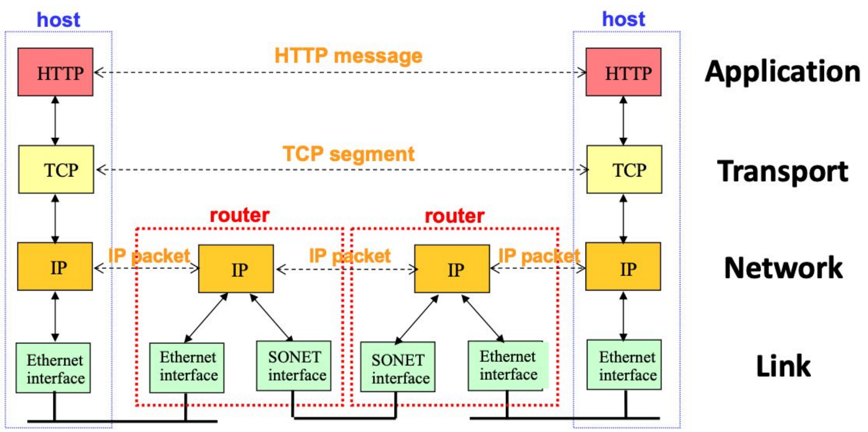
Table of contents

- Terminology
- Repeaters and hubs
- Bridges and switches
- Routers
- Broadcast and collision domain
- Summary and conclusions

Terminology

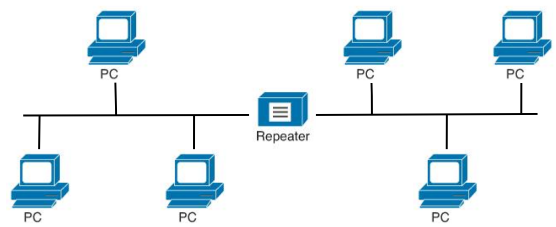
- Repeaters and hubs
  - Broadcast: All signals are sent out to all physical ports except input port
- Bridges and switches
  - Only send frames to selected physical port based on destination MAC address
- Routers
  - Only send packet to selected physical port based on destination IP address

Terminology



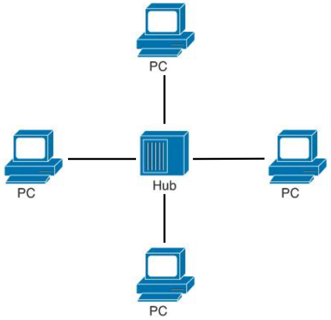
Repeaters and hubs

- Repeaters
  - Distance limitation in LAN
    - Electrical signal becomes weaker as it travels
    - Imposes a limit on length of a LAN
  - Repeaters join LANs together
    - Typically, analog electronic device
    - Continuously monitors electrical signals
    - Transmits an amplified copy
  - Layer 1 device



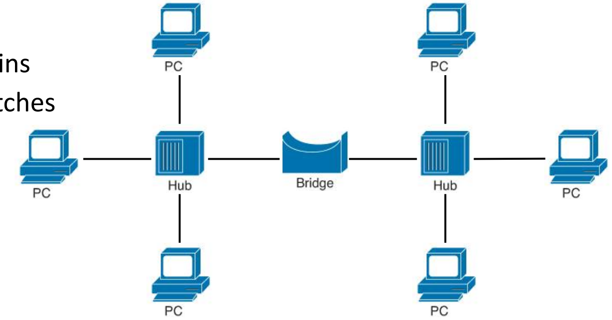
Repeaters and hubs

- Hubs
  - Layer 1 device
    - No knowledge of addresses
  - Half-duplex
    - Can't send and receive data at same time, causing collisions
  - 1 collision domain
  - Wasted bandwidth
  - Security risks
  - Multiport repeater
  - Replaced by switches



Bridges and switches

- Bridges
  - Layer 2 device
    - Can learn MAC addresses
  - Segments lans
    - Each segment can carry its own traffic
  - Fewer ports
    - Usually, two
  - 2 collision domains
  - Replaced by switches

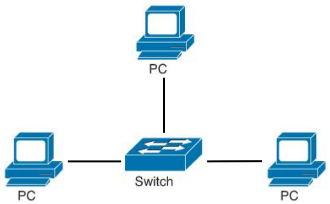


Repeaters and hubs

- Limitations of repeaters and hubs
  - One large shared link
    - Each bit is sent everywhere
      - So, aggregate throughput is limited
  - Cannot support multiple LAN technologies
    - Can’t interconnect between different rates/formats
      - Does not buffer or interpret frames
  - Limitations on maximum nodes and distances
    - Shared medium imposes length limits
      - E.g., cannot go beyond 2500 meters on Ethernet

Bridges and switches

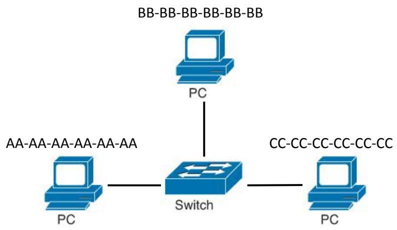
- Switches
  - Layer 2 device
    - Switch = Hub + Bridge
  - Full-duplex
  - Multiple collision domains
    - Each port has its own collision domain
  - Saves bandwidth
  - Increased security



Bridges and switches

- Switches
  - Self learning
    - Initially, MAC address table is empty

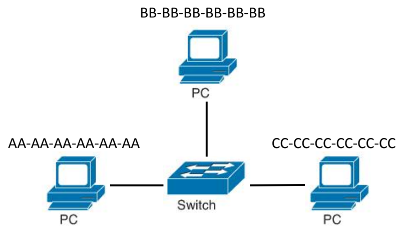
MAC address	Port



Bridges and switches

- Switches
  - Self learning: Building the table
    - C sends data to A
      - Send data to A
        - And, learn C's port number
    - ...

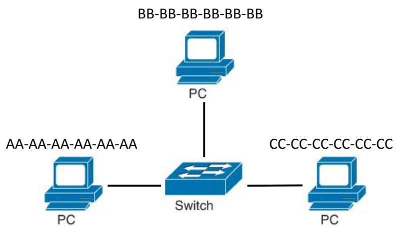
MAC address	Port
AA-AA-AA-AA-AA-AA	1
CC-CC-CC-CC-CC-CC	3
...	...



Bridges and switches

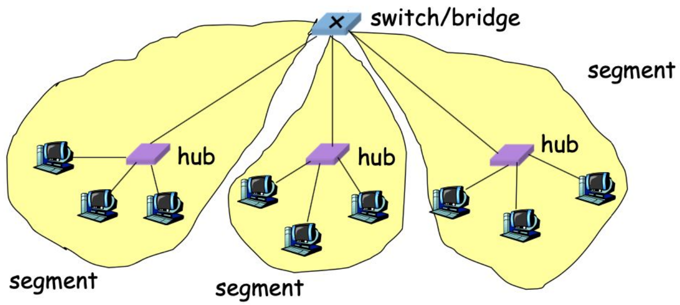
- Switches
  - Self learning: Handling misses
    - A sends data to C
      - **Broadcast** data received from A
        - And, learn A's port number

MAC address	Port
AA-AA-AA-AA-AA-AA	1



Bridges and switches

- Bridges/switches
  - Traffic isolation
    - Frame only forwarded to necessary segments
    - Segments can support separate transmissions

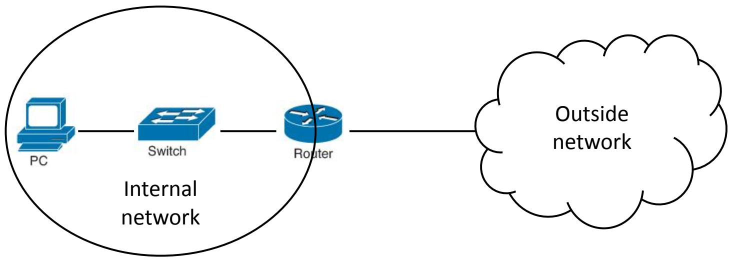


Bridges and switches

- Switches vs hubs
  - Compared to hubs, switches provide
    - (A) Higher load on links
    - (B) Less privacy
    - (C) Traffic isolation
  - Compared to hubs, Ethernet switches support
    - (A) Larger geographic span
    - (B) Similar span
    - (C) Smaller span

Routers

- Router
  - Layer 3 device
    - Routes traffic between networks
  - Fewer ports
  - Each router/switch has a forwarding table
    - Maps destination address to outgoing interface

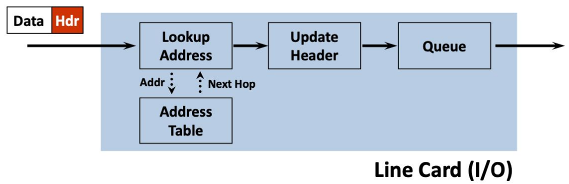


Bridges and switches

- Switches vs hubs
  - Compared to hubs, switches provide
    - (A) Higher load on links
    - (B) Less privacy
    - **(C) Traffic isolation**
  - Compared to hubs, Ethernet switches support
    - **(A) Larger geographic span**
    - (B) Similar span
    - (C) Smaller span

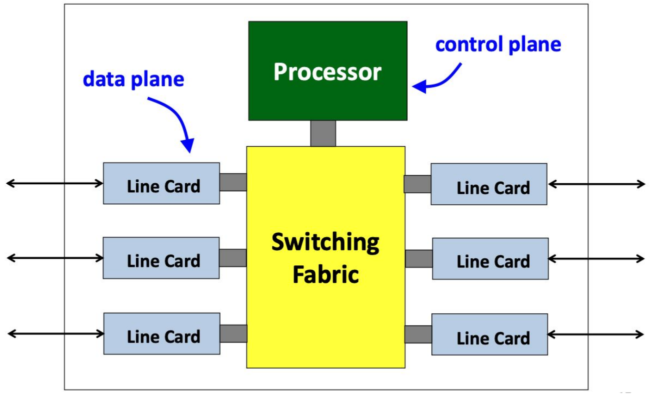
Routers

- Basic operation
  - Receive packet
  - Look at header to determine destination address
  - Look in forwarding table to determine output interface
  - Modify packet header (e.g., decrease TTL, update checksum)
  - Send packet to output interface



Routers

- Key aspects
  - Placement, behaviour of lookup tables
  - Design of switching fabric



Routers

- Longest Prefix Match (LPM)
  - Each packet has destination IP address
    - Router looks up table entry that matches address
      - Exact match is difficult to keep
    - Use longest address prefix that matches destination address
      - Finds most specific route (that has fewest hosts)

Routers

- Lookup algorithm depends on protocol

Protocol	Mechanism	Techniques
Ethernet (48-bit) MPLS ATM	Exact match	Direct lookup Associative lookup Hashing Binary tree
IPv4 (32-bit) IPv6 (128-bit)	Longest-prefix match	Radix trie Compressed trie TCAM

Routers

- Longest Prefix Match (LPM)
  - Example 1

Destination	Output
11001000 00010111 00010*** *********	1
11001000 00010111 00011000 *********	2
11001000 00010111 00011*** *********	3

11001000 00010111 00010110 10100001

Routers

- Longest Prefix Match (LPM)
  - Example 1

Destination	Output
11001000 00010111 00010*** *********	1
11001000 00010111 00011000 *********	2
11001000 00010111 00011*** *********	3

11001000 00010111 00010110 10100001

Routers

- Longest Prefix Match (LPM)
  - Example 2

Destination	Output
11001000 00010111 00010*** *********	1
11001000 00010111 00011**000** *********	2
11001000 00010111 00011*** *********	3

11001000 00010111 00011**000** 10101010

Routers

- Longest Prefix Match (LPM)
  - Example 1

Destination	Output
11001000 00010111 0001**0**** *********	1
11001000 00010111 00011000 *********	2
11001000 00010111 00011*** *********	3

11001000 00010111 0001**0**110 10100001

Routers

- Longest Prefix Match (LPM)
  - Example 3

Prefix	Output
192.168.20.16/28	1
192.168.0.0/16	2

192.168.20.191

Routers

- Longest Prefix Match (LPM)
  - Example 3

Prefix	Binary	Output
192.168.20.16/28	11000000.10101000.000**10100**.00010000	1
192.168.0.0/16	11000000.10101000.00000000.00000000	2
192.168.20.191	11000000.10101000.000**10100**.10111111	

Routers

- Longest Prefix Match (LPM)
  - Example 4

Prefix	Binary	Output
68.208.0.0/12	01000100.11010000.00000000.00000000	1
68.211.0.0/17	01000100.11010011.**00000**000.00000000	1
68.211.128.0/19	01000100.11010011.10000000.00000000	2
68.211.160.0/19	01000100.11010011.10100000.00000000	2
68.211.192.0/18	01000100.11010011.11000000.00000000	1
68.211.6.120	01000100.11010011.**00000**110.01111000	

Routers

- Longest Prefix Match (LPM)
  - Example 4

Prefix	Output
68.208.0.0/12	1
68.211.0.0/17	1
68.211.128.0/19	2
68.211.160.0/19	2
68.211.192.0/18	1

68.211.6.120

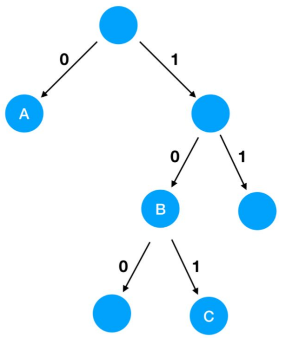
Routers

- Benefits of CIDR allocation and LPM
  - Efficiency
    - Prefixes can be allocated at much finer granularity
  - Hierarchical aggregation
    - Upstream ISP can aggregate 2 contiguous prefixes from downstream ISPs to shorter prefix

Routers

- Software LPM lookup using binary trie
  - Routing table and corresponding binary trie

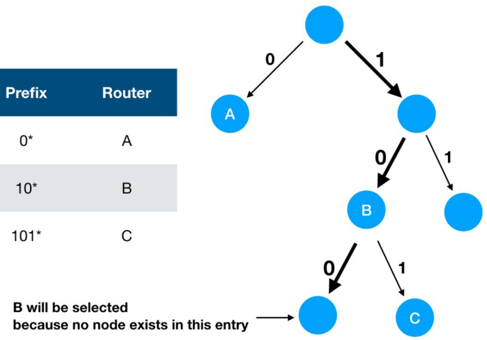
Prefix	Router
0*	A
10*	B
101*	C



Routers

- Software LPM lookup using binary trie
  - Input: 1000
  - Output: B

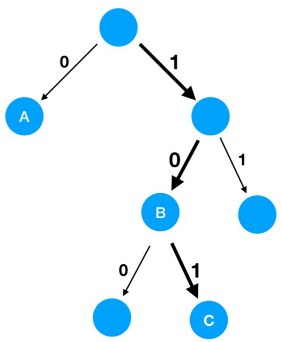
Prefix	Router
0*	A
10*	B
101*	C



Routers

- Software LPM lookup using binary trie
  - Input: 1011
  - Output: C

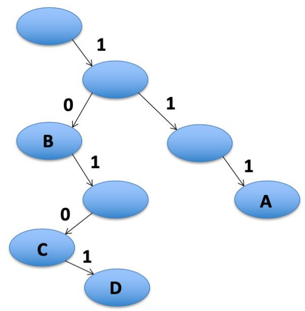
Prefix	Router
0*	A
10*	B
101*	C



Routers

- Software LPM lookup using binary trie
  - Prefixes spelled out by following path from root
    - Looking up is very fast
  - Inefficient due to number of memory access
    - Visit up to 32 nodes for an IPv4 address

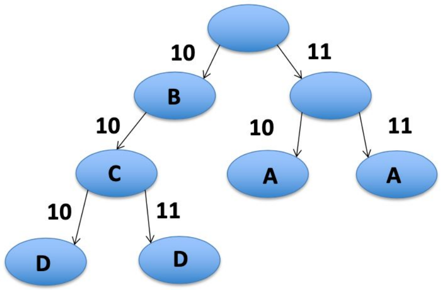
	Prefixes
A	111*
B	10*
C	1010*
D	10101



Routers

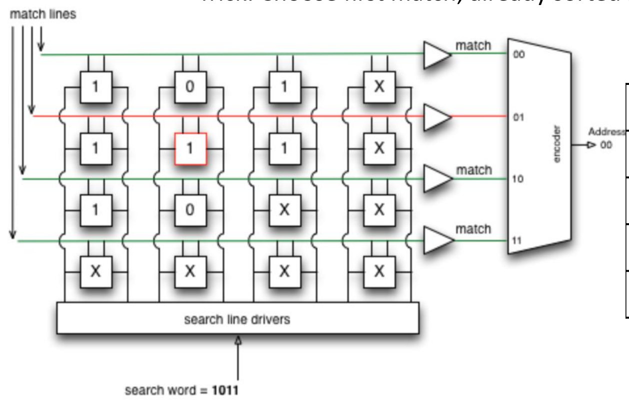
- Software LPM lookup using direct trie
  - Multiple bits are represented by one node
    - Reduces depth, looking up needs fewer memory accesses
  - Reduces number of memory access
    - But, consumes a significant amount of memory

	Prefixes
A	111*
B	10*
C	1010*
D	10101



Routers

- LPM via TCAM
  - In parallel, search all prefixes for all matches
    - Then, choose longest match
      - Trick: Choose first match; already sorted by prefix length



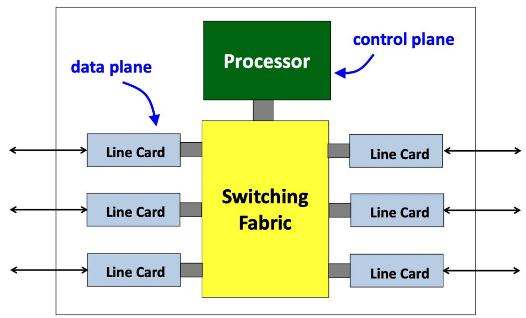
Prefix	TCAM format
101/3	101X
111/3	111X
10/2	10XX
0/0	XXXX

Routers

- Hardware for LPM lookup
  - Content-Addressable Memory (CAM)
    - Input: Tag (address)
    - Output: Value (port)
    - Exact match, but O(1) in hardware
  - Ternary Content-Addressable Memory (TCAM)
    - Can have wildcards: 0, 1, X (don't care)
    - “value” memory cell and “mask” (care/don’t care) cell
      - E.g., If stored word is "10XX0"
        - Then, TCAM will match any of four search words
          - "10000", "10010", "10100", or "10110"

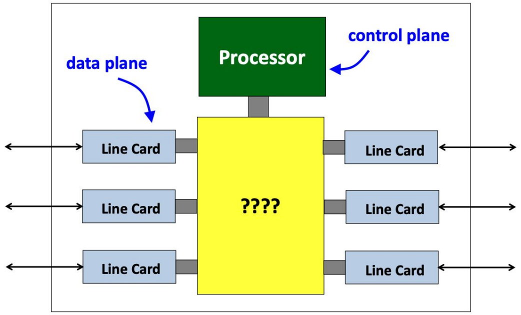
Routers

- Decision
  - Forwarding tables on line cards
    - Each line card has its own forwarding table copy
      - Prevents central table bottleneck



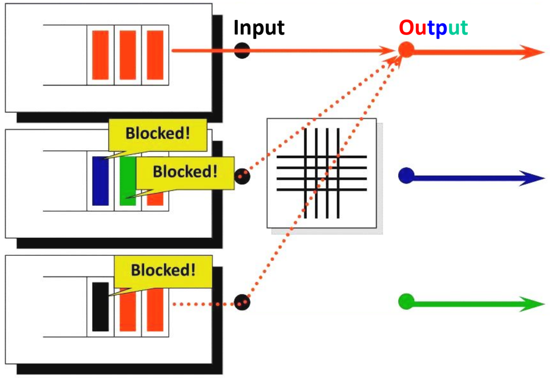
Routers

- Decision
  - Switching fabric
    - Shared bus
      - Only one input can speak to one output at a time
    - Crossbar switch/switched backplane
      - Input/output pairs that don't compete can send in same time slot



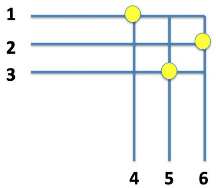
Routers

- Crossbar switching
  - Problem: Head-of-Line (HoL) blocking
    - Packet in front of queue blocks packets behind it from being processed

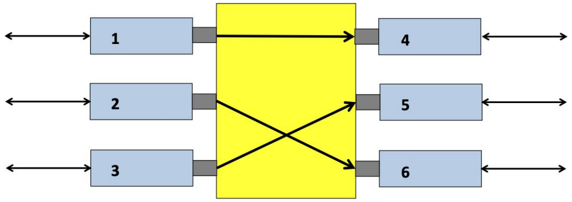


Routers

- Crossbar switching
  - Every input port has connection to every output port
    - Good parallelism
    - Needs scheduling

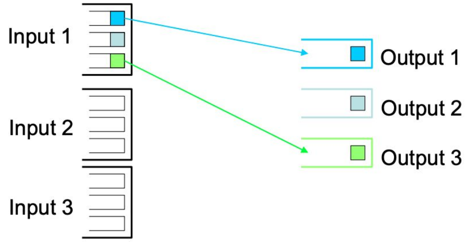


  - In each time slot, each input connected to zero or more outputs



Routers

- Crossbar switching
  - Problem: Head-of-Line (HoL) blocking
    - Solution: Virtual output queues
      - Instead of a single queue at each input
        - Maintain N (one per output) virtual queues at each input

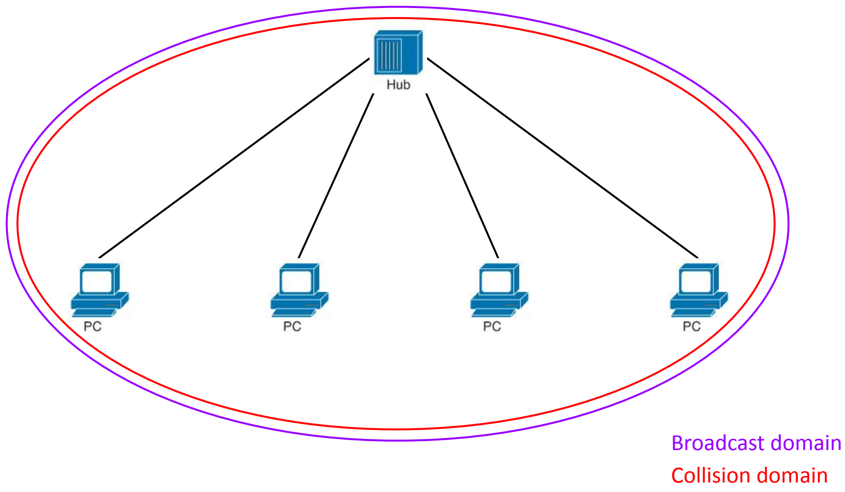


Broadcast and collision domain

- Broadcast domain
  - Includes all hosts that a broadcast frame transmitted by a single host can reach
- Collision domain
  - Section of network where packet collisions can occur if two nodes attempt to communicate at same time

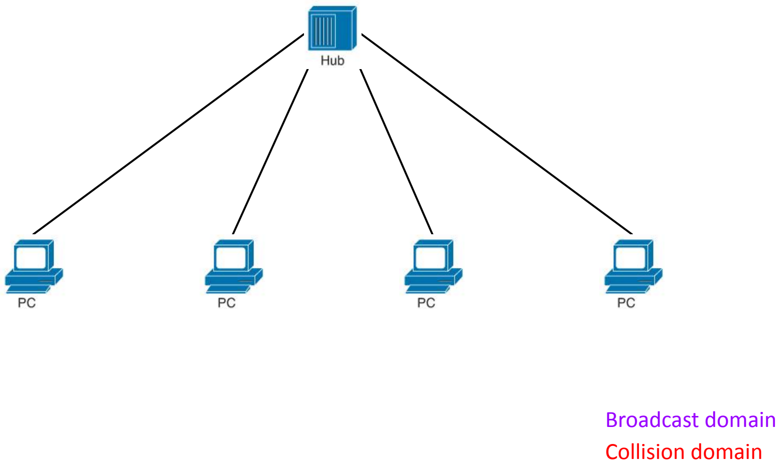
Broadcast and collision domain

- Hub



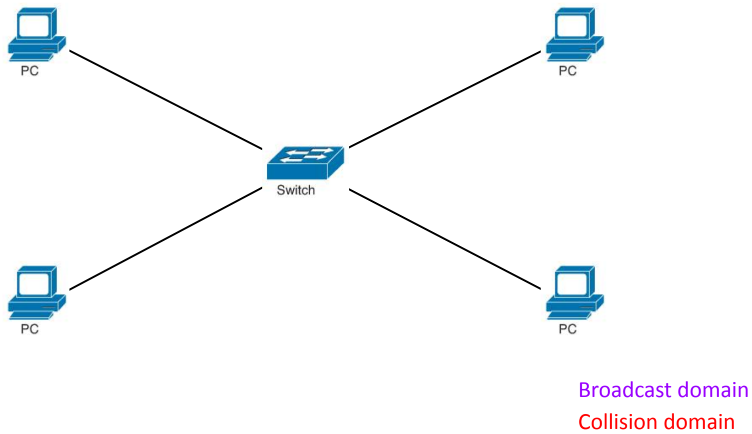
Broadcast and collision domain

- Hub



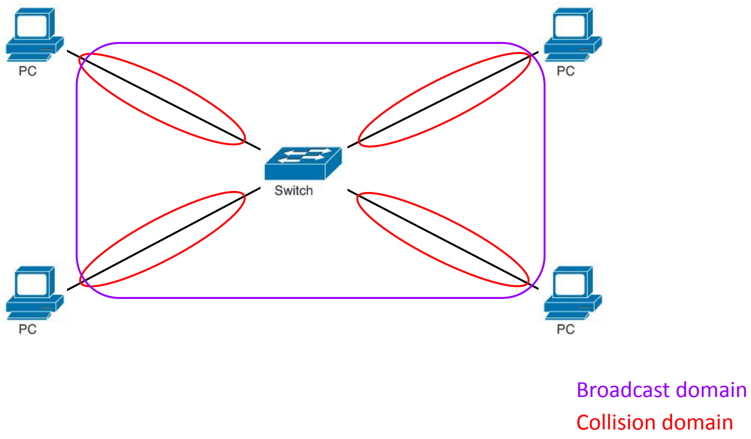
Broadcast and collision domain

- Switch



Broadcast and collision domain

- Switch



Broadcast and collision domain

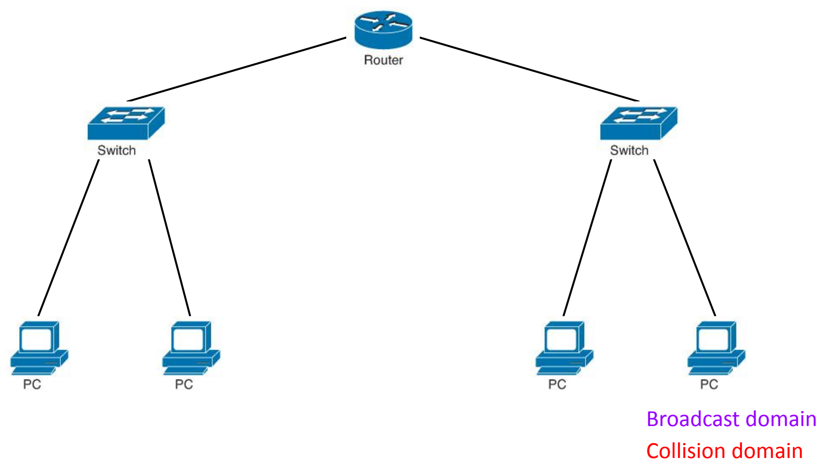
- Router
  - Breaks/separates collision as well as broadcast domains
    - Broadcast message from one network will never reach other one

Broadcast and collision domain

- Switch
  - All switch interfaces can be full-duplex
    - We can’t get collisions
      - CSMA/CD is disabled on these interfaces
  - Why do we still care about collision domains?
    - A hub may be connected to switch
      - Switch interface to hub will be in half-duplex and CSMA/CD
        - Leading to collisions
    - A defective network card or interface sending trash
  - Switch is a collision domain separator
    - Each switch interface is a separate collision domain

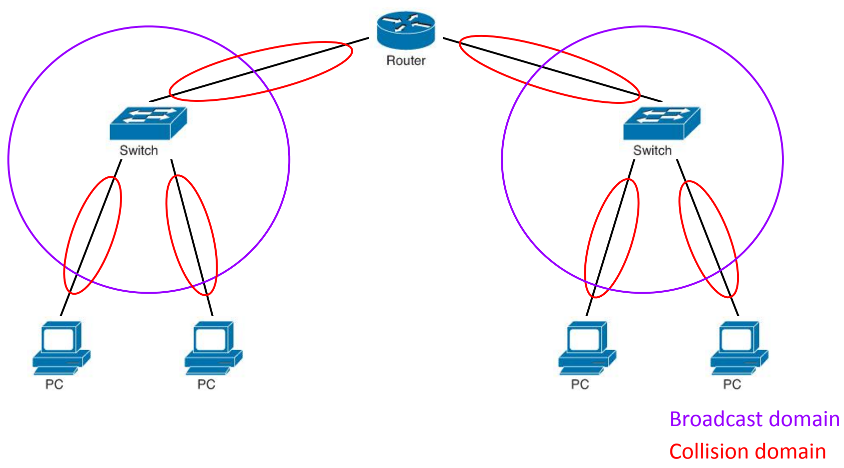
Broadcast and collision domain

- Switch and router



Broadcast and collision domain

- Switch and router



Broadcast and collision domain

- Modern networks
  - Wired networks use switches to reduce or eliminate collisions
    - Half-duplex links
      - Each switch port becomes its own collision domain
    - Full-duplex links
      - Possibility of collisions is eliminated entirely
    - For Gigabit Ethernet and faster
      - No hubs or repeaters exist
      - All devices require full-duplex links
  - Collision domains are also found in other shared medium networks
    - E.g., wireless networks such as Wi-Fi

Summary and conclusions

- L2 and L3 physical devices have many common features
  - Forward table lookups
  - Queueing and backplane switching
  - Fast vs slow paths
    - Switches and routers separate control plane from data plane
- High speed necessitates innovation
  - Specialized hardware
  - Software algorithms