

Ethernet

Kai Shen

1/30/2013 CSC 257/457 - Spring 2013 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

1/30/2013 CSC 257/457 - Spring 2013 2

Ethernet

- Developed by Metcalfe 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

1/30/2013 CSC 257/457 - Spring 2013 3

Ethernet Frame Structure

Sender encapsulates data payload in **Ethernet frame**

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:** length of the body field
 - max 1500bytes standard, higher in jumbo frames
- **CRC:** error detection
 - what to do when an error is detected?

1/30/2013 CSC 257/457 - Spring 2013 4

Not Supported

- **Connectionless:**
 - no handshaking between sender and receiver
 - no connection-based resource management, flow control
- **Unreliable:**
 - corrupted frames are dropped, no attempt to retransmit
 - lost frames are ignored (in fact, not even bother to detect)
 - no guarantee of in-order delivery
 - **Rationale:**
 - chance of corruption/loss is small
 - corrupted or lost packets may be recovered by higher layer protocols, e.g. TCP

1/30/2013

CSC 257/457 - Spring 2013

5

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

1/30/2013

CSC 257/457 - Spring 2013

6

Ethernet's CSMA/CD Backoff

Load-based backoff:

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

Specifically:

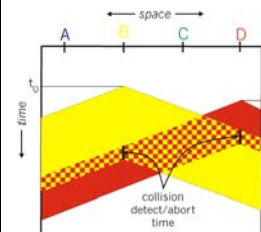
- First collision: delay K units of time, K is randomly chosen from {0,1}
 - a unit of time separation should be sufficient for carrier sense to avoid conflict (e.g., 5 microsec)
- After second collision: choose K from {0,1,2,3}
-
- After tenth collisions, choose K from {0,1,2,3,4,...,1023}
- Also called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

1/30/2013

CSC 257/457 - Spring 2013

7

Ethernet Efficiency



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

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

- Higher efficiency at lower t_{prop} or higher t_{trans}

More efficient than ALOHA, but still decentralized and simple.

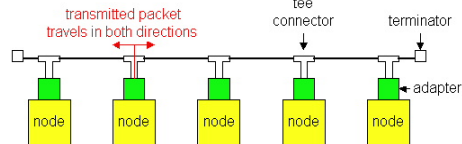
1/30/2013

CSC 257/457 - Spring 2013

8

Ethernet Technologies: 10Base2

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



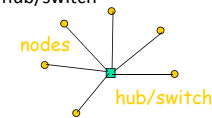
1/30/2013

CSC 257/457 - Spring 2013

9

10BaseT and 100BaseT

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



- Hubs/switches propagate signals:
 - bits coming in one link go out on other links

1/30/2013

CSC 257/457 - Spring 2013

10

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??

1/30/2013

CSC 257/457 - Spring 2013

11

Gbit Ethernet

- 1Gbps and even 10Gbps
 - if latency doesn't change, can all apps benefit?

- efficiency of CSMA/CD

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

- short node distances, large frames to be efficient!
- fortunately today's Ethernet links are not shared

1/30/2013

CSC 257/457 - Spring 2013

12

Where is link layer implemented?

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

1/30/2013 CSC 257/457 - Spring 2013 13

Interaction with the Host

- Control from the host
 - read/write control status registers
- Notifying the host
 - interrupts
- Moving data
 - direct memory access (DMA)

1/30/2013 CSC 257/457 - Spring 2013 14

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”

1/30/2013 CSC 257/457 - Spring 2013 15

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!)

1/30/2013 CSC 257/457 - Spring 2013 16

A Problem

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

The diagram shows a central LAN (purple oval) with several yellow squares representing user hosts. To the right, a green vertical bar represents router B, which is connected to a large blue cloud representing the Internet. A yellow square labeled 'Web Server' is also connected to the Internet.

1/30/2013 CSC 257/457 - Spring 2013 17

ARP: Address Resolution Protocol

- Each node (host, router) on LAN has **ARP** table
 - IP/Link-layer address mappings for LAN nodes <IP addr; Link-layer addr>
- Use ARP table to find link-layer address

The diagram shows a central blue circle labeled 'LAN'. Four yellow squares represent nodes connected to the LAN via green squares labeled 'adapter'. Each node has an IP address and a MAC address:

- Top node: IP 222.222.222.220, MAC 1A-23-F9-CD-06-9B
- Left node: IP 222.222.222.223, MAC 5C-66-AB-90-75-B1
- Right node: IP 222.222.222.221, MAC 88-B2-2F-54-1A-0F
- Bottom node: IP 222.222.222.222, MAC 48-FD-D7-C7-56-2A

1/30/2013 CSC 257/457 - Spring 2013 18

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.


Link-layer protocol or network-layer protocol?

1/30/2013 CSC 257/457 - Spring 2013 19

Two addresses

- Co-existence of link-layer and network-layer addresses
- Why not just one?

1/30/2013 CSC 257/457 - Spring 2013 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).

1/30/2013 CSC 257/457 - Spring 2013 21