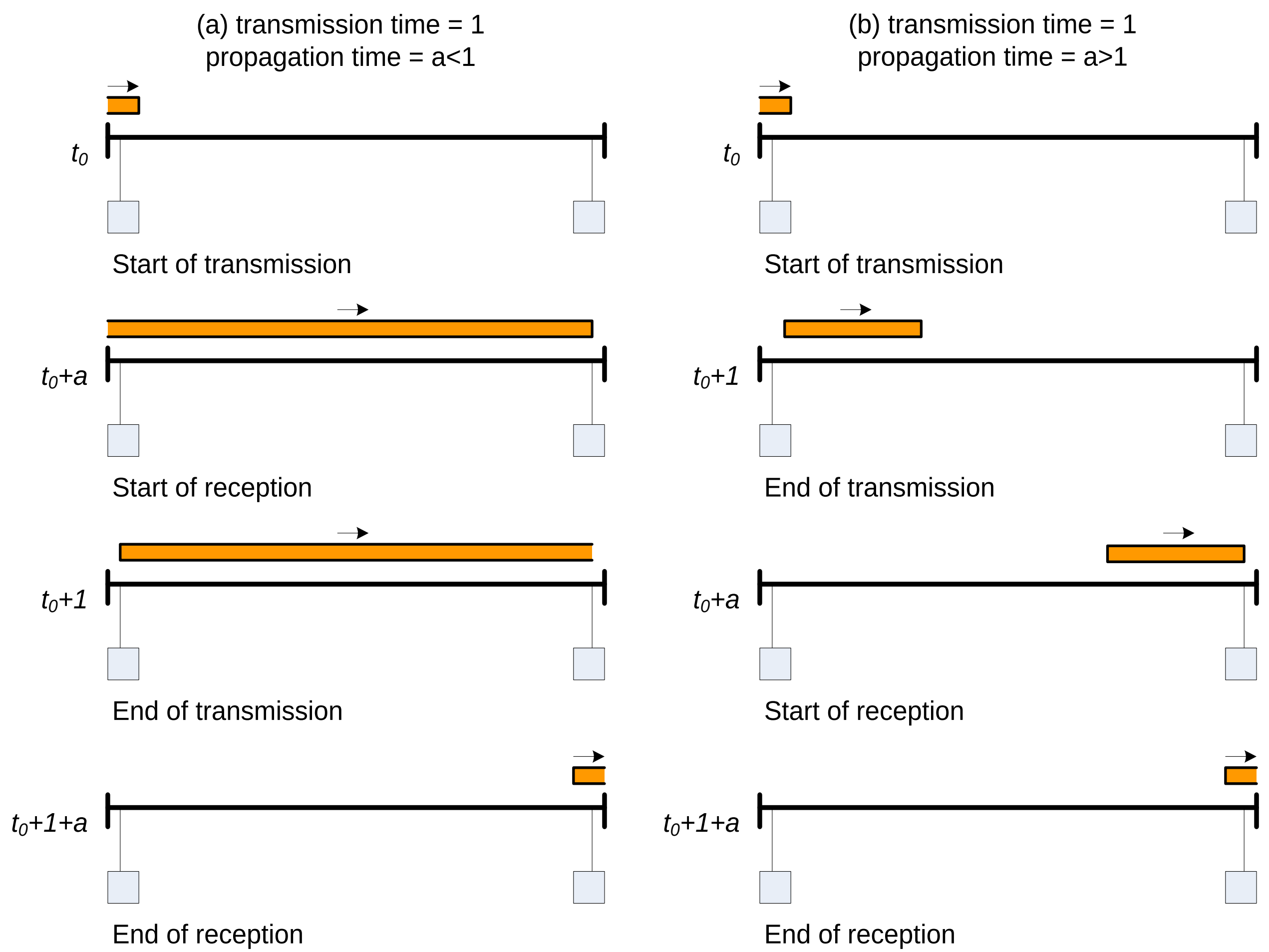
- Contention free
  - Point-to-Point Link
  - Ring LAN
- Random access
  - ALOHA, slotted ALOHA
  - CSMA/CD

Point-to-Point Link with No ACK

$$U = \frac{\text{Time for frame transmission}}{\text{total time for a frame}}$$

Large frame

Small frame



Define

1: normalized frame transmission time

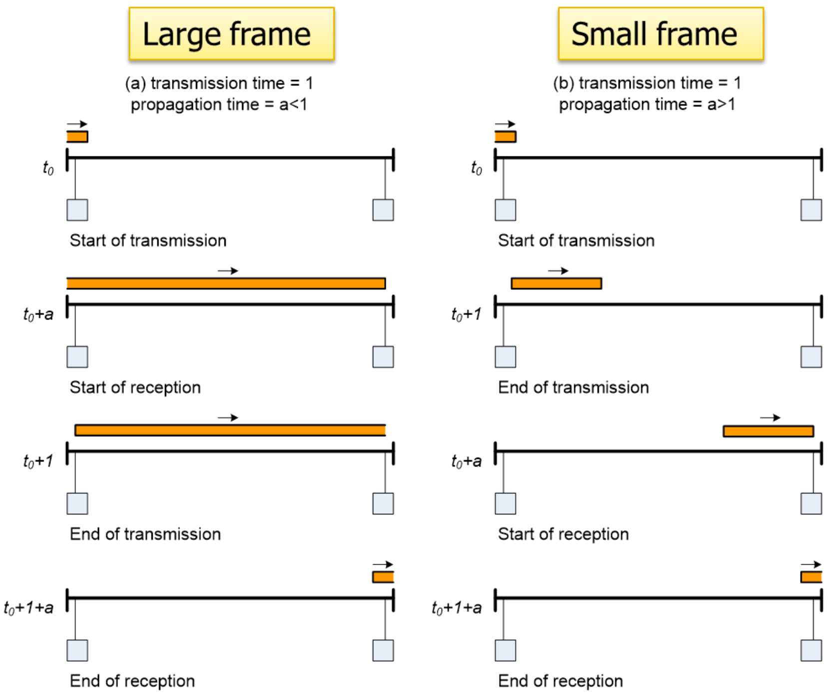
a: end to end propagation delay

N: number of stations

Q:
Max Utilization U=?

Max Utilization for Point-to-Point Link

- Parameters and assumptions
  - 1: normalized frame transmission time
  - a: end to end propagation delay
  - N: number of stations
- Each station has frames to transmit
- Total frame time=transmission delay + propagation delay: $1+a$
- Max Utilization:



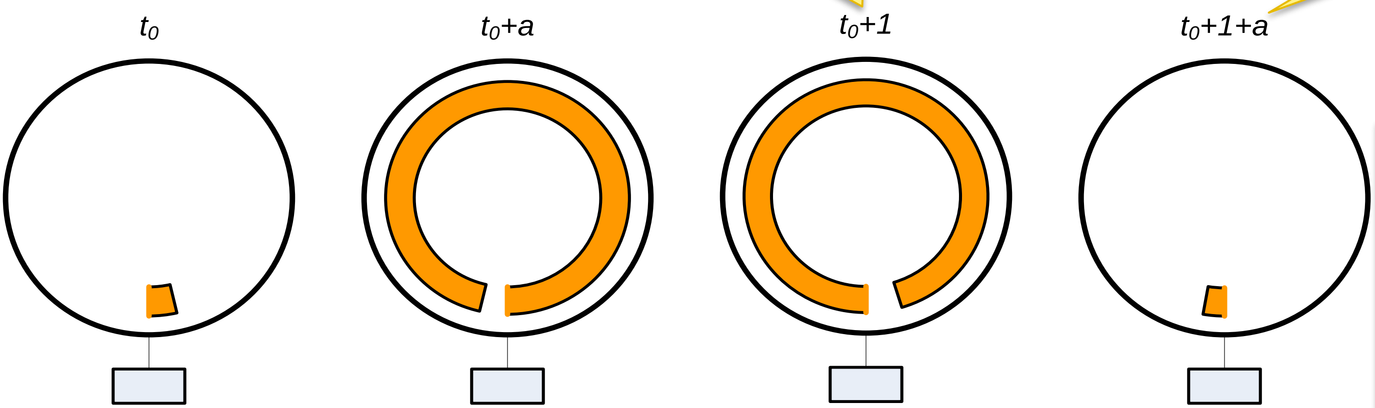
$$U = \frac{1}{1+a}$$

$$U = \frac{\text{Time for frame transmission}}{\text{total time for a frame}}$$

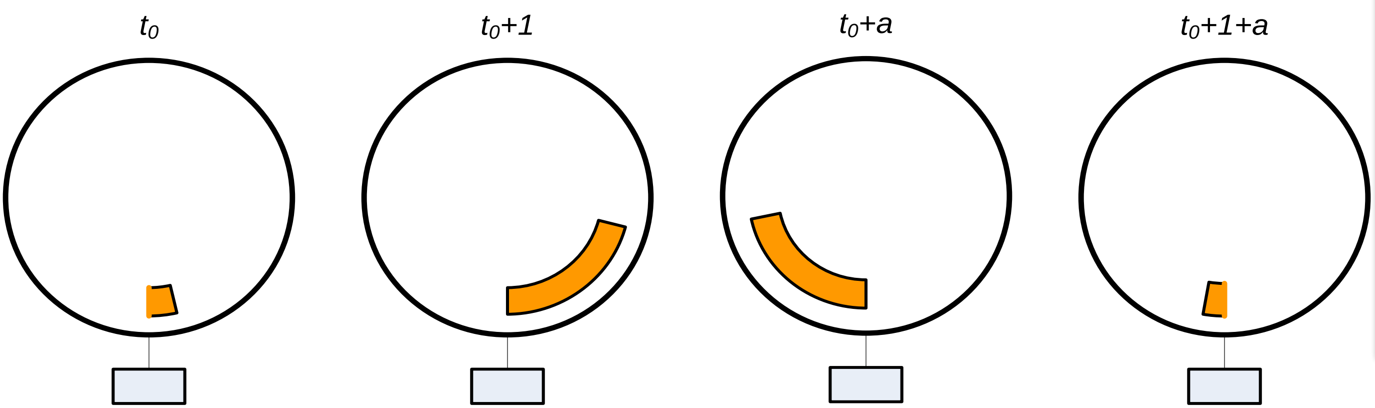
Ring

Token is released at t_0+1, and it will arrive the next station at t_0+1+a/N (next transmission starts).

End of the previous transmission at t_0+1+a.



(a) transmission time = 1, propagation time = a<1



(b) transmission time = 1, propagation time = a>1

Token is released at t_0+a, and it will arrive the next station at t_0+a+a/N (next transmission starts).

End of the previous transmission at t_0+1+a.

Define:

T_1: Average time to transmit a frame, i.e. $T_1 = 1$

T_2: Average time to pass the token after frame transmission

N: number of stations

Q:

Max Utilization: $U = ?$

Max Utilization for Ring LAN

- Define
  - T_1: Average time to transmit a frame, i.e. $T_1 = 1$
  - T_2: Average time to pass the token after frame transmission
- Max Utilization: $U = T_1/(T_1+T_2)$

2 cases

- Case 1: $a<1$ (frame longer than ring)
  - T_2 = time to pass token to the next station = a/N
- Case 2: $a>1$ (frame shorter than ring)
  - T_2 = sender wait for frame returns after transmission = $a-1+a/N$

$$U = \begin{cases} \dfrac{1}{1+a/N} & a<1 \\ \dfrac{1}{a+a/N} & a>1 \end{cases}$$

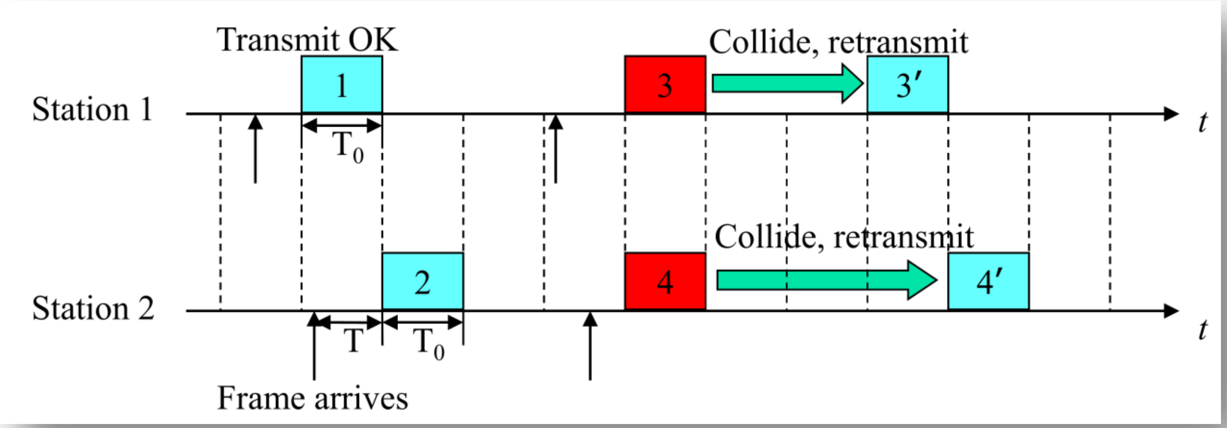
Slotted ALOHA

- All frames have same size
- Time is uniformly slotted
- Nodes are synchronized
- Transmission begins at slot boundary
- Frames either miss or overlap totally

Operation:

- N nodes with many frames to send
- Each transmits in each slot with probability p until success

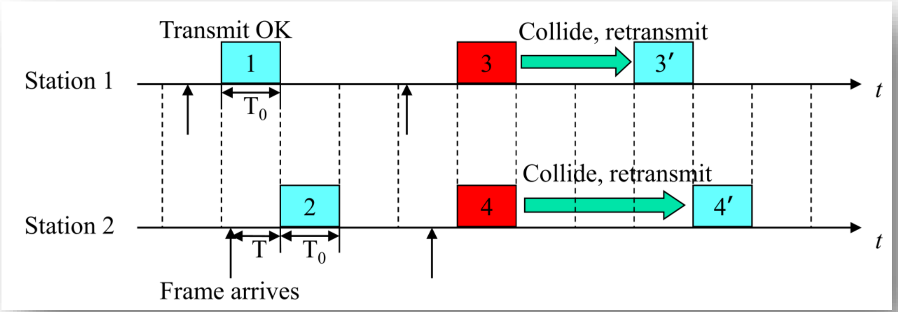
- Q:
- Max Utilization: $U = ?$



Slotted ALOHA

- Suppose:
  - N nodes with many frames to send, each transmits in slot with probability p
- Probability of successful transmission
  - One node has success in a slot $= p(1-p)^{N-1}$
  - Any node has a success $A = Np(1-p)^{N-1}$
- Maximize value of A (let A'(p)=0)

$$p = \frac{1}{N} \Longrightarrow A = \left(1 - \frac{1}{N}\right)^{N-1}$$



Slotted ALOHA Efficiency

- Utilization if a slot is successfully used

$$U_s = \frac{1}{1+2a} \approx 1 \quad (a \ll 1)$$

Before data transmission, it takes a to detect collision;
After transmission, it takes a to make sure the transmission of the last bit

- Since A is the rate of success slot

efficiency: long-run fraction of successful slots (many nodes, all with many frames to send)

$$U = U_s \times A \approx \left(1 - \frac{1}{N}\right)^{N-1}$$

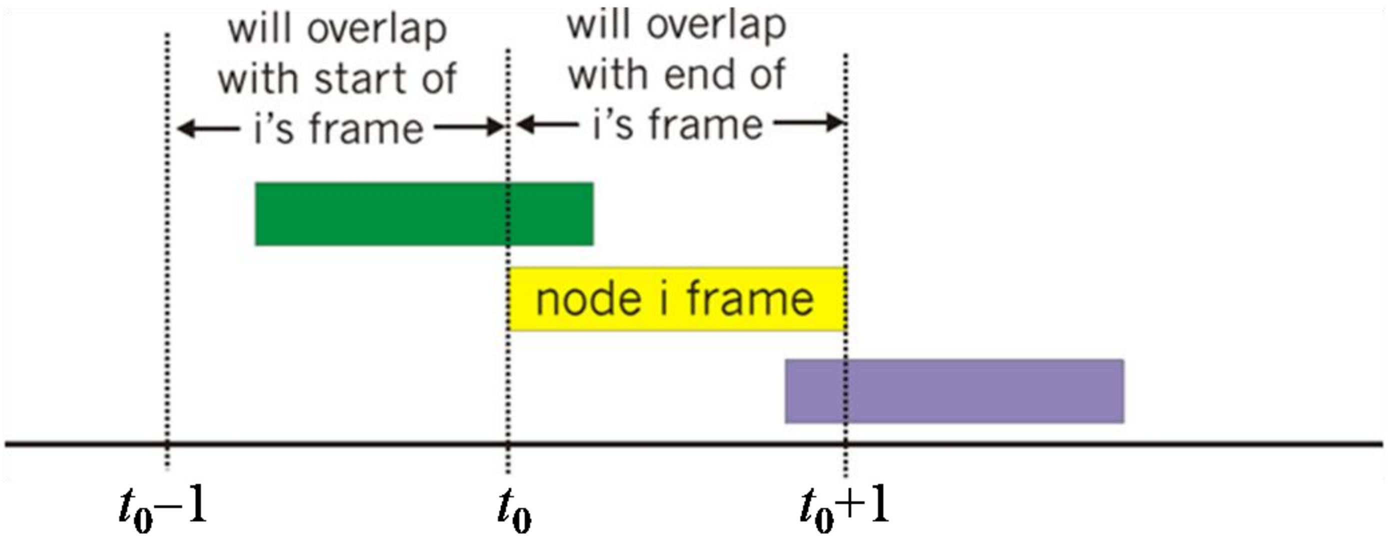
$$\text{Let } N \to \infty$$

$$U \approx e^{-1} = 0.367879$$

$$\lim_{x\to\infty} (1 - \frac{1}{x})^x = e^{-1}$$

Pure ALOHA

- Simpler but collision probability increases
  - Frame sent at t_0 collides with other frames sent in $[t_0-1, t_0+1]$



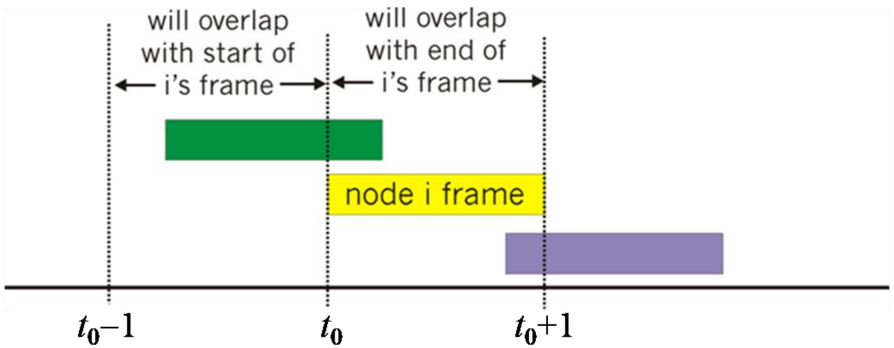
Suppose:

N nodes with many frames to send, each transmits in any time with probability p

Q: Max Utilization: U =?

Pure ALOHA Efficiency

- Probability of successful transmission



$$A = N \cdot P(\text{one transmits in the slot}) \cdot$$
$$P(\text{no other node transmits in } [t_{0-1}, t_0] \cdot$$
$$P(\text{no other node transmits in } [t_0, t_{0+1}]$$

$$\mathrm{U} \approx \mathrm{A} = \mathrm{Np} \cdot (1-p)^{2(N-1)}$$
$$\approx \frac{1}{2}\left(1 - \frac{1}{2N}\right)^{2(N-1)} \quad \left(p = \frac{1}{2N}\right)$$
$$\approx \frac{1}{2e} = 0.183940 \quad (N \to \infty)$$

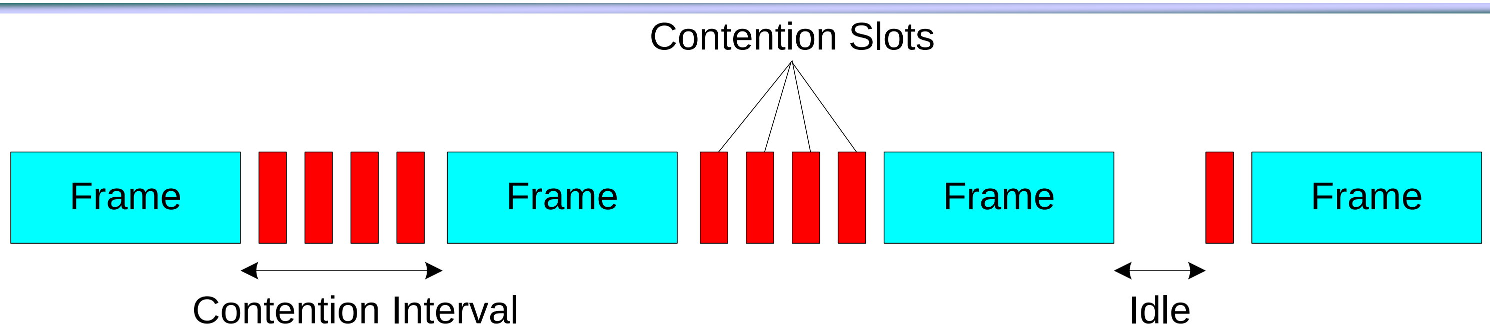
CSMA/CD

- With CSMA, collision occupies medium for duration of transmission
  - Colliding transmissions aborted once detected

- Stations listen whilst transmitting
  1. If medium idle, transmit; otherwise, step 2
  2. If busy, listen for idle, then transmit immediately

  3. If collision detected, send jam signal then abort
  4. After jam, wait random time then start from step 1

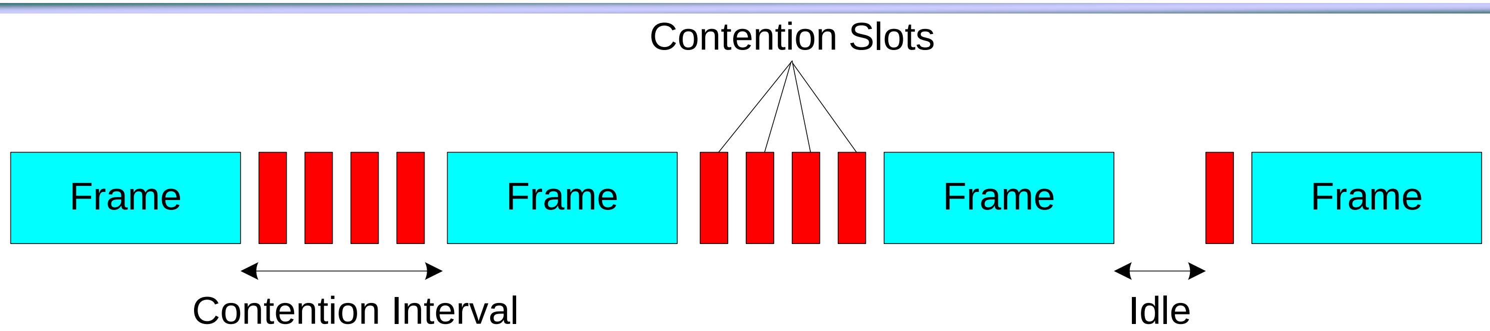
CSMA/CD (p-persistent): Max Utilization



- Contention slots end in a collision
- Contention interval is a sequence of contention slots
  - Length of a slot in contention interval is $2a$

    (in worst case it takes $2a$ time to detect contention)
- Assume p-persistent:
  - The probability that a station attempts to transmit in a slot is p
- Q: Max Utilization: $U = ?$

Max Utilization for CSMA/CD (1)



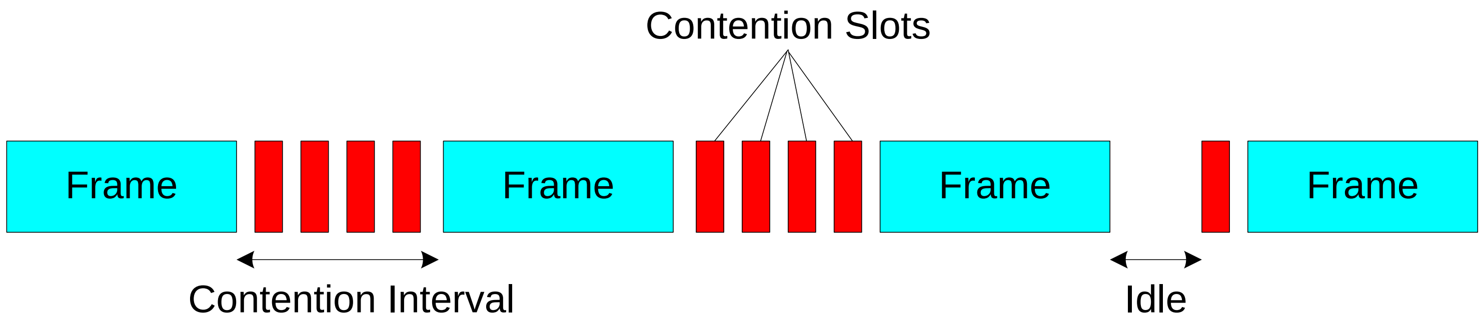
- Let A be the probability that some station can successfully transmit in a slot, then:

$$A = \binom{N}{1} p^1 (1-p)^{N-1} = Np(1-p)^{N-1}$$

- In above formula, A is maximized when $p=1/N$, thus:

$$A = \left(1 - \frac{1}{N}\right)^{N-1}$$

Max Utilization for CSMA/CD (2)



- Probability of a contention interval with j slots

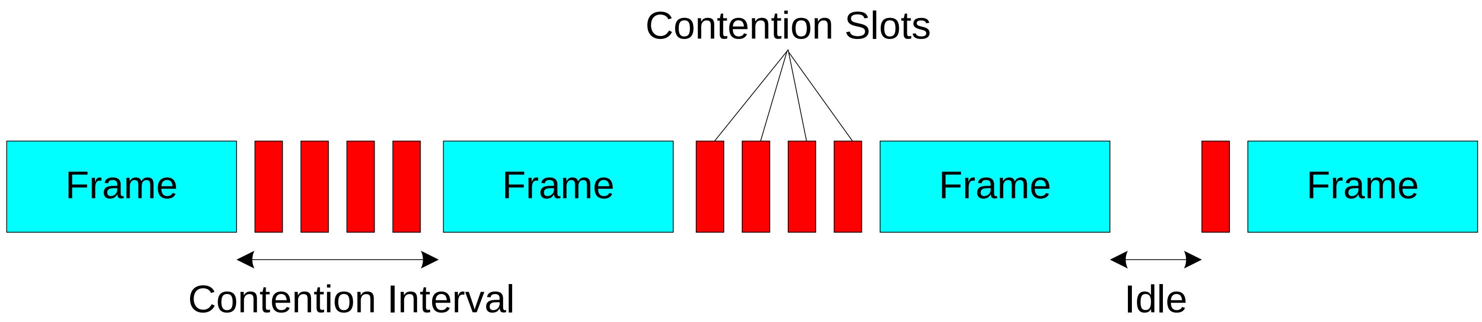
$$Prob[\,j \text{ unsuccessful attempts}] \times Prob[1 \text{ successful attempt}] =$$

$$A(1-A)^j$$

- The expected number of slots in a contention interval is then calculated as (Geometric distribution, mean=(1-p)/p):

$$\sum_{j=1}^{\infty} jA(1-A)^j = \frac{1-A}{A}$$

Max Utilization for CSMA/CD (3)



- Maximum Utilization

$$U = \frac{\text{Frame time}}{\text{Frame time+Propagation time+Average contention interval}}$$

$$= \frac{1}{1 + a + 2a\dfrac{1-A}{A}} = \frac{1}{1 + \dfrac{2-A}{A}a}$$

- Let $N \to \infty$, $A = (1-1/N)^{N-1} = 1/e$ (e=2.718)

$$U = \frac{1}{1 + \dfrac{2-A}{A}a} = \frac{1}{1 + (2e-1)a} \approx \frac{1}{1 + 4.44a}$$

MAC Address and Discovery

What is MAC Address?

- **Medium Access Control (MAC) Address**
  - Numerical address associated with a network adapter
  - Flat name space of 48 bits (e.g., 00-15-C5-49-04-A9 in HEX)
  - Unique, hard-coded in the adapter when it is built
- Hierarchical Allocation
  - Blocks: assigned to vendors (e.g., Dell) by the IEEE
    - First 24 bits (e.g., 00-15-C5-**-**-**)
  - Adapter: assigned by the vendor from its block
    - Last 24 bits

Address Configuration

```
C:\WINDOWS\System32\cmd.exe

C:\Documents and Settings\X129>ipconfig /all

Windows IP Configuration

        Host Name . . . . . . . . . . . . : X129-T
        Primary Dns Suffix  . . . . . . . : seven.parkland.cc.il.us
        Node Type . . . . . . . . . . . . : Hybrid
        IP Routing Enabled. . . . . . . . : No
        WINS Proxy Enabled. . . . . . . . : No
        DNS Suffix Search List. . . . . . : seven.parkland.cc.il.us
                                            parkland.cc.il.us
                                            cc.il.us
                                            il.us

Ethernet adapter Local Area Connection:

        Connection-specific DNS Suffix  . : parkland.edu
        Description . . . . . . . . . . . : Intel(R) PRO/100 VE Network Connection
        Physical Address. . . . . . . . . : 00-07-E9-6F-37-53
        Dhcp Enabled. . . . . . . . . . . : Yes
        Autoconfiguration Enabled . . . . : Yes
        IP Address. . . . . . . . . . . . : 10.10.2.111
        Subnet Mask . . . . . . . . . . . : 255.255.0.0
        Default Gateway . . . . . . . . . : 10.10.1.1
        DHCP Server . . . . . . . . . . . : 216.125.249.50
        DNS Servers . . . . . . . . . . . : 216.125.249.48
                                            206.166.53.20
                                            206.166.49.21
        Primary WINS Server . . . . . . . : 216.125.249.50
        Secondary WINS Server . . . . . . : 216.125.249.51
        Lease Obtained. . . . . . . . . . : Wednesday, October 22, 2003 10:29:37 PM
        Lease Expires . . . . . . . . . . : Thursday, November 06, 2003 10:29:37 PM

C:\Documents and Settings\X129>_
```

MAC Address: only can be seen in the same subnet

IP Address

Subnet mask: define the range of a subnet

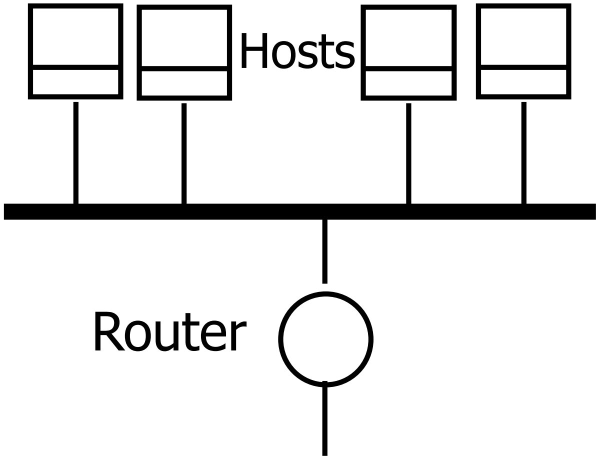
Default Gateway: all packets to the IPs in the same subnet will be broadcasted; all packets to the other IPs will be sent to the default gateway (if no other route rule is given)

Discovery

- A host is "born" knowing only its MAC address
- Must discover lots of information before it can communicate with a remote host B
  - What is my IP address?
  - What is B's IP address? (remote)
  - What is B's MAC address? (if B is local)
  - What is my first-hop router's address? (if B is not local)
  - …

ARP and DHCP

- Link layer discovery protocols
  - ARP → Address Resolution Protocol
  - DHCP → Dynamic Host Configuration Protocol
  - Confined to a single local-area network (LAN)
  - Rely on broadcast capability



ARP and DHCP

- Link layer discovery protocols
- Serve two functions
  - Discovery of local end-hosts
    - For communication between hosts on the same LAN
  - Bootstrap communication with remote hosts
    - What's my IP address?
    - Who/where is my local DNS server?
    - Who/where is my first hop router?

Address Configuration

```
C:\WINDOWS\System32\cmd.exe

C:\Documents and Settings\X129>ipconfig /all

Windows IP Configuration

        Host Name . . . . . . . . . . . . : X129-T
        Primary Dns Suffix  . . . . . . . : seven.parkland.cc.il.us
        Node Type . . . . . . . . . . . . : Hybrid
        IP Routing Enabled. . . . . . . . : No
        WINS Proxy Enabled. . . . . . . . : No
        DNS Suffix Search List. . . . . . : seven.parkland.cc.il.us
                                            parkland.cc.il.us
                                            cc.il.us
                                            il.us

Ethernet adapter Local Area Connection:

        Connection-specific DNS Suffix  . : parkland.edu
        Description . . . . . . . . . . . : Intel(R) PRO/100 VE Network Connection
        Physical Address. . . . . . . . . : 00-07-E9-6F-37-53
        Dhcp Enabled. . . . . . . . . . . : Yes
        Autoconfiguration Enabled . . . . : Yes
        IP Address. . . . . . . . . . . . : 10.10.2.111
        Subnet Mask . . . . . . . . . . . : 255.255.0.0
        Default Gateway . . . . . . . . . : 10.10.1.1
        DHCP Server . . . . . . . . . . . : 216.125.249.50
        DNS Servers . . . . . . . . . . . : 216.125.249.48
                                            206.166.53.20
                                            206.166.49.21
        Primary WINS Server . . . . . . . : 216.125.249.50
        Secondary WINS Server . . . . . . : 216.125.249.51
        Lease Obtained. . . . . . . . . . : Wednesday, October 22, 2003 10:29:37 PM
        Lease Expires . . . . . . . . . . : Thursday, November 06, 2003 10:29:37 PM

C:\Documents and Settings\X129>_
```

MAC Address: only can be seen in the same subnet

IP Address

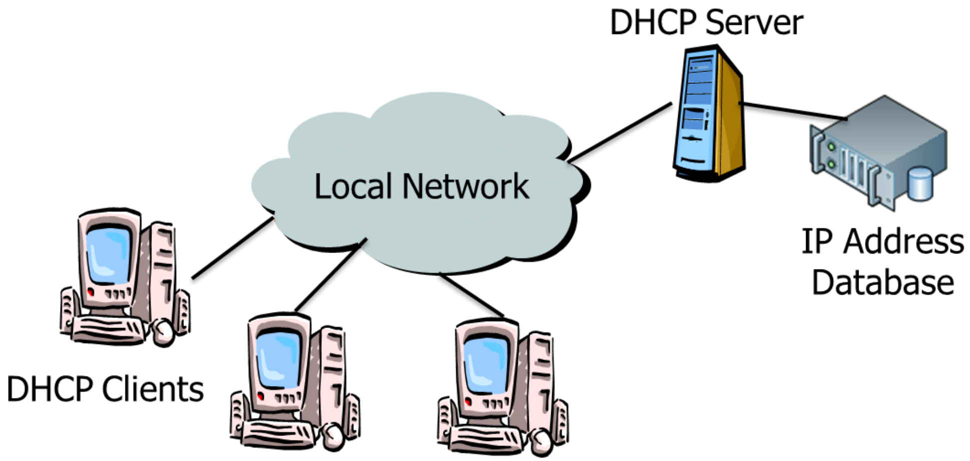
Subnet mask: define the range of a subnet

Default Gateway

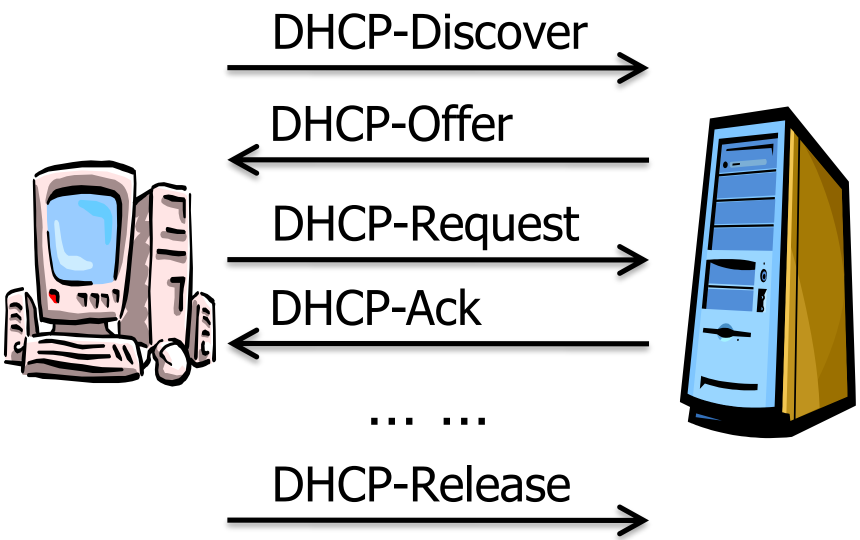
DHCP Automatic address configuration

DHCP

- Dynamic Host Configuration Protocol
  - Defined in RFC 2131
- A host uses DHCP to discover
  - Its own IP address
  - Its netmask
  - IP address(es) for its local DNS name server(s)
  - IP address(es) for its first-hop “default” router(s)

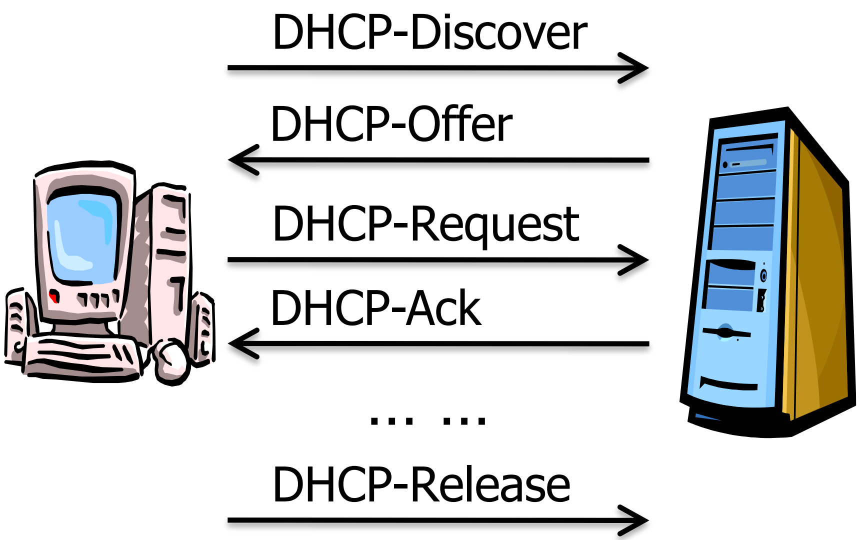


Typical Procedure of DHCP



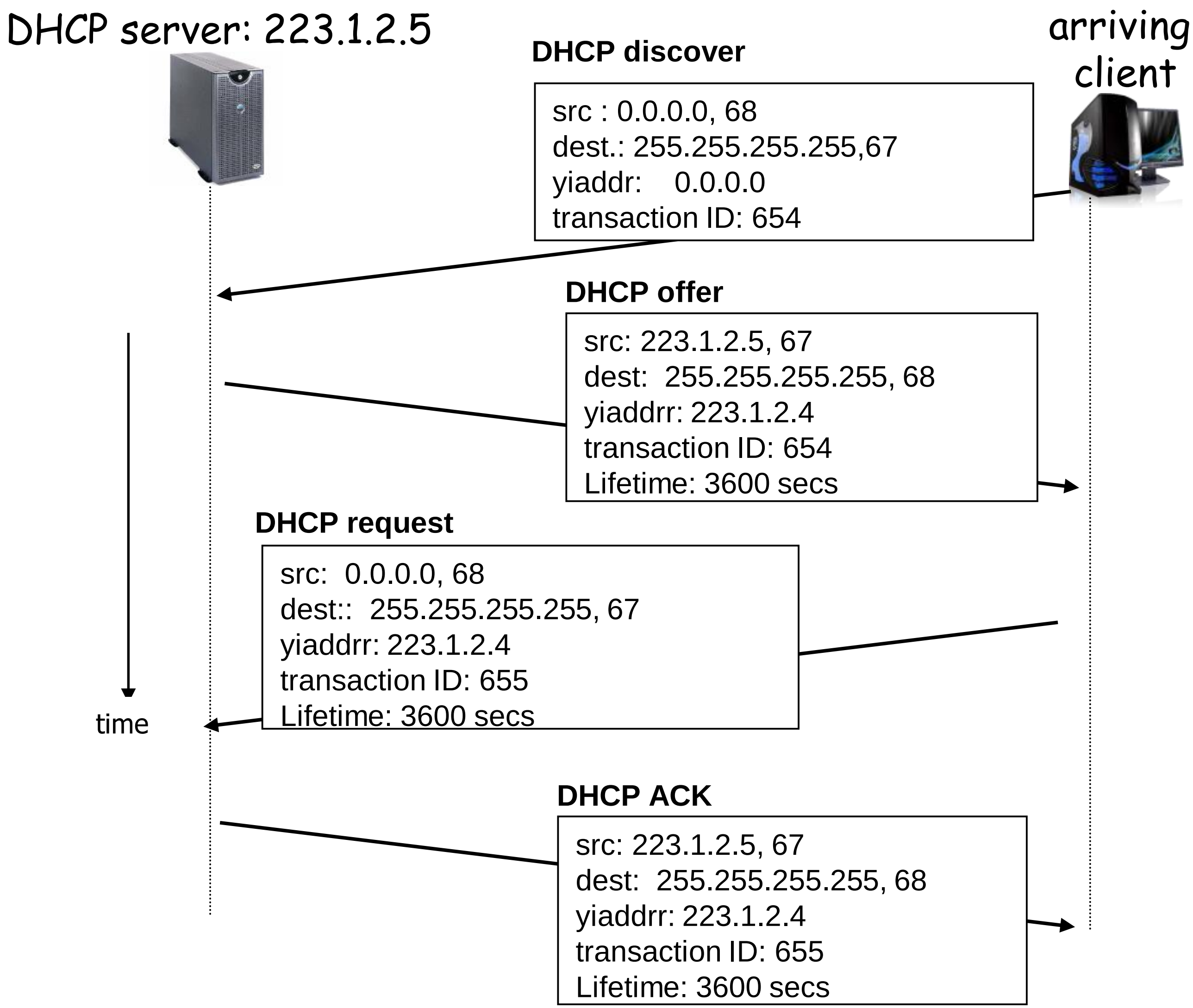
- The client broadcasts a DHCP-DISCOVER message on its subnet
- Each server may respond with a DHCP-OFFER message
- The client chooses one server, broadcasts a DHCP-REQUEST message including server IP
- The selected server commits the binding, responds with a DHCP-ACK message

Typical Procedure of DHCP

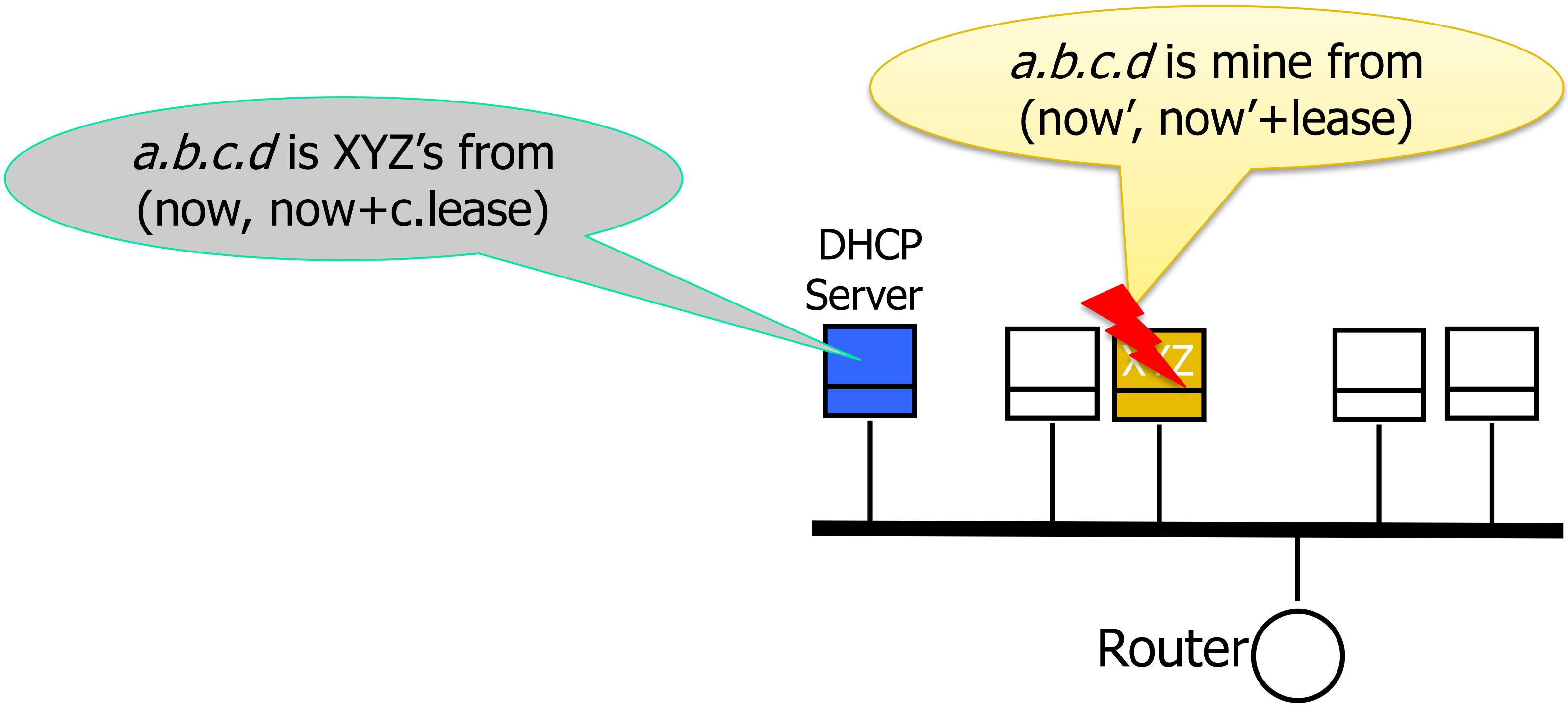


- The client set its configuration parameters within the DHCP-ACK
- The client relinquish the binding by a DHCP-RELEASE message
- The binding will be expired if the client does not renew (rebind) the binding before

DHCP Messages

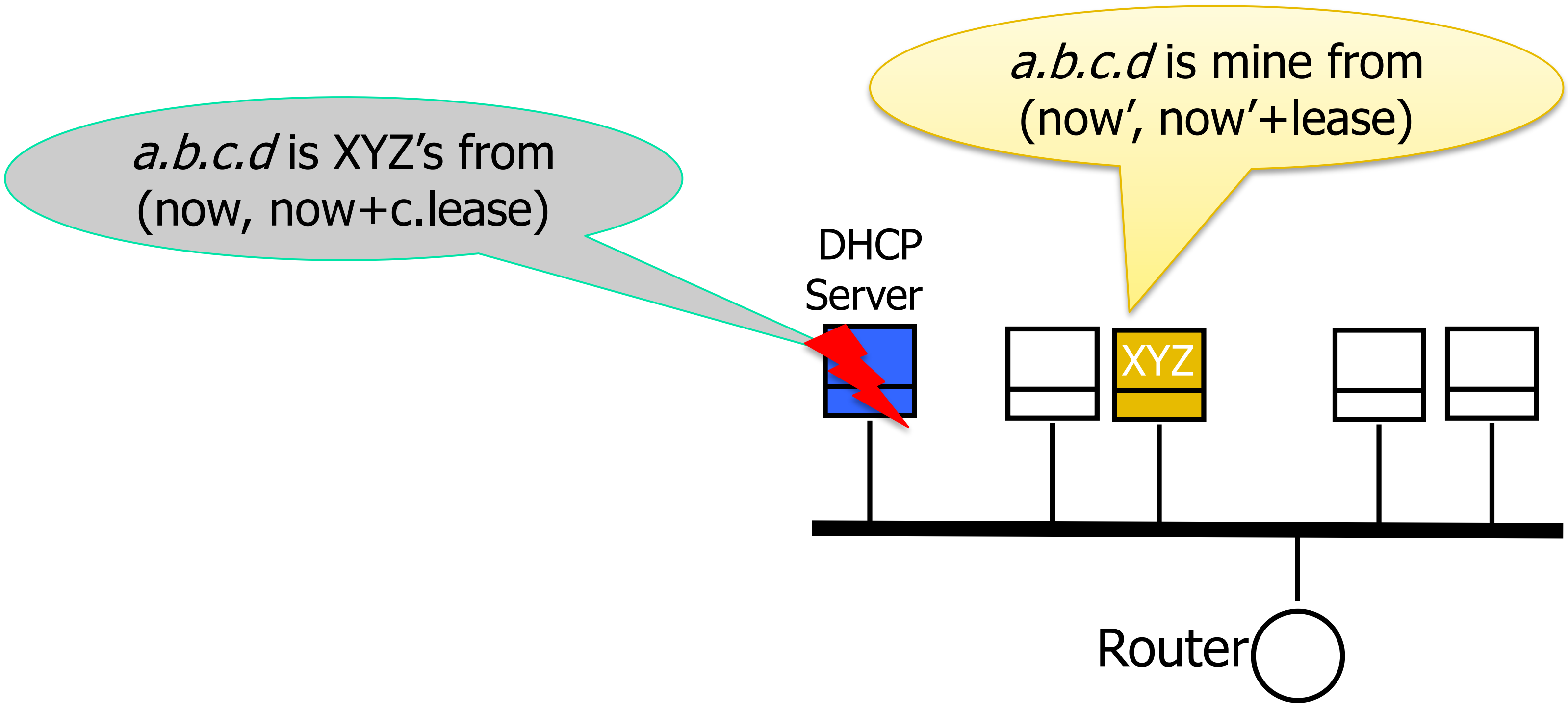


Soft state under failure



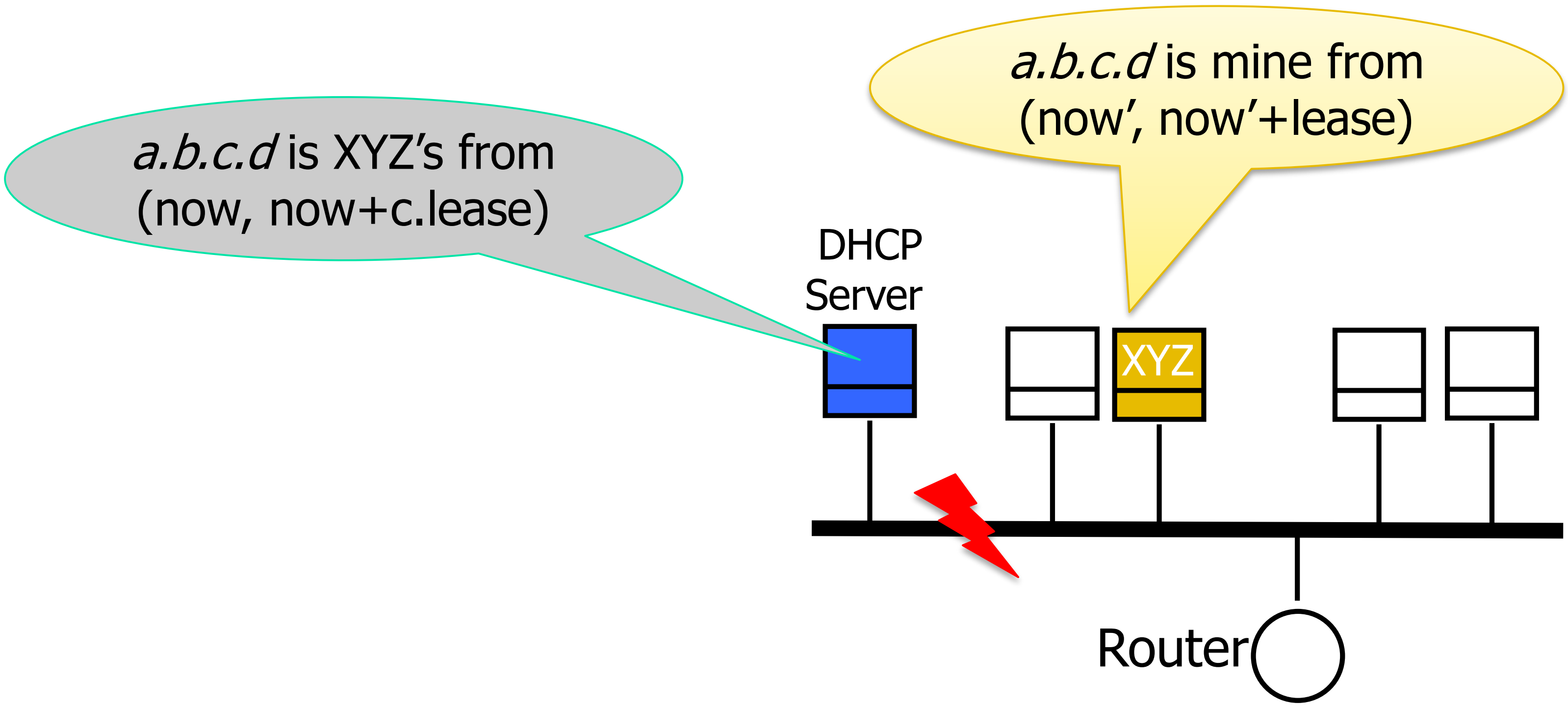
- What happens when host XYZ fails?
  - Refreshes from XYZ stop
  - Server reclaims a.b.c.d after O(lease period)

Soft state under failure



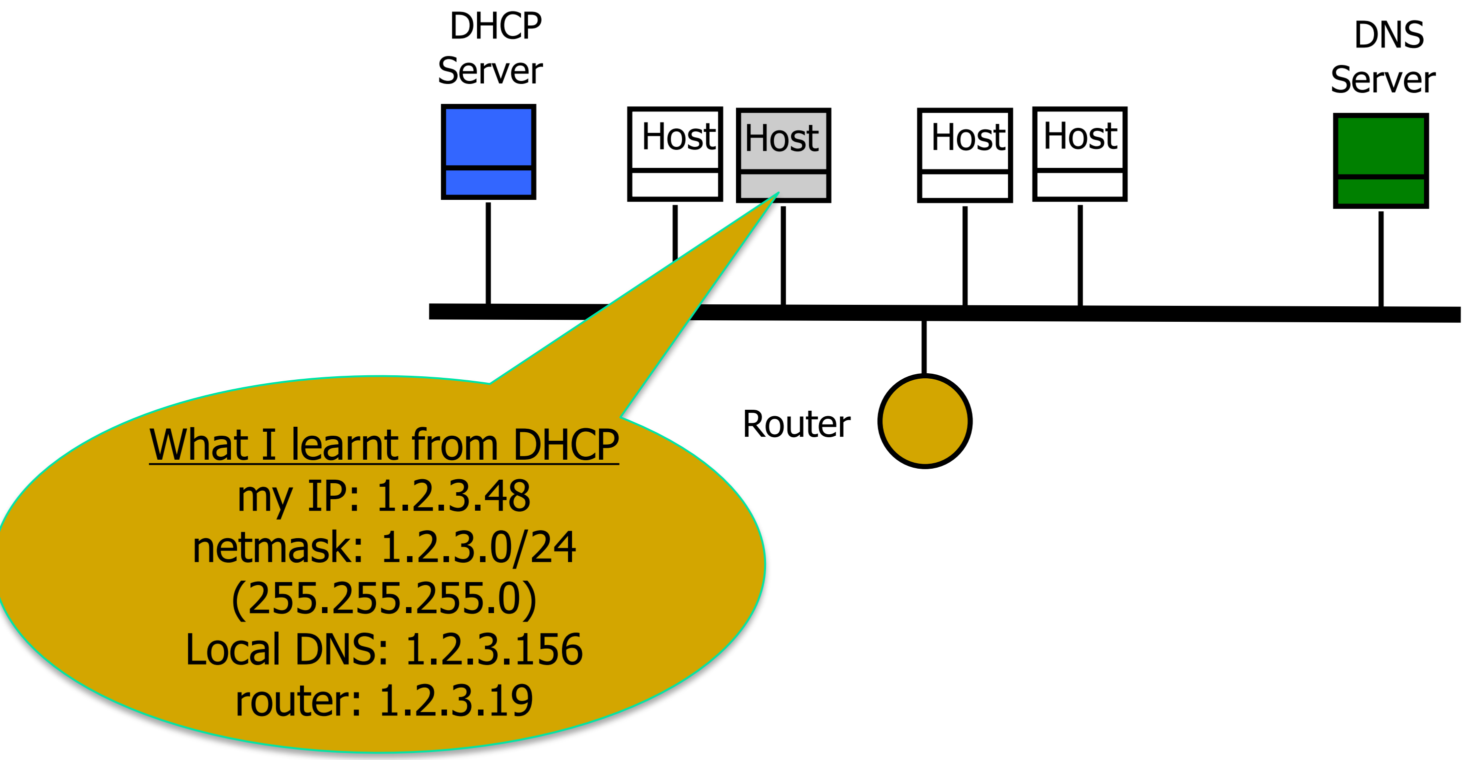
- What happens when server fails?
  - ACKs from server stop
  - XYZ releases address after O(lease period); send new request
  - A new DHCP server can come up from a `cold start' and we are back on track in ~lease time

Soft state under failure



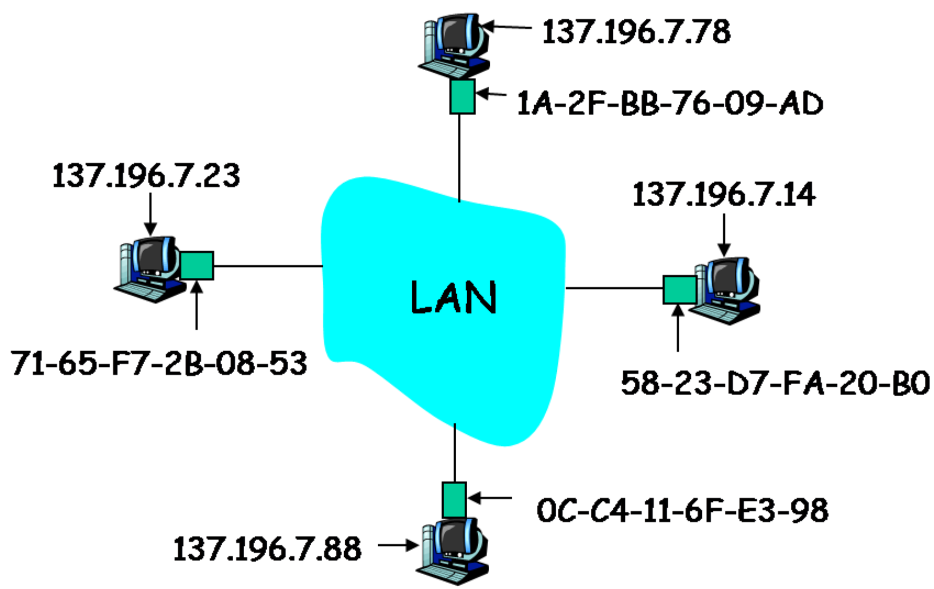
- What happens if the network fails?
  - Refreshes and ACKs don't get through
  - XYZ release address; DHCP server reclaims it

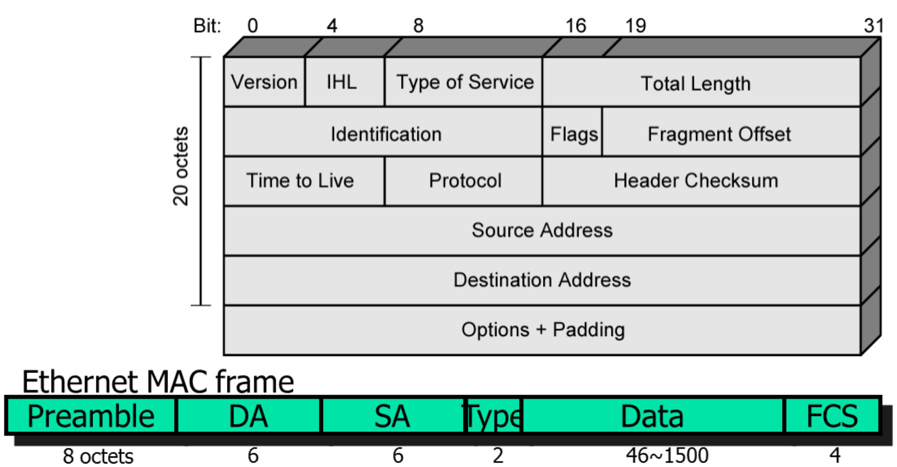
Are we there yet?



MAC Address Resolution Problem

- User 137.196.7.23 want to Ping 137.196.7.88
  - Source IP: 137.196.7.23
  - Destination IP: 137.196.7.88
  - Source MAC: 71-65-F7-2B-08-53
  - Destination Mac: ?
- Its MAC address is needed to deliver the data
- On LAN, ARP is used get a host/router’s MAC given its IP address



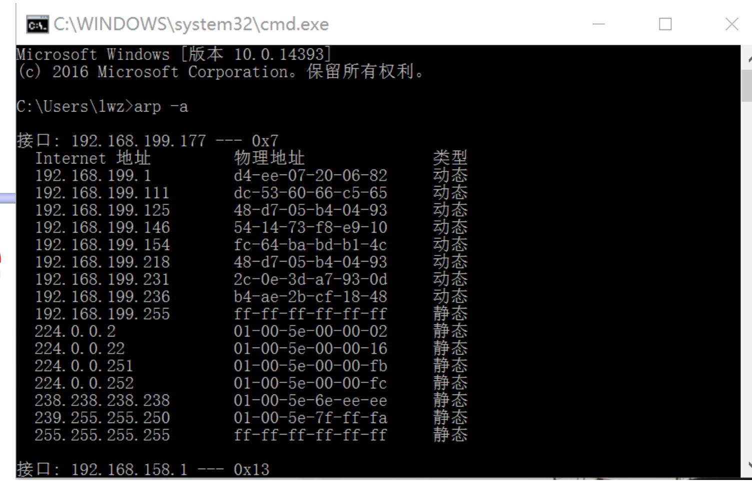


Address Resolution Protocol

- ARP (Address Resolution Protocol)
  - Map IP address to MAC address
  - 192.168.1.2 -> 00-15-C5-49-04-A9
  - Only works in a LAN

- Compare: DNS (domain name system)
  - Map domain name to IP address
  - Baidu.com -> 220.181.38.148

ARP Procedure



- Every host maintains an ARP table
  - List of (IP address → MAC address)
- Sender
  - Looks into local cache first, if none
  - Constructs ARP request, insert <*sender IP*, *sender MAC*, *destination IP*>
  - Broadcasts using MAC frame
  - Caches destination's <*MAC*, *IP*> pair with timestamp
- Receiver
  - Checks the destination IP, if OK
  - Constructs ARP reply, insert <*destination IP*, *destination MAC*>
  - Sends to sender MAC using MAC frame
  - Caches sender's <*MAC*, *IP*> pair with timestamp

ARP protocol in action

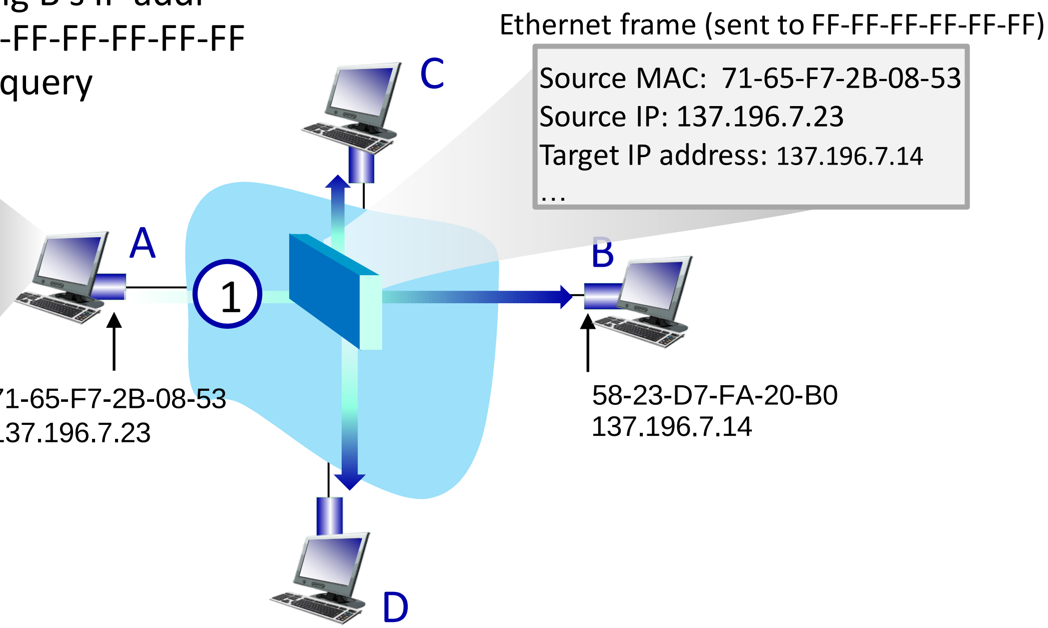
example: A wants to send datagram to B

- B's MAC address not in A's ARP table, so A uses ARP to find B's MAC address

1. A broadcasts ARP query, containing B's IP addr
   - destination MAC address = FF-FF-FF-FF-FF-FF
   - all nodes on LAN receive ARP query

ARP table in A

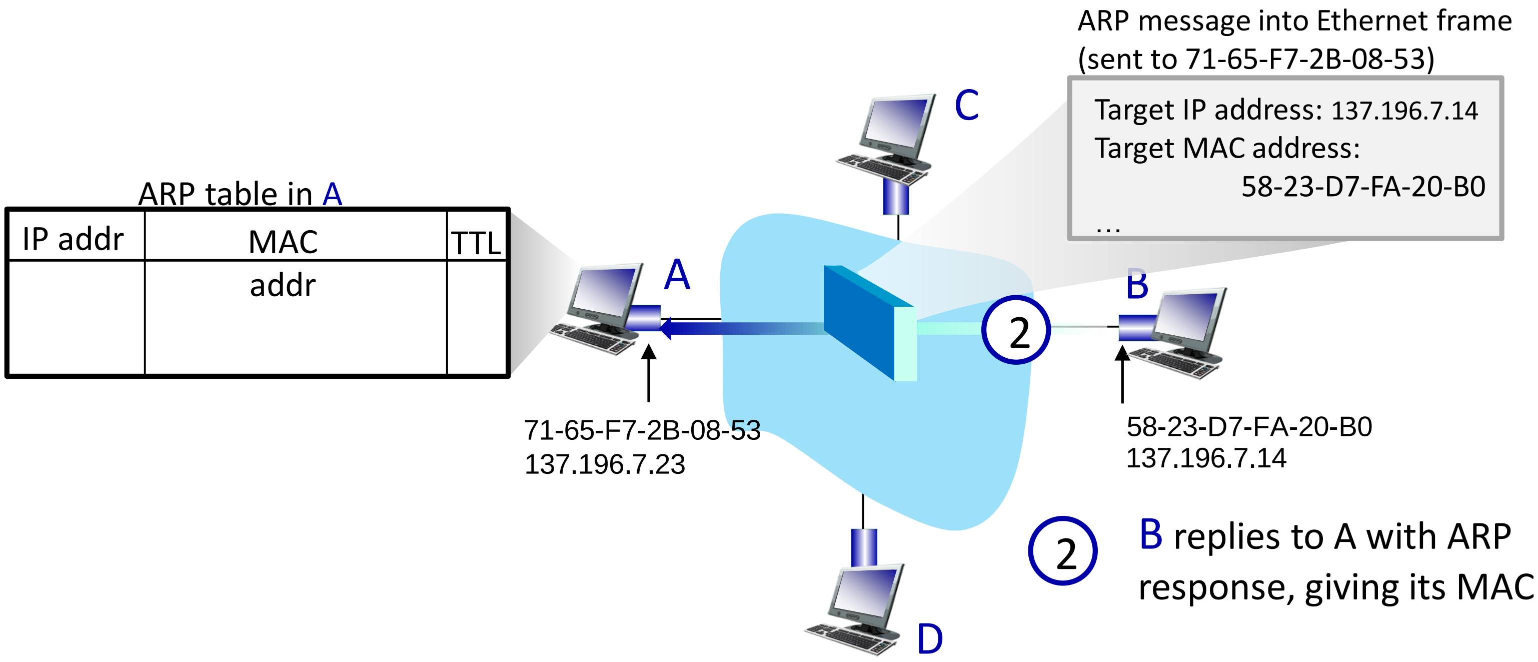
IP addr	MAC addr	TTL



ARP protocol in action

example: A wants to send datagram to B

- B's MAC address not in A's ARP table, so A uses ARP to find B's MAC address



2 B replies to A with ARP response, giving its MAC address

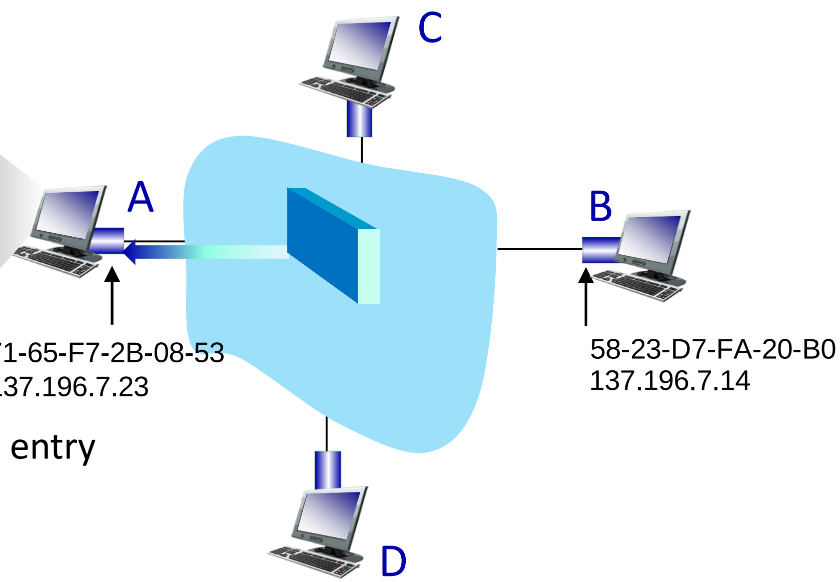
ARP protocol in action

example: A wants to send datagram to B

- B's MAC address not in A's ARP table, so A uses ARP to find B's MAC address

ARP table in A

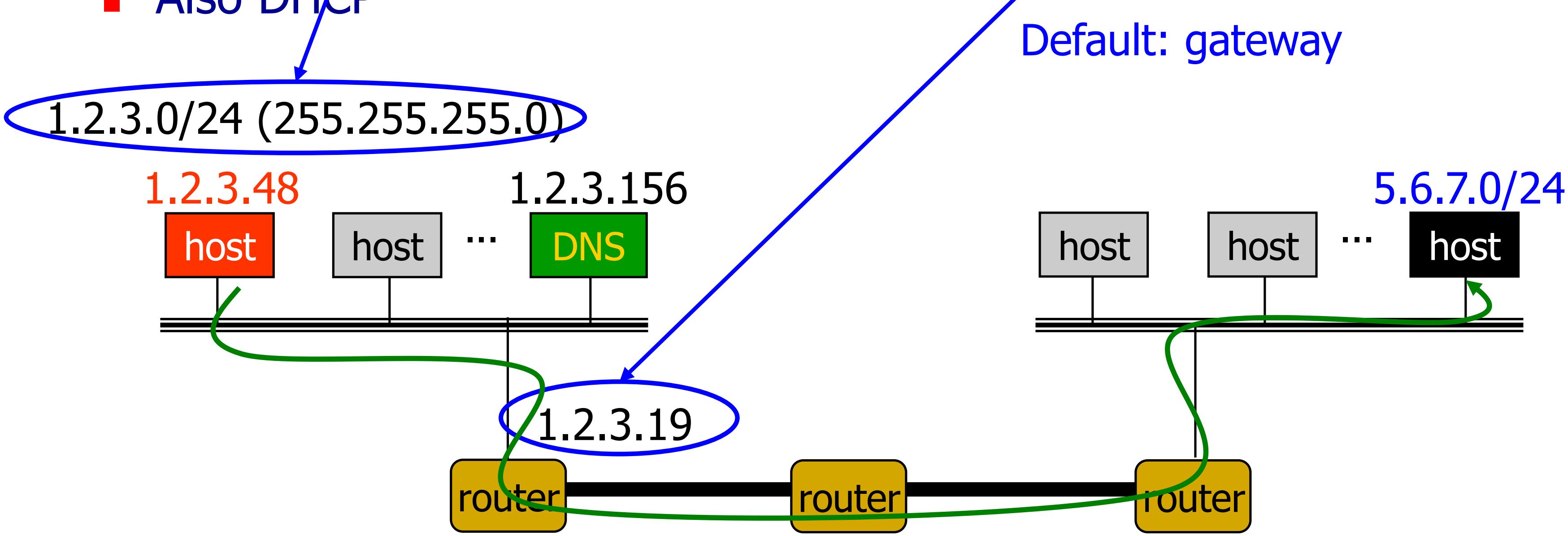
IP addr	MAC addr	TTL
137.196.7.14	58-23-D7-FA-20-B0	500



(3) A receives B's reply, adds B entry into its local ARP table

What if the destination is remote?

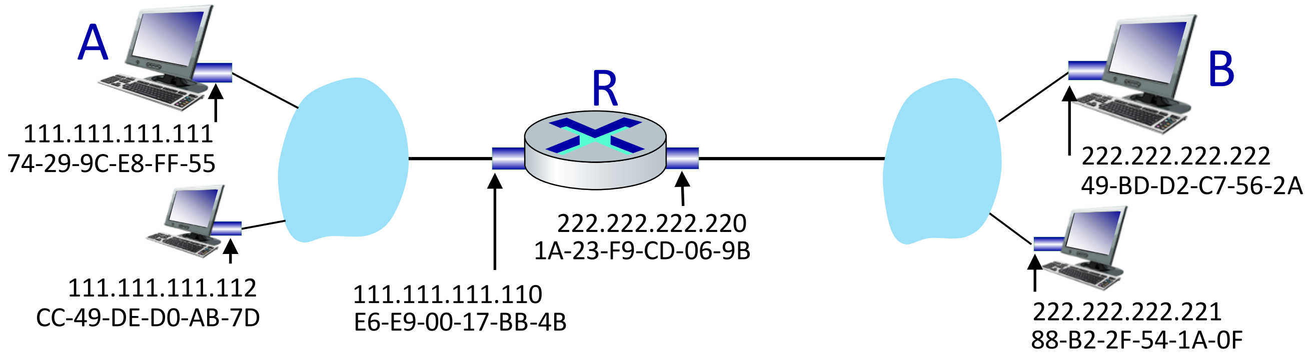
- Look up the MAC address of the first hop router
  - 1.2.3.48 uses ARP to find MAC address for first-hop router **1.2.3.19** rather than ultimate destination IP address
- How does the red host know the destination is not local?
  - Uses netmask (discovered via DHCP)
- How does the red host know about 1.2.3.19?
  - Also DHCP



Routing to another subnet

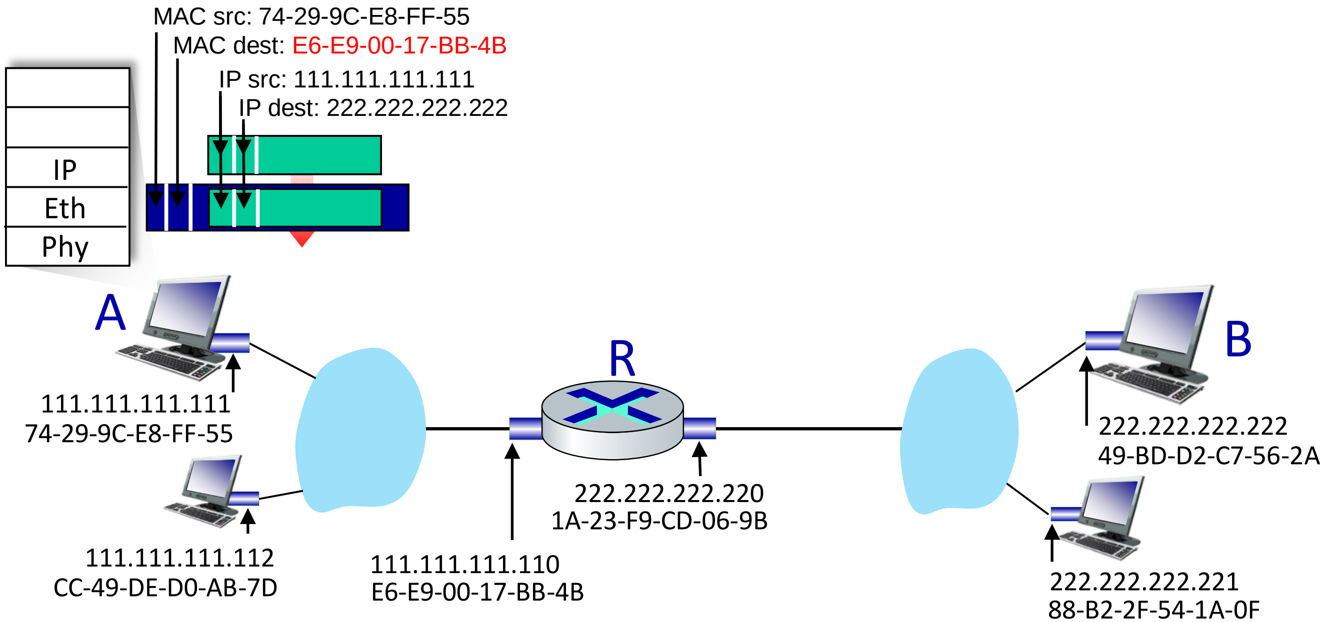
walkthrough: sending a datagram from *A* to *B* via *R*

- focus on addressing – at IP (datagram) and MAC layer (frame) levels
- assume that:
  - A knows B's IP address
  - A knows IP address of first hop router, R (how?)
  - A knows R's MAC address (how?)



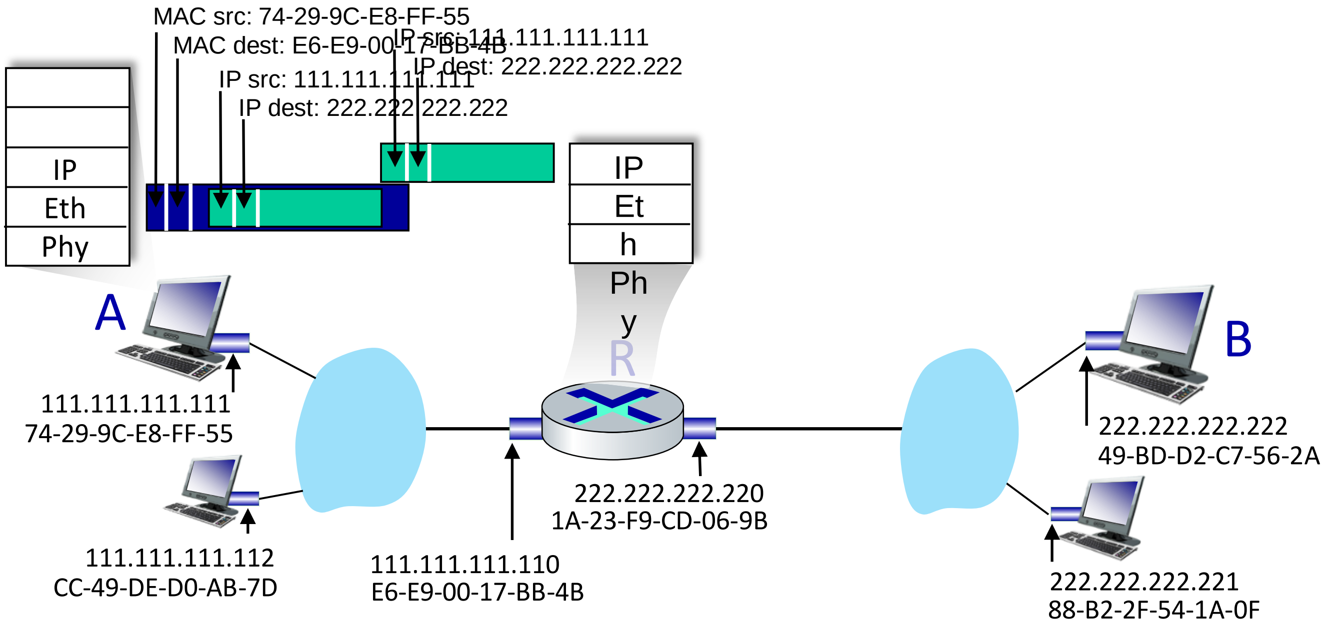
Routing to another subnet

- A creates IP datagram with IP source A, destination B
- A creates link-layer frame containing A-to-B IP datagram
  - R's MAC address is frame's destination



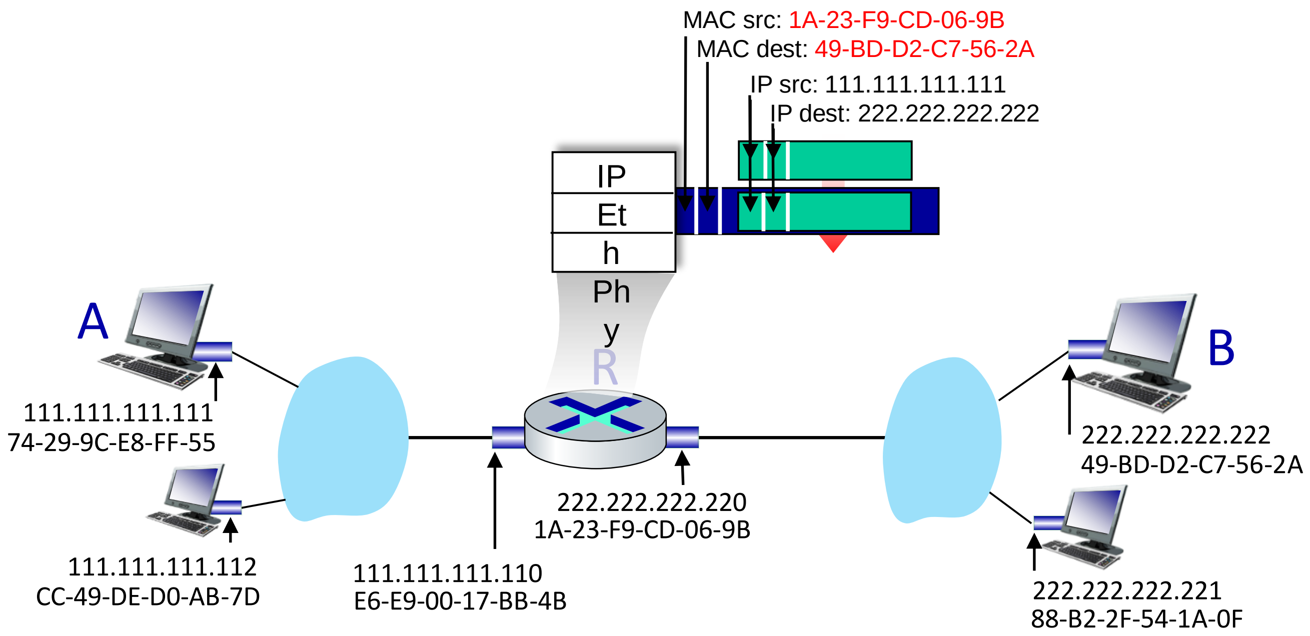
Routing to another subnet

- frame sent from A to R
- frame received at R, datagram removed, passed up to IP



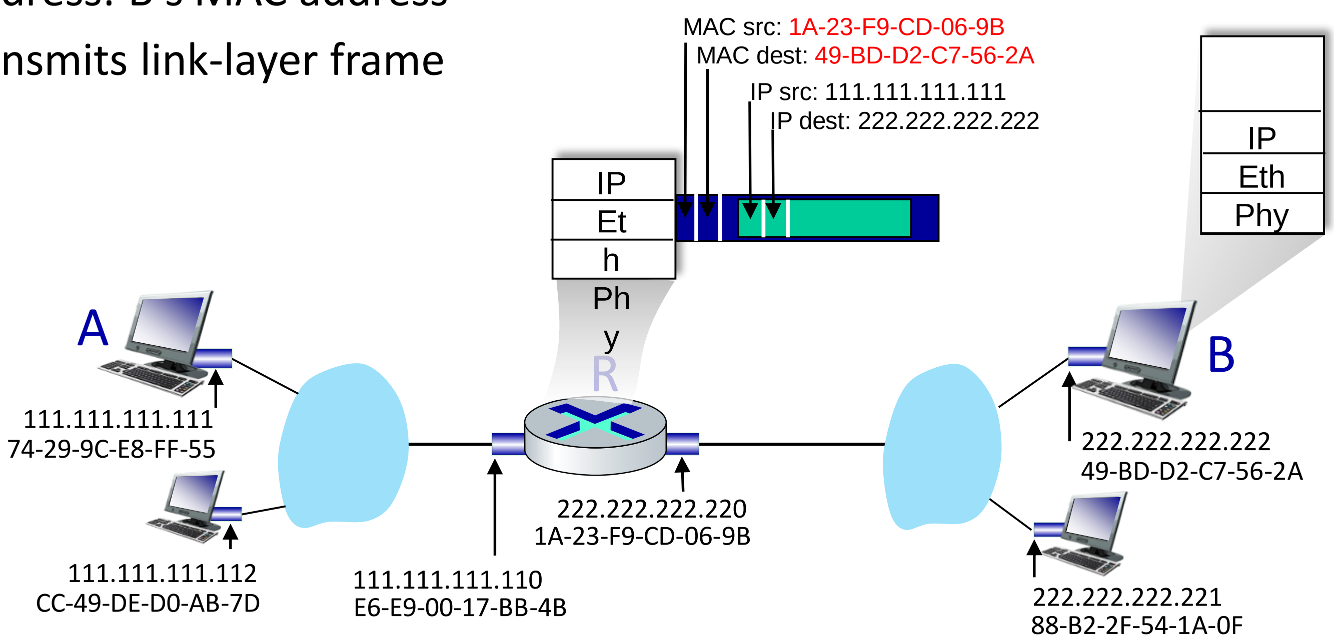
Routing to another subnet

- R determines outgoing interface, passes datagram with IP source A, destination B to link layer
- R creates link-layer frame containing A-to-B IP datagram. Frame destination address: B's MAC address



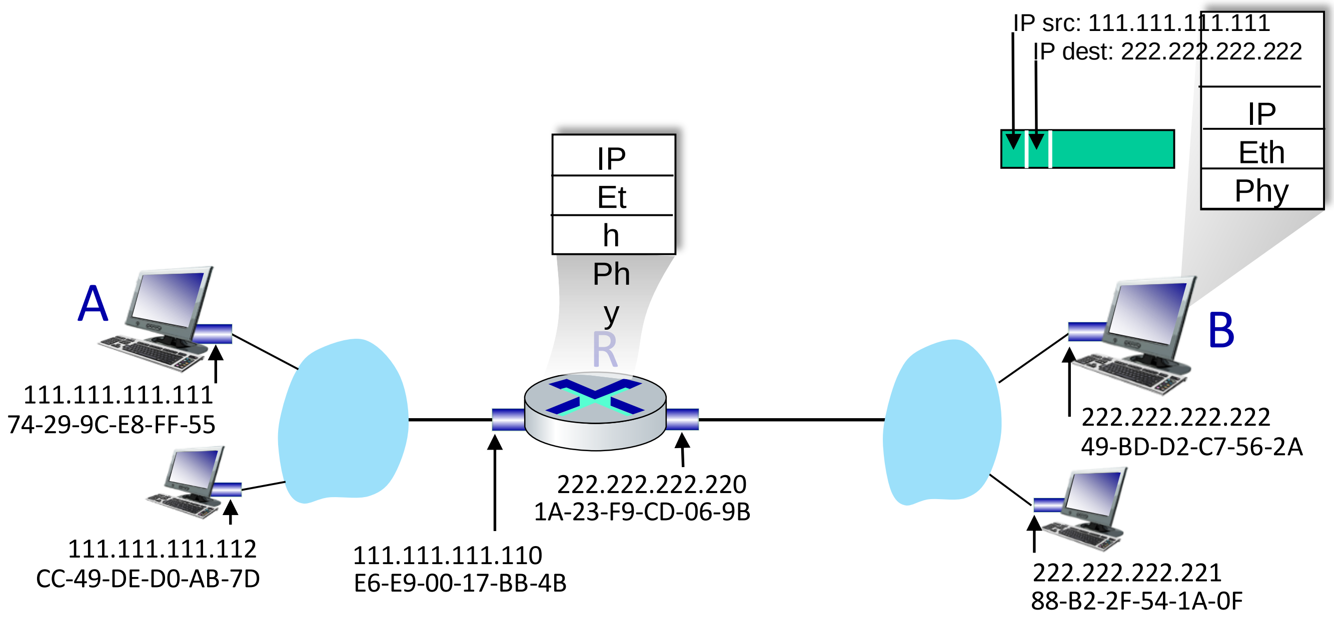
Routing to another subnet

- R determines outgoing interface, passes datagram with IP source A, destination B to link layer
- R creates link-layer frame containing A-to-B IP datagram. Frame destination address: B's MAC address
- transmits link-layer frame



Routing to another subnet

- B receives frame, extracts IP datagram destination B
- B passes datagram up protocol stack to IP



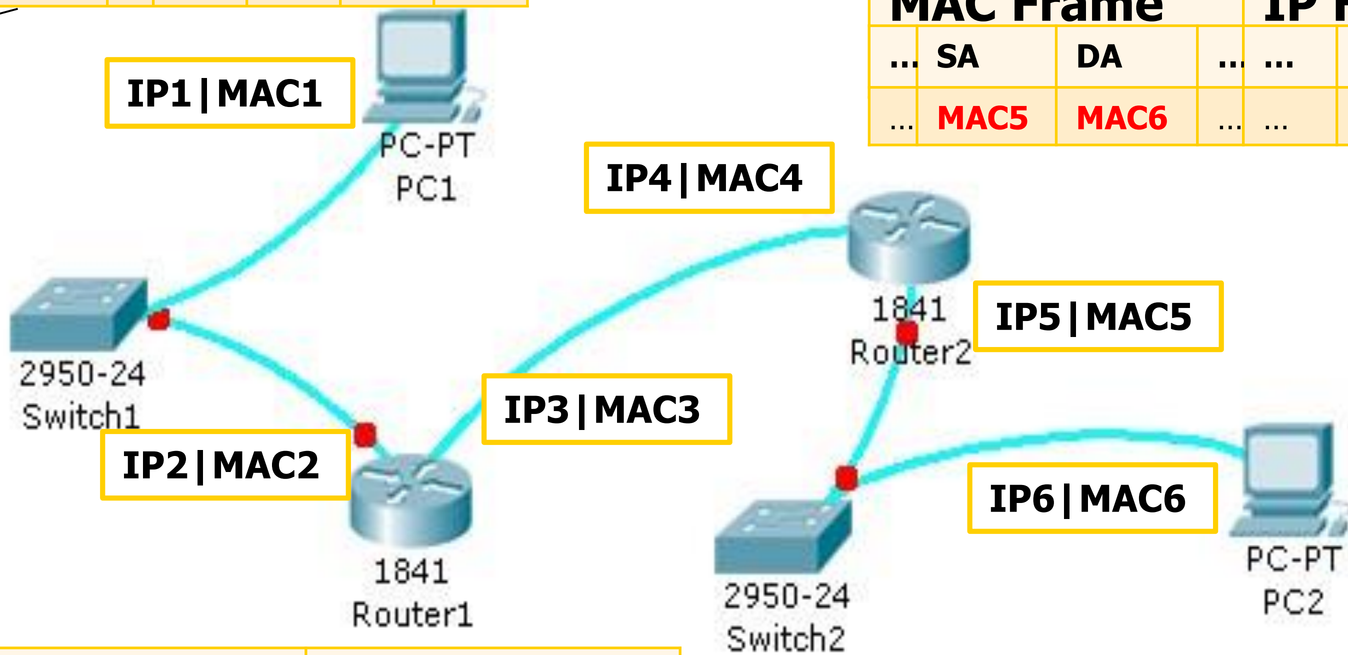
Example: PC1 Ping PC2

MAC Frame				IP Head			
..	SA	DA	...	...	SA	DA	...
...	MAC1	?	...	...	IP1	IP6	...

MAC2

Use ARP to find the next-hop MAC!

MAC Frame				IP Head			
...	SA	DA	...	...	SA	DA	...
...	MAC5	MAC6	...	...	IP1	IP6	...



MAC Frame				IP Head			
...	SA	DA	...	...	SA	DA	...
...	MAC3	MAC4	...	...	IP1	IP6	...

Key ideas in both ARP and DHCP

- Broadcasting: Can use broadcast to make contact
  - Scalable because of limited size
- Caching: remember the past for a while
  - Store the information you learn to reduce overhead
- Soft state: eventually forget the past
  - Associate a time-to-live field with the information
  - … and either refresh or discard the information
  - Key for robustness in the face of unpredictable change

ID resolution in the networking stack

Layer	Examples	Structure	Configuration	Resolution Service
App. Layer	cse.umich.edu	Organizational hierarchy	~ manual	DNS
Network Layer	123.45.6.78	topological hierarchy	DHCP	ARP
Link layer	45-CC-4E-12-F0-97	vendor (flat)	hard-coded	

Discovery mechanisms

- We have seen two approaches
  - Broadcast (ARP, DHCP)
    - Flooding does not scale
    - No centralized point of failure
    - Zero configuration
  - Directory service (DNS)
    - No flooding / scalable
    - Root of the directory is vulnerable (caching is key)
    - Needs configuration to bootstrap (local, root servers, etc.)

Summary

- MAC机制性能分析
  - Point-to-point link
  - Ring LAN
  - ALOHA, Slotted ALOHA
  - CSMA/CD （ p-persistent ）
- MAC地址发现
  - 自动地址配置：DHCP
  - MAC地址解析：ARP