CSCE 463/612
Networks and Distributed Processing
Fall 2020

Data-link Layer II

Dmitri Loguinov
Texas A&M University

November 19, 2020

Link Layer

5.1 Introduction and services

5.2 Error detection and correction

5.3 Multiple access protocols

5.4 Link-Layer Addressing

5.5 Ethernet

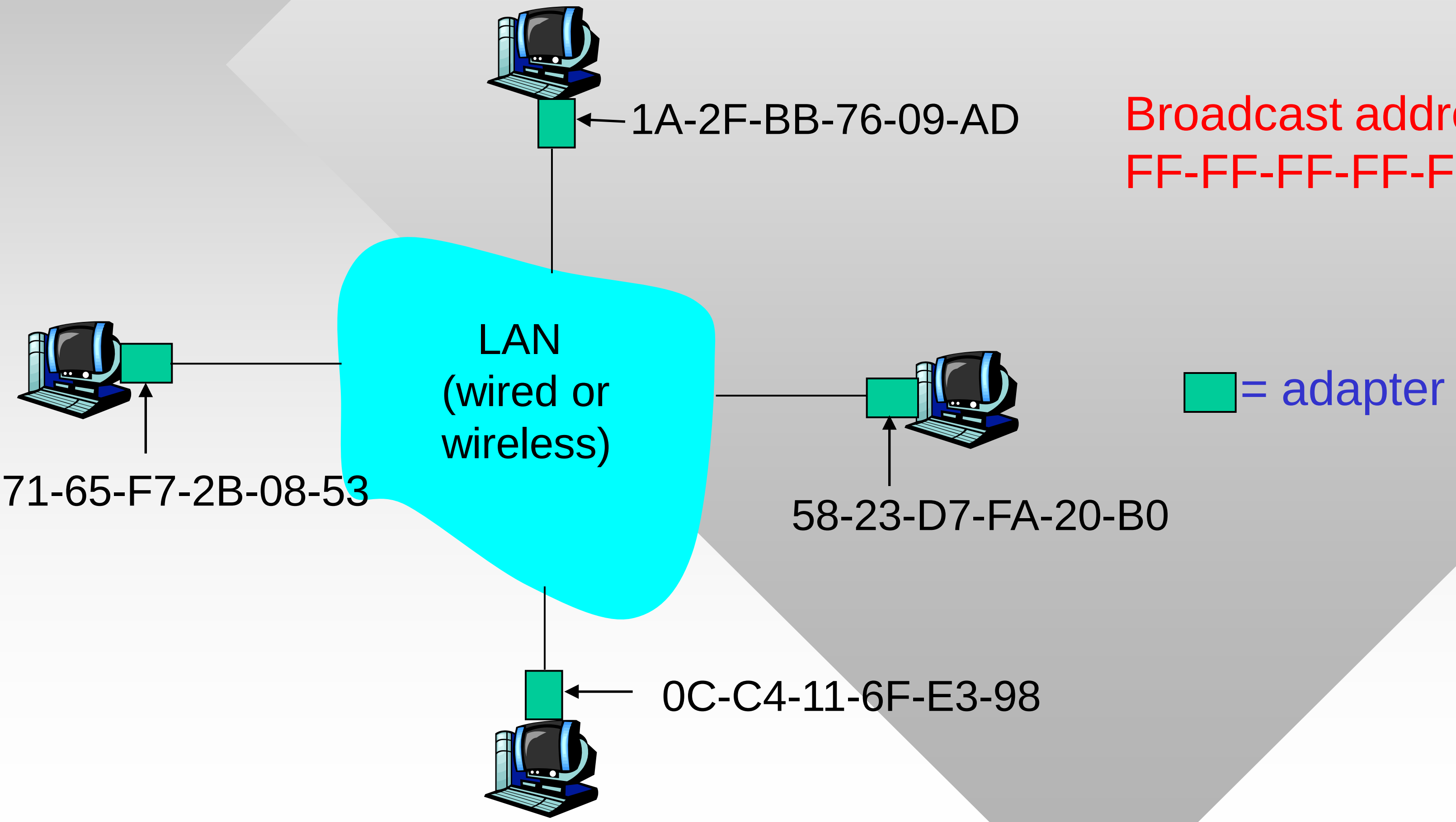
5.6 Hubs and switches

Network Addresses Revisited

- Transport-layer address:
  - 16-bit port number
  - Find correct application within a host
- Network-layer address:
  - 32-bit (IPv4) or 128-bit (IPv6)
  - Find correct subnet & host on the Internet
- MAC (or LAN, or physical, or Ethernet) address:
  - 48 bit number burned in the adapter ROM
  - Find correct interface within a subnet

LAN Addresses

Each adapter in a LAN must have a unique LAN address



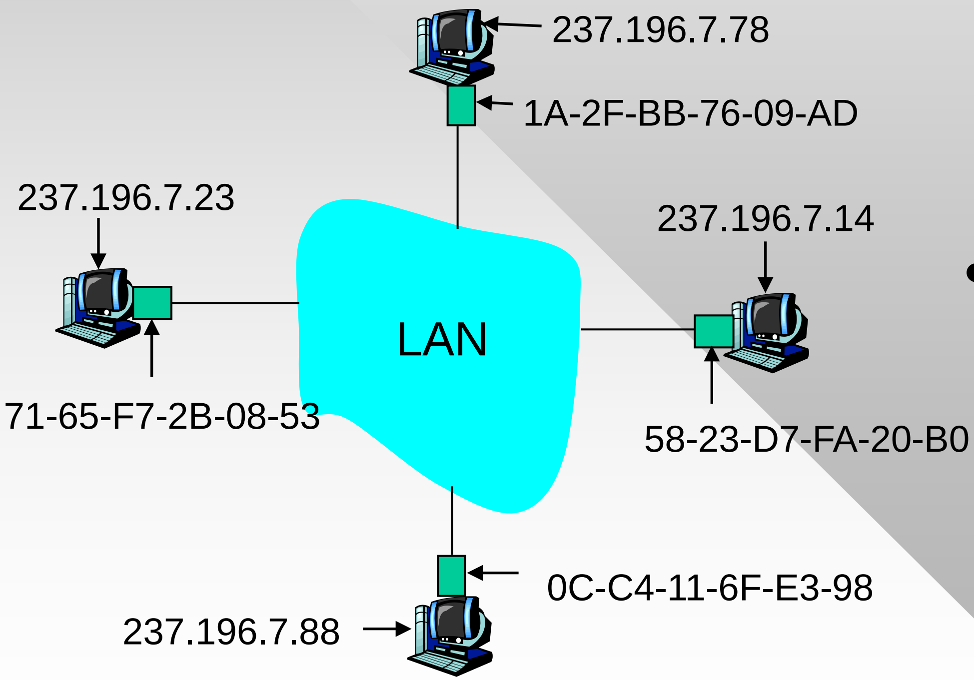
Broadcast address = FF-FF-FF-FF-FF-FF

LAN Addresses

- MAC address allocation administered by IEEE
  - $660 for 36-bit prefix, $1595 for 28-bit; $2655 for 24-bit
- Manufacturer buys portion of MAC address space (to assure uniqueness)
- Analogy:
  - MAC address: like Social Security Number
  - IP address: like postal address
- Flat MAC addresses achieve portability
  - Can move NIC from one LAN to another
- Hierarchical IP addresses NOT portable
  - IP depends on subnet to which node is attached

ARP: Address Resolution Protocol (1984)

Question: how to determine MAC address of a host knowing its IP address?



- Each IP node (host, router) on LAN has an ARP table
  - Contains IP/MAC address mappings for known LAN nodes

< IP address; MAC address; TTL>

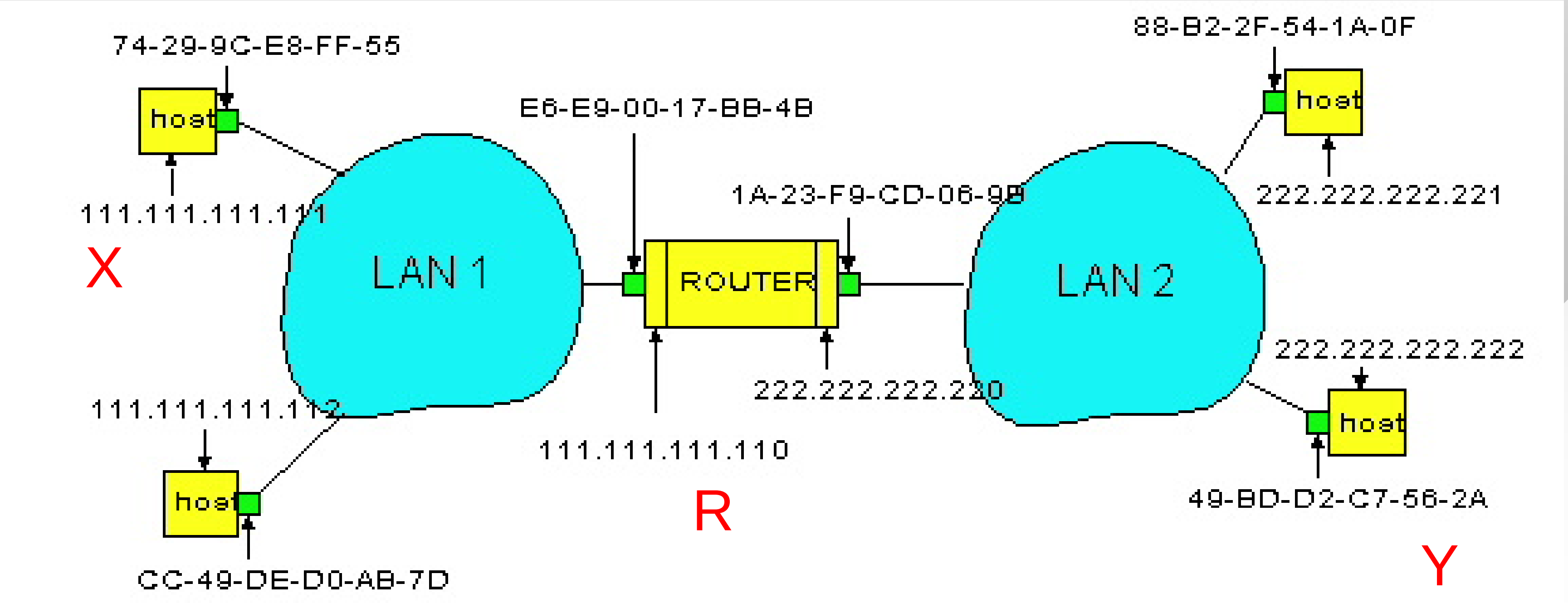
- TTL (Time To Live): time after which address mapping will be forgotten (typically 20 min)

ARP Protocol: Same LAN/Subnet

- X wants to send datagram to Y, but doesn't know Y's MAC address yet
  - X broadcasts ARP query packet, containing Y's IP address
  - All machines on LAN process ARP query
- Y receives ARP packet, replies to X with its MAC address
  - Frame sent to X's MAC address (unicast)
- X caches (saves) IP-to-MAC address pair in its ARP table until this information becomes outdated
  - Soft state: information times out (goes away) unless refreshed
- ARP is "plug-and-play"
  - Nodes create their ARP tables without intervention from net administrator

Routing to Another LAN

- Walkthrough: send datagram from X to Y via R
  - Suppose 1) X knows Y's IP address; and 2) X knows its default router R (111.111.111.110)



ARP Example

```
C:\>arp -a
Interface: 128.194.135.66
  Internet Address      Physical Address      Type
  128.194.135.65        00-0c-f1-ad-9b-d9     dynamic
  128.194.135.72        00-0c-f1-ad-9b-d9     dynamic
  128.194.135.73        00-e0-18-91-68-9c     dynamic
  128.194.135.74        00-08-74-ce-97-60     dynamic
  128.194.135.76        00-04-23-ab-be-50     dynamic
  128.194.135.81        00-04-23-ab-bc-7a     dynamic
  128.194.135.92        00-0c-f1-ad-9b-d9     dynamic
```

- Why do 3 hosts in bold have the same MAC?
- Which hosts have NICs from the same manufacturer?

DHCP

- DHCP (Dynamic Host Configuration Protocol, 1993)
  - Assigns IP address, netmask, DNS server, default router, and other parameters to end-hosts
- DHCP runs over UDP (ports 67-68), using MAC-layer broadcasts to find available servers
  - Discovery→Offer→Request→Lease (4 packets exchanged)
  - Client may receive multiple offers, must choose one
  - Leased IPs carry some TTL (expiration time)
  - Routers and switches may implement DHCP
- Routers may be configured to forward broadcast DHCP packets between subnets
  - One DHCP server may cover multiple LANs

DHCP Example

```
C:\> ipconfig /all
Ethernet adapter Wireless Network Connection 6:
Connection-specific DNS Suffix  . : tamu.edu
Description . . . . . . . . . . . : DWL-G650M  Super G MIMO
   Wireless Adapter #4
   Physical Address. . . . . . . . . : 00-11-95-6A-96-8A
   Dhcp Enabled. . . . . . . . . . . : Yes
   Autoconfiguration Enabled . . . . : Yes
   IP Address. . . . . . . . . . . . : 10.32.39.138
   Subnet Mask . . . . . . . . . . . : 255.255.254.0
   Default Gateway . . . . . . . . . : 10.32.38.1
   DHCP Server . . . . . . . . . . . : 10.32.38.1
   DNS Servers . . . . . . . . . . . : 128.194.254.3
                                       128.194.254.2
                                       128.194.254.1
   Lease Obtained. . . . . . . . . . : Tuesday, November 21, 2006
   3:22:13 PM
   Lease Expires . . . . . . . . . . : Tuesday, November 21, 2006
   4:22:13 PM
```

Link Layer

5.1 Introduction and services

5.2 Error detection and correction

5.3 Multiple access protocols

5.4 Link-Layer Addressing

5.5 Ethernet

5.6 Hubs and switches

Ethernet

- Dominant wired LAN technology
  - Inspired by ALOHAnet, based on Robert Metcalfe's PhD thesis in 1973
  - Xerox patented in 1976, first standardized in 1980
  - In the early 1980s, competed with Token Ring (IBM) and Token Bus (GM), eventually overpowering both
- Ethernet data rates
  - Experimental: 2.94 Mbps (1973)
  - 10Base5 thick coax [IEEE 802.3]: 10 Mbps (1983)
  - Fast Ethernet [IEEE 802.3u]: 100 Mbps (1995)
  - Gigabit Ethernet (GE) [IEEE 802.3ab]: 1 Gbps (1999)
  - 10 GE [802.3ae]: 10 Gbps (2003 fiber, 2006 twisted pair)
  - 40/100 GE [IEEE 802.3ba]: 2010-2015

Ethernet Frame Structure

- Sending adapter encapsulates IP datagram (or other network-layer protocol packet) in Ethernet frame

preamble	dest MAC	src MAC	type	data	CRC	gap

- 8-byte preamble (physical layer):
  - 7 bytes 10101010 followed by one byte 10101011, synchronizes receiver-sender clock rates
- 6-byte MAC addresses:
  - If adapter receives frame with a matching or broadcast address, it passes data to the network layer
  - Otherwise, adapter discards frame

Ethernet Frame Structure

preamble	dest MAC	src MAC	type	data	CRC	gap

- 2-byte protocol type:
  - Indicates the higher-layer protocol
  - Examples are IPv4 (0x800), IPv6 (0x86DD), ARP (0x806), Wake-On-LAN (0x842), AppleTalk, Novell IPX, MPLS
- 32-bit CRC checksum:
  - Checked at receiver, if error is detected, frame is dropped
- Minimum payload 46 bytes, inter-frame gap 12 bytes
  - Frames shorter than minimum are interpreted as collisions
  - Resulting smallest frame time is 8+14+46+4+12 = 84 bytes
  - 1 Gbps max rate is 1.488M pps (packets per second)

Ethernet CSMA/CD

- No slots
- Adapter doesn't transmit if it senses that some other adapter is transmitting, that is, carrier sense (CS)
- Transmitting adapter aborts when it senses that another adapter is transmitting, that is, collision detection (CD)
- Before attempting a retransmission, adapter waits a random time, that is, random access
- Connectionless: no handshaking between sending and receiving adapter
- Unreliable: receiving adapter doesn't send ACKs or NAKs to sending adapter

Ethernet's CSMA/CD

- Bit time = 1 / speed
  - 100 nanosec for 10 Mbps Ethernet
  - 1 nanosec for GE
  - 100 picosec for 10 GE

TCP's exponential timer backoff during congestion is similar

Exponential Backoff:

- *Goal*: adapt retransmission attempts to estimated load
  - Heavy load: random wait should be longer
- After m-th collision in a row
  - Choose integer D $\in$ $[0, 2^m-1]$; then wait $512 \cdot D$ bit times before retx
- Example:
  - After second collision: choose D from {0, 1, 2, 3}
  - After ten collisions, choose D from {0, 1, 2, 3, 4, …,1023}

Ethernet Efficiency

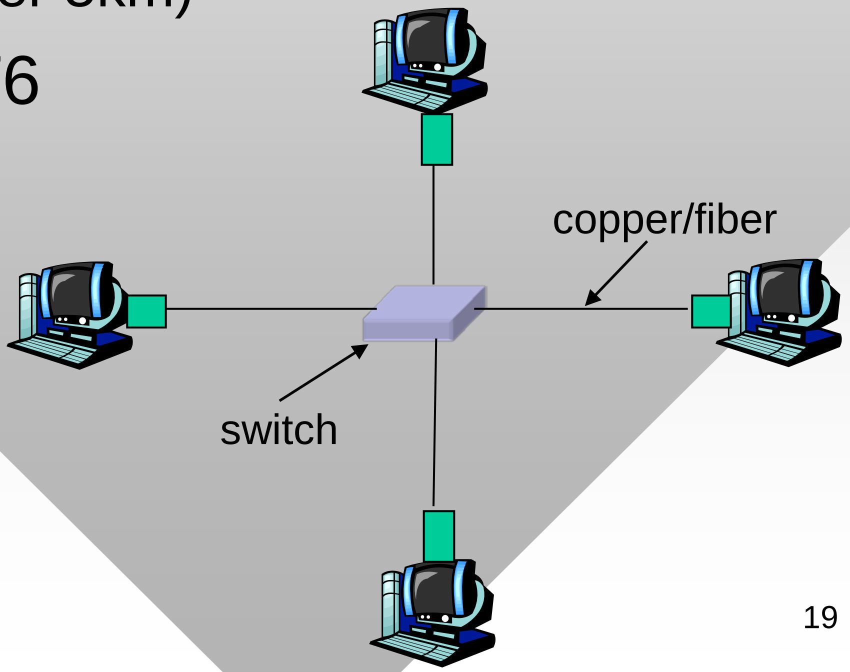
- $d =$ max propagation delay between any two nodes in LAN
- $T =$ time to transmit frame

$$\text{efficiency} = \frac{1}{1 + 5\frac{d}{T}}$$

- Efficiency goes to 1 as d goes to 0 (less collision probability)
- It also goes to 1 as T goes to infinity (less frequent switching between hosts)
- 1GE with 10m link and 1500 byte MTU: 98.6% utilization
  - Much better than slotted ALOHA (where efficiency is 37%), but still decentralized, simple, and cheap

Ethernet Technology

- Notation: [speed]Base[medium]
- Examples
  - 10base5 (thick coax 500m), 10base2 (thin coax 200m), 10baseT (twisted-pair/copper CAT3 with 8 wires and RJ45 connectors 100m), 1000BaseSX (short-range fiber 550m), 1000BaseLX (long-range fiber 5km)
- Now: 10GBaseT over CAT6 (55m), CAT6a (100m)
- Coax networks were daisy-chained, while copper and fiber run the star topology

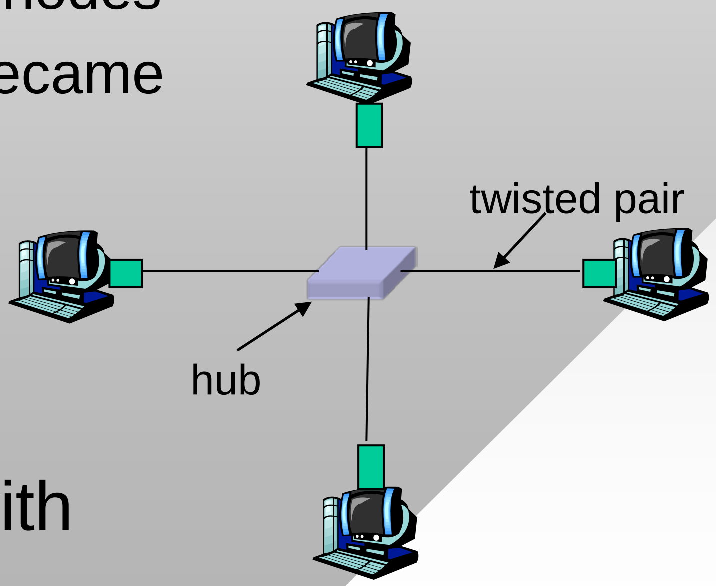


Link Layer

5.1 Introduction and services
5.2 Error detection and correction
5.3 Multiple access protocols
5.4 Link-Layer Addressing
5.5 Ethernet
5.6 Hubs and switches

Hubs

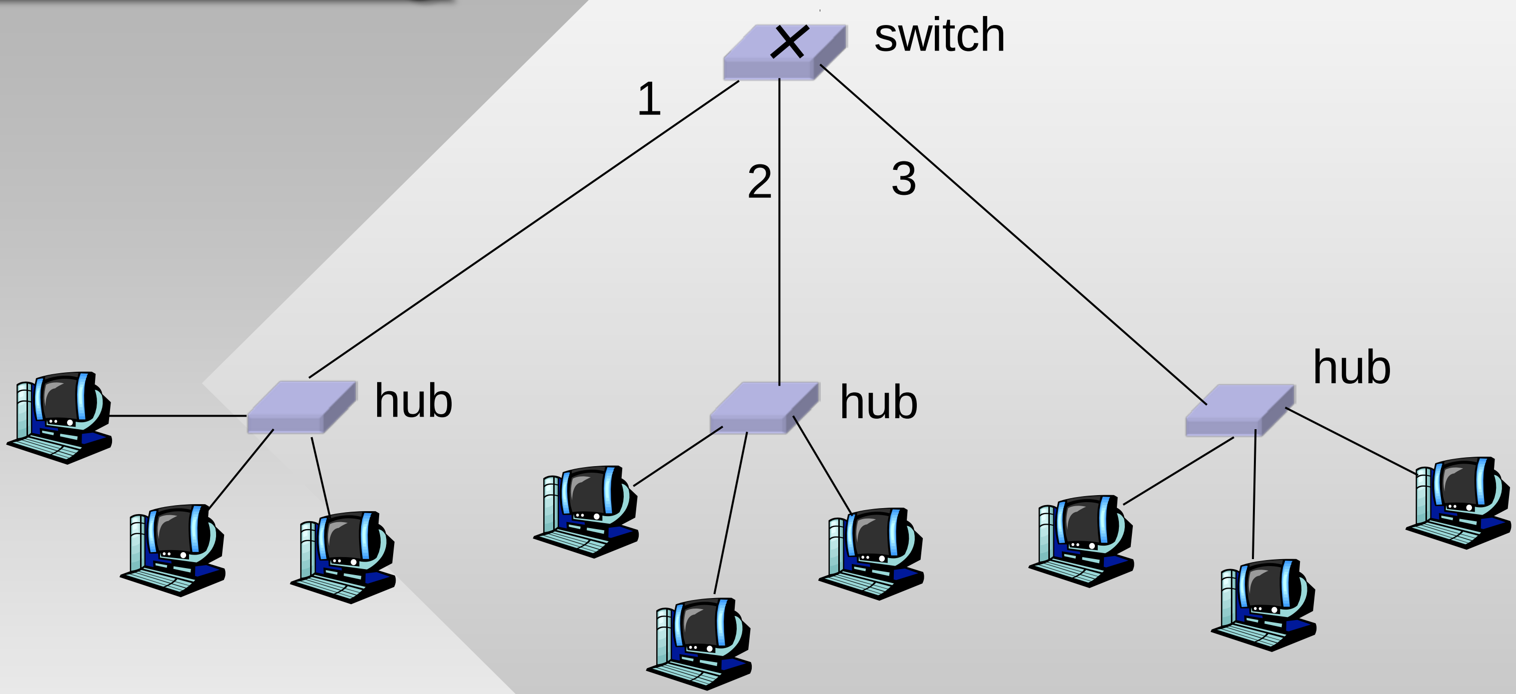
- Hubs were physical-layer repeaters:
  - Bits coming from one link went out all other links
- No frame buffering
  - No CSMA/CD at hub: host adapters must detect collisions
- Backbone hubs interconnected other hubs
  - Increased max distance between nodes
  - But individual collision domains became one large collision domain
- Additional limitations
  - No management functionality
  - All ports had to be same speed
- Similar to daisy-chaining, but with all connections in one place



Switches

- Link layer device
  - Stores and forwards Ethernet frames
  - Examines frame header and selectively forwards frame based on MAC dest address
  - When frame is to be forwarded on segment, uses CSMA/CD to access segment
- Transparent
  - Hosts are unaware of presence of switches
- Plug-and-play, self-learning
  - Switches can function without manual configuration

Forwarding



- How to determine onto which LAN segment to forward frame?
  - Looks like a routing problem…
- Most LAN networks are trees and flooding is permitted, which makes the problem much simpler!

Self Learning

- A switch has a switch table
- Entry in switch table:
  - (MAC Address, Interface, TTL)
  - Stale entries in table dropped (TTL can be 60 min)
- Switch *learns* which hosts can be reached through which interfaces
  - When frame received, switch “learns” location of sender: incoming LAN segment
  - Records sender/location pair in switch table

Filtering/Forwarding

When switch receives a frame:

index switch table using destination MAC address

if entry found for destination **then {**

    if dest on segment from which frame arrived
    then drop the frame
    else forward the frame on interface indicated

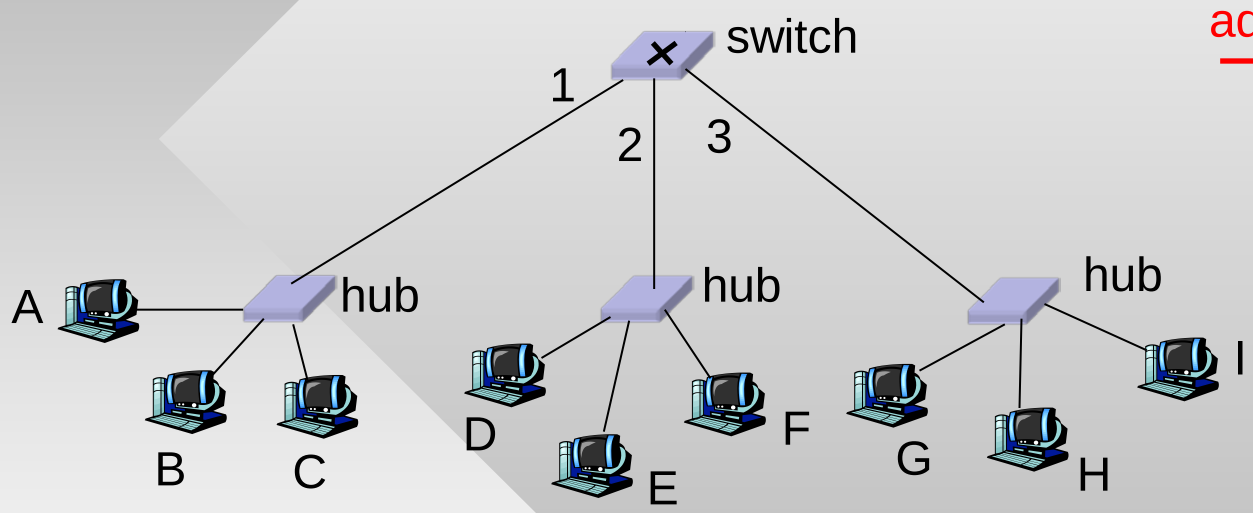
}

else flood

Forward on all but the interface on which the frame arrived

Switch Example

Suppose C sends a frame to D

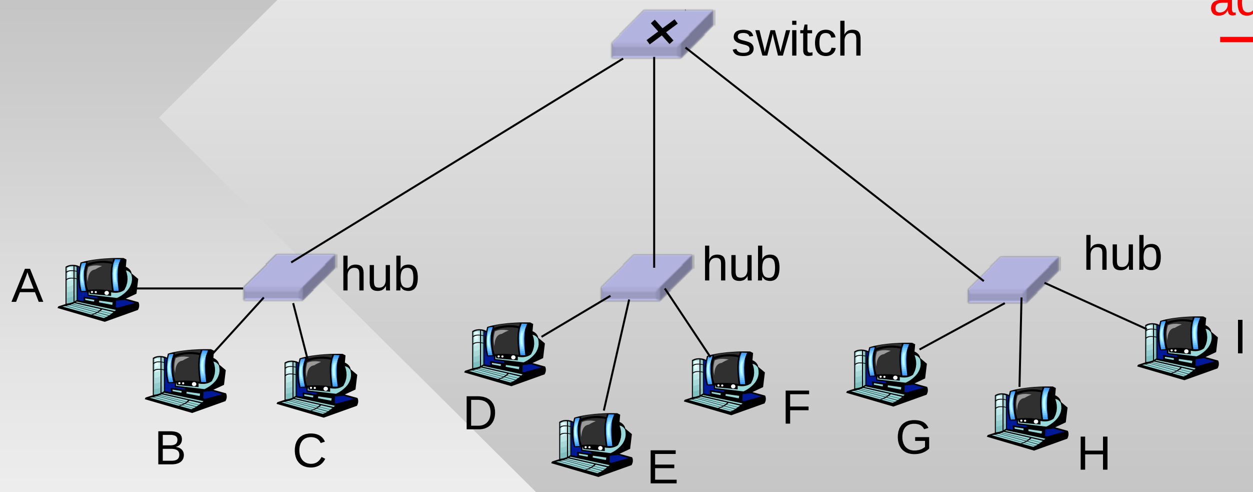


address	interface
A	1
B	1
E	2
G	3
C	1

- Switch receives frame from C
  - Notes in its table that C is on interface 1
  - Because D is not in table, switch forwards frame into interfaces 2 and 3
- Frame received by D

Switch Example

Suppose D replies back with frame to C

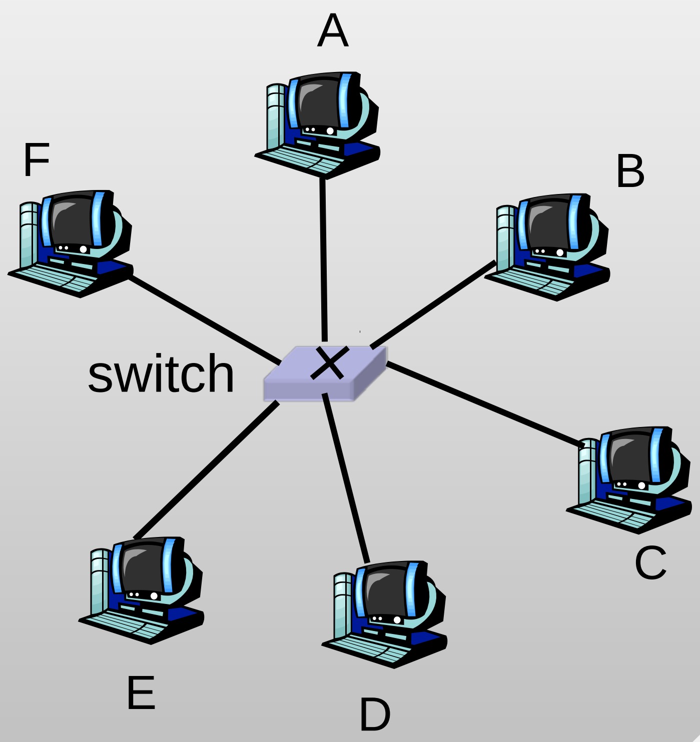


address	interface
A	1
B	1
E	2
G	3
C	1
D	2

- Switch receives frame from D
  - Notes in the table that D is on interface 2
  - Because C is in table, switch forwards only to interface 1
- Frame received by C

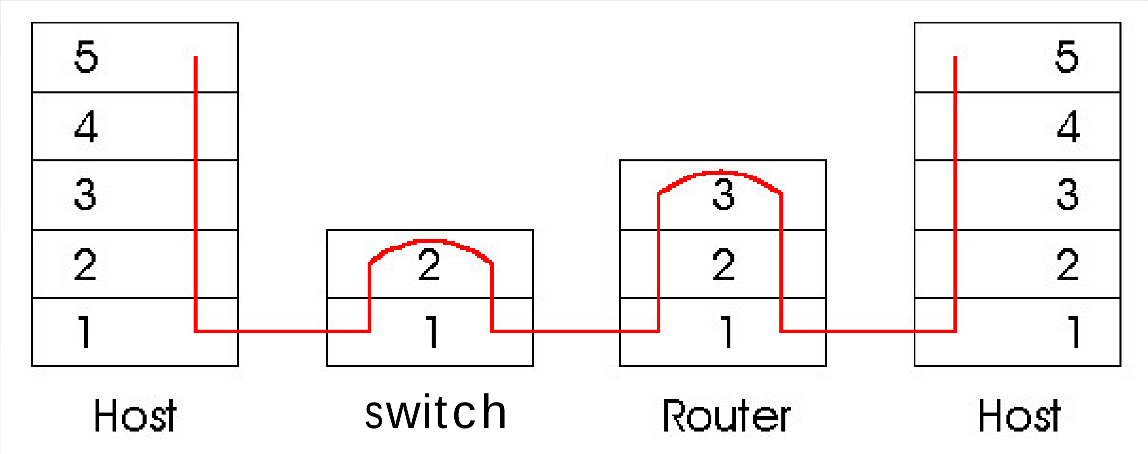
Switches: Dedicated Access

- Dedicated: hosts have direct connection to switch
  - No collisions; full duplex
- Switching: A-to-D and B-to-E simultaneous, no collisions
- Buffering: A-to-D and C-to-D simultaneous, no collisions
- Combinations of shared/dedicated and diverse (10/100/1000 Mbps) interfaces are possible



Switches vs. Routers

- Both store-and-forward devices
  - Routers: network-layer devices
  - Switches: link-layer devices



- Modern switches can perform some IP functionality
  - This violates the end-to-end principle, but makes network administration easier