

Ethernet

Kai Shen

9/16/2009 CSC 257/457 - Fall 2009 1

Review of Multiple Access Protocols

- **Channel partitioning**
 - divide channel into smaller "pieces" (time slots, frequency, code), allocate piece to node for exclusive use
 - Time Division, Frequency Division, Code Division
- **Random access**
 - channel not divided, allow collisions, recover from collisions
 - ALOHA, S-ALOHA, CSMA, CSMA/CD
- **Centrally controlled/coordinated**
 - coordinate shared access to avoid collisions

9/16/2009 CSC 257/457 - Fall 2009 2

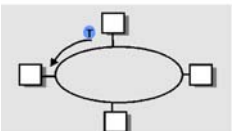
"Centrally Controlled" MAC protocols

Polling:

- master node "invites" slave nodes to transmit in turn
- concerns:
 - polling overhead (bandwidth & latency)
 - single point of failure (master)

Token passing:

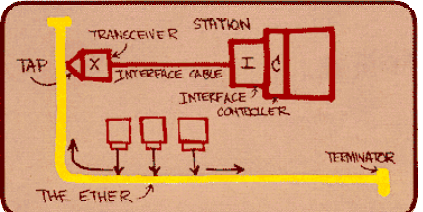
- control **token** passed from one node to next sequentially.
- transmit only when holding the token
- concerns:
 - token overhead (bandwidth & latency)
 - single point of failure (token)



9/16/2009 CSC 257/457 - Fall 2009 3

Ethernet

- Developed at Xerox PARC in 1970s
- First widely used technology for local-area networks (now dominant)
- Standardized as IEEE 802.3
- For shared access link



Metcalfe's Ethernet sketch

9/16/2009 CSC 257/457 - Fall 2009 4

Ethernet Frame Structure

Sender encapsulates data payload in **Ethernet frame**

bits	64	48	48	16		32
	Preamble	Dest addr	Src addr	Len	Body	CRC

- Preamble:**
 - 7 bytes with pattern 10101010 followed by one byte with pattern 10101011
 - indicating the beginning of a frame, synchronizing receiver/sender clocks
- Addresses:**
 - link layer addresses, different from IP addresses
- Length:** indicates the length of the body field (max 1500bytes)
- CRC:** checked at receiver
 - if error is detected, the frame is simply dropped

9/16/2009 CSC 257/457 - Fall 2009 5

Connectionless, Unreliable

- Connectionless:**
 - no handshaking between sender and receiver
- Unreliable:**
 - corrupted frames are simply dropped, no attempt is made on the receiver to detect and recover lost frames
 - Rationale:**
 - chance of corruption/loss is small
 - corrupted or lost packets may be recovered by higher layer protocols, e.g. TCP

9/16/2009 CSC 257/457 - Fall 2009 6

Ethernet uses CSMA/CD

- Random access, no synchronized clocks
- One doesn't transmit if it senses that another is transmitting, that is, **carrier sense**
- Transmitter aborts when it senses that another is transmitting, that is, **collision detection**
- Before attempting a retransmission, one waits a non-deterministic delay

9/16/2009 CSC 257/457 - Fall 2009 7

Ethernet's CSMA/CD Backoff

Load-based backoff:

- Longer backoff when more are competing for the link
- Previous conflicts indicate high load

Specifically:

- First collision: delay K units of time, K is randomly chosen from {0,1}
 - a unit of time separation should be sufficient for carrier sense to avoid conflict (e.g., 5 microsec)
- After second collision: choose K from {0,1,2,3}...
- After tenth collisions, choose K from {0,1,2,3,4,...,1023}
- Also called **exponential backoff**

9/16/2009 CSC 257/457 - Fall 2009 8

Ethernet Efficiency

- t_{prop} = max prop between 2 nodes
- t_{trans} = time to transmit a frame

$$\text{efficiency} \approx \frac{1}{1 + 5t_{prop} / t_{trans}}$$

- Efficiency goes to 1 as t_{prop} goes to 0 or as t_{trans} goes to infinity

Much more efficient than ALOHA, but still decentralized and simple

9/16/2009 CSC 257/457 - Fall 2009 9

Ethernet Technologies: 10Base2

- 10: 10Mbps; 2: under 200 meters max cable length
- thin coaxial cable in a bus topology

- has become a legacy technology

9/16/2009 CSC 257/457 - Fall 2009 10

10BaseT and 100BaseT

- 10/100 Mbps rate; latter called "fast ethernet"
- T stands for Twisted Pair
- Nodes connect to a hub: "star topology"; 100 m max distance between nodes and hub

- Hubs are essentially physical-layer repeaters:
 - bits coming in one link go out all other links
 - no collision detection, no frame buffering

9/16/2009 CSC 257/457 - Fall 2009 11

Encoding

- 10BaseT, 10Base2 - Manchester encoding
- 100BaseT - 4B5B encoding
- Allows clocks in sending and receiving nodes to synchronize to each other
 - no need for a centralized, global clock among nodes
- Still remember their efficiencies??

9/16/2009 CSC 257/457 - Fall 2009 12

Gbit Ethernet

- use standard Ethernet frame format
- efficiency of CSMA/CD

$$\text{efficiency} = \frac{1}{1 + 5t_{prop} / t_{trans}}$$

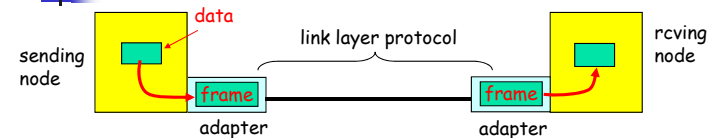
- short distances between nodes to be efficient!
- fortunately today's Ethernet links are not shared
- 10 Gbps now!
 - optical fiber as well as copper cable
 - if latency doesn't change, what apps can benefit?

9/16/2009

CSC 257/457 - Fall 2009

13

Where is link layer implemented?



- link layer implemented in "adaptor" (aka NIC)
 - Ethernet card, wireless card
- receiving side
 - looks for errors, rdt, flow control, etc
 - extracts data, passes to OS at rcvng node
- sending side:
 - encapsulates data in a frame
 - adds error checking bits, rdt, flow control, etc.
- adapter is semi-autonomous
 - has its own processor/buffer

9/16/2009

CSC 257/457 - Fall 2009

14

Interaction with the Host

- Control from the host
 - read/write control status registers
- Notifying the host
 - interrupts, or polling
- Moving data
 - CPU reading/writing, direct memory access (DMA)
- Device drivers (device-specific software executed by host)
 - initialize the adapter
 - manipulate the adapter (e.g. send a frame)
 - handle interrupts

9/16/2009

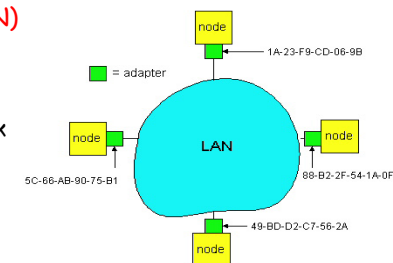
CSC 257/457 - Fall 2009

15

Link-layer Addresses

Link-layer (or MAC, LAN) address:

- uniquely identifying each node (or adapter) connected to a shared link
 - ⇒ integral to multiple access protocols
- 48 bits
- broadcast address - "FF-FF-FF-FF-FF-FF"



9/16/2009

CSC 257/457 - Fall 2009

16

Properties of Link-layer Addresses

- Per-device static address
 - Factory pre-set address for each adapter
 - Change possible (OS may instruct the adapter to use a different MAC address), but not usual
- Universal uniqueness:
 - No two adapters anywhere have the same pre-set address
 - How is this achieved? MAC address allocation administered by IEEE and a manufacturer gets portion of address space.
 - Portability: can move adaptor from one network to another (no configuration required!)

Coexistence of link-layer addresses and network-layer addresses.

9/16/2009 CSC 257/457 - Fall 2009 17

A Problem

Question: how to determine B's link-layer address knowing B's IP address?

9/16/2009 CSC 257/457 - Fall 2009 18

ARP: Address Resolution Protocol

- Specific to Internet Architecture
- Each node (host, router) on LAN has ARP table
- ARP Table: IP/Link-layer address mappings for some LAN nodes
<IP addr; Link-layer addr>

9/16/2009 CSC 257/457 - Fall 2009 19

How does ARP work?

First principle: no manual configuration - "plug-and-play"; must deal with dynamic network changes

- A wants to send datagram to B with an IP address.
- Suppose B's link-layer address is not in A's ARP table.
- A broadcasts ARP query packet, containing B's IP address.
- B receives ARP packet, replies to A with its (B's) link-layer address.
- A caches (saves) IP-to-link-layer address pair in its ARP table until it times out
 - when can a mapping change?

Go to any lab Linux machine:

- "/sbin/arp" shows the current ARP table;
- ping any node not in the table, then see the ARP table again.

9/16/2009 CSC 257/457 - Fall 2009 20



Disclaimer

- Parts of the lecture slides contain original work of James Kurose, Larry Peterson, and Keith Ross. The slides are intended for the sole purpose of instruction of computer networks at the University of Rochester. All copyrighted materials belong to their original owner(s).