Introduction to Ethernet

Outline

- Ethernet Intro
- CSMA/CD protocol
- Ethernet Framing
- Exponential backoff
- Hub, Bridge, Switch

Introduction

- History
  - Developed by Bob Metcalfe and others at Xerox PARC in mid-1970s
  - Standardized by Xerox, DEC, and Intel in 1978
  - LAN standards define MAC and physical layer connectivity
    - IEEE 802.3 (CSMA/CD - Ethernet) standard – originally 2Mbps
    - IEEE 802.3u standard for 100Mbps Ethernet
    - IEEE 802.3z standard for 1,000Mbps Ethernet
- CSMA/CD: Ethernet's Media Access Control (MAC) policy
  - CS = carrier sense
    - Send only if medium is idle
  - MA = multiple access
  - CD = collision detection
    - Stop sending immediately if collision is detected

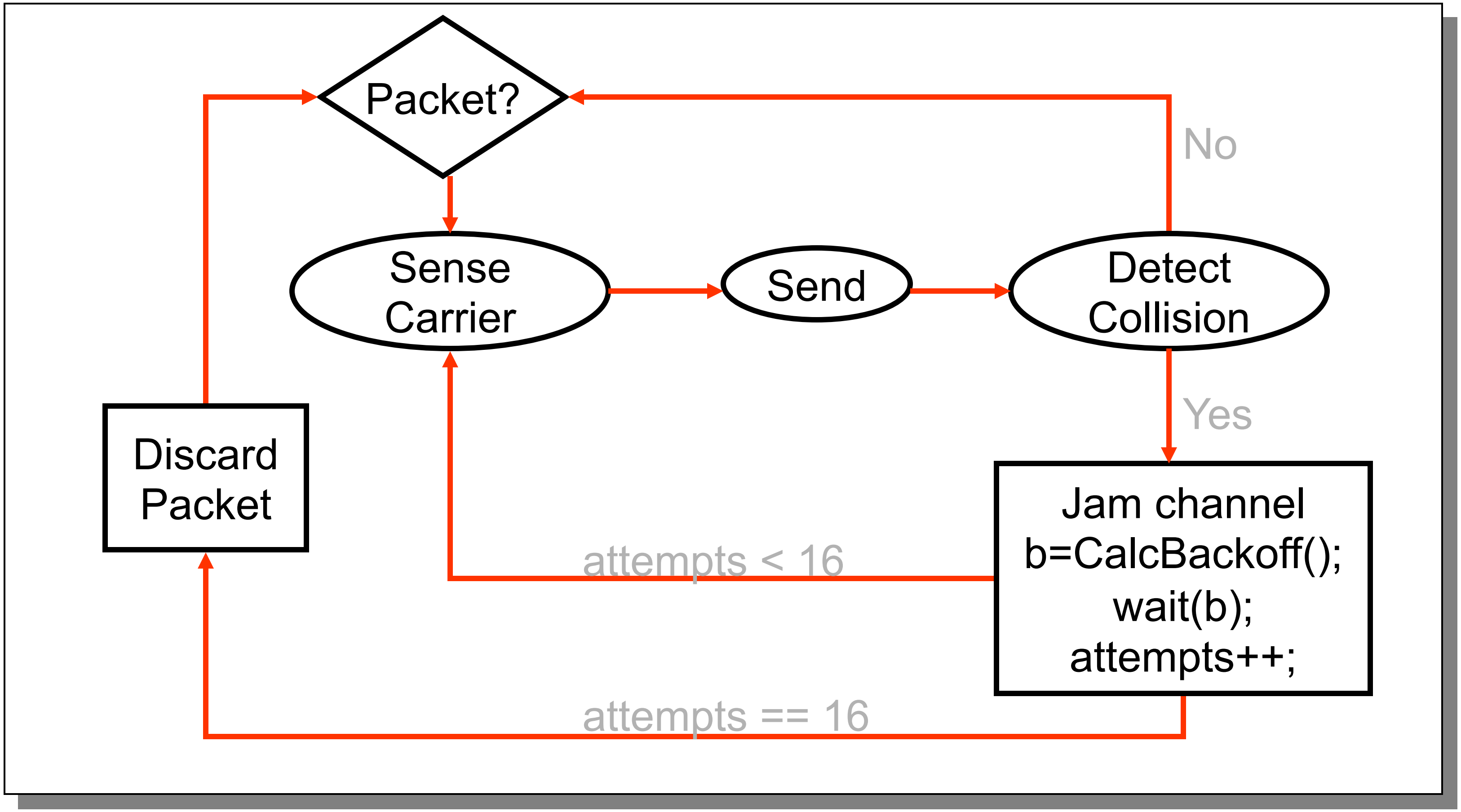
Ethernet Overview

- Most popular packet-switched LAN technology
- Bandwidths: 10Mbps, 100Mbps, 1Gbps
- Max bus length: 2500m
  - 500m segments with 4 repeaters
- Bus and Star topologies are used to connect hosts
  - Hosts attach to network via Ethernet transceiver or hub or switch
    - Detects line state and sends/receives signals
  - Hubs are used to facilitate shared connections
  - All hosts on an Ethernet are competing for access to the medium
    - Switches break this model
- Problem: Distributed algorithm that provides fair access

Ethernet's MAC Algorithm

- Ethernet uses CSMA/CD – listens to line before/during sending
- If line is idle (no carrier sensed)
  - send packet immediately
  - upper bound message size of 1500 bytes
  - must wait 9.6us between back-to-back frames
- If line is busy (carrier sensed)
  - wait until idle and transmit packet immediately
    - called *1-persistent* sending
- If collision detected
  - Stop sending and jam signal
  - Try again later

State Diagram for CSMA/CD

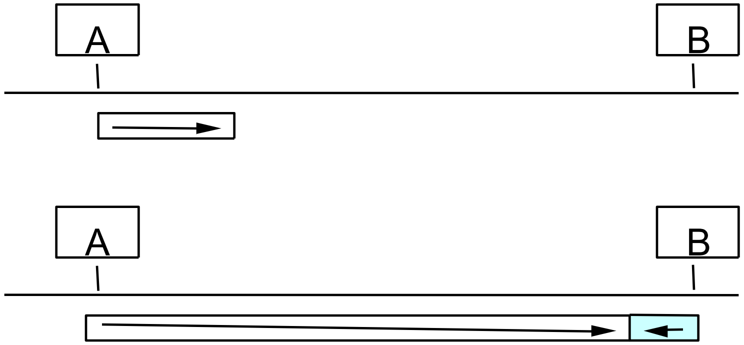


Collisions

Collisions are caused when two adaptors transmit at the same time (adaptors sense collision based on voltage differences)

- Both found line to be idle
- Both had been waiting to for a busy line to become idle

A starts at time 0



Message almost there at time T when B starts – collision!

How can we be sure A knows about the collision?

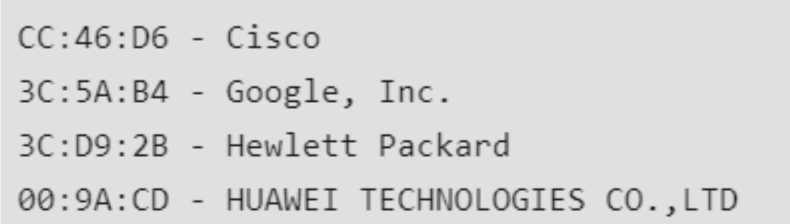
Exponential Backoff

- If a collision is detected, delay and try again
- Delay time is selected using binary exponential backoff
  - 1st time: choose K from {0,1} then delay = K * 51.2us
  - 2nd time: choose K from {0,1,2,3} then delay = K * 51.2us
  - *nth* time: delay = K x 51.2us, for K=0..2^n – 1
    - Note max value for k = 1023
  - give up after several tries (usually 16)
    - Report transmit error to host
- If delay were not random, then there is a chance that sources would retransmit in lock step
- Why not just choose from small set for K
  - This works fine for a small number of hosts
  - Large number of nodes would result in more collisions

Ethernet address (1)

- Also called "MAC address"
- Globally unique ID for each device
- Burnt into ROM, cannot be modified
- 6 Bytes in which manufacturer, device model and serial number are coded
- Readable with many auxiliary tools e.g. ipconfig /all, ifconfig

- More specifically;
  - A 12-digit hexadecimal number (6-Byte binary number), which is mostly represented by Colon-Hexadecimal notation (e.g., **3C:D9:2B:DA:71:13**)
  - First 6-digits (say 00:40:96) of MAC Address identifies the manufacturer, called as OUI (Organizational Unique Identifier).
  - The rightmost 6 digits represents **Network Interface Controller**, which is assigned by manufacturer.



```
CC:46:D6 - Cisco
3C:5A:B4 - Google, Inc.
3C:D9:2B - Hewlett Packard
00:9A:CD - HUAWEI TECHNOLOGIES CO.,LTD
```

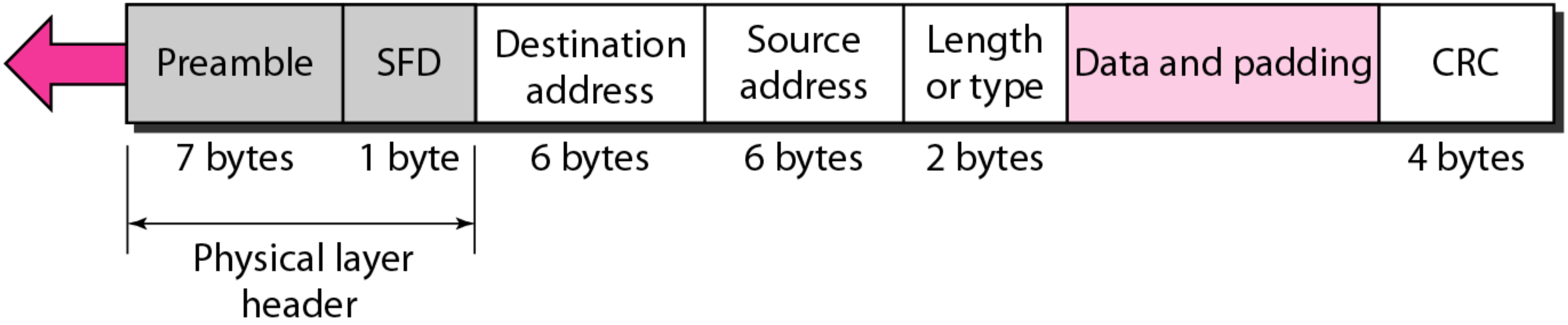
Ethernet address (2)

- The Layer 2 traffic can be classified as
  - unicast (one to one),
  - multicast (one to many), and
  - broadcast (one to all).
- Unicasts, Multicasts and Broadcasts are different types of network communication and are required for the normal operation of the network.
- MAC addresses for broadcast and multicast are given below.
  - Broadcast Destination MAC address - FF:FF:FF:FF:FF:FFF
  - Multicast Destination MAC addresses - 01:00:5E:00:00:00 to 01:00:5E:7F:FF:FF
- In case of a broadcast and multicast switch need to forward the Ethernet frame out all its ports.

Ethernet *802.3 MAC frame*

Preamble: 56 bits of alternating 1s and 0s.

SFD: Start frame delimiter, flag (10101011)



Ethernet frame

- **Preamble**
  Trailer consisting of the bit sequence “0101010101...” serving the bit synchronization of the receiver.
- **SFD** (Start Frame Delimiter)
  Start character consisting of the bit pattern “10101011” showing the recipient that the actual information will follow now.
- **DA** (Destination Address)
  Evaluated by the recipient‘s address filter; only data frames destined for this recipient will be passed on to the communication software.
- **SA** (Source Address)
  Sender‘s address
- **LEN** (Length) or **EtherType**
  LEN Indicates the length of the subsequent data field in Bytes according to IEEE 802.3. Similarly, EtherType is a two-octet field used to indicate which protocol is encapsulated in the payload of the frame and is used at the receiving end by the data link layer to determine how the payload is processed.

Ethernet frame

- **Data and Pad**
  The data field may contain 46 to 1500 user data bytes. Are there less than 46 bytes the Ethernet controller independently adds padding bytes, until the total amount (data + pad) is 46. This miminum length is crucial for the CSMA/CD procedure to work faultlessly. The data field can be used at will, it only has to contain complete bytes.

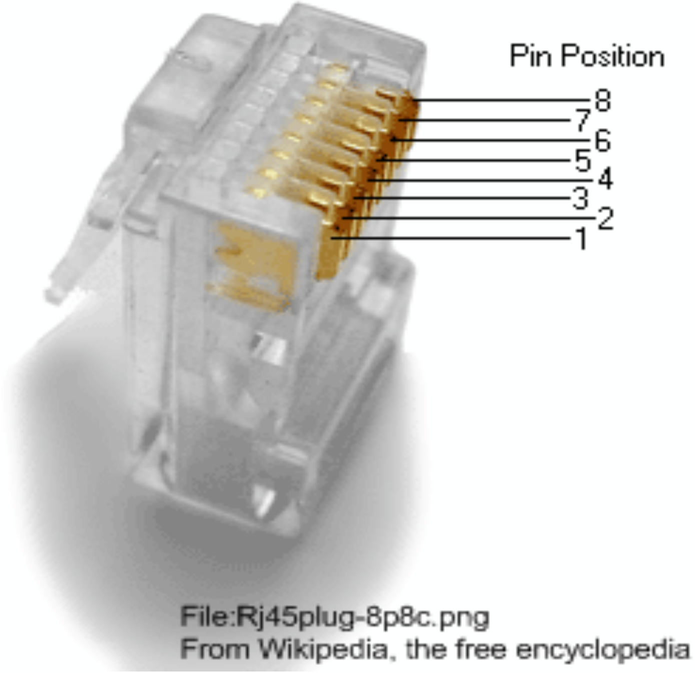
- **CRC** (cyclic redundancy check)
  A check character. It is obtained by taking the rest of the division operation from the formula representing the wide-spread cyclic- redundancy-check procedure. This formula is applied to the bit sequence including the address field through to the padding field. In case of en error the whole frame is ignored, i.e. not passed on to the application program.

Experiences with Ethernet

- Ethernets work best under light loads
  - Utilization over 30% is considered heavy
    - Network capacity is wasted by collisions
- Most networks are limited to about 200 hosts
  - Specification allows for up to 1024
- Most networks are much shorter
  - 5 to 10 microsecond RTT
- Transport level flow control helps reduce load (number of back to back packets)
- Ethernet is inexpensive, fast and easy to administer!

Ethernet Cable

Color		Pin (T568B)
⬜🟧	White/Orange	1
🟧	Orange	2
⬜🟩	White/Green	3
🟦	Blue	4
⬜🟦	White/Blue	5
🟩	Green	6
⬜🟫	White/Brown	7
🟫	Brown	8



File:Rj45plug-8p8c.png
From Wikipedia, the free encyclopedia

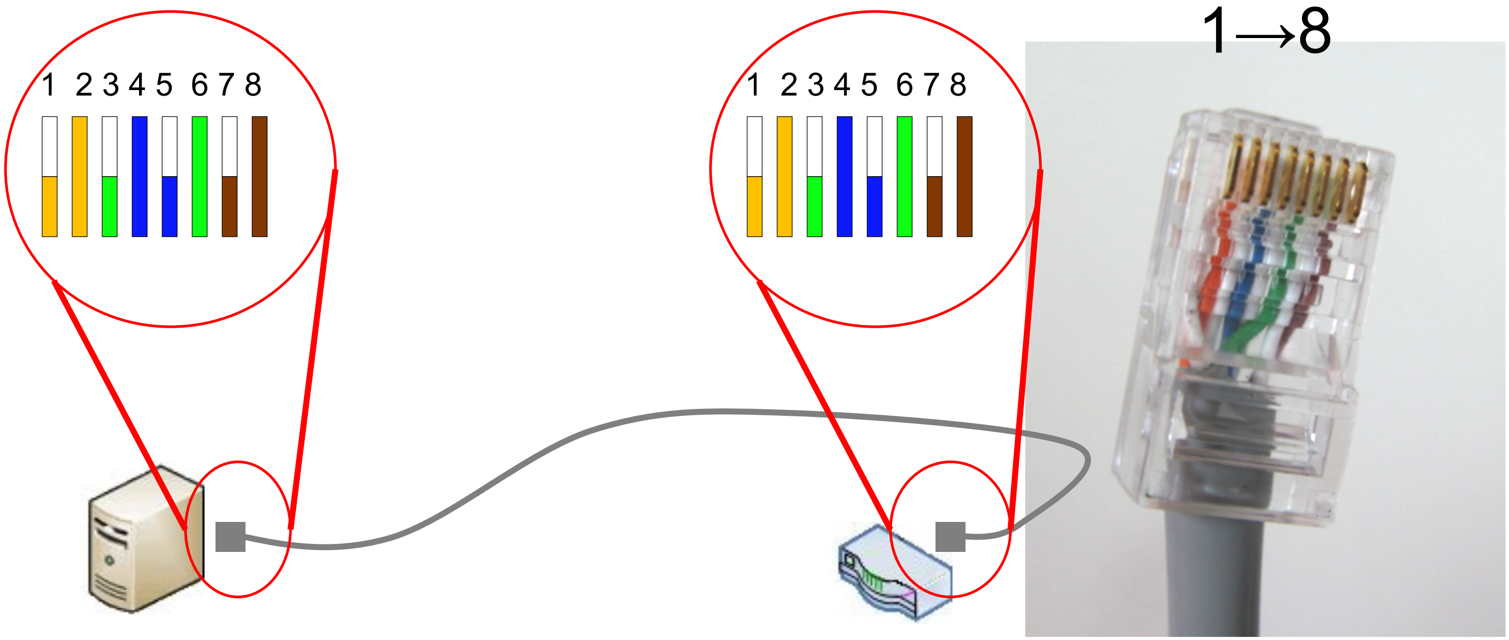
- You can use the order of rainbow colors to memorize the order of this wiring.

Ethernet Cabling (1)

Straight-Through Cable

- Host to Switch or hub
- Router to Switch or hub

- Straight Through
  - All order of the wirings is the same as the other side.

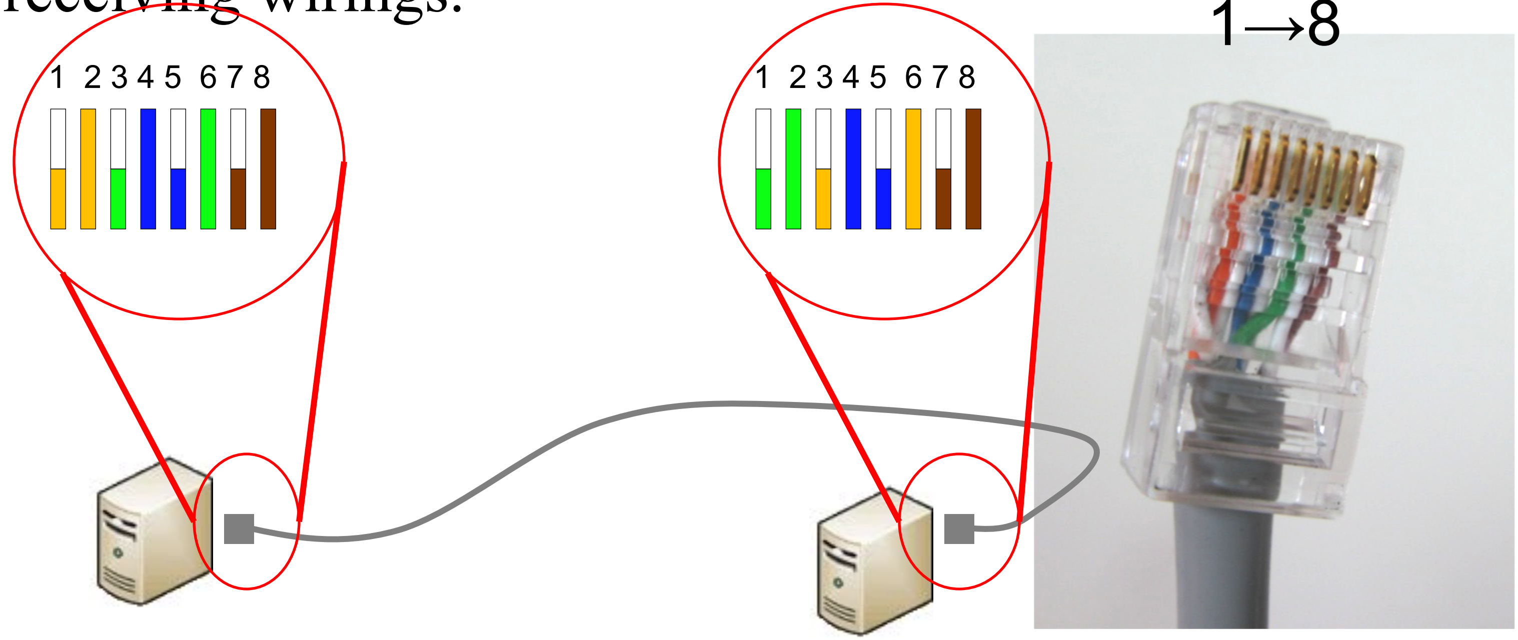


Ethernet Cabling (2)

Crossover Cable

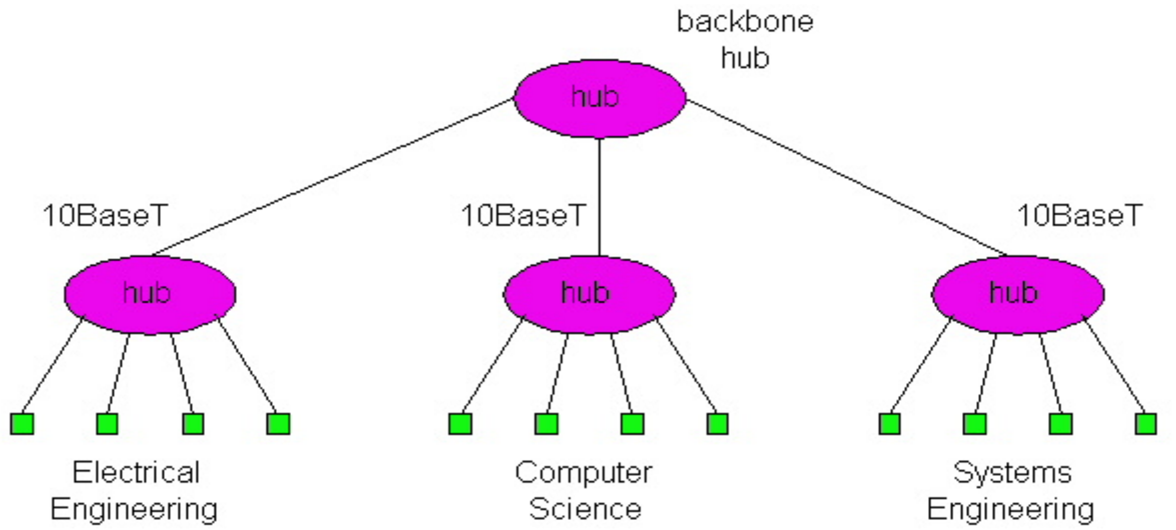
- Switch to Switch
- Hub to Hub
- Host to Host

- Crossover
  - We need to change the order of the transmission and receiving wirings.



Hubs (Multiport Repeaters) (1)

- Physical Layer devices: essentially repeaters operating at bit levels: repeat received bits on one interface to all other interfaces
- Hubs can be arranged in a hierarchy (or multi-tier design), with backbone hub at its top



Hubs (Multiport Repeaters) (2)

- Each connected LAN referred to as LAN **segment**
- Hubs do not isolate collision domains: node may collide with any node residing at any segment in LAN
- Hub Advantages:
  - simple, inexpensive device
  - Multi-tier provides graceful degradation: portions of the LAN continue to operate if one hub malfunctions
  - extends maximum distance between node pairs (100m per Hub)

Bridges (1)

- Link Layer devices: operate on Ethernet frames, examining frame header and selectively forwarding frame based on its destination
- Bridge isolates collision domains since it buffers frames
- When frame is to be forwarded on segment, bridge uses CSMA/CD to access segment and transmit

Bridges (2)

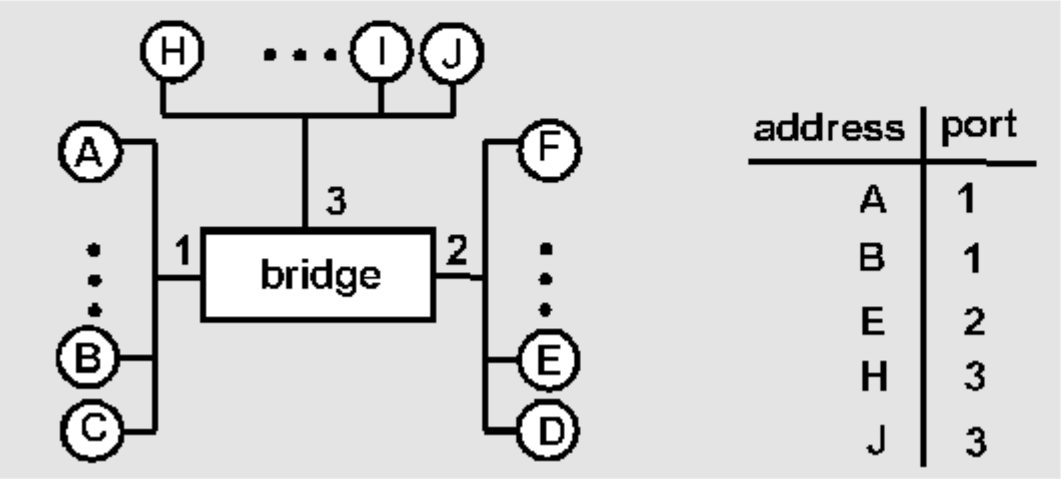
- Bridge advantages:
  - Isolates collision domains resulting in higher total max throughput, and does not limit the number of nodes nor geographical coverage
  - Transparent: no need for any change to hosts LAN adapters

Bridge Filtering

- Bridges *learn* which hosts can be reached through which interfaces: maintain filtering tables
  - when frame received, bridge “learns” location of sender: incoming LAN segment
  - records sender location in filtering table
- Filtering table entry:
  - (Node LAN Address, Bridge Interface, Time Stamp)
  - stale entries in Filtering Table dropped

Bridge Learning: example

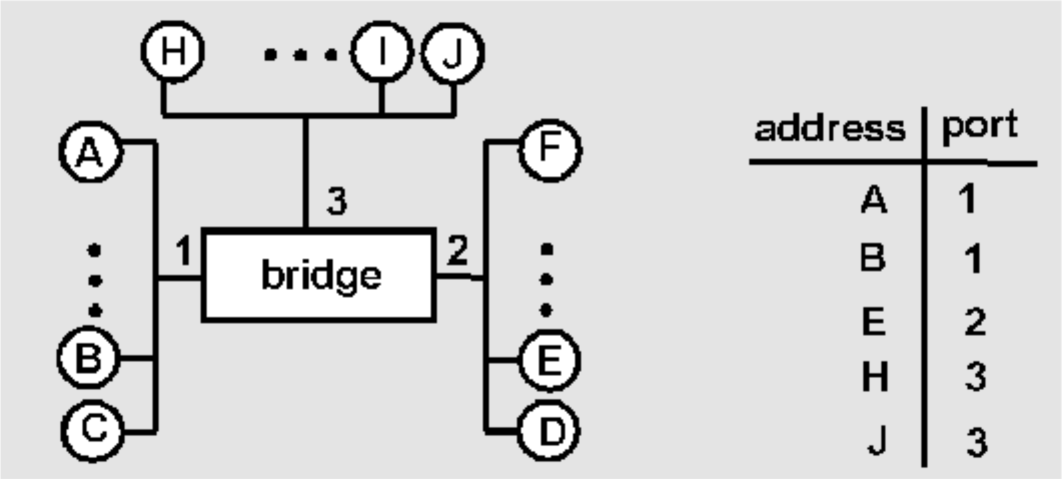
Suppose C sends frame to D and D replies back with frame to C



address	port
A	1
B	1
E	2
H	3
J	3

- C sends frame, bridge has no info about D, so broadcasts/floods to both LANs
  - bridge notes that C is on port 1
  - frame ignored on upper LAN
  - frame received by D

Bridge Learning: example

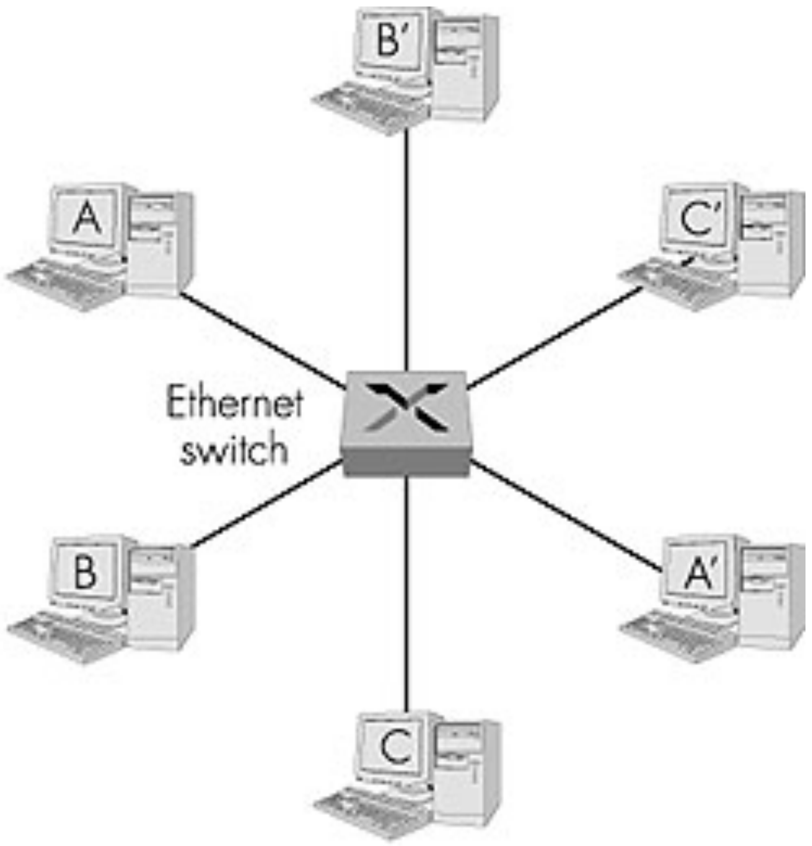


address	port
A	1
B	1
E	2
H	3
J	3
C	1

- D generates reply to C, sends
  - bridge sees frame from D
  - bridge notes that D is on interface 2
  - bridge knows C on interface 1, so *selectively* forwards frame out via interface 1

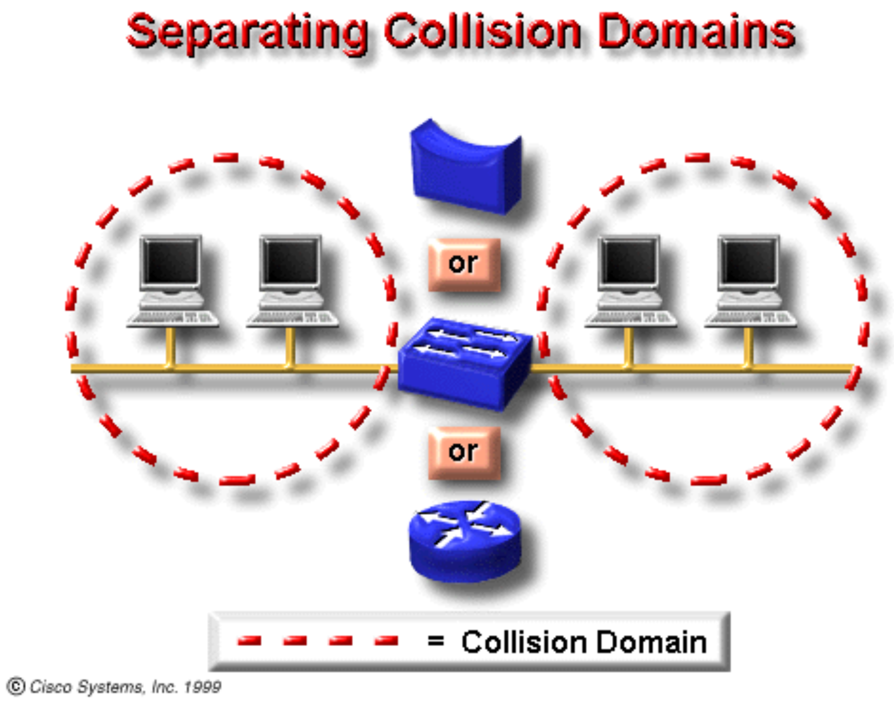
Ethernet Switches (Multiport Bridges)

- Layer 2 (frame) forwarding, filtering using LAN addresses
- Switching: A-to-B and A'-to-B' simultaneously, no collisions
- Large number of interfaces
- Often: individual hosts, star-connected into switch
  - Ethernet, but no collisions!



Collision Domain

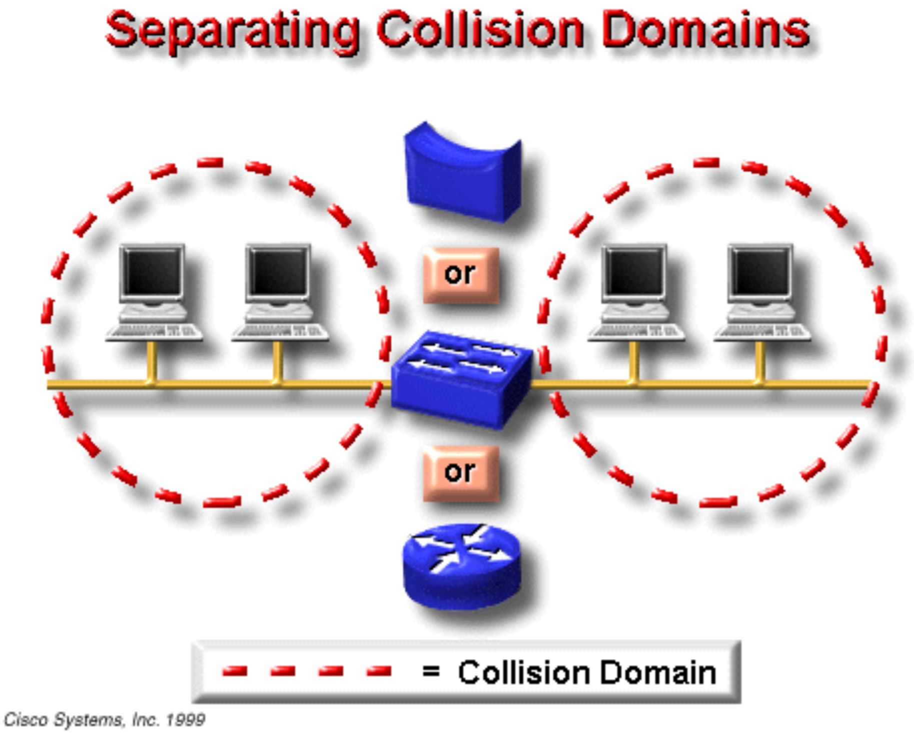
- Network region in which collisions are propagated.
- Repeaters and hubs propagate collisions.
- Bridges, switches and routers do not.



Reducing Collisions

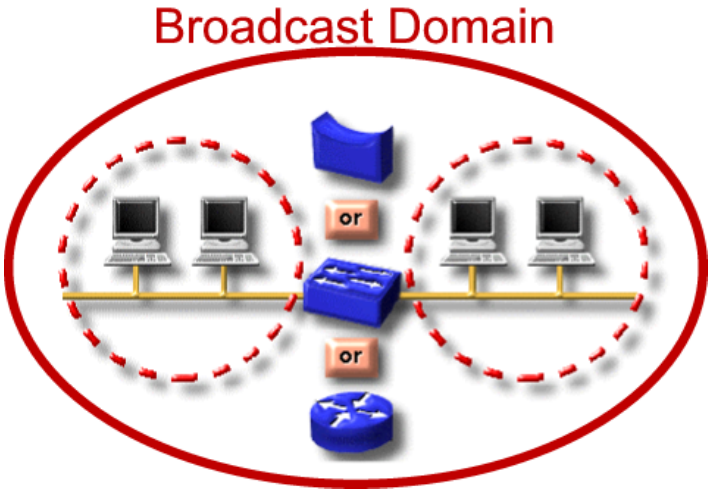
Collision frequency can be kept low by breaking the network into segments bounded by:

- bridges
- switches
- routers



Broadcast Domain

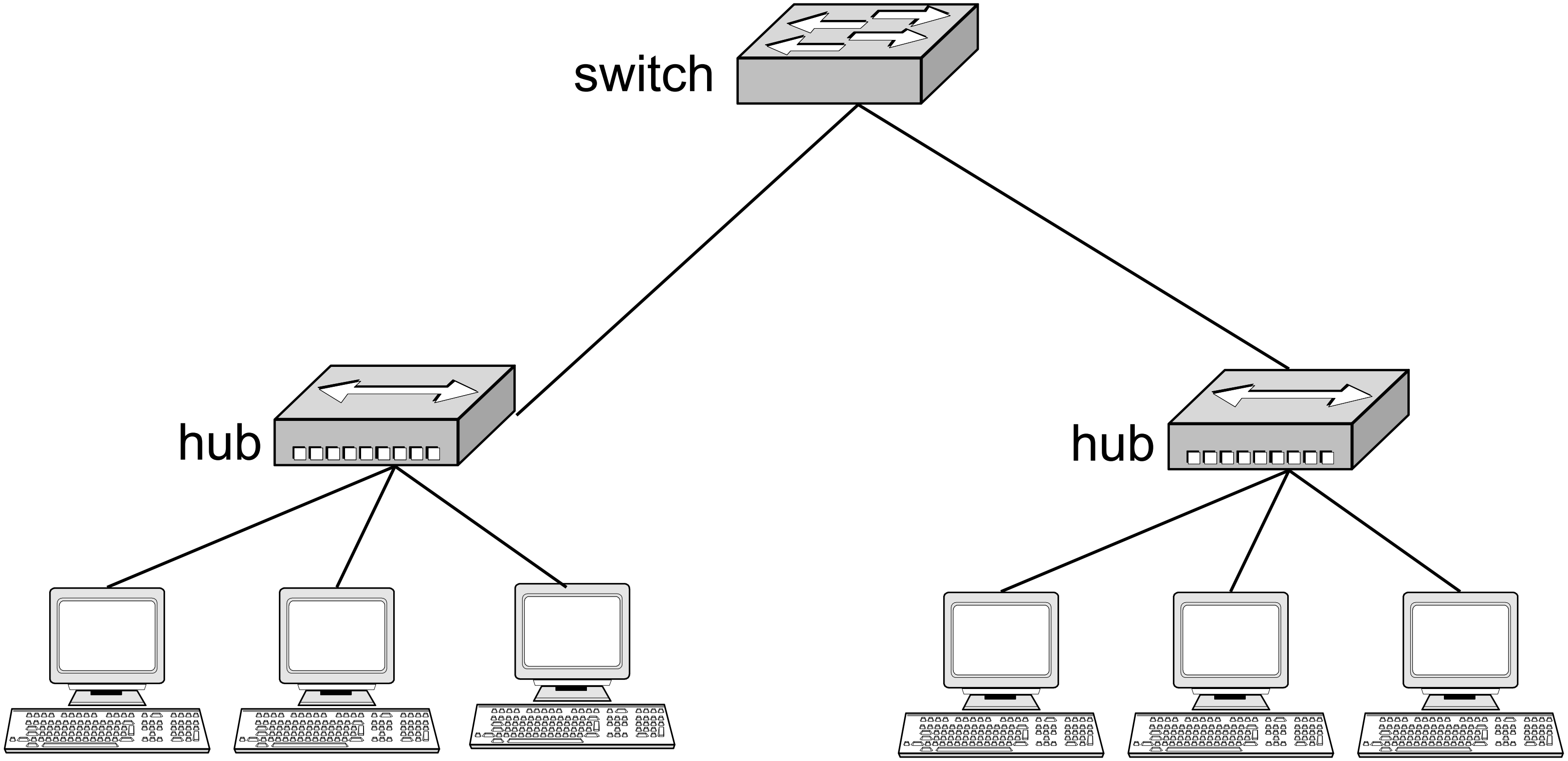
- Network region in which broadcast frames are propagated.
- Repeaters, hubs, bridges, & switches propagate broadcasts.
- Routers either do or don’t, depending on their configuration.



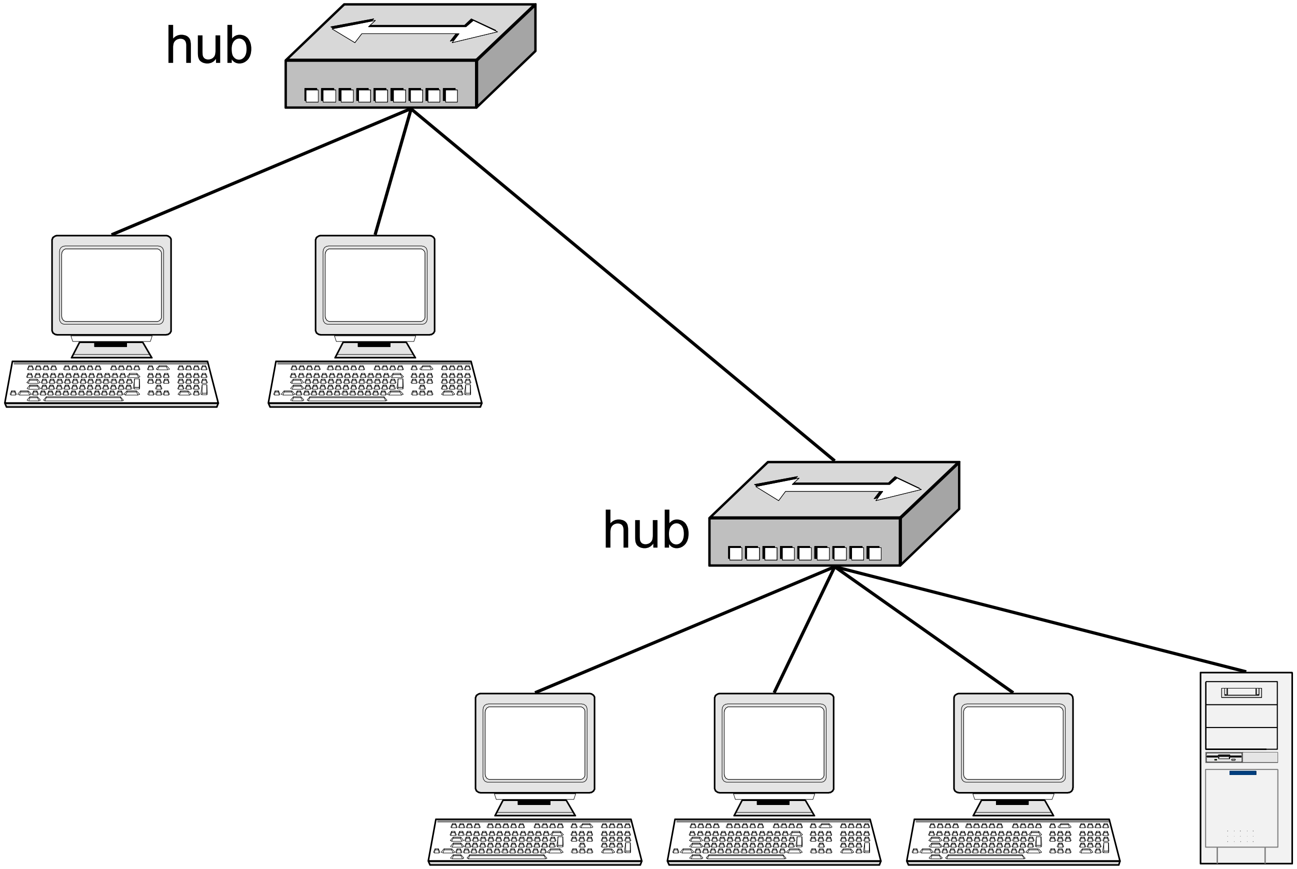
Reducing Broadcasts

- Broadcasts are necessary for network function.
- Some devices and protocols produce lots of broadcasts; avoid them.
- Broadcast frequency can be kept manageable by limiting the LAN size.
- LANs can then be cross-connected by routers to make a larger internetwork.

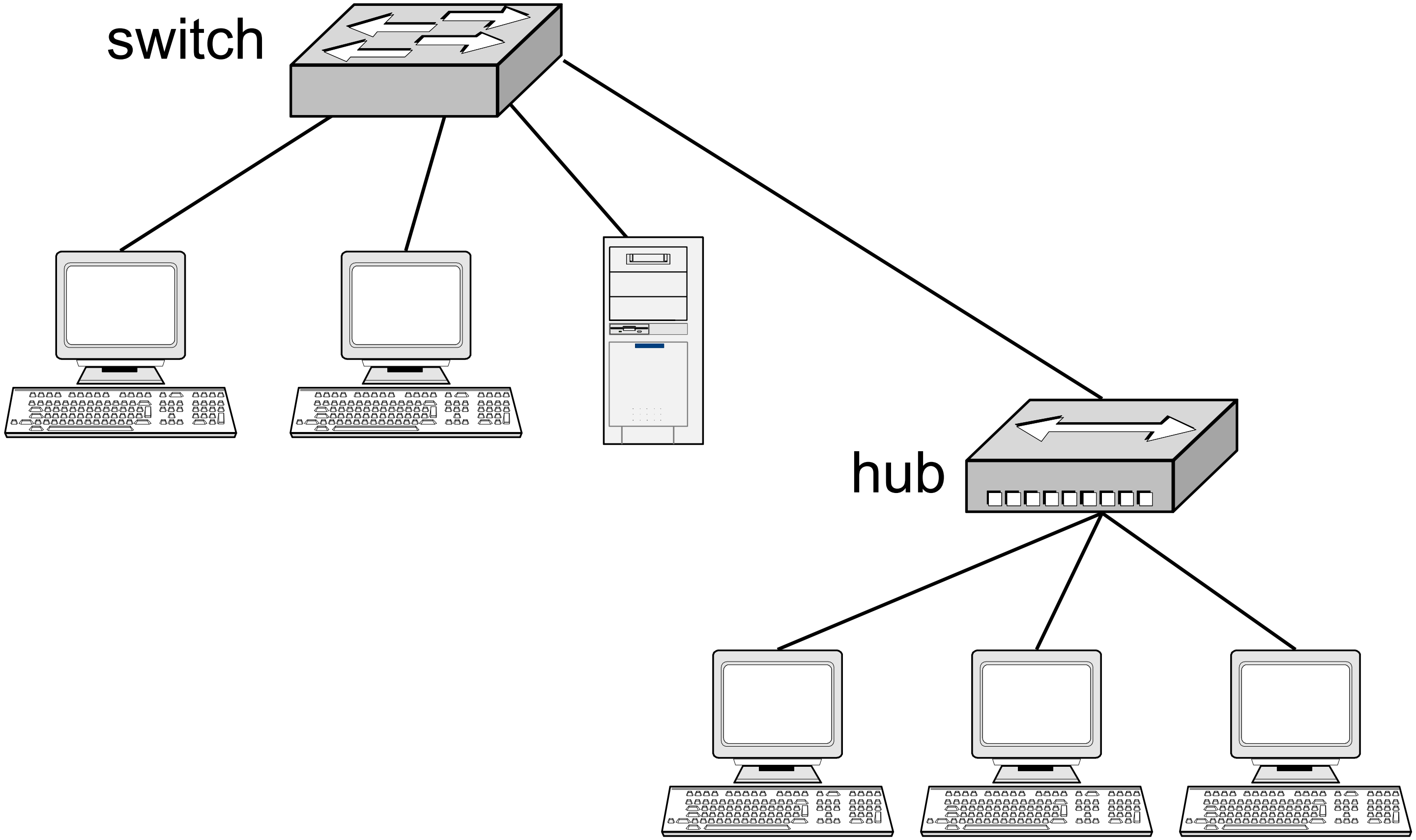
Identify the collision domains & broadcast domains:



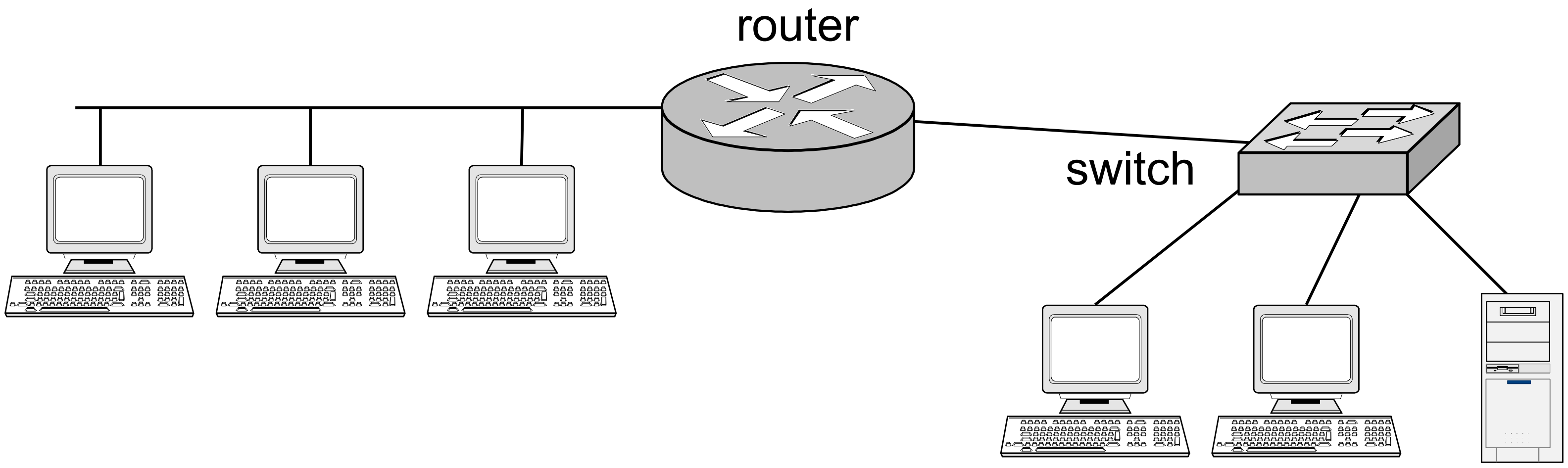
Identify the collision domains & broadcast domains:



Identify the collision domains & broadcast domains:



Identify the collision domains & broadcast domains:

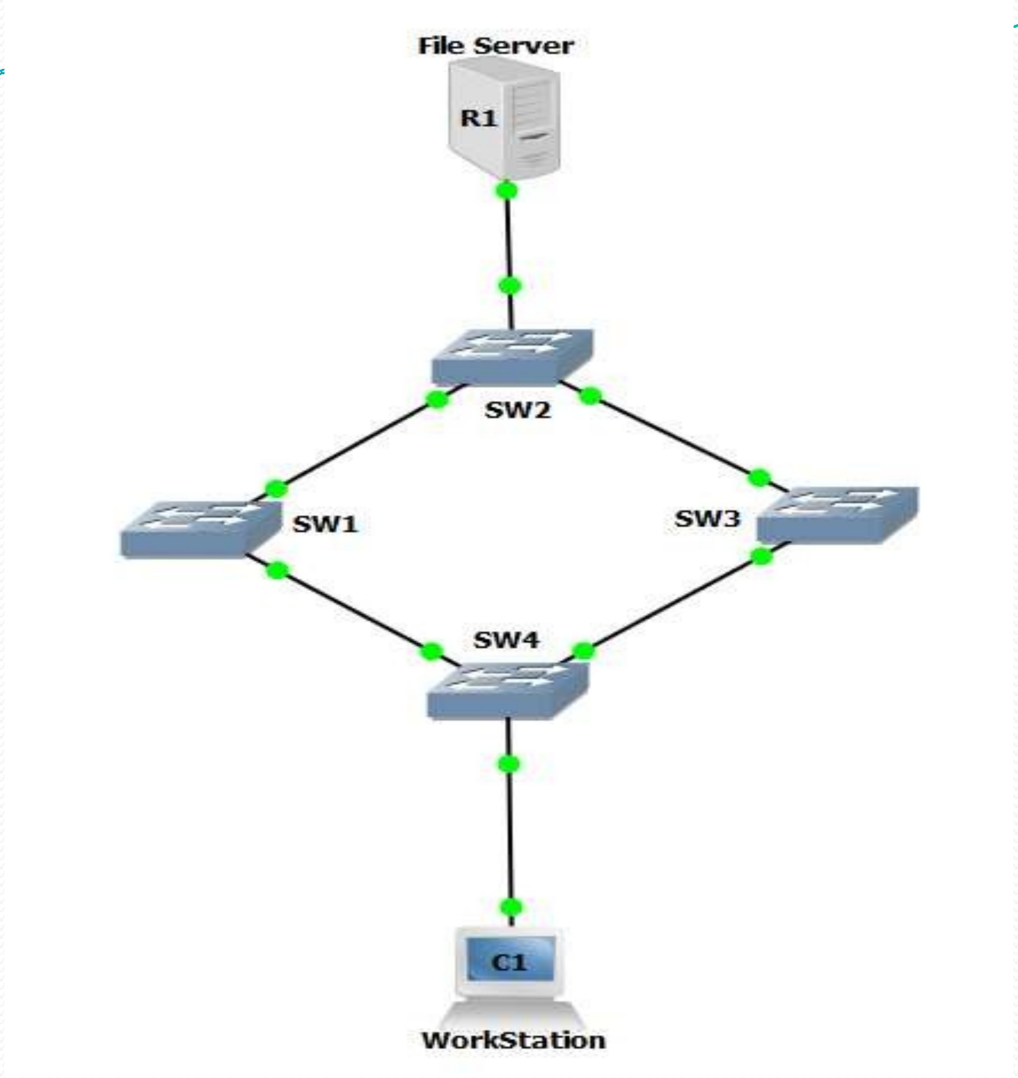


Router connects separate networks.
One broadcast domain per router interface.

Problem: Layer 2 Switching Loop (1)

- In practical Local Area Networking, it is common that the switches are interconnected for redundancy.
- When switches are interconnected, the network will not fail completely even one if the connected link fails.
- When switches are interconnected for redundancy as shown in the topology, another serious network problem can occur, which is known as Layer 2 Switching loop.

Layer 2 Switching Loop (2)



Layer 2 Switching Loop (3)

- A Ethernet frame originating from Workstation to the File Server, first reaches the Switch 4.
- Switch 4 will forward the packet to all its ports (except the source port) since the MAC address of the destination device (File Server) may not be available in its MAC address table (File Server is attached to Switch 2).

Layer 2 Switching Loop (4)

- Both Switch 1 and Switch 3 will receive a copy of the Ethernet frame. Now the Switch 1 and Switch 3 will search for the destination MAC address in its MAC address table and if they fail to find the destination MAC address in their MAC address tables, both the Switches will forward the Ethernet frame to all the ports (except the source port). This may cause the Ethernet frame to reach back the Switch 4 via path Switch 1 – Switch 3 – Switch 4 or Switch 3 – Switch 1 – Switch 4.
- This may lead to a switching loop and the Ethernet frame will start circulating the network in a loop.

Layer 2 Switching Loop (5)

- Another problem is that the File Server can receive multiple copies of the same Ethernet frame arriving via different paths, which leads to additional overhead.
- Layer 2 Switching loops may cause serious problem to network performance.
- Layer 2 Switching loops are prevented in networks using Spanning Tree Protocol.

Spanning Tree Algorithm

- Allow a path between every LAN without causing loops (*loop-free environment*)
- Bridges communicate with special configuration messages (*BPDUs*)
- Standardized by IEEE 802.1D

Note: redundant paths are good, active redundant paths are bad (they cause loops)