Week 9: From Ethernet to IEEE 802.3

EE3017/IM2003 Computer Communications

School of Electrical and Electronic Engineering

Prof. Cheng Tee Hiang

Room: S1-B1a-29

Email: ethcheng@ntu.edu.sg

Phone: 6790-4534

Topic Outline
(Updated in January 2020)

Topic Outline

From Ethernet to IEEE 802.3

- Ethernet
- IEEE 802.3
- Bridge/Switch Learning
- Virtual Local Area Network (VLAN)

Recommended reading:
Section 5.5, Pages 501 to 523 of the recommended textbook
(Page numbers are based on 5th or 2010 Edition.)

Learning Objectives

Learning Objectives

By the end of this topic, you should be able to:

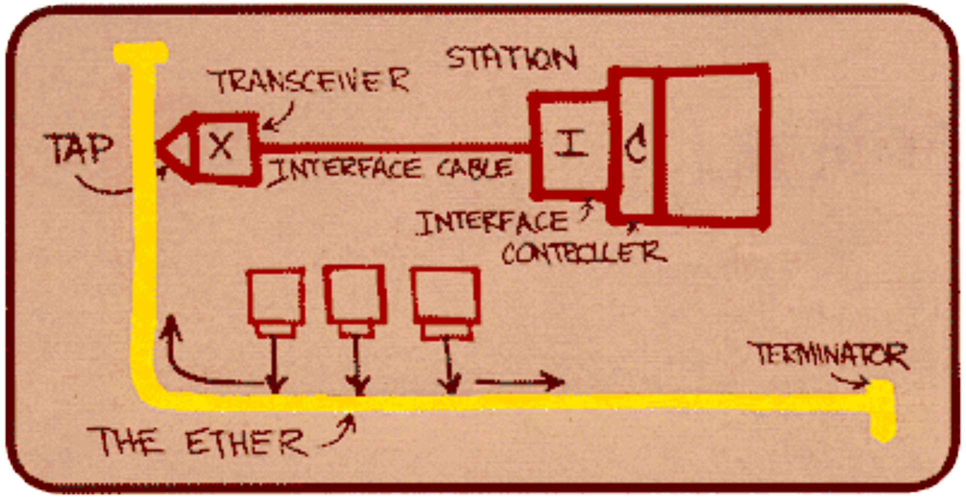
- Explain the old-fashion Ethernet.
- Explain the functions of the various fields in the Ethernet frame structure.
- Explain the CSMA/CD algorithm.
- Compute the efficiency of CSMA/CD network.
- Explain the differences between hubs and switches.
- Explain about the forwarding function and self-learning function of a switch.
- Explain the virtual local area networks (VLAN).

Ethernet

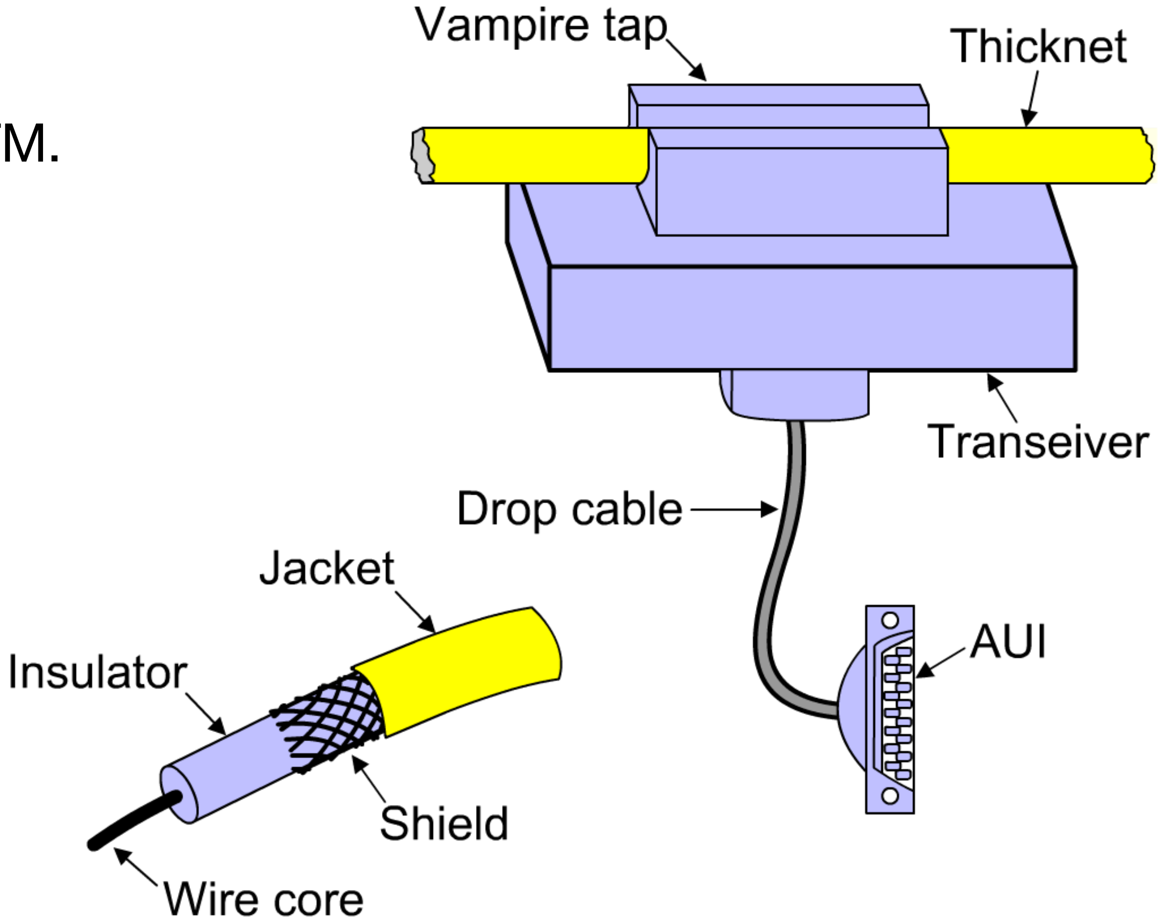
Ethernet

"Dominant" wired LAN technology:

- Cheap $20 for NIC.
- First widely used LAN technology.
- Simpler, cheaper than token LANs, FDDI and ATM.
- Kept up with speed race: 10 Mbps – 10 Gbps.

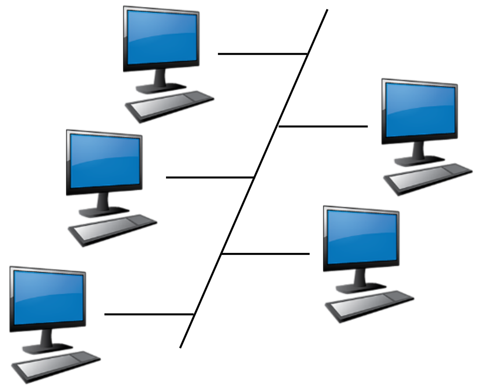


Metcalfe's Ethernet sketch

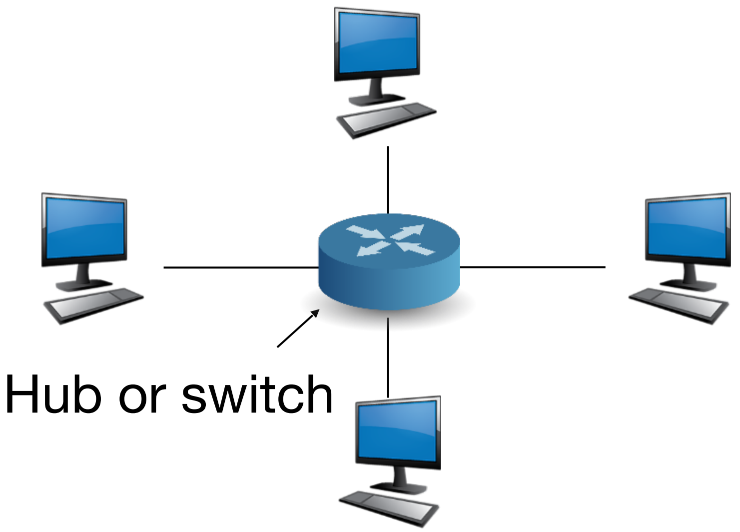


Ethernet Topology

- Bus topology popular from 70s to mid 90s.
  - all nodes in same collision domain (can collide with each other).
- Hub-based star topology emerged in 80s.
- Today, Switch-based star topology prevails
  - active switch in centre.
  - each “spoke” runs a (separate) Ethernet protocol (nodes do not collide with each other).



Bus: coaxial cable



Star

Notes: Ethernet and Ethernet Topology

- Ethernet is by far the most prevalent wired LAN technology.

- There are many reasons for Ethernet's success. First, it was the first widely deployed high-speed LAN and network administrators became intimately familiar with it and were reluctant to switch to other LAN technologies.

- Second, token ring, FDDI, and ATM were more complex and expensive than Ethernet, which further discouraged network administrators from switching over.

- Third, the most compelling reason to switch to another LAN technology was usually higher data rate of the new technology; however, Ethernet always fought back, producing versions that operated at equal data rates or higher.

- The original Ethernet was invented in the mid-1970s by Bob Metcalf and David Boggs. The original version used a coaxial bus to interconnect the nodes. Today's Ethernet is very different from the original version and adopts star topology instead of bus topology.

Notes: Ethernet and Ethernet Topology

- The original Ethernet was based on bus topology.

- In 1990s, hub-based Ethernet based on star topology emerged. A hub is a physical layer device that simply re-creates the bit so that any noise and distortion are cleaned up, and the weakened signal strength is restored to the nominal level. In effect, a hub-based Ethernet is still a single broadcast medium with the shared medium contained within the hub. Hence, Ethernet nodes that are connected via a hub are still within the same collision domain, meaning that when more than one node transmit at about the same time, their signals will interfere with each other and result in collision.

- In 2000s, hubs at the centre of the star topology were replaced by switches. Unlike a hub that simply acts on bits, a switch acts on frames and partitions each Ethernet connection between a node and a switch port into a separate collision domain, which removes collision completely from the Ethernet.

Ethernet Frame Structure

Ethernet Frame Structure

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

Preamble	Dest. Address	Source Address	Type	Data	CRC

Preamble:

- Seven bytes with pattern 10101010 followed by one byte with pattern 10101011.
- Used to synchronise receiver, sender clock rates.

$$\underbrace{\boxed{10101010101010101010101010101010101010101010101010101010}}_{\text{Preamble (56 bits – 7 bytes)}}\underbrace{\boxed{10101011}}_{\text{SOF (8 bits – 1 byte)}}$$

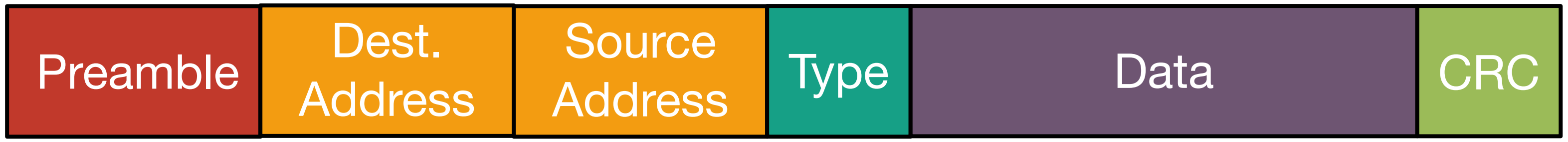
Ethernet Frame Structure

Addresses: 6 bytes (48 bits)

- If adapter receives frame with matching destination address, or with broadcast address (FF-FF-FF-FF-FF-FF), it passes data in frame to network layer protocol.
- Otherwise, adapter discards frame.

Type: Indicates higher layer protocol (mostly IP but others possible).

CRC: Cyclic redundancy check at receiver. If error is detected, frame is dropped.



Ethernet: Unreliable, Connectionless

Connectionless

No handshaking between sending and receiving network interface cards (NIC)s.

Unreliable

Receiving NIC doesn't send acks or nacks to sending NIC.

- Stream of datagrams passed to network layer can have gaps (missing datagrams).
- Gaps will be filled if application is using TCP.
- Otherwise, application will see gaps.

Ethernet's MAC protocol: Unslotted CSMA/CD

Slot Time in CSMA/CD Network

Slot time is an important parameter in a CSMA/CD network.

The default slot time is 512 bits. This means that the default slot time of a 10 Mbps CSMA/CD network is 51.2 μs and that of a 100 Mbps CSMA/CD network is 5.12 μs.

Slot time must be at least twice the time it takes for an electronic pulse to travel the length of the maximum theoretical distance between two nodes.

Due to how the slot time is set, any collision will be detected by all the hosts within the slot time.

Minimum Slot Time in CSMA/CD

- Assume that there are only two nodes A and B connected to the two extreme ends of a CSMA/CD network.

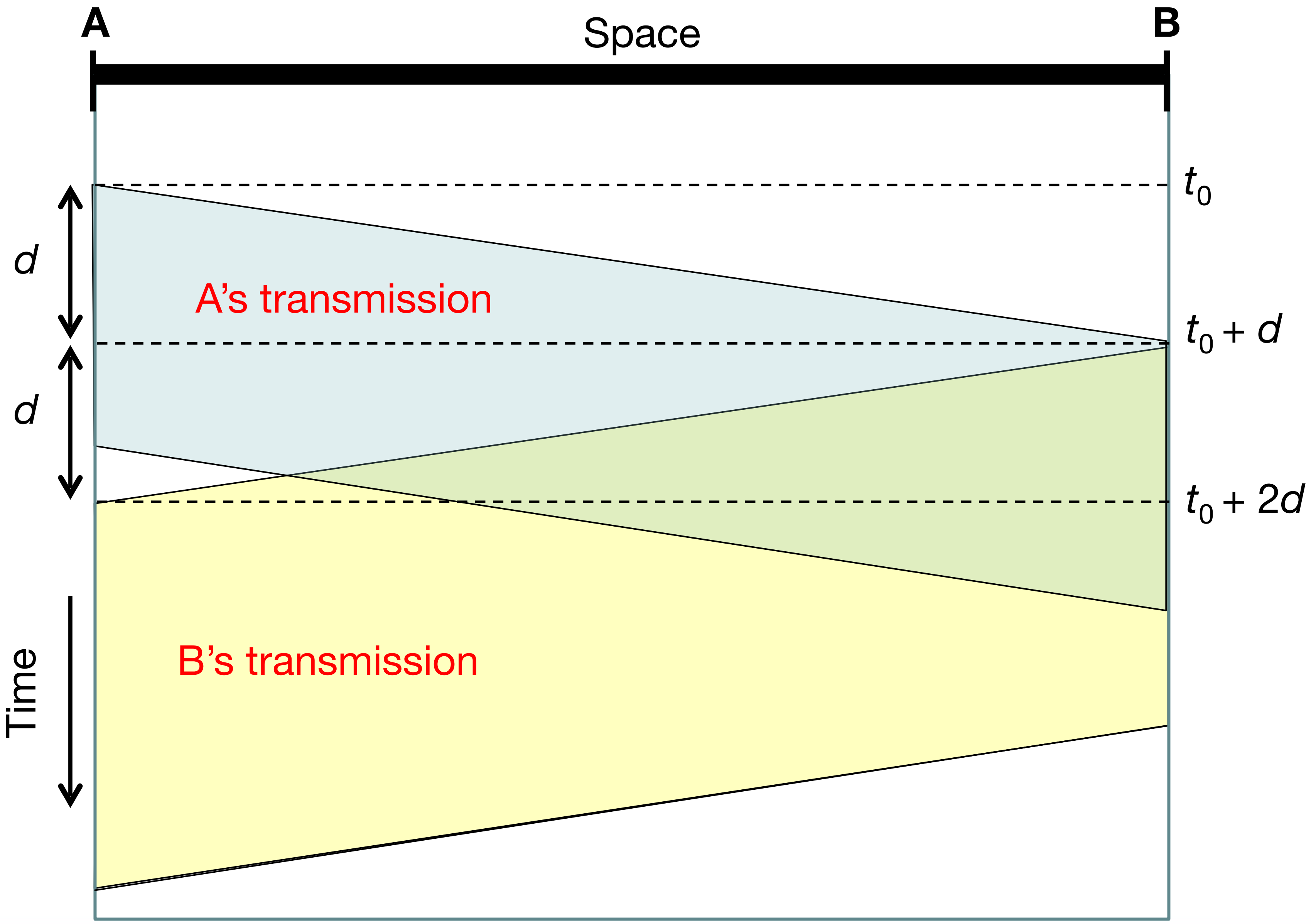
- Assume that A starts to transmit a frame at time t_0 and due to propagation delay, A's transmission will only be detected by B at time $t_0 + d$, where d is the one way propagation for the signal to travel the whole length of the shared medium.

- If B starts to transmit a frame just before $t_0 + d$, there will be a collision. While this collision will be detected by B immediately, it will take another d units of time for A to detect it.

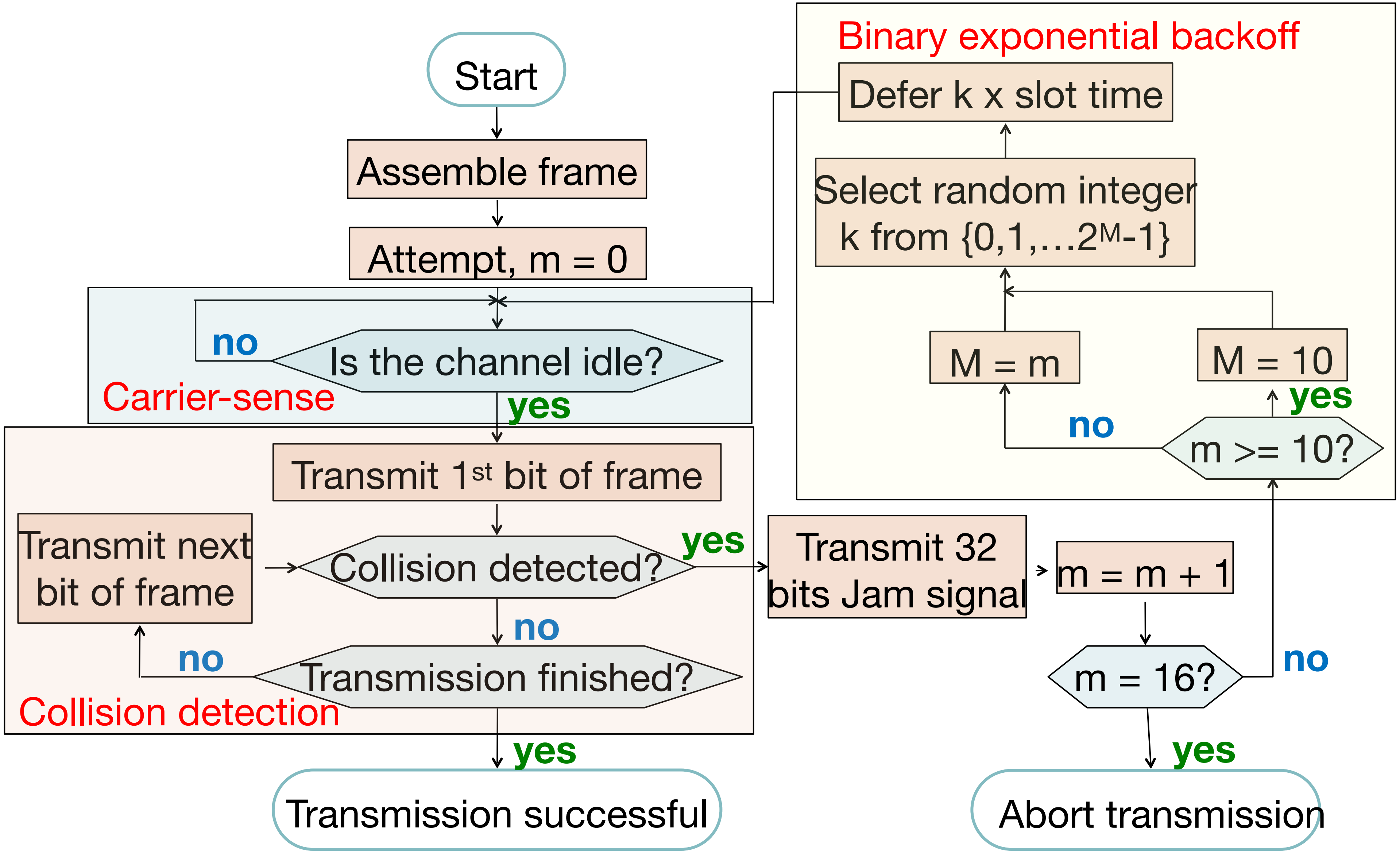
- Hence, after A has transmitted, if there is a collision, A will be able to detect it between t_0 and $t_0 + 2d$, or within a period of $2d$ after it starts to transmit.

- It can be seen that the slot time has to be set at least equal to $2d$, which is also known as the minimum slot time.

Minimum Slot Time in CSMA/CD



Ethernet CSMA/CD Algorithm



Ethernet CSMA/CD Algorithm

1. NIC receives datagram from network layer, creates frame.
2. If NIC senses channel idle, starts frame transmission. If NIC senses channel busy, waits until channel idle, then transmits.
3. If NIC transmits entire frame without detecting another transmission, NIC is done with frame!
4. If NIC detects collision while transmitting, aborts and sends jam signal.
5. If 16 attempts to transmit the same frame have been reached, transmission is aborted and the frame is discarded.
6. If the maximum of 16 attempts has not been reached, NIC enters binary exponential backoff: after m^{th} collision, NIC chooses k at random from $\{0, 1, 2, \ldots, 2^M-1\}$, where $M = m$ if $m < 10$ and $M = 10$ if $M >= 10$. NIC waits for $k \cdot$ Slot Time, returns to Step 2.

Ethernet CSMA/CD Algorithm (more)

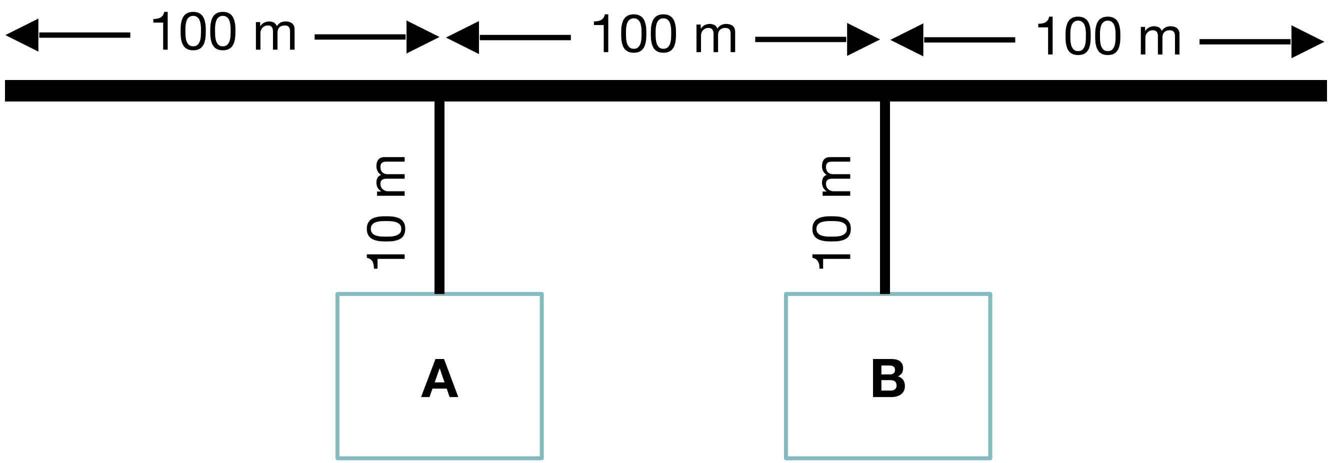
Jam Signal	Make sure all other transmitters are aware of collision; 32 bits
Bit Time	0.1 microsec for 10 Mbps Ethernet
Exponential Backoff	• **Goal**: Adapt retransmission attempts to estimated current load. ○ heavy load: random wait will be longer. • After 1st collision: choose k from {0, 1}; delay is k · slot time. Default slot time is 512 bits. • After 2nd collision: choose k from {0, 1, 2, 3}… • After 10th collisions, choose k from {0, 1, 2, 3, 4, …, 1023}

For k = 1023, wait time is

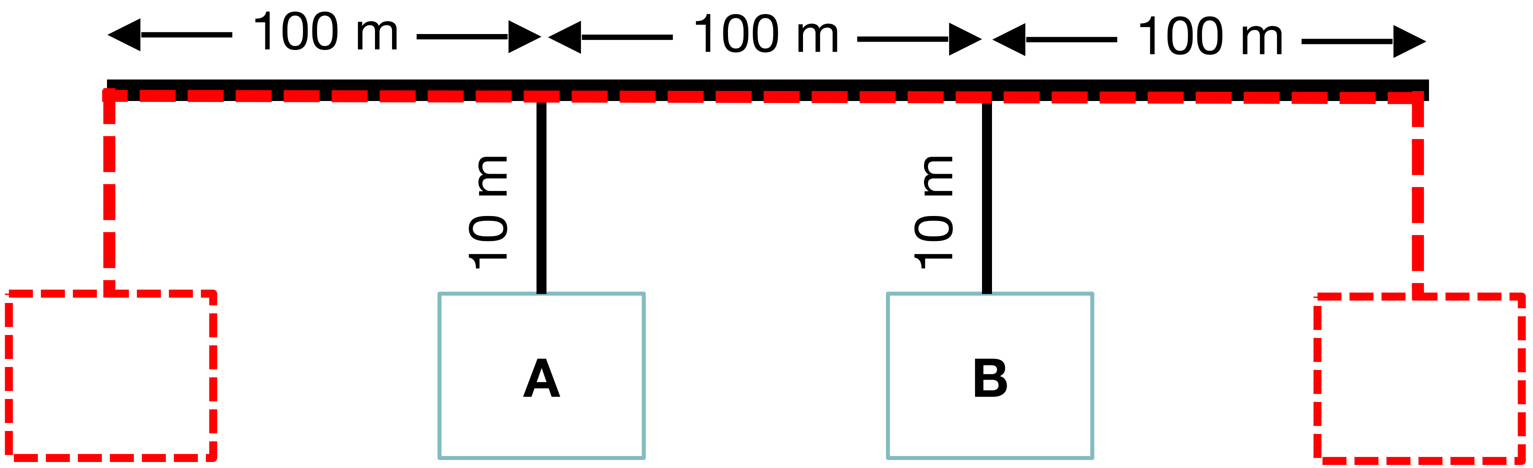
$1023 * 512 \text{ bits} * 0.1\ \mu s = 52.4 \text{ msec}$

Exercise

For the 10 Mbps CSMA/CD network shown below, assume that signal propagates on the network at a speed of 2×10^8 m/s. Apart from Hosts A and B shown in the figure, other hosts may also be similarly connected anywhere in the LAN segment of the network. Assume that the drop cables connecting the hosts are each 10 m in length. What is the minimum slot time for the network?



Answers



In the worst case, two hosts could be connected to both ends of the LAN segment and each with a 10 m drop cable.

Minimum slot time = 2 * (Maximum delay between any two nodes in the network)

$$= 2 * [320 \text{ m} / (2 \times 10^8 \text{ m/s})] = 3.2 \ \mu\text{s}$$

Exercise

For a CSMA/CD network, after the 5^{th} collision, what is the probability that a node chooses k = 4? How much delay will k = 4 correspond to on a 10 Mbps Ethernet? You may assume that the slot time is 512 bits.

After m^{th} (m <= 10) collision, k is chosen randomly from {0, 1, 2, …, 2^m-1}

For m = 5, the range of possible values of k are {0, 1, 2, …, 31}. Hence, the chance that any value of k (including k = 4) is chosen is 1/32 = 0.03125.

Slot time = 512 bits = 512 / 10 Mbps = 51.2 μs

Hence, k = 4 corresponds to a delay of 4 x 51.2 μs = 204.8 μs = 0.2048 ms

Exercise

A CSMA/CD network has only two stations A and B. The standard exponential backoff rule is followed after collision with kT (k random, T = 512 bit times) as the actual random backoff delay that is chosen. Suppose A tries to send a frame after 1st collision while B tries to send a frame after 2nd collision and these two frames collided again; i.e., A's frame has encountered 2nd collision while B's frame has encountered 3rd collision. What are the ranges of backoff delay for A and B? What is the probability that A and B will collide in their next transmission?

Answers

A has encountered 2nd collision; hence, the random integer k_A will be chosen from {0, 1, …, 2^2 - 1} = {0, 1, 2, 3}; hence, the range of backoff delay is {0, T, 2T, 3T}.

Similarly, B has encountered 3rd collision; hence, the random integer k_B will be chosen from {0, 1, …, 2^3 - 1} = {0, 1, 2, …, 7}; hence, the range of backoff delay is {0, T, 2T, …, 7T}.

Let the chosen delay combination for A and B be (k_A, k_B). All the possible combinations of (k_A, k_B) are as follow:

(0, 0)	(0, 1)	(0, 2)	(0, 3)	(0, 4)	(0, 5)	(0, 6)	(0, 7)
(1, 0)	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)	(1, 7)
(2, 0)	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)	(2, 7)
(3, 0)	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)	(3, 7)

There are 4 x 8 = 32 possible combinations for (k_A, k_B) and the 4 combinations that will result in another collision are: (0, 0), (1, 1), (2, 2), (3, 3) and (4, 4).
Hence, the probability of a collision is 4/32 = 0.125.

CSMA/CD efficiency

- t_{prop} = max prop delay between 2 nodes in LAN
  t_{trans} = time to transmit max-size frame

- Researchers have shown that the efficiency can be approximated as

$$efficiency = \frac{1}{1 + 5t_{prop}/t_{trans}}$$

- Efficiency goes to 1
  - as t_{prop} goes to 0, or
  - as t_{trans} goes to infinity

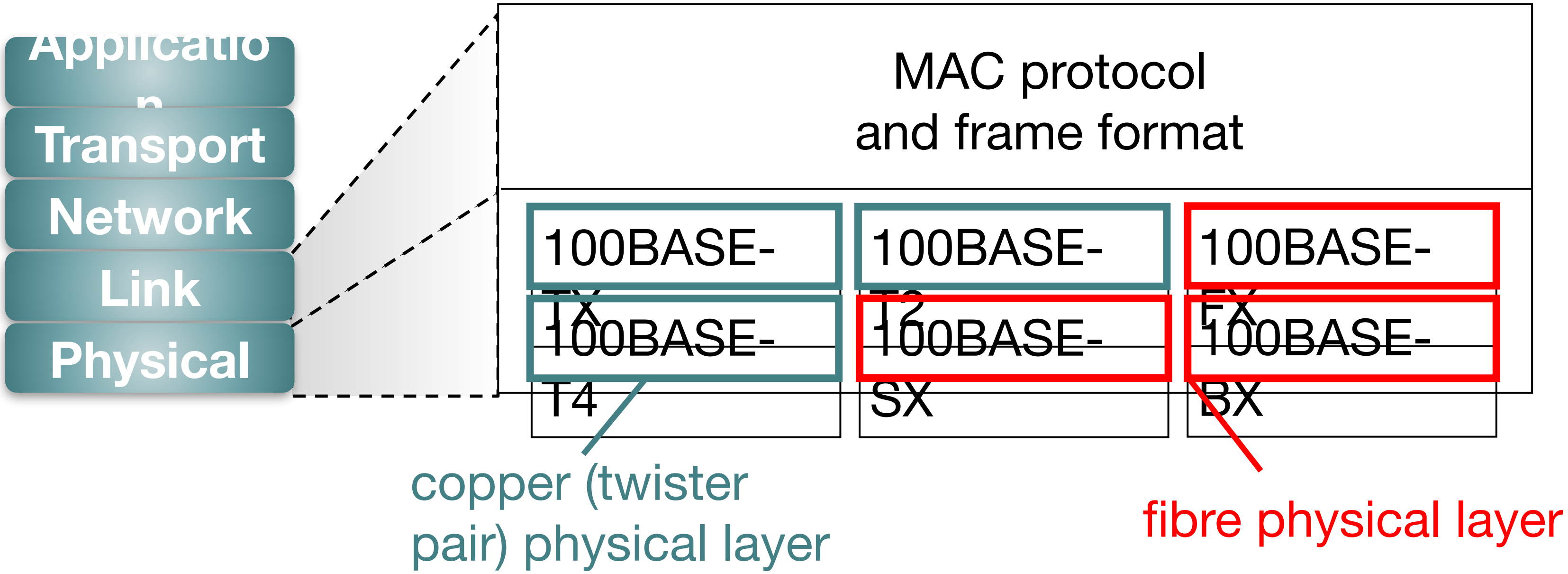
- Better performance than ALOHA: and simple, cheap, decentralised!

IEEE 802.3

802.3 Ethernet Standards: Link and Physical Layers

Many different Ethernet standards

- Common MAC protocol and frame format
- Different speeds: 2 Mbps, 10 Mbps, 100 Mbps, 1 Gbps, 10 Gbps
- Different physical layer media: fibre, cable

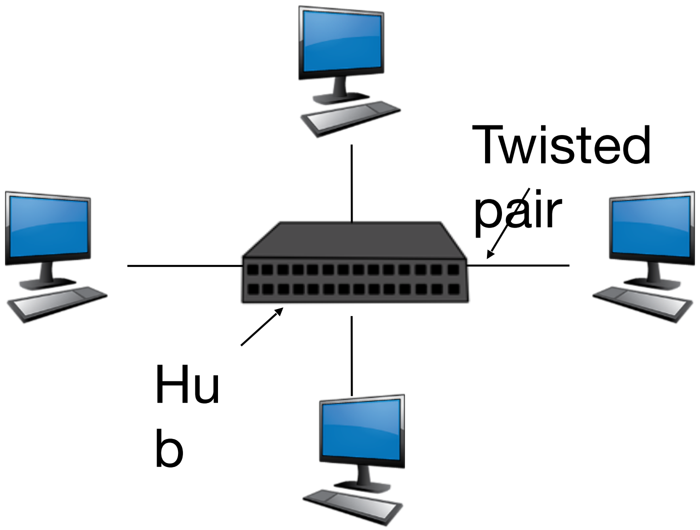


Details of the various different physical layer are not covered in this course.

Hubs

… physical-layer ("dumb") repeaters:

- Bits coming in one link and go out all other links at same rate
- All nodes connected to hub can collide with one another
- No frame buffering
- No CSMA/CD at hub: host NICs detect collisions



Exercise

Consider a 100 Mbps 100BASE-T Ethernet with all nodes directly connected to a hub. Assume a frame length of 64 bytes and that there are no repeaters. You may assume that signal propagates on the copper medium at a speed of 2×10^8 m/s. To have an efficiency of 0.5, what should be the maximum distance between a node and the hub? How does your maximum distance compare with the actual 100 Mbps standard that specifies the maximum host to hub distance of 100 m?

Answers

Efficiency of the CSMA/CD network can be approximated to be 1 / (1 + 5*a*) where *a* = end-to-end propagation delay (t_{prop}) / average transmission delay (t_{trans}).

$$t_{trans} = 64 \times 8 \text{ bits} / 100 \text{ Mbps} = 5.12\ \mu s$$

For efficiency to be 0.5,

$$0.5 = 1 / (1 + 5a) \quad \Rightarrow \quad 1 + 5a = 2 \quad \Rightarrow \quad a = 1/5 = 0.2$$

Hence, $t_{prop} = a \cdot t_{trans} = 0.2 * 5.12 = 1.024\ \mu s$

End-to-end distance between two nodes

$$= 1.024\ \mu s * (2 \times 10^8 \text{ m/s}) = 204.8 \text{ m}$$

Hence, maximum distance between a node and the hub is 102.4 m. This is very close to the 100BASE-T standards of 100 m.

Bridge/Switch Learning

Switch

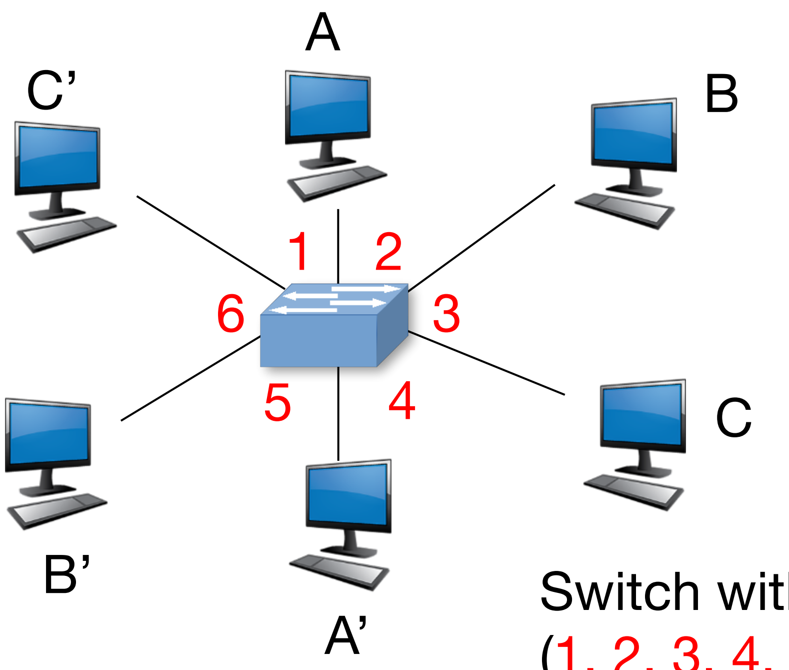
- Link-layer device: smarter than hubs, take active role
  - Store and forward Ethernet frames.
  - Examine incoming frame's MAC address, selectively forward frame to one-or-more outgoing links when frame is to be forwarded on segment, uses CSMA/CD to access segment.
  - A bridge is essentially the same as a switch but have a very small number of ports.

- Transparent
  - Hosts are unaware of presence of switches.

- Plug-and-play, self-learning
  - Switches do not need to be configured.

Switch: Allows multiple simultaneous transmissions

- Hosts have dedicated, direct connection to switch.
- Switches buffer packets.
- Ethernet protocol used on each incoming link, but no collisions; full duplex.
- Each link is its own collision domain.
- Switching: A-to-A’ and B-to-B’ simultaneously, without collisions .
- Not possible with dumb hub.



Switch with six interfaces (1, 2, 3, 4, 5, 6)

Notes: Switch

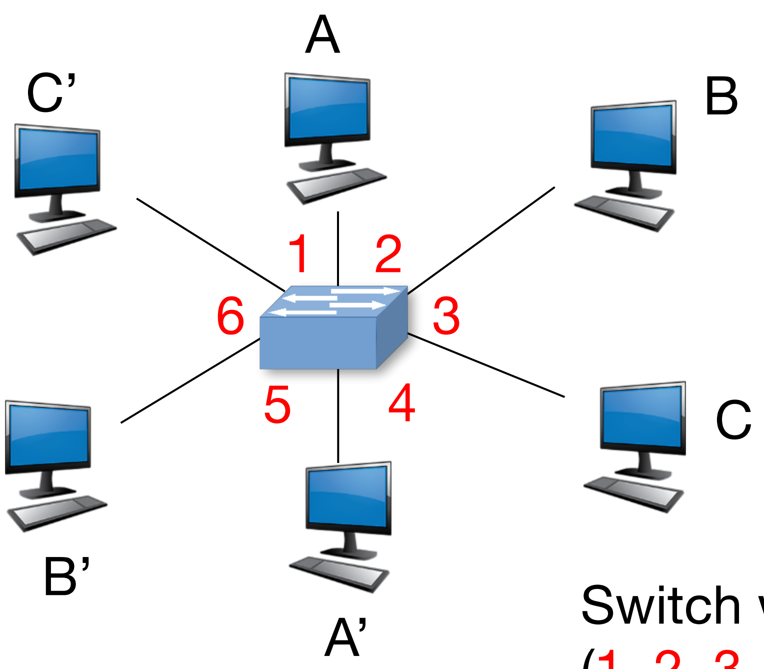
- The role of the switch is to receive incoming link-layer frames and forward them to outgoing links.
- The switch is transparent to the nodes; i.e., a node addresses a frame to another node and sends the frame onto the LAN, unaware that a switch will be receiving the frame and forwarding it to other nodes.
- The rate at which frames arrive to any one of the switch's output interfaces may temporarily exceed the link capacity of that interface. To accommodate this problem, switch output interfaces have buffers.
- Forwarding is the switch function that determines the interfaces to which a frame should be directed, and then moves the frames to those interfaces.
- Switch forwarding is done with a switch table.

Switch Table

How does switch know that A' is reachable via interface 4 and B' is reachable via interface 5?

A Each switch has a switch table, each entry contains:
(MAC address of host, interface to reach host, time stamp)

Looks like a routing table!



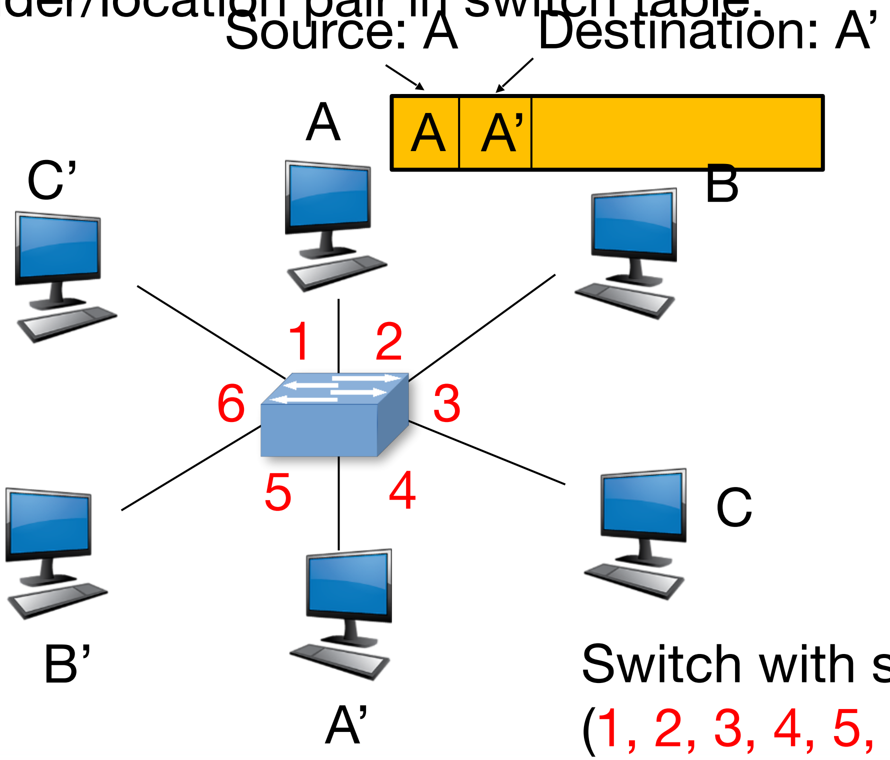
Switch with six interfaces
(1, 2, 3, 4, 5, 6)

Switch Learning

? How are entries created and maintained in the switch table?

A 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.



Switch with six interfaces (1, 2, 3, 4, 5, 6)

MAC Address	Interface	TTL
A	1	60

Switch table (initially empty)

Notes: Switch: Self-Learning

A switch builds its switch table automatically by self-learning. This capability is accomplished as follows:

- The switch table is initially empty.
- For each incoming frame received on an interface, the switch stores in its table (1) the MAC address in the frame's source address field, (2) the interface from which the frame arrived, and (3) the default time-to-live (TTL) value. If every node in the LAN eventually sends a frame, then every node will eventually get recorded in the table.
- The switch deletes an address in the table if no frames received with that address as the source address the TTL count down to zero. In this manner, if a PC is replaced by another PC (with a different adapter), the MAC address of the original PC will eventually be purged from the switch table.

Switch: Frame Filtering/Forwarding

When frame received:

1. Record link associated with sending host.

2. Index switch table using MAC destination address.

3. ```
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 not the interface on which the frame arrived.
```

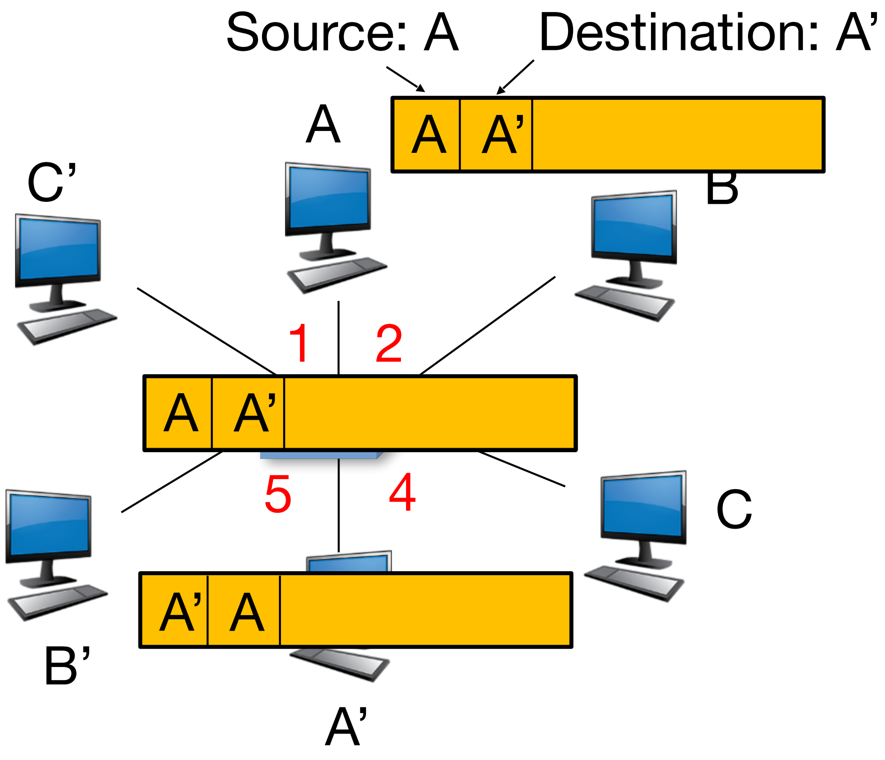
Self-learning Forwarding (Example)

- Frame destination unknown: flood
- Destination A location known: selective send

Note that when the switch floods, it broadcasts the frames to all the interfaces except the interface where it receives the frame from.

MAC Address	Interface	TTL
A	1	60
A’	4	60

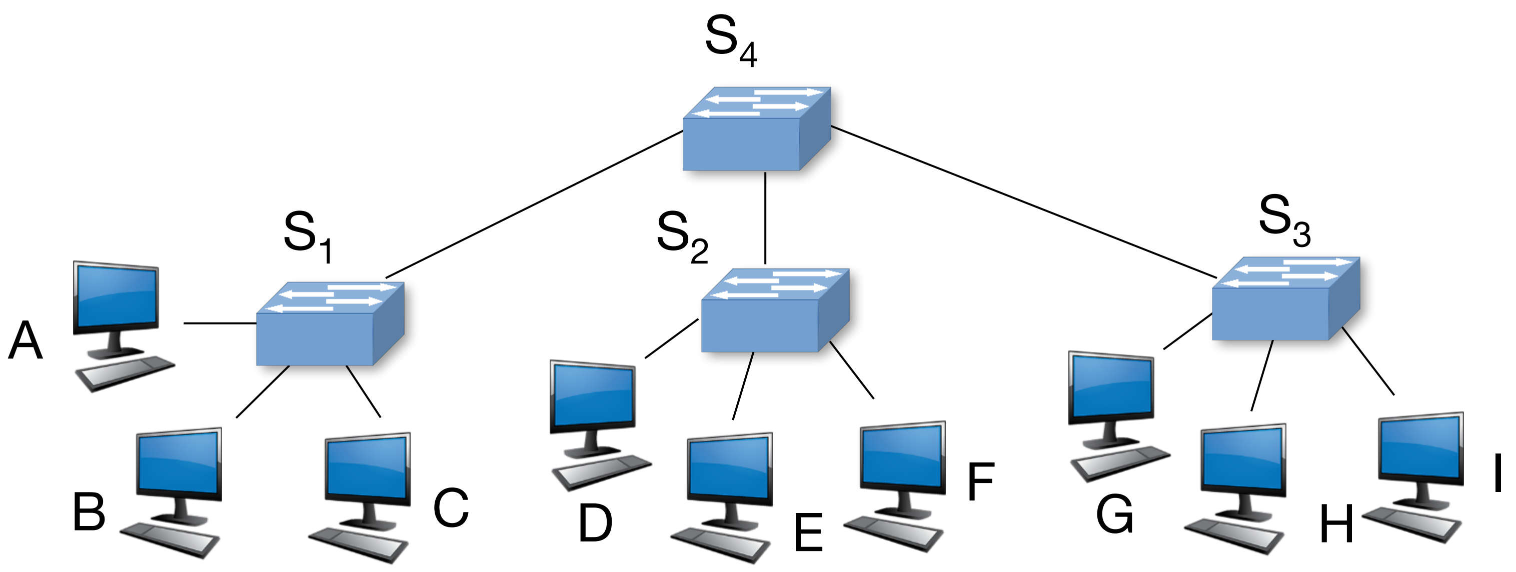
Switch table
(initially empty)



Switch with six interfaces
(1, 2, 3, 4, 5, 6)

Interconnecting Switches

Switches can be connected together.

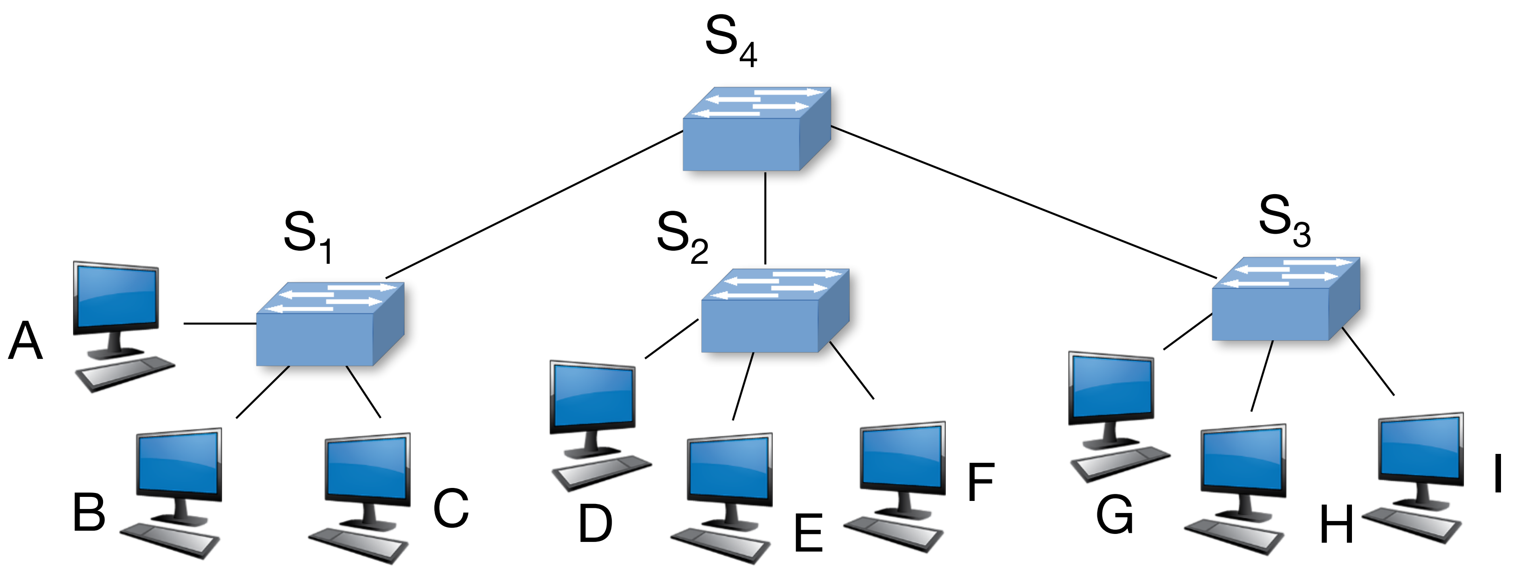


? Sending from A to F - How does S_1 know how to forward frame destined to F via S_4 and S_2?

A Self learning! (Works exactly the same as in single-switch case!)

Self-learning Multi-Switch (Example)

Suppose C sends frame to I, I responds to C.



? Show switch tables and packet forwarding in S_1, S_2, S_3 and S_4.

A (View answers in the animation on the next slide.)

Self-learning Multi-Switch (Example)

Virtual Local Area Networks (VLANS)

Virtual Local Area Networks (VLANS)

VLAN

A switch that supports VLANs allows multiple virtual local area networks to be defined over a single physical local area network infrastructure.

VLANs reduce the need to have routers deployed on a network to contain broadcast traffic.

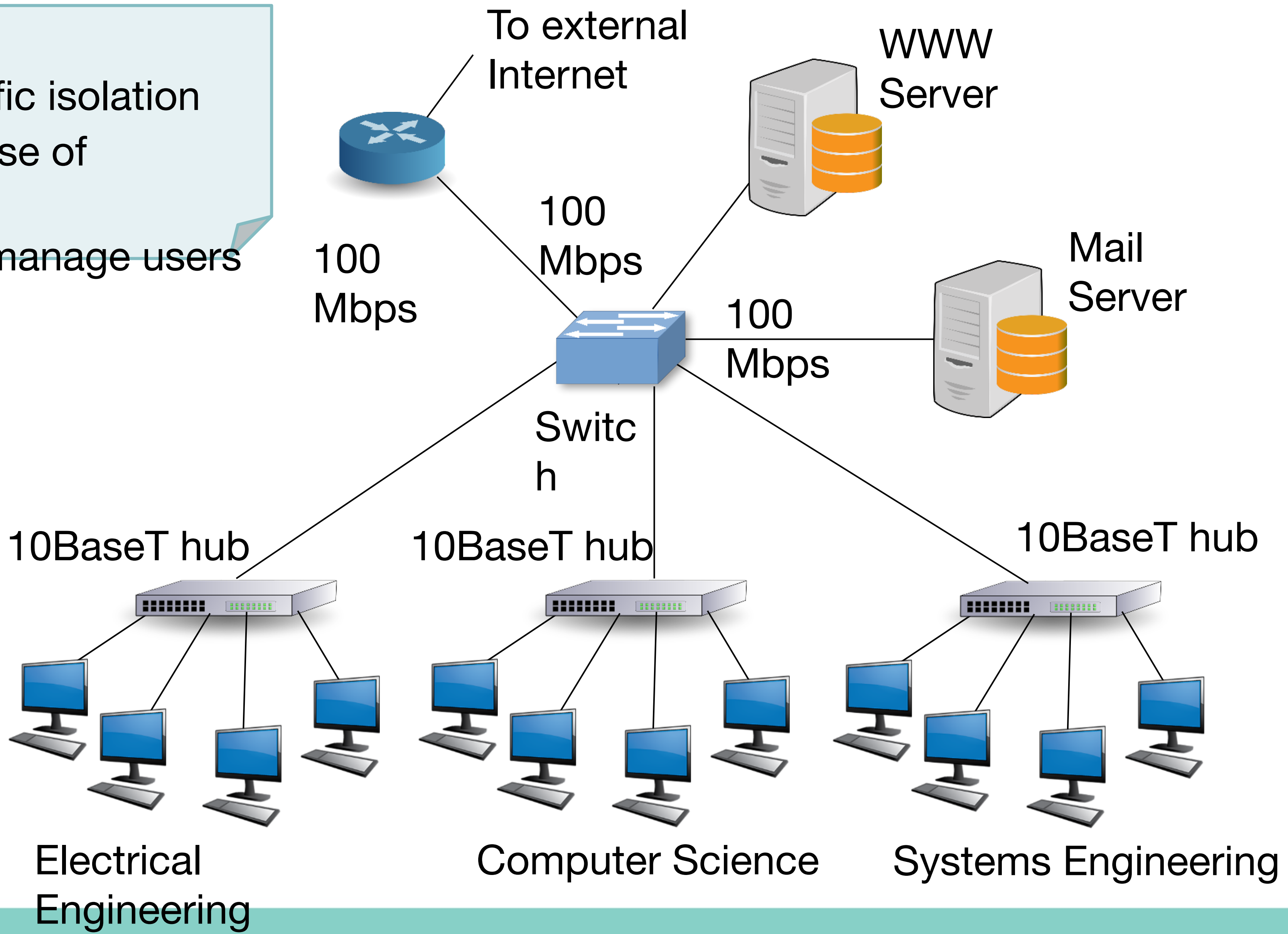
VLANs segregate the physical LAN into several smaller broadcast domains and reduce the broadcast traffic.

When users change departments, no physical re-wiring and other changes are required; hence, VLANs provide a lot of flexibility.

Hierarchical Institutional LANs

Drawbacks

- Lack of traffic isolation
- Inefficient use of switches
- Difficult to manage users

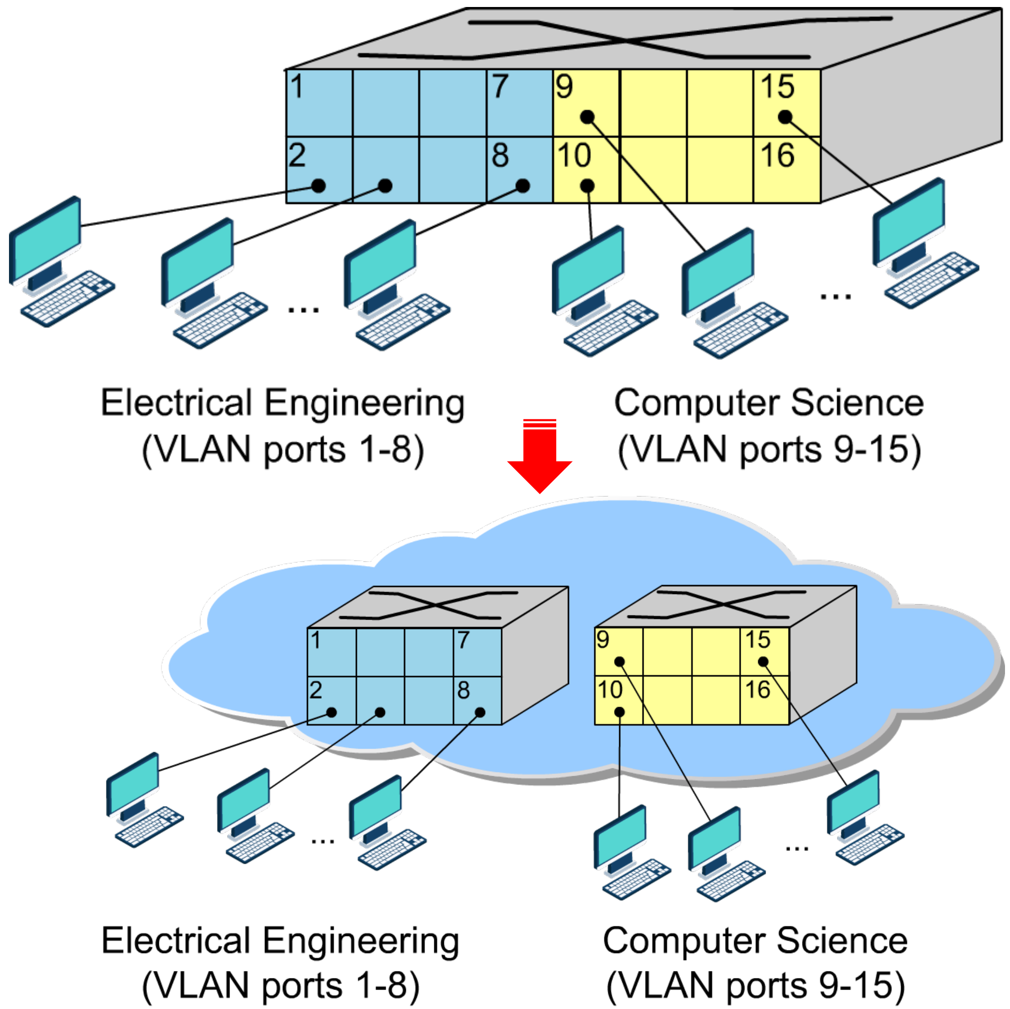


VLANs

Switch(es) supporting VLAN capabilities can be configured to define multiple virtual LANS over single physical LAN infrastructure.

Port-based VLAN:

Switch ports are grouped (by switch management software) so that a single physical switch …



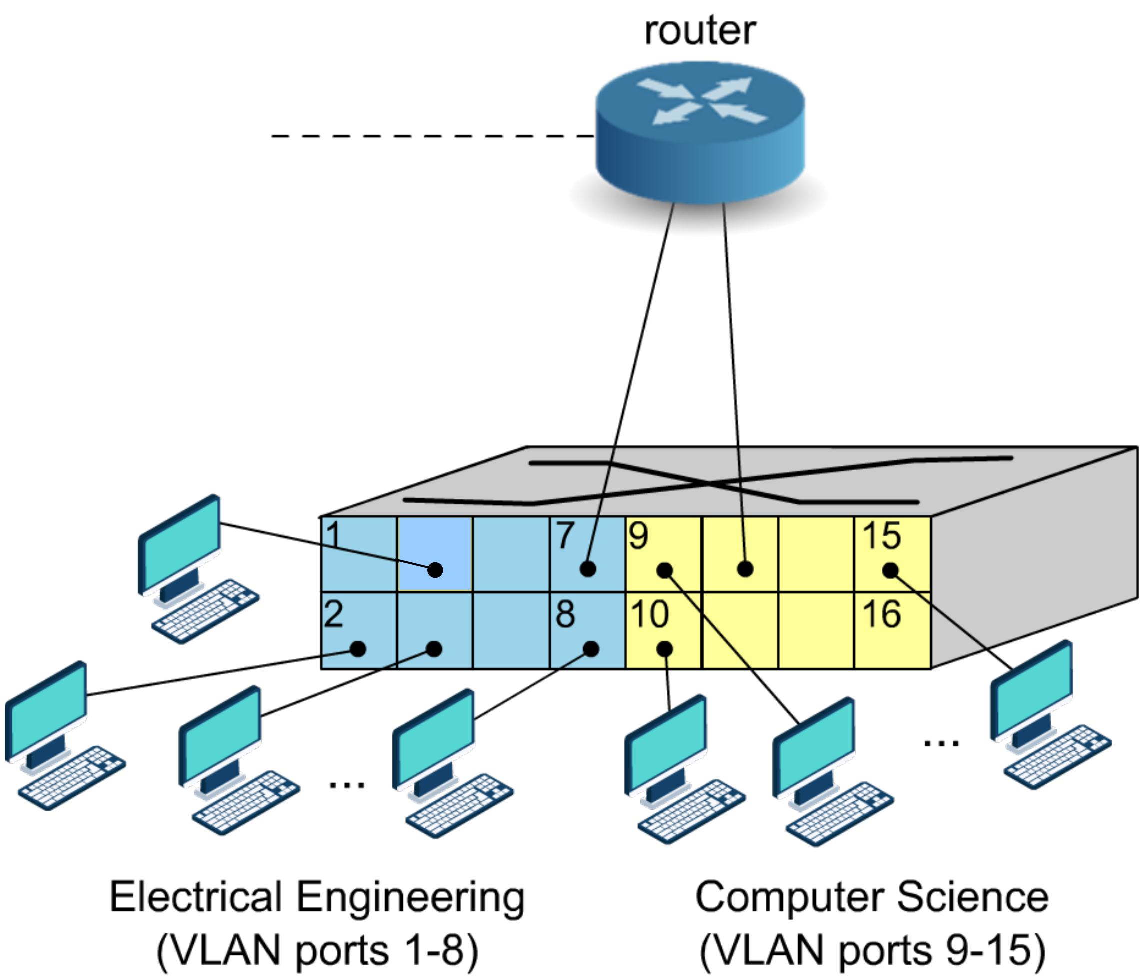
… operates as multiple virtual switches.

Port-based VLAN

- **Traffic isolation**: frames to/from ports 1-8 can only reach ports 1-8.
  - Can also define VLAN based on MAC addresses of endpoints and applications rather than switch port.

- **Dynamic membership**: ports can be dynamically assigned among VLANs.

- **Forwarding between VLANs**: done via routing (just as with separate switches)
  - In practice vendors sell switch and router integrated into one piece of network equipment (multi-layer switch).



Notes: VLANs: Motivation

- Modern institutional LANs are often configured hierarchically, with each workgroup (department) having its own switched LAN, which is connected to the switched LANs of other groups via a switch hierarchy. However, such a network has three drawbacks.

Lack of traffic isolation	Inefficient use of switches	Managing users
Although the hierarchy localises group traffic to within a single switch, broadcast traffic must still traverse the entire institutional network. Limiting the scope of such broadcast traffic would improve LAN performance, security and privacy.	Modern switches could have up to 96 ports and a few switches would be enough to accommodate a medium size organisation with a couple of hundreds employees. Having a hierarchical network with many layers would need many switches but having just a few switches will not provide traffic isolation.	If an employee moves between groups, the physical cabling must be changed.

Notes: VLANs: Motivation

- Each of these difficulties could be handled by a switch that supports VLANs.

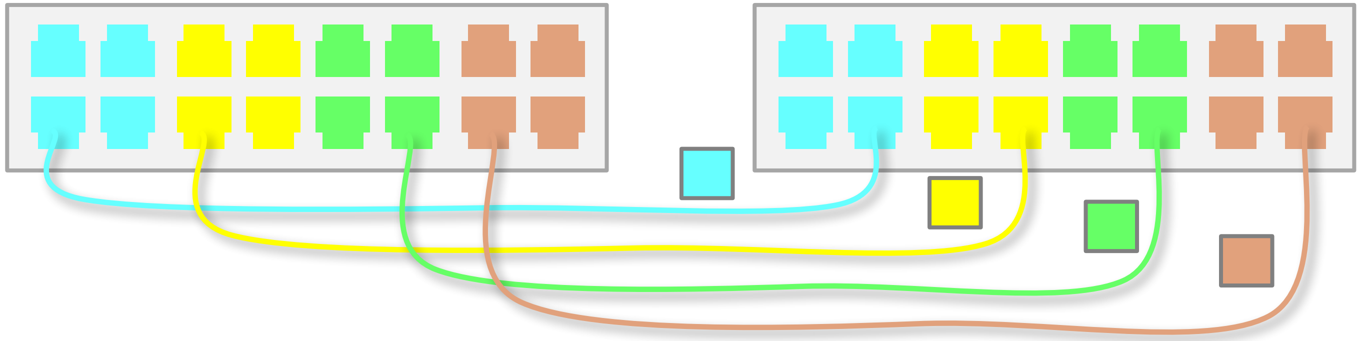
- A switch that supports VLANs allows multiple virtual LANs to be defined over a single physical LAN infrastructure. Hosts within the same VLAN communicate with each other as if they were connected to the switch.

- In a port-based VLAN, the switch's ports (interfaces) are divided into groups by the network manager. Each group constitutes a VLAN, with the ports in each VLAN forming a broadcast domain.

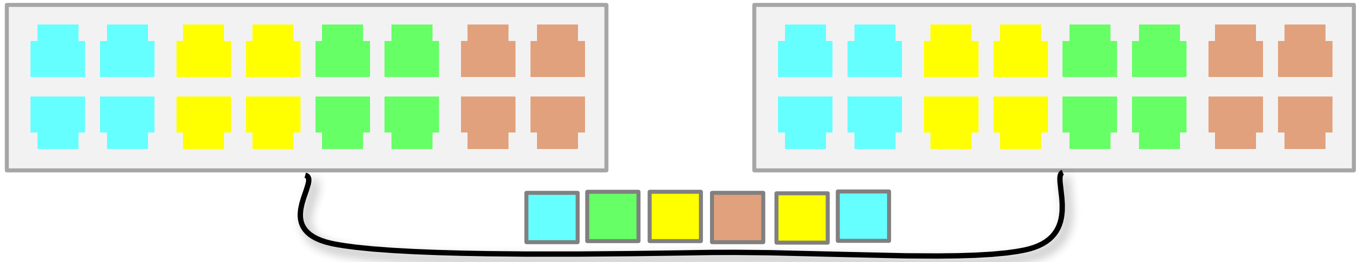
- By completely isolating two VLANs, we will need a router to interconnect the two VLANs even though these two VLANs may be configured from ports on the same physical switch. Fortunately, switch vendors make such configurations easy for network manager by building a single device that contains both a VLAN switch and a router, so a separate external router is not needed.

VLAN Trunking

Without VLAN Trunking



With VLAN Trunking

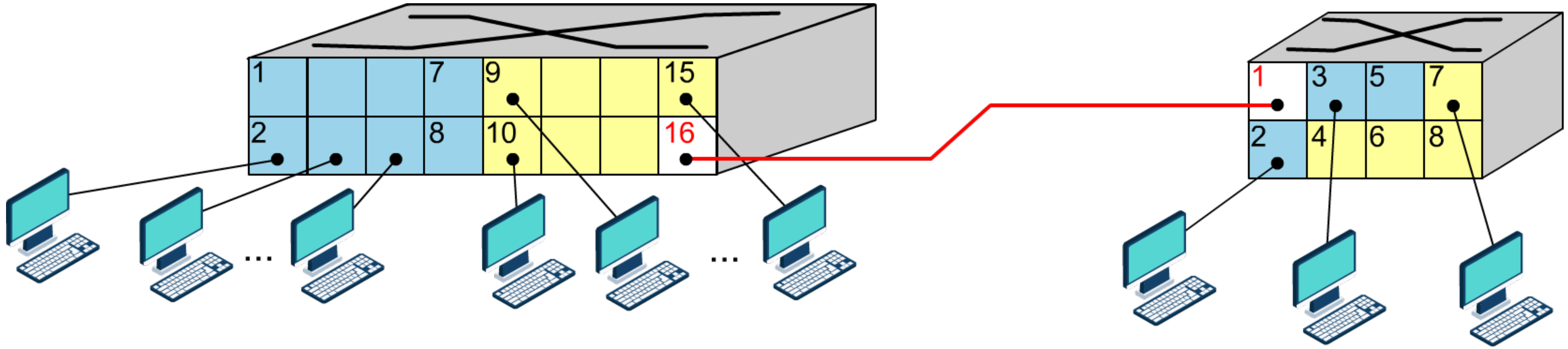


Dest. Address	Source Address	802.1Q VLAN Tag	Type	Data	CRC

4 Bytes

VLAN tag contains VLAN ID and other information

VLANs Spanning Multiple Switches



Electrical Engineering
(VLAN ports 1-8)

Computer Science
(VLAN ports 9-15)

Ports 2, 3, 5 belong to EE VLAN
Ports 4, 6, 7, 8 belong to CS VLAN

- **Trunk port:** carries frames between VLANs defined over multiple physical switches.

Dest. Address	Source Address	802.1Q VLAN Tag	Type	Data	CRC

$\underbrace{\text{802.1Q VLAN Tag}}_{\text{4 Bytes}}$

VLAN tag contains VLAN ID and other information

Notes: VLAN Trunking and Tag

- When multiple VLANs span over multiple switches, a scalable approach to connect VLAN switches is known as VLAN trunking. In this approach, a special port on each switch is configured as a trunk port to interconnect two VLAN switches. The trunk port belongs to all VLANs, and frames sent to any VLAN are forwarded over the trunk link to the other switch.

- How does a switch know that a frame arriving on a trunk port belongs to which VLAN? The IEEE has defined an extended Ethernet frame format, 802.1Q, for frame crossing a VLAN trunk.

- The 802.1Q frame consists of the standard Ethernet frame with a four-byte VLAN tag added into the header that carries the identity of the VLAN to which the frame belongs. The VLAN tag is added into a frame by the switch at the sending side of a VLAN trunk, parsed and removed by the switch at the receiving side of the trunk.

Summary

Summary

Key points discussed in this topic:

- Ethernet is by far the most prevalent wired LAN technology.
- Ethernet frame structure consists of the preamble, destination address, source address, type, data and CRC.
- The MAC protocol used in Ethernet is unslotted CSMA/CD.
- A hub is a layer 1 network device and all the hosts connected to a hub belong to the same collision domain.
- A switch is a layer 2 network device, which has self-learning capability. Hosts connected to the switch belong to different collision domains.

Summary

Key points discussed in this topic (cont’d):

- The two key functions of a switch are self-learning and forwarding.
- Forwarding is the switch function that determines the interfaces to which a frame should be directed, and then moves the frames to those interfaces.
- Switch self-learning means it learns which hosts can be reached through which interfaces. When a frame is received, the switch “learns” location of sender: incoming LAN segment, and records sender/location pair in switch table.
- A switch that supports VLANs allows multiple virtual local area networks to be defined over a single physical local area network infrastructure.