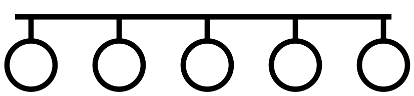
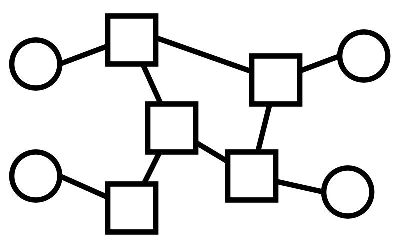
Lecture 7
Datalink – Bridging and Switching

Peter Steenkiste
Departments of Computer Science and
Electrical and Computer Engineering
Carnegie Mellon University

15-441 Networking, Spring 2006
http://www.cs.cmu.edu/~prs/15-441

Datalink Layer Architectures





- Multiple access networks.
  - » Contention-based
  - » Taking turns
  - » Partitioning
- Switched networks.
  - » Switch design
  - » Finding a path

Today's Lecture

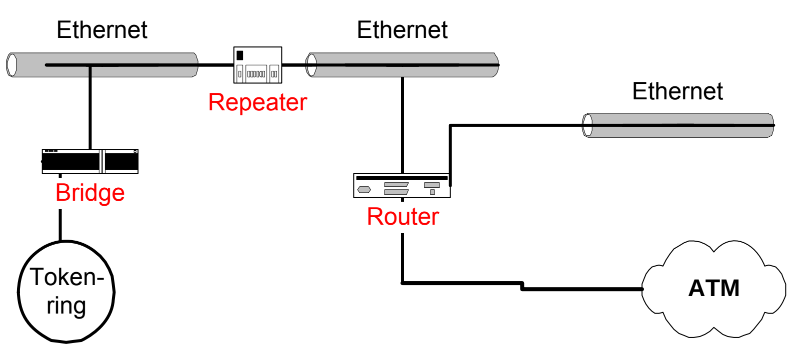
- **Bridges.**
- **Spanning tree protocol.**
- **Switching.**
- **Connectivity to the home.**

Peter A. Steenkiste, SCS, CMU

3

Internetworking

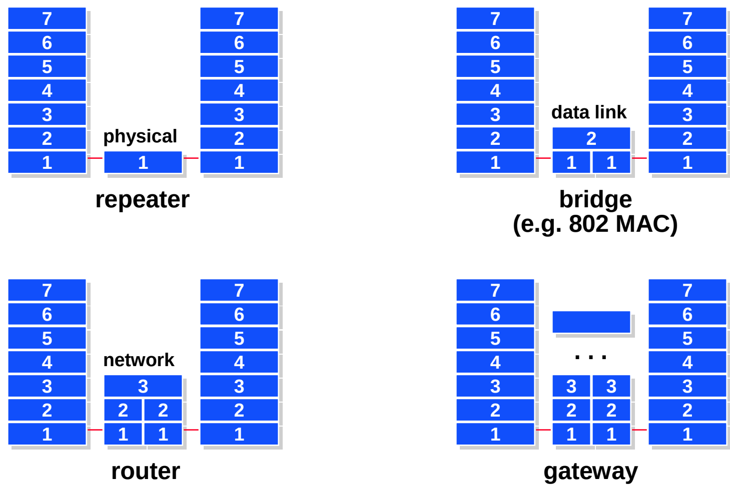
- There are many different devices for interconnecting networks.



Peter A. Steenkiste, SCS, CMU

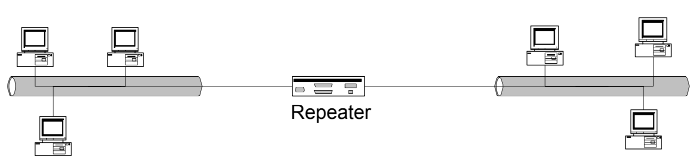
4

Internetworking Options



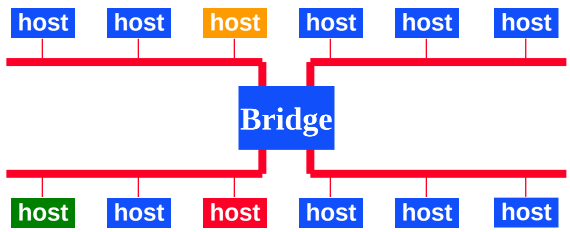
Repeaters

- Used to interconnect multiple Ethernet segments
- Merely extends the baseband cable
- Amplifies all signals including collisions



Building Larger LANs: Bridges

- Bridges connect multiple IEEE 802 LANs at layer 2.
  - » Only forward packets to the right port
  - » Reduce collision domain compared with single LAN
- In contrast, repeaters rebroadcast packets.



Transparent Bridges

- Design goals:
  - » "Plug and play" capability
  - » Self-configuring without hardware or software changes
  - » Bridge do not impact the operation of the individual LANs
- Three parts to making bridges transparent:
  1) Forwarding of frames
  2) Learning of addresses
  3) Spanning tree algorithm

The Forwarding Database

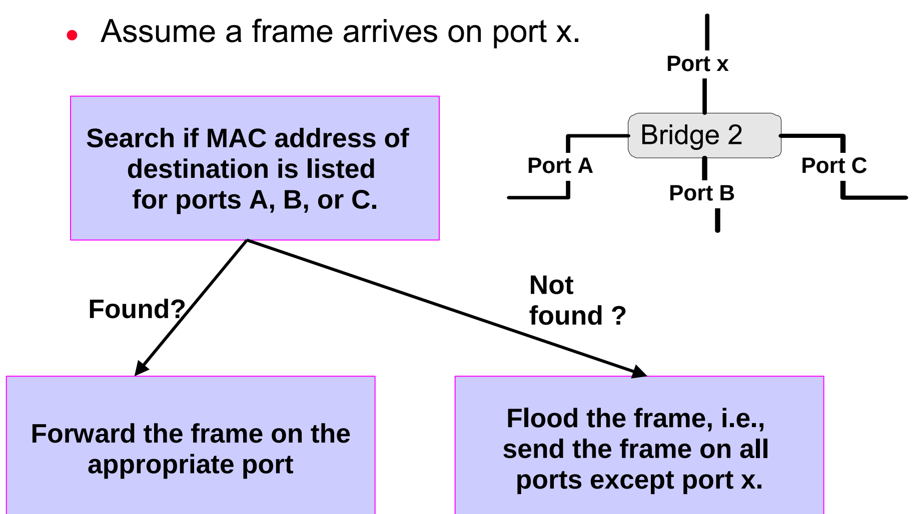
- **Each switch maintains a forwarding database:**

```
<MAC address, port, age>
    MAC address: host or group address
    Port: port number on the bridge
    Age: age of the entry
```

- **Meaning: A machine with MAC address lies in the direction of number port of the bridge**
- **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 broadcasted – why?**

Frame Forwarding

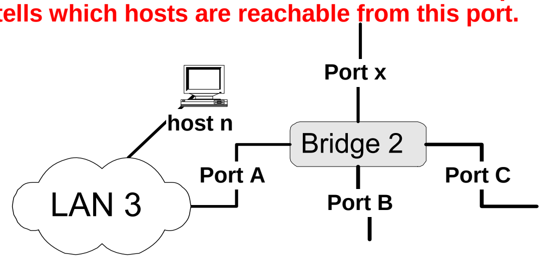
- Assume a frame arrives on port x.



Address Learning

- In principle, the forwarding database could be set statically (=static routing)
- In the 802.1 bridge, the process is made automatic with a simple heuristic:

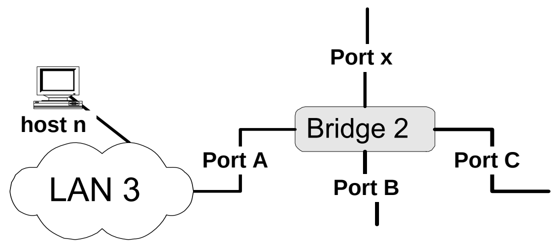
  The source field of a frame that arrives on a port tells which hosts are reachable from this port.



Address Learning 2

Algorithm:

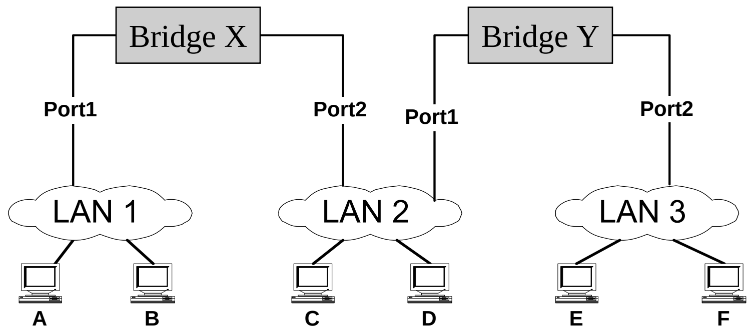
- For each frame received, the source stores the source field in the forwarding database together with the port where the frame was received.
- All entries are deleted after some time.



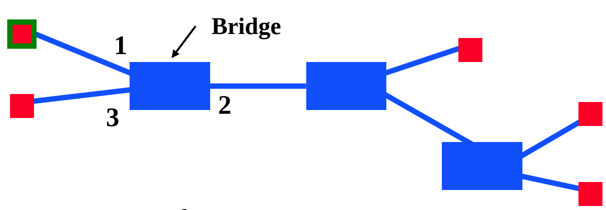
Example

- Consider the following packets:
  <Src=A, Dest=F>, <Src=C, Dest=A>, <Src=E, Dest=C>

- What have the bridges learned?



Address Lookup

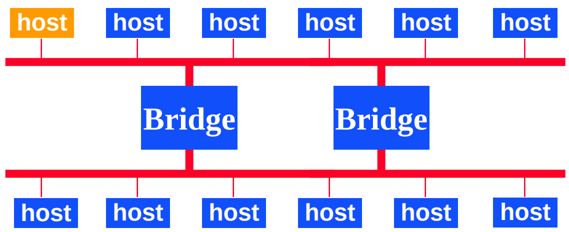


Address	Next Hop	Info
A21032C9A591	1	8:36
99A323C90842	2	8:01
8711C98900AA	2	8:15
301B2369011C	2	8:16
695519001190	3	8:11

- **Address is a 48 bit IEEE MAC address.**
- **Next hop: output port for packet.**
- **Timer is used to flush old entries (15 second default)**
- **Size of the table is equal to the number of hosts.**

Spanning Tree Bridges

- More complex topologies can provide redundancy.
  - » But can also create loops.
- What is the problem with loops?

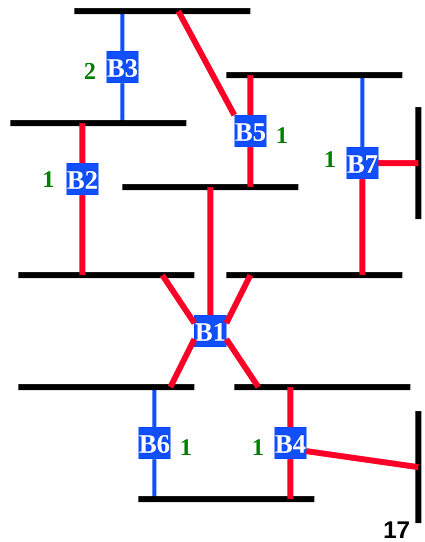


Solution: Spanning Trees

- A solution to the loop problem is to embed a tree into the topology
- IEEE 802.1 has an algorithm that builds and maintains a spanning tree in a dynamic environment.
- Want distributed, fully automated protocol.
  - » Bridges exchange messages to configure the bridge
  - » Configuration BPDUs: Configuration Bridge Protocol Data Units

Spanning Tree Example

- 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



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) Select the ports that form the spanning tree

Concepts

- Each bridge has a unique identifier:

  Bridge ID = <MAC address + priority level>

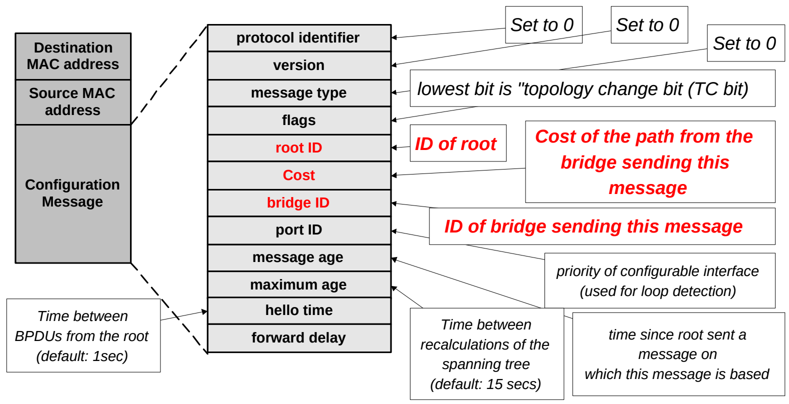
  Note that a bridge has several MAC addresses (one for each port), but only one ID

- Each port within a bridge has a unique identifier (port ID).
- Root Bridge: The bridge with the lowest identifier is the root of the spanning tree.
- Path Cost: Cost of the least cost path to the root from the port of a transmitting bridge; Assume it is measured in #Hops to the root.

More Concepts

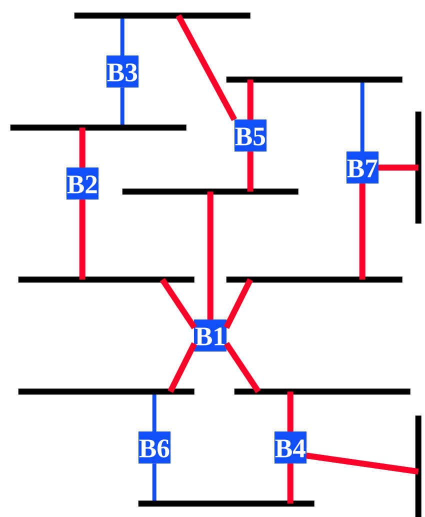
- Root Port: Each bridge has a root port which identifies the next hop from a bridge to the root.
- Designated Bridge, Designated Port: Single bridge on a LAN that provides the minimal cost path to the root for this LAN:
  - if two bridges have the same cost, select the one with highest priority
  - if the min-cost bridge has two or more ports on the LAN, select the port with the lowest identifier

Configuration BPDUs



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 its 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.**



Ordering of Messages

- We can order BPDU messages with the following ordering relation "∠":

M1 | ID R1 | C1 | ID B1 | ∠ | ID R2 | C2 | ID B2 | **M2**

If (R1 < R2)

  M1 ∠ **M2**

elseif ((R1 == R2) and (C1 < C2))

  - M1 ∠ M2

elseif ((R1 == R2) and (C1 == C2) and (B1 < B2))

  - M1 ∠ M2

Peter A. Steenkiste, SCS, CMU

23

Determine the Root Bridge

- Initially, all bridges assume they are the root bridge.
- Each bridge B sends BPDUs of this form on its LANs:

| B | 0 | B |

- Each bridge looks at the BPDUs received on all its ports and its own transmitted BPDUs.
- Root bridge is the smallest received root ID that has been received so far (Whenever a smaller ID arrives, the root is updated)

Peter A. Steenkiste, SCS, CMU

24

Calculate the Root Path Cost

- <u>At this time</u>: A bridge B has a belief of who the root is, say R.
- Bridge B determines the Root Path Cost (Cost) as follows:
  - **If B = R :** **Cost = 0.**
  - **If** $B \angle P$ **Cost = {Smallest Cost in any of BPDUs that were received from R} + 1**
- B's root port is the port from which B received the lowest cost path to R (in terms of relation "∠").
- Knowing R and Cost, B can generate its BPDU (but will not necessarily send it out):

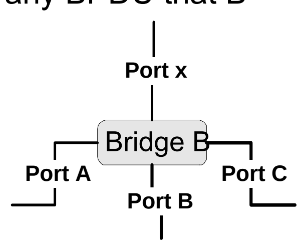
R	Cost	B

Determine the Root Port

- <u>At this time</u>: B has generated its BPDU

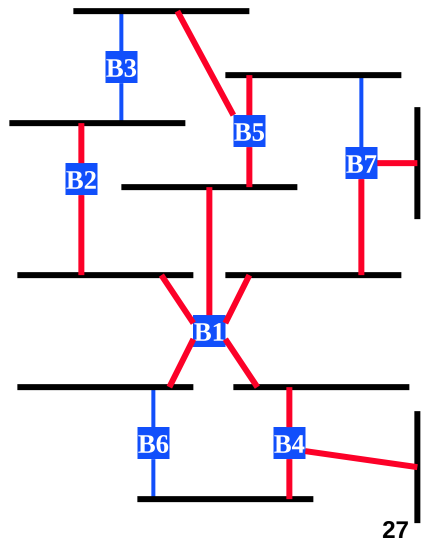
R	Cost	B

- B will send this BPDU on one of its ports, say port x, only if its BPDU is lower (via relation ∠) than any BPDU that B received from port x.
- In this case, B also assumes that it is the designated bridge for the LAN to which the port connects.



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 now 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.

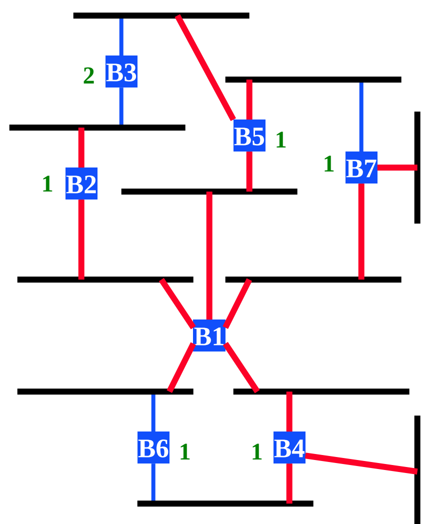


Peter A. Steenkiste, SCS, CMU

27

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"



Peter A. Steenkiste, SCS, CMU

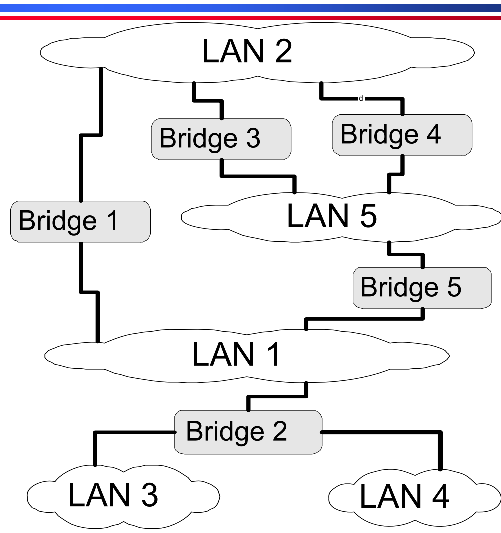
28

Ethernet Switches

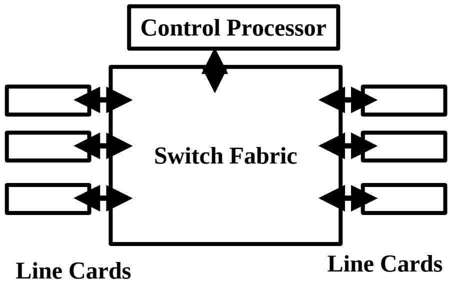
- Bridges make it possible to increase LAN capacity.
  - » Packets are no longer broadcasted - they are only forwarded on selected links
  - » Adds a switching flavor to the broadcast LAN
- Ethernet switch is a special case of a bridge: each bridge port is connected to a single host.
  - » Can make the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

Scalability of Bridging

- Complex topology
  - » avoid loop
  - » use alternate path
- Explosion forwarding table and routing cost
  - » Forwarding table grows with number of nodes
  - » Convergence time grows with longest path
- Solution:
  - » hierarchical addresses



Structure of A Generic Communication Switch



- Switches
  - circuit switch
  - Ethernet switch
  - ATM switch
  - IP router

- **Switch fabric**
  - high capacity interconnect
- **Line card**
  - address lookup in the data path (forwarding)
- **Control Processor**
  - load the forwarding table (routing or signaling)

Peter A. Steenkiste, SCS, CMU

31

Addressing and Look-up

- Flat address
  - **Ethernet: 48 bit MAC address**
  - **ATM: 28 bit VPI/VCI**
  - **DS-0: timeslot location**
- Limited scalability
- High speed lookup

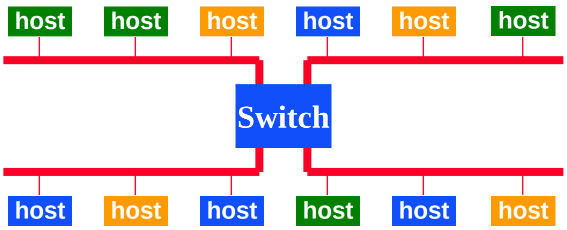
- Hierarchical address
  - **IP <network>.<subnet>.<host>**
  - **Telephone: country.area.home**
- Scalable
- Easy lookup if boundary is fixed
  - **telephony**
- Difficult lookup if boundary is flexible
  - **longest prefix match for IP**

Peter A. Steenkiste, SCS, CMU

32

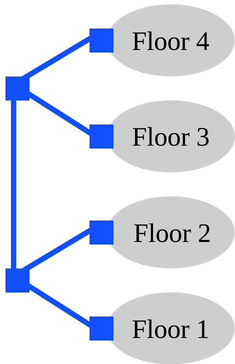
Virtual LANs

- Single physical LAN infrastructure that carries multiple "virtual" LANs simultaneously.
- Each virtual LAN has a LAN identifier in the packet.
  - » Switch keeps track of what nodes are on each segment and what their virtual LAN id is
- Can bridge and route appropriately.
- Broadcast packets stay within the virtual LAN.
  - » Limits the collision domain for the packet



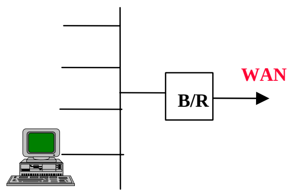
Example LAN Configuration

- 10 or 100 Mbit/second connectivity to the desk top using switch or hubs in wiring closets.
- 100 or 1000 Mbit/second switch fabric between wiring closets or floors.
- Management simplified by having wiring based on star topology with wiring closet in the center.
- Network manager can manage capacity in two ways:
  - » speed of individual links
  - » hub/bridge/switch tradeoff



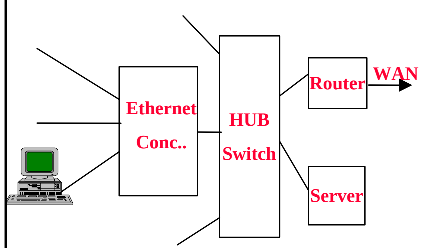
Ethernet – Anything but Name and Framing

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



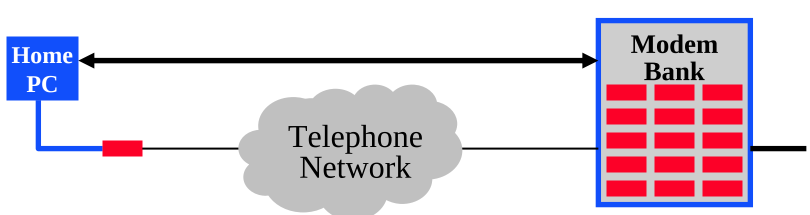
- Switched solution
- Little use for collision domains
- 80% of traffic leaves the LAN
- Servers, routers 10 x station speed
- 10/100/1000 Mbps, 10gig coming: Copper, Fiber
- 95% of new LANs are Ethernet

CSMA - Carrier Sense Multiple Access

CD - Collision Detection

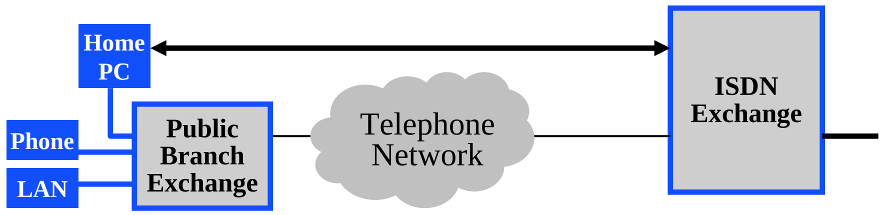
Modems

- Modem offers a bit stream.
  - » Aggressive signal processing has dramatically increased the available throughput - beats the Nyquist limit!
- SLIP: Serial Line IP.
  - » Protocol to sent IP packets with minimum framing
  - » Lacks authentication, error detection, non-IP support, ..
- PPP: Point-to-Point Packets.
  - » Better framing, error control, and testing support
  - » Can negotiate choice of higher layer protocols, IP address
  - » Can support unreliable and reliable transmission



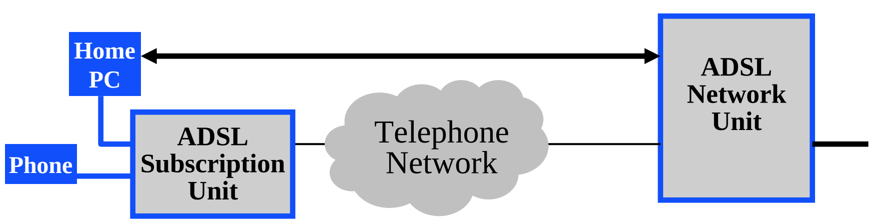
Integrated Services Digital Network (ISDN)

- ISDN integrates voice and data services.
- Provides a set of bit pipes that can be used for voice, data, signaling.
  - » Implemented by using time multiplexing
- Basic rate ISDN offers to 64Kbs data bit pipes and one 16 Kbs signaling channel.



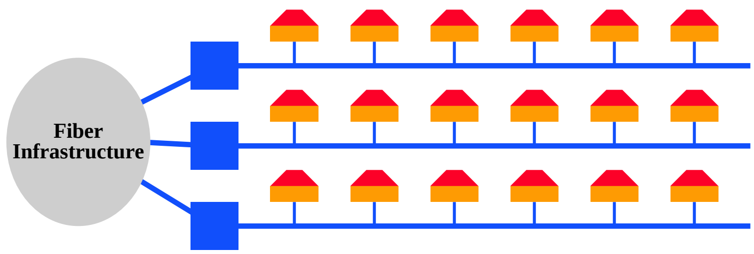
Digital Subscriber Line

- Squeeze more bandwidth out of the telephone line using advanced signal processing.
- Asymmetric digital subscriber line (ADSL).
  - » More "download" bandwidth, e.g. video on demand or web surfing
  - » Example: T1 incoming path, 64 Kbs outgoing path
- (Symmetric) digital subscriber line (DSL).
  - » Same bandwidth both ways, e.g. 768 Kbs



Cable Modem

- Use cable infrastructure for data service.
  - » Inherently has more bandwidth
- The last mile is a shared infrastructure that was designed for broadcasting.
  - » Meaning: the bandwidth is shared by users
  - » Example: 27 Mbs shared incoming path; 768 Kbs common outgoing path



Comparison

- Modems use “worst case” technology.
  - » Has to fit within any voice channel so encoding suboptimal
- ISDN can be more aggressive.
  - » But it is old technology by now
- DSL is highly optimized for the transmission medium.
  - » But there are some constraints on distance
- Cablemodem uses a transmission medium that has inherently a higher bandwidth, but the network architecture will limit throughput.
  - » Designed for broadcasting, not for point-point connections