Computer Networks
ICS 651

- Ethernet and IP
- Ethernet NIC
- Hubs
- Learning Bridges and Ethernet Switches
- Ethernet Network Design

Ethernet and IP

- use ethernet to carry IP packet
- IP host/router knows: IP address of next hop, interface on which to send it
- must fragment if header + data > 1500 bytes
- must pad packet if header + data < 46 bytes
- could broadcast packet, let all but one IP on the network discard it
- instead, broadcast a request, let the specific host reply: Address Resolution Protocol

Ethernet Evolution and Summary

- Xerox
- 3 Mb/s system with 2- or 6-byte addresses
- 10-Base 5: coaxial 10Mb/s with vampire taps
- 10-Base 2 (thinwire): coaxial 10Mb/s with point-to-point
- 10-Base T: CAT-5 to hubs, no actual shared medium
- 100- and 1000 Mb/s Ethernet(s) over twisted pair or fiber
- Gigabit and 10-Gigabit Ethernet
- Ethernet is usually classified as: Carrier Sense Multiple Access (CSMA) with Collision Detection: CSMA/CD

Network Interface Card

- Analog portion can:
  - read and write bits
  - detect and generate jamming signal
  - synchronize to incoming preamble
- Digital receive section:
  - read destination address and discard unless for us or broadcast (configurable)
  - discard if CRC does not match
  - DMA to main memory
- Digital transmit section:
  - DMA, buffer, and send frame
  - compute and send CRC
  - retransmit in case of collision

Hub operation

- A hub interconnects one or more NICs or hubs
- receives a frame F from one of its ports
- the hub must immediately forward F to all its other ports
- if a transmission is in progress on another port, hub must drop both frames and jam all the ports
- if a frame ends with a jamming signal, we must jam all other ports
- a hub is a layer-1 device

Learning Bridge

- like a hub, but we look at the frame header
- if we get frames from A on port I, remember that
- any future frames **for** A (within 60 seconds) can be sent only on port I, do not need to be sent on the other ports
- advantages: breaks the collision domain
- disadvantages: must buffer at least 6 bytes (may have to buffer entire frame)
- a bridge is a layer-2 device
- a bridge could interconnect different LAN technologies, by translating headers

Ethernet Switch

- like a learning bridge
- hardware forwarding of frames from one interface to the next
- buffering and queueing of frames for each interface
- distributed spanning-tree algorithm used among switches to determine where to send broadcast frames
- advantages: breaks the collision domain, reduces the number of collisions
- disadvantages: may have to buffer entire frames, more complex
- a switch is a layer-2 device

Ethernet Switches and Spanning Tree

- like hubs, etherswitches must avoid forwarding loops
- to allow redundant links, some etherswitches run the **Spanning Tree Protocol** (STP)
- STP figures out which links must be used in forwarding, and which must be blocked, by building a spanning tree
- the spanning tree is rooted at the node with the lowest priority or the lowest Ethernet address
- the process to determine the spanning tree is automatic
- in case a link or switch goes down, the STP quickly (usually less than a minute for original STP, under 10s for RSTP) starts using previously blocked interfaces
- STP must usually be enabled by the operator

Names and Layers

1. Hub: forward to all (physical layer)

2. Bridge: something interconnecting LANs (data link layer).

   Switch: a bridge that uses hardware to forward packets (data link layer)

3. Gateway: old name for router (network layer).

   Router: a box that forwards packets among networks (network layer).

   IP switch: a router with hardware forwarding (network layer).

4. Firewall or NAT box: a forwarding engine that looks at and rewrites network and transport layer headers (transport layer).

Ethernet Network Design

- star topology using hubs
- at most 2500 m. (for 10Mb/s) with up to 4 repeaters between any two hosts (one hub to a central hub)
- avoid collisions if possible:
  - split into segments connected by switches, routers
  - switch to 100Mb/s, Gb Ethernet
  - connect 100Mb/s, Gb directly to switch (full-duplex mode)

Ethernet Collision Probability

- likelihood that two hosts are waiting to send at end of current packet (assume all packets are the same size)
- load (probability of 1 host wanting to send) is p
- probability of collision is p2
- cost of collision is: 1 packet lost to collision (plus latency), giving additional load: delta = $\sum_{i = 1 \ldots \infty} p^{2i}$

load p	p^2	delta
0.1	0.01	0.01
0.3	0.09	0.1
0.5	0.25	0.3

Splitting the Collision Domain

- Suppose a network is to be split into parts
- Etherswitch:
  - security considerations: who gets to see what packets? (e.g. faculty vs. students, accounting vs. engineering)
  - performance: is it all-to-all, partitionable, or one server to many clients?
  - single server case: put the server on its own high-speed link to the switch, every one else on slower shared links
  - cost: more ports, or higher speeds?
- router: all of the above, plus address (re)assignment considerations (easier with DHCP)

Security, Firewalls, Rogue IPs

- anyone with root access can read all the packets on the network
- not all of the network traffic is encrypted – but these days, most of the security-conscious network traffic is encrypted
- no ideal solution, try to make broadcasts smaller so only mutually trusted individuals can see each other's broadcasts
- firewall (L4 switch, NAT box): appropriate combinations of inside/outside and accessible/inaccessible
- visitor finds data jack, plugs in, doesn't work, so visitor picks an IP address, works! :(
  - other host using the same IP address has intermittent problems
  - solved by DHCP

IP address assignments

- if using a router to split a network, all addresses in one subnet must be on one side, addresses in the other subnet must be on the other side
- assign IPs by location: "these addresses for the POST building"
- assign IPs by security split: "these addresses for the faculty"
- assign IPs dynamically: DHCP

Dynamic Host Configuration Protocol

- Dynamic assignment, on demand, of temporary IP address "leases"
- a lease can be renewed, but no guarantees
- hard to allow TCP connections initiated outside the network, since the IP address may change:
  - bad for servers
  - good for security
  - or, could hard-code some of the DHCP addresses to assign only to specific MAC addresses
- a DHCP block can be configured in the firewall, so no incoming connections would be allowed for DHCP addresses
- harder to automatically snoop on specific host (again, both good and bad)

Finding Rogue Hosts

- a host that has a duplicate IP address (or that offers illegitimate ARP responses) is hard to detect on a broadcast network
- partition the network and ping on both sides (very painful -- the network MUST stay up)
- see (using the ARP cache) which host(s) you are reaching on which side of a switch (ping will usually tell if it is getting duplicate replies)
- in a small group, find all your visitors and anyone who has bought or is bringing in a laptop...

Ethernet-based home setup

- one host, H, connected to the internet (the “router”)
- internally, 100Mb/s or 1Gb/s switch-based Ethernet
- externally cable modem, ADSL, or other technology
- H does Network Address Translation (L4 translation) so:
  - IP packets going out are rewritten to have H as source address, and usually a different source port
  - incoming packets are rewritten to have the correct destination host and port number
  - for ICMP Echo, can rewrite identifier
  - for ICMP Error messages, need to look up port number in original header
- H can also:
  - perform firewall functions
  - allocate addresses via DHCP
  - be the default router for all internal computers

Network Address Translation and Firewall

- for TCP we know when the connection starts and ends, so can de/allocate state in the NAT box
  - and we generally disallow inbound connections
- for UDP/ICMP we don't know when the connection is done, so we must cache: allocate when the first packet between (pair of IPs and ports) is sent from inside to outside, deallocate after a timeout
  - drop inbound UDP/ICMP that do not match a translation
- it generally doesn’t matter what local port we use for outgoing connections
- to run a server, must “poke a hole” in the firewall, i.e. tell the firewall to accept inbound TCP connections to a given port