Data Link Protocols

Link Layer Services

Framing, Addressing, link access:

- encapsulate datagram into frame, adding header, trailer
- □ channel access if shared medium
- "MAC" addresses used in frame headers to identify source, dest
 - different from IP address!

Error Detection:

- errors caused by signal attenuation, noise.
- □ receiver detects presence of errors:
 - signals sender for retransmission or drops frame

Medium Access Control (MAC) Protocols

MAC Protocols: a taxonomy

Three broad classes:

- Channel Partitioning
 - divide channel into smaller "pieces" (time slots, frequency, code)
 - allocate piece to node for exclusive use

Random Access

- channel not divided, allow collisions
- □ "recover" from collisions

"Taking turns"

Nodes take turns, but nodes with more to send can take longer turns

Channel Partitioning MAC protocols: TDMA

TDMA: time division multiple access

- access to channel in "rounds"
- each station gets fixed length slot (length = pkt trans time) in each round
- unused slots go idle
- example: 6-station LAN, 1,3,4 have pkt, slots
 2,5,6 idle



Channel Partitioning MAC protocols: FDMA

FDMA: frequency division multiple access

- channel spectrum divided into frequency bands
- each station assigned fixed frequency band
- unused transmission time in frequency bands go idle
- example: 6-station LAN, 1,3,4 have pkt, frequency bands 2,5,6 idle



- TDM (Time Division Multiplexing): channel divided into N time slots, one per user; inefficient with low duty cycle users and at light load.
- FDM (Frequency Division Multiplexing): frequency subdivided.

Random Access Protocols

- When node has packet to send
 - □ transmit at full channel data rate R.
 - □ no *a priori* coordination among nodes
- two or more transmitting nodes \rightarrow "collision",
- random access MAC protocol specifies:
 - how to detect collisions
 - □ how to recover from collisions (e.g., via delayed retransmissions)
- Examples of random access MAC protocols:
 - □ slotted ALOHA
 - ALOHA
 - □ CSMA, CSMA/CD, CSMA/CA

Ethernet uses CSMA/CD

- No slots
- adapter doesn't transmit if it senses that some other adapter is transmitting, that is, carrier sense
- transmitting adapter aborts when it senses that another adapter is transmitting, that is, collision detection
- Before attempting a retransmission, adapter waits a random time, that is, random access

Ethernet CSMA/CD algorithm

- Adaptor receives datagram from net layer & creates frame
- 2. If adapter senses channel idle, it starts to transmit frame. If it senses channel busy, waits until channel idle and then transmits
- 3. If adapter transmits entire frame without detecting another transmission, the adapter is done with frame !

- If adapter detects another transmission while transmitting, aborts and sends jam signal
- 5. After aborting, adapter enters exponential backoff: after the mth collision, adapter chooses a K at random from {0,1,2,...,2^m-1}. Adapter waits K⁵12 bit times and returns to Step 2

IEEE 802.11 Wireless LAN

802.11b

- 2.4-5 GHz unlicensed radio spectrum
- □ up to 11 Mbps
- direct sequence
 spread spectrum
 (DSSS) in physical
 layer
 - all hosts use same chipping code
- widely deployed, using base stations

802.11a

- □ 5-6 GHz range
- □ up to 54 Mbps
- 802.11g
 2.4-5 GHz range
 up to 54 Mbps
- All use CSMA/CA for multiple access
- All have basestation and ad-hoc network versions

802.11 LAN architecture



- wireless host communicates with base station
 - base station = access
 point (AP)
- Basic Service Set (BSS) (aka "cell") in infrastructure mode contains:
 - wireless hosts
 - access point (AP): base station
 - ad hoc mode: hosts only

802.11: Channels, association

- 802.11b: 2.4GHz-2.485GHz spectrum divided into 11 channels at different frequencies
 - □ AP admin chooses frequency for AP
 - interference possible: channel can be same as that chosen by neighboring AP!
- host: must associate with an AP
 - scans channels, listening for *beacon frames* containing AP's name (SSID) and MAC address
 - selects AP to associate with
 - □ may perform authentication [Chapter 8]
 - will typically run DHCP to get IP address in AP's subnet

IEEE 802.11: multiple access

- avoid collisions: 2⁺ nodes transmitting at same time
- 802.11: CSMA sense before transmitting
 - don't collide with ongoing transmission by other node
- 802.11: no collision detection!
 - difficult to receive (sense collisions) when transmitting due to weak received signals (fading)
 - □ can't sense all collisions in any case: hidden terminal, fading
 - □ goal: *avoid collisions:* CSMA/C(ollision)A(voidance)





IEEE 802.11 MAC Protocol: CSMA/CA

802.11 sender

 1 if sense channel idle for DIFS then transmit entire frame (no CD)
 2 if sense channel busy then
 start random backoff time timer counts down while channel idle transmit when timer expires if no ACK, increase random backoff interval, repeat 2

802.11 receiver

- if frame received OK

return ACK after **SIFS** (ACK needed due to hidden terminal problem)



Avoiding collisions (more)

- *idea:* allow sender to "reserve" channel rather than random access of data frames: avoid collisions of long data frames
- sender first transmits *small* request-to-send (RTS) packets to BS using CSMA
 - □ RTSs may still collide with each other (but they're short)
- BS broadcasts clear-to-send CTS in response to RTS
- RTS heard by all nodes
 - sender transmits data frame
 - other stations defer transmissions

Avoid data frame collisions completely using small reservation packets!

Collision Avoidance: RTS-CTS exchange



802.11 frame: addressing



802.11 frame: addressing



802.11 frame: more





MAC Addresses and ARP

32-bit IP address:

□ network-layer address

□ used to get datagram to destination IP subnet

MAC (or LAN or physical or Ethernet) address:

used to get datagram from one interface to another physically-connected interface (same network)

□ 48 bit MAC address (for most LANs) burned in the adapter ROM

LAN Addresses and ARP

Each adapter on LAN has unique LAN address



LAN Address (more)

- MAC address allocation administered by IEEE
- manufacturer buys portion of MAC address space (to assure uniqueness)
- Analogy:

(a) MAC address: like Social Security Number

(b) IP address: like postal address

- MAC flat address → portability
 □ can move LAN card from one LAN to another
- IP hierarchical address NOT portable
 depends on IP subnet to which node is attached

ARP: Address Resolution Protocol

Question: how to determine MAC address of B knowing B's IP address?



- Each IP node (Host, Router) on LAN has
 ARP table
- ARP Table: IP/MAC address mappings for some LAN nodes
 < IP address; MAC address; TTL>
 - TTL (Time To Live): time after which address mapping will be forgotten (typically 20 min)

ARP protocol: Same LAN (network)

- A wants to send datagram to B, and B's MAC address not in A's ARP table.
- A broadcasts ARP query packet, containing B's IP address
 - Dest MAC address = FF-FF-FF-FF-FF
 - all machines on LAN receive ARP query
- B receives ARP packet, replies to A with its (B's) MAC address
 - frame sent to A's MAC address (unicast)

- A caches (saves) IP-to-MAC address pair in its ARP table until information becomes old (times out)
 - soft state: information that times out (goes away) unless refreshed
- ARP is "plug-and-play":
 - nodes create their ARP tables without intervention from net administrator

Routing to another LAN

walkthrough: send datagram from A to B via R

assume A know's B IP address



- Two ARP tables in router R, one for each IP network (LAN)
- In routing table at source Host, find router 111.111.111.110
- In ARP table at source, find MAC address E6-E9-00-17-BB-4B, etc

- A creates datagram with source A, destination B
- A uses ARP to get R's MAC address for 111.111.111.110
- A creates link-layer frame with R's MAC address as dest, frame contains A-to-B IP datagram
- A's adapter sends frame
- R's adapter receives frame
- R removes IP datagram from Ethernet frame, sees its destined to B
- R uses ARP to get B's MAC address
- R creates frame containing A-to-B IP datagram sends to B



Ethernet Frame Structure

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



Preamble:

- 7 bytes with pattern 10101010 followed by one byte with pattern 10101011
- used to synchronize receiver, sender clock rates

Ethernet Frame Structure (more)

Addresses: 6 bytes

- if adapter receives frame with matching destination address, or with broadcast address (eg ARP packet), it passes data in frame to net-layer protocol
- otherwise, adapter discards frame
- Type: indicates the higher layer protocol (mostly IP but others may be supported such as Novell IPX and AppleTalk)
- CRC: checked at receiver, if error is detected, the frame is simply dropped



Ethernet Hubs vs. Ethernet Switches

• An Ethernet switch is a packet switch for Ethernet frames

- Buffering of frames prevents collisions.
- Each port is isolated and builds its own collision domain
- An Ethernet Hub does not perform buffering:
 - Collisions occur if two frames arrive at the same time.



.....



Switch



Hubs

... physical-layer ("dumb") repeaters:

- bits coming in one link 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



Switch

link-layer device: smarter than hubs, take active role

□ store, 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
- 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)

Switch Table

- <u>Q</u>: how does switch know that A' reachable via interface 4, B' reachable via interface 5?
- <u>A</u>: each switch has a switch table, each entry:
 - (MAC address of host, interface to reach host, time stamp)
- Iooks like a routing table!
- Q: how are entries created, maintained in switch table?

□ something like a routing protocol?



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

Self learning

- A switch has a switch table
- entry in switch table:
 - □ (MAC Address, Interface, Time Stamp)
 - □ stale entries in table dropped (TTL can be 60 min)
- 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

Filtering/Forwarding

When switch receives a frame:

index switch table using MAC dest address 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,

forward on all but the interface on which the frame arrived

else flood

Switch example

Suppose C sends frame to D



- Switch receives frame from C
 - □ notes in bridge table that C is on interface 1
 - because D is not in table, switch forwards frame into interfaces 2 and 3
- frame received by D

Switch example

Suppose D replies back with frame to C.



- Switch receives frame from from D
 notes in bridge table that D is on interface 2
 because C is in table, switch forwards frame only to interface 1
- frame received by C

Institutional network



Switches vs. Routers

- both store-and-forward devices
 - routers: network layer devices (examine network layer headers)
 - □ switches are link layer devices
- routers maintain routing tables, implement routing algorithms
- switches maintain switch tables, implement filtering,

learning algorithms



5: DataLink Layer

VLANs: motivation



What's wrong with this picture?

What happens if:

- CS user moves office to EE, but wants connect to CS switch?
- single broadcast domain:
 - all layer-2 broadcast traffic (ARP, DHCP) crosses entire LAN (security/privacy, efficiency issues)
- each lowest level switch has only few ports in use

VLANs

Virtual Local Area Network

Switch(es) supporting VLAN capabilities can be configured to define multiple <u>virtual</u> LANS over single physical LAN infrastructure. Port-based VLAN: switch ports grouped (by switch management software) so that *single* physical switch



5: DataLink Layer

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, 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 combined switches plus routers



VLANS spanning multiple switches



- trunk port: carries frames between VLANS defined over multiple physical switches
 - frames forwarded within VLAN between switches can't be vanilla 802.1 frames (must carry VLAN ID info)
 - 802.1q protocol adds/removed additional header fields for frames forwarded between trunk ports

802.1Q VLAN frame format



Data center networks

- 10's to 100's of thousands of hosts, often closely coupled, in close proximity:
 - e-business (e.g. Amazon)
 - □ content-servers (e.g., YouTube, Akamai, Apple, Microsoft)
 - □ search engines, data mining (e.g., Google)
- ✤ challenges:
 - multiple applications, each serving massive numbers of clients
 - managing/balancing load, avoiding processing, networking, data bottlenecks



Inside a 40-ft Microsoft container, Chicago data center

Data center networks

load balancer: application-layer routing



Data center networks

- rich interconnection among switches, racks:
 - increased throughput between racks (multiple routing paths possible)
 - increased reliability via redundancy



Link Layer Encapsulation

Two types of networks at the data link layer

- Broadcast Networks: All stations share a single communication channel
- Point-to-Point Networks: Pairs of hosts (or routers) are directly connected



Broadcast Network

Point-to-Point Network

 Typically, local area networks (LANs) are broadcast and wide area networks (WANs) are point-to-point

Local Area Networks

- Local area networks (LANs) connect computers within a building or a enterprise network
- Almost all LANs are broadcast networks
- Typical topologies of LANs are bus or ring or star
- We will work with Ethernet LANs. Ethernet has a bus or star topology.



Bus LAN

Ring LAN

MAC and LLC

- In any broadcast network, the stations must ensure that only one station transmits at a time on the shared communication channel
- The protocol that determines who can transmit on a broadcast channel are called Medium Access Control (MAC) protocol
- The MAC protocol are implemented in the MAC sublayer which is the lower sublayer of the data link layer
- The higher portion of the data link layer is often called Logical Link Control (LLC)



IEEE 802 Standards

 IEEE 802 is a family of standards for LANs, which defines an LLC and several MAC sublayers

IEEE 802 standard



Ethernet

- Speed: 10Mbps -10 Gbps
- Standard: 802.3, Ethernet II (DIX)
- Most popular physical layers for Ethernet:
 - 10Base5
 - 10Base2
 - 10Base-T
- Thick Ethernet: 10 Mbps coax cable
- Thin Ethernet: 10 Mbps coax cable
- 10 Mbps Twisted Pair
- 100Base-TX 100 Mbps over Category 5 twisted pair
- 100Base-FX 100 Mbps over Fiber Optics
- 1000Base-FX 1Gbps over Fiber Optics
 - 1Gbps over Fiber Optics (for wide area links)
- 10000Base-FX

Bus Topology

10Base5 and 10Base2 Ethernets has a bus topology



Star Topology

Starting with 10Base-T, stations are connected to a hub in a star configuration



Ethernet and IEEE 802.3: Any Difference?

- There are two types of Ethernet frames in use, with subtle differences:
- "Ethernet" (Ethernet II, DIX)
 - An industry standards from 1982 that is based on the first implementation of CSMA/CD by Xerox.
 - Predominant version of CSMA/CD in the US.
- **802.3**:
 - IEEE's version of CSMA/CD from 1985.
 - Interoperates with 802.2 (LLC) as higher layer.

 Difference for our purposes: Ethernet and 802.3 use different methods to encapsulate an IP datagram.

Ethernet II, DIX Encapsulation (RFC 894)

destination address	source address	type	data	CRC
6	6	2	46-1500	4

0800	IP datagram						
2	38-1492						
0806	ARP request/reply	PAD					
2	28	10					
0835	RARP request/reply	PAD					
2	28	10	-				

IEEE 802.2/802.3 Encapsulation (RFC 1042)

destination address	source address	length	DSAP AA	SSAP AA	cntl 03	org code 0	type	data	CRC
6	6	2	1	1	1	3	2	38-1492	4

- **destination address, source address:** MAC addresses are 48 bit

- length: frame length in number of bytes
- DSAP, SSAP: always set to 0xaa
- Ctrl: set to 3
- org code: set to 0
- **type field** identifies the content of the data field
- **CRC:** cylic redundancy check



Point-to-Point (serial) links

- Many data link connections are pointto-point serial links:
 - Dial-in or DSL access connects hosts to access routers
 - Routers are connected by high-speed point-to-point links
- Here, IP hosts and routers are connected by a serial cable
- Data link layer protocols for point-topoint links are simple:
 - □ Main role is encapsulation of IP datagrams
 - □ No media access control needed



Dial-Up Access



Data Link Protocols for Point-to-Point links

SLIP (Serial Line IP)

- First protocol for sending IP datagrams over dial-up links (from 1988)
- Encapsulation, not much else

PPP (Point-to-Point Protocol):

- Successor to SLIP (1992), with added functionality
- Used for dial-in and for high-speed routers

HDLC (High-Level Data Link) :

- Widely used and influential standard (1979)
- Default protocol for serial links on Cisco routers
- Actually, PPP is based on a variant of HDLC

PPP - IP encapsulation

The frame format of PPP is similar to HDLC and the 802.2 LLC frame format:

flag	addr	ddr ctrl proto		data	CRC	flag	
7E	FF	03				7E	
1	1	1	2	<= 1500	2	1	
			0021	IP datagram			
			C021	link control data			
			8021	network control data			

- PPP assumes a duplex circuit
- Note: PPP does not use addresses
- Usual maximum frame size is 1500



Parity Checking

Detect single bit errors
 ▲ d data bits → parity bit
 0111000110101011 0

Single Bit Parity:

Two Dimensional Bit Parity:

Detect and correct single bit errors





101011 101100 parity 011101 001010 parity error correctable single bit error

Internet checksum

<u>Goal</u>: detect "errors" (e.g., flipped bits) in transmitted segment (note: used at transport layer *only*)

Sender:

- treat segment contents as sequence of 16-bit integers
- checksum: addition (1's) complement sum) of segment contents
- sender puts checksum value into UDP checksum field

Receiver:

- compute checksum of received segment
- check if computed checksum equals checksum field value:
 - NO error detected
 - YES no error detected. But maybe errors nonetheless? More later

Checksumming: Cyclic Redundancy Check

- view data bits, D, as a binary number
- choose r+1 bit pattern (generator), G
- goal: choose r CRC bits, R, such that
 - C <D,R> exactly divisible by G (modulo 2)
 - receiver knows G, divides <D,R> by G. If non-zero remainder: error detected!
 - □ can detect all burst errors less than r+1 bits
- widely used in practice (ATM, HDCL)
 d bits r bits —

formula

CRC Example

Want:

 $D^{2^{r}} XOR R = nG$ equivalently: $D^{-2^{r}} = nG XOR R$ equivalently: if we divide D.2r by G, want remainder R R = remainder[-

