15-441
15-641 Computer Networking

Lecture 3: Packet-Switched Networks

Peter Steenkiste

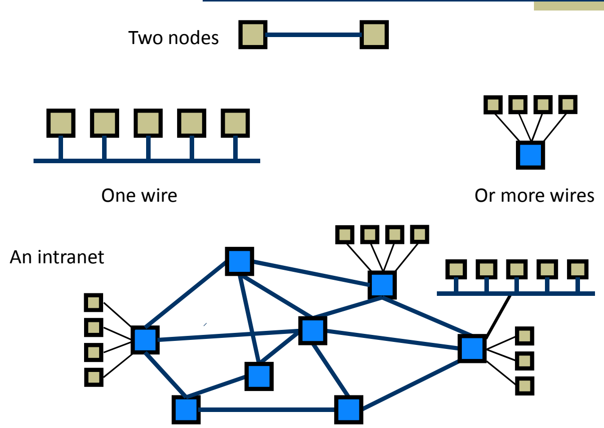
Fall 2016

www.cs.cmu.edu/~prs/15-441-F16

Goal and Outline

- Goal: gain a basic understanding of how you can build a (small) packet switched network
  - Focus is to convince you that this is feasible
  - A bit more detail later in the course for Ethernet and WiFi
- Physical and Datalink functions
- Physical layer: Modulation
- Datalink
  - Medium access control
  - Scaling up

Today's Story



What Do We Need?

- Physical layer:
  - Modulation: send a stream of bits to a receiver using an electromagnetic signal
  - Coding: add redundancy for error detection, meet electrical constraints, …
- Datalink layer:
  - Framing: identify packet boundaries and headers
  - Error control: error detection and correction
  - Media access control: arbitrating access to the "link"
  - Bridging, switching, …: extending network size
- Described "by example"

Outline

- PHY and DL functions
- Modulation
- Datalink layer
  - Media access control
  - Scaling up

Transferring Information

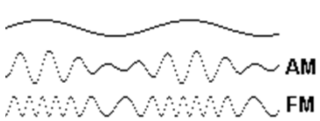
- Information transfer is a physical process

"The wireless telegraph is not difficult to understand.
The ordinary telegraph is like a very long cat.
You pull the tail in New York, and it meows in Los Angeles.
The wireless is exactly the same, only without the cat."

- In this class, we generally care about
  - Electrical signals (on a wire)
  - Optical signals (in a fiber)
  - RF signals (wireless)
  - More broadly: electromagnetic signals

What is Modulation?

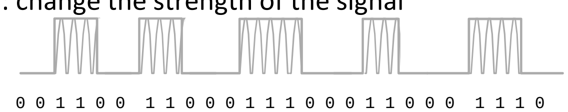
- The sender changes a signal in a way that the receiver can recognize - conveys information
- Ways to modulate a signal (think: sinusoidal wave)
  - Change frequency, phase, or amplitude
- Similar to AM/FM radio:
  - But we encode bits!



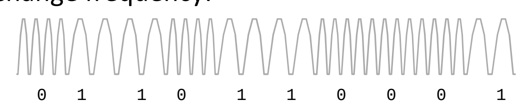
- Analogy from music:
  - Volume: Amplitude Modulation (AM)
  - Pitch: Frequency Modulation (FM)
  - Timing: Phase Modulation (PM)

Binary Modulation

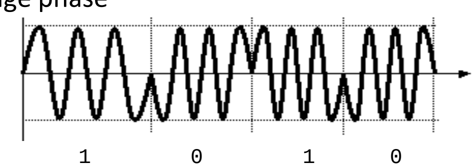
- AM: change the strength of the signal



- FM: change frequency:

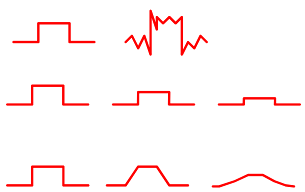


- PM: change phase



Looks Straightforward, but …

- Bad things happen to the signal as it travels to receiver:
- Noise: “random” energy is added to the signal
- Attenuation: some of the signal’s energy leaks away
- Dispersion: signal is distorted due to frequency-dependent effects distorts the signal



- These effects get worse with distance and depend on the transmission medium

What is the impact of a Bad Signal?

- The receiver may no longer be able to determine what bits were sent, resulting in bit errors
  - Bit error rate increases with the bit rate
- The result is that we need to limit the bit rate and the length of the links.
- For wired network, that standard specifies both
  - E.g., standards for 10 Mbs, 100 Mbs, .. Ethernet
- For wireless networks many other factors impact the bit error rate – requires more complex solutions
  - Wait for wireless lectures

Sketch of Solution

- Solutions for optimizing bandwidth and recovering from errors fall in two classes:

1. Retransmission by a higher layer protocol
2. Coding: add redundancy to the bit stream so the receiver can recover from the errors (FEC)

- Can be used in any layer of the stack, but a common approach is:
  1. Retransmission in datalink or transport protocol
  2. FEC in PHY layer

Outline

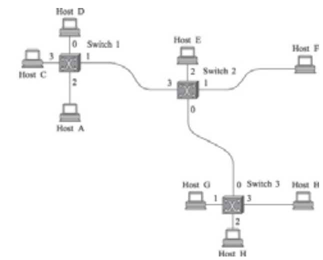
- PHY and DL functions
- Modulation
- Datalink layer
  - Media access control
  - Scaling up

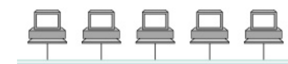
Datalink Functions

- Framing: encapsulating a network layer datagram into a bit stream.
  - Add header, mark and detect frame boundaries
- Flow control: avoid that sender outruns the receiver
- Error control: error detection and correction to deal with bit errors.
  - May also include other reliability support, e.g. retransmission
- Media access: controlling which frame should be sent next over a link.
- Bridging, switching: extend the size of the network

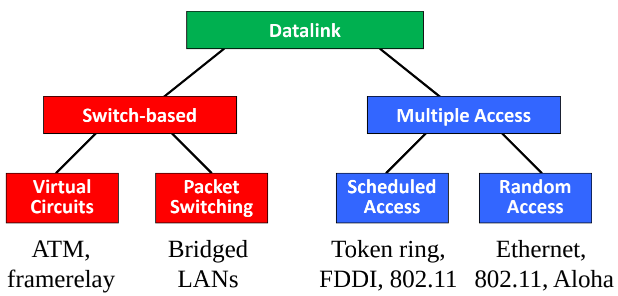
Datalink Architectures

- Switches connected by point-to-point links -- store-and-forward.
  - Used in WAN, LAN, and for home connections
  - Conceptually similar to "routing"
    - But at the datalink layer instead of the network layer
  - MAC = (local) scheduling
- Multiple access networks -- contention based.
  - Multiple hosts are sharing the same transmission medium
  - Used in LANs and wireless
  - Access control is distributed and much more complex

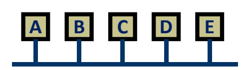




Datalink Classification

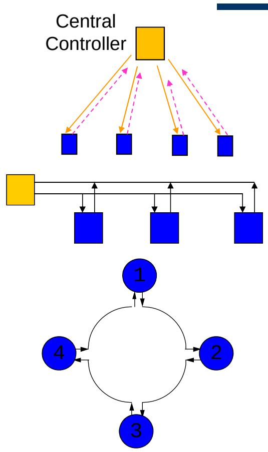


Multiple Access: How to Share a Wire (or the wireless ether)



- Problem: how do you prevent nodes from "talking" at the same time – causes "collision"
- Two classes of solutions:
  - Explicit coordination: schedule transmissions sequentially
  - Randomly access medium: send and hope you get lucky

Scheduled Access MACs

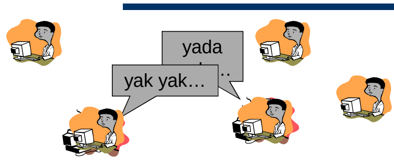


- Reservation systems
  - Central controller
  - Distributed algorithm, e.g. using reservation bits in frame
- Polling: controller polls each nodes
- Token ring: token travels around ring and allows nodes to send one packet
  - Distributer version of polling
  - FDDI, …

Random Access Protocols

- When a node has a packet to send
  - Transmit at full channel data rate R
  - No *a priori* coordination among nodes
- If you are lucky, receiver will receive packet, but ..
- Multiple simultaneous transmissions → "collision"
- Random access MAC protocol specifies:
  - How to avoid and/or detect collisions
  - How to recover from collisions (e.g., via retransmissions)
- Examples of random access MAC protocols:
  - Slotted ALOHA and ALOHA
  - CSMA/CD (~Ethernet) and CSMA/CA (~WiFi)

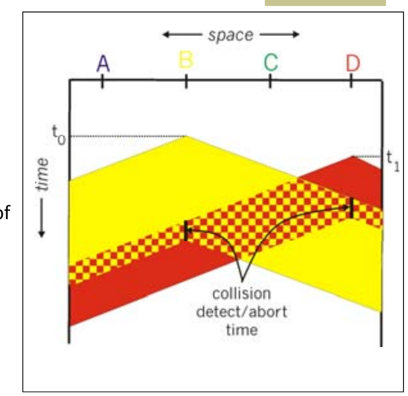
How Can We Avoid Collisions?



- Natural scheme – listen before you talk…
  - Works well in practice
  - A cheap form of coordination
- But sometimes this breaks down
  - When?
  - How do we fix/prevent this?

Collision Example

- B and D transmit before they can hear the other transmission
  - Unavoidable!
- Solution is to detect collisions:
  - B and D will hear the sum of the signals = garbage
  - They stop sending right away and retransmit
- Lots of details left out
  - E.g., backoff, packet and wire length, ..

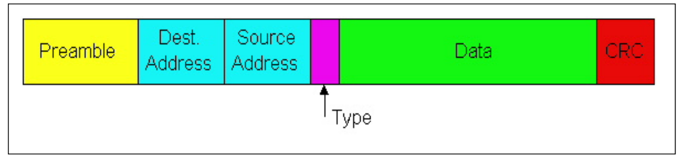


Ethernet MAC Features

- Carrier Sense: listen before you talk
  - Avoid collision with an ongoing transmission
- Advantage is that it is very efficient
  - No coordination overhead or transmission delay
- But it does not always work: simultaneous transmissions can happen
  - Speed of light is "only" 1 foot/nsec
- Collision Detection during transmission
  - Listen while transmitting
  - If you notice interference → assume collision
  - Abort transmission immediately and schedule a retransmission

Ethernet Frame Structure

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



- Addresses are 48 bit IEEE MAC addresses
  - Used by all IEEE 802 LAN standards, including WiFi
  - In practice used as a flat address – no structure

How Well Does Ethernet Work?

- The protocol is broken, right?
  - You would not design a traffic light this way!
- The protocol is very effective in practice
  - Most LANs are under-utilized
  - Scheduled access protocols have high overhead
- Transmission is fairly reliable in practice
  - Collisions can be detected reliably and corrupted packets are transmitted
  - No need for acknowledgements – low overhead!
  - Error rates due to random bit errors are very low in practice

Other Datalink Technologies

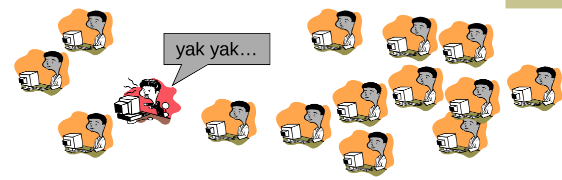
- WiFi is sometimes called "wireless Ethernet"
  - Same "listen before you talk" concepts
- But the details are very different!
  - Collision detection does not work, attenuation is much higher, bit error rates are much higher – life is rough
- WAN has used a variety of technologies
  - Early days: framerelay – based on virtual circuits
  - SONET: very widely over several generations of fiber
    - Supports both voice and data effectively
  - Today: Ethernet (of course)

24

Outline

- PHY and DL functions
- Modulation
- Datalink layer
  - Media access control
  - Scaling up
    - Number of nodes
    - Bit rate

Scaling Up the Number of Nodes

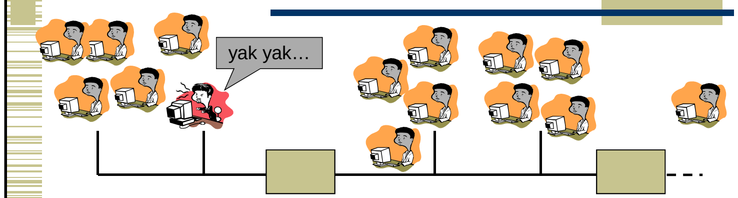


- What breaks when we keep adding people to the same wire?

Scaling Up the Ethernet Speed

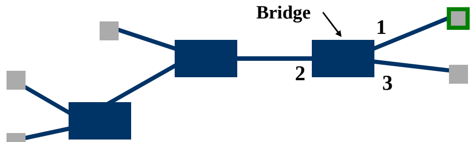
- Technology improvements lead to higher bit rates: 10Mbps, 100Mbps, 1Gbps, 40 Gbps, ...
- Problem: carrier sense becomes completely ineffective
  - For example, for 40 Gps links
  - →0.3 microsec to send a maximum sized Ethernet frame
  - →forget about carrier sense
- Solution: use a bridge or switch-based design
  - And call it Ethernet!

Scaling Up Solution



- Break up the network in smaller networks
- Smaller "collision domain" - fewer nodes per network
  - Also shorter wires
- Networks can transmit packet in parallel – more capacity
- Uses "bridges" (switches) to connect the networks
  - Bridges must forward the packets when needed
- Challenge: how do you know which packets to copy and where?

Frame Forwarding



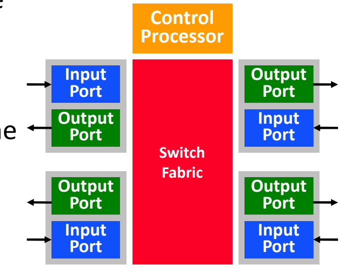
MAC Address	Port	Age
A21032C9A591	1	36
99A323C90842	2	01
8711C98900AA	2	15
301B2369011C	2	16
695519001190	3	11

- Bridge/switch has a table that shows for each MAC Address which port to use for forwarding
- For every packet, the bridge "looks up" the entry for the packets destination MAC address and forwards the packet on that port.
  - Other packets are broadcast – why?
- Timer is used to flush old entries

29

Switch Architecture

- Packets come in one interface, forwarded to output interface based on address.
  - Same idea for bridges, switches, routers: address look up differs
- Control processor manages the switch and executes higher level protocols.
  - E.g. routing, management, ...
- The switch fabric directs the traffic to the right output port.
- The input and output ports deal with transmission and reception of packets.



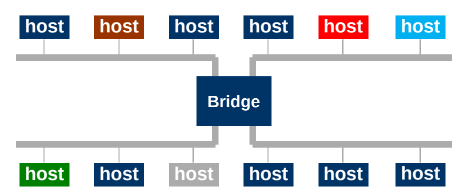
Transparent Bridges

- Design goals:
  - Self-configuring without hardware or software changes
  - Bridge does not impact the operation of the individual LANs, i.e., a set of bridged LANs acts as a single LAN
- Three parts to making bridges transparent:
  1) Forwarding frames
  2) Learning addresses/host locations
  3) Spanning tree algorithm

31

Learning Bridges

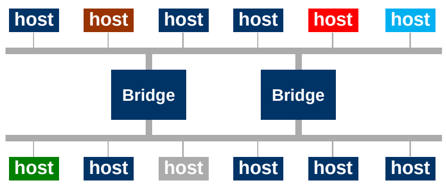
- Manually filling in bridge tables?
  - Time consuming, error-prone
- Keep track of source address of packets arriving on every link, showing what segment hosts are on
  - Fill in the forwarding table based on this information



32

But Does it Scale?

- More complex topologies can provide redundancy.
  - Especially important in larger networks
- But this creates a problem: loops!
- Solution: spanning tree



33

Spanning Tree Protocol Overview

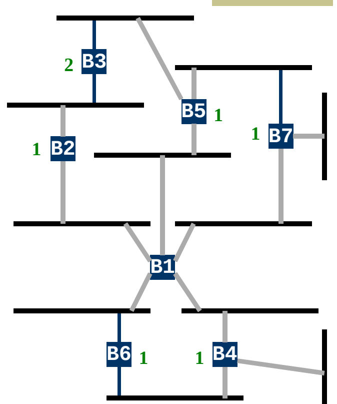
Embed a tree that provides a single unique path to each destination:

1) Elect a single bridge as a root bridge
2) Each bridge calculates the distance of the shortest path to the root bridge
3) Each LAN identifies a *designated bridge*, the bridge closest to the root. It will forward packets to the root.
4) Each bridge determines a *root port*, which will be used to send packets to the root
5) Identify the ports that form the spanning tree

34

Spanning Tree Algorithm Steps

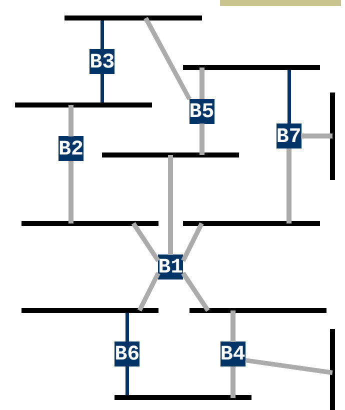
- Root of the spanning tree is the bridge with the lowest identifier.
  - All ports are part of tree
- Each bridge finds shortest path to the root.
  - Remembers port that is on the shortest path
  - Used to forward packets
- Select for each LAN the designated bridge that has the shortest path to the root.
  - Identifier as tie-breaker
  - Responsible for that LAN



35

Spanning Tree Algorithm

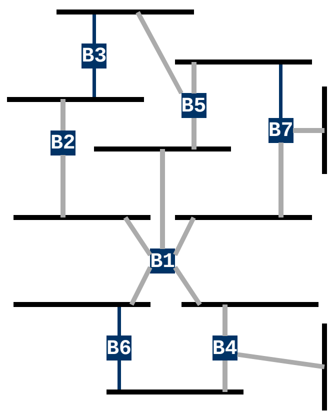
- Each node sends configuration message to all neighbors.
  - Identifier of the sender
  - Id of the presumed root
  - Distance to the presumed root
  - E.g. B5 sends (B5, B5, 0)
- When B receive a message, it decide whether the solution is better than their local solution.
  - A root with a lower identifier?
  - Same root but lower distance?
  - Same root, distance but sender has lower identifier?
- After convergence, each bridge knows the root, distance to root, root port, and designated bridge for each LAN.



36

Spanning Tree Algorithm (part 2)

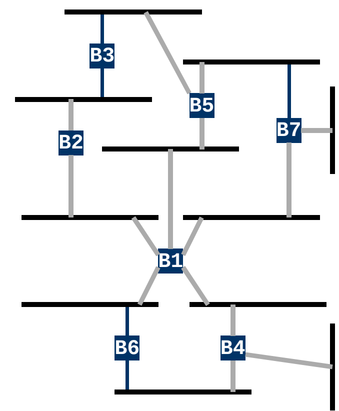
- Each bridge B can now select which of its ports make up the spanning tree:
  - B's root port
  - All ports for which B is the designated bridge on the LAN
- Bridges can not configure their ports.
  - *Forwarding state* or *blocked state*, depending on whether the port is part of the spanning tree
- Root periodically sends configuration messages and bridges forward them over LANs they are responsible for.



37

Spanning Tree Algorithm Example

- Node B2:
  - Sends (B2, B2, 0)
  - Receives (B1, B1, 0) from B1
  - Sends (B2, B1, 1) "up"
  - Continues the forwarding forever
- Node B1:
  - Will send notifications forever
- Node B7:
  - Sends (B7, B7, 0)
  - Receives (B1, B1, 0) from B1
  - Sends (B7, B1, 1) "up" and "right"
  - Receives (B5, B5, 0) - ignored
  - Receives (B5, B1, 1) - better
  - Continues forwarding the B1 messages forever to the "right"



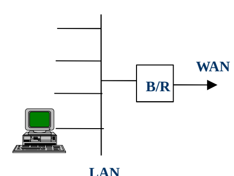
38

Ethernet Switches

- Bridges make it possible to increase LAN capacity.
  - Packets are no longer broadcasted - they are only forwarded on selected links
  - Adds a switching flavor to the broadcast LAN
- Ethernet switch is a special case of a bridge: each bridge port is connected to single host.
  - Simplifies the protocol and hardware used (only two stations on the link) – no longer full CSMA/CD
  - Can make the link full duplex (really simple protocol!)
  - Can have different port speeds on the same switch

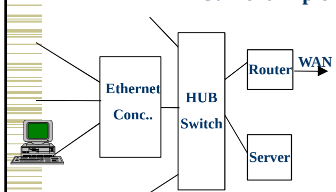
Ethernet Evolution

Early Implementations



- A Local Area Network
- MAC addressing, non-routable
- BUS or Logical Bus topology
- Collision Domain, CSMA/CD
- Bridges and Repeaters for distance/capacity extension
- 1-10Mbps: coax, twisted pair (10BaseT)

Current Implementations

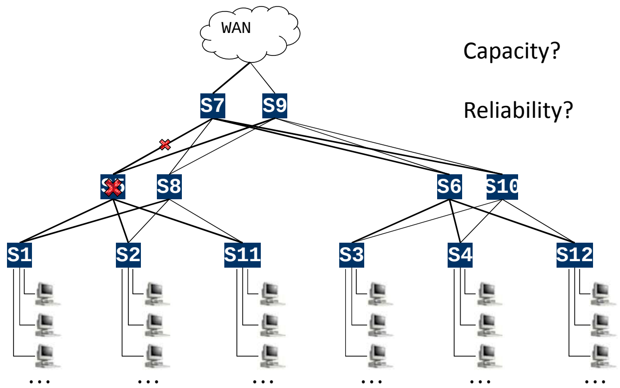


- Switched solution
- Little use for collision domains
- 80% of traffic leaves the LAN
- Servers, routers 10 x station speed
- 10/100/1000 Mbps, 10gig coming: Copper, Fiber
- 95% of new LANs are Ethernet

CSMA Carrier Sense Multiple Access

CD - Collision Detection

Typical Campus Topology



Capacity?

Reliability?