ICS 351: Today's plan

- review of packet forwarding: switching and Internet routing

packets

- data is divided into units of finite size, called packets or frames, datagrams, PDUs, etc.
- the finite size limits the delay when multiplexing data from different sources
- to accomodate packets, longer data must be split up into smaller units: fragmentation (IP), segmentation (TCP)
- each packet has one or more headers, and possibly one or more trailers

network structure

- networks include point-to-point links (most of today's Internet, including the lab), or physical broadcast media (older technologies)
- networks are interconnected by a variety of switching technologies: hubs, ethernet bridges or switches, routers (IP switches), NAT
- a switch implies hardware forwarding

Ethernet switching

- every packet on an Ethernet has an Ethernet header
- a hub works on the physical layer, and rebroadcasts the packet without regard to the header
- an Ethernet switch works on the data-link layer, and rebroadcasts the packet unless it has better information:
- o a learning switch that has seen this destination address as the source address of previous packets, only forwards on the interface the prior packet came from
- o a switch that implements the Spanning Tree Protocol (STP), only forwards on interfaces that are part of the spanning tree
- data-link forwarding does not scale to really large networks, since it requires broadcasting at least some of the time
- special provisions (e.g. STP) are needed to prevent forwarding loops in Ethernet networks, and redundant links cannot be used for load balancing

Internet addresses

- for scaling, it is important that network addresses be assigned based on their point of connection to the Internet (MAC addresses are encoded in the hardware)
- in IP, there is a network part of the address, and a host part of the address: all hosts on the same network should have the same network part of the address
- IP addresses and netmasks must be configured (perhaps automatically, via DHCP) for every interface on the Internet!
- the netmask determines the number of bits in the IPv4 address that are in the network part of the address
- different netmasks can be used in different parts of a network, allowing for subnetting

Internet Routing

- where there is broadcasting, it is limited, e.g. to routers within an OSPF network
- each router processes the data it receives, updates its routing table accordingly
- Distance-Vector (RIP): the data is essentially the neighbor's routing table, better routes are added to the routing table, and worse routes from that neighbor may be discarded or time out
- Path-Vector (BGP) is similar, but the entire path (the list of each AS in the path) is distributed, rather than just the cost
- Link-State (OSPF): the data is the router's neighborhood information. This is integrated in a graph representation of the network, and Dijkstra's shortest path algorithm gives the actual routes

Exactly what happens when an IP packet is received over Ethernet?

- hardware error checking (CRC) and destination address verification
- interrupt and dispatch to the IP module
- variety of checks, including IP version, header length, and header checksum
- IP destination address lookup:
- o is it for local delivery? If so, check protocol, port number(s), match an existing socket, and deliver to the application
- o otherwise, check (routing cache and) routing table, obtain interface and next hop: gateway if given, otherwise destination address
- decrement TTL
- if necessary, queue for transmission on the selected interface
- if outgoing interface is Ethernet, check ARP table for translation, if not available, broadcast ARP request for next hop IP address, wait for response
- create an Ethernet frame using the MAC address given by ARP, and give to NIC for transmission
- recycle buffer space