

Chapter 5 Data Link Layer

Dott.ssa Chiara Petrioli



*Computer Networking:
A Top Down Approach
Featuring the Internet,
2nd edition.*
Jim Kurose, Keith Ross
Addison-Wesley, July
2002.

All material copyright 1996-2002
J.F Kurose and K.W. Ross, All Rights Reserved

5: DataLink Layer 5a-1

Chapter 5: The Data Link Layer

Our goals:

- understand principles behind data link layer services:
 - error detection, correction
 - sharing a broadcast channel: multiple access
 - link layer addressing
 - reliable data transfer, flow control: *done!*
- instantiation and implementation of various link layer technologies

5: DataLink Layer 5a-2

Chapter 5 outline

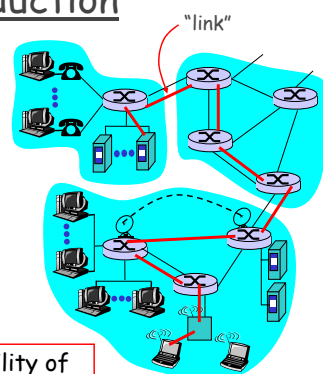
- 5.1 Introduction and services
- 5.2 Error detection and correction
- 5.3 Multiple access protocols
- 5.4 LAN addresses and ARP
- 5.5 Ethernet
- 5.6 Hubs, bridges, and switches
- 5.7 Wireless links and LANs
- 5.8 PPP
- 5.9 ATM
- 5.10 Frame Relay

5: DataLink Layer 5a-3

Link Layer: Introduction

Some terminology:

- hosts and routers are **nodes** (bridges and switches too)
- communication channels that connect adjacent nodes along communication path are **links**
 - wired links
 - wireless links
 - LANs
- 2-PDU is a **frame**, encapsulates datagram



data-link layer has responsibility of transferring datagram from one node to adjacent node over a link

5: DataLink Layer 5a-4

Link layer: context

- Datagram transferred by different link protocols over different links:
 - e.g., Ethernet on first link, frame relay on intermediate links, 802.11 on last link
 - Each link protocol provides different services
 - e.g., may or may not provide rdt over link
- transportation analogy
- trip from Princeton to Lausanne
 - limo: Princeton to JFK
 - plane: JFK to Geneva
 - train: Geneva to Lausanne
 - tourist = datagram
 - transport segment = communication link
 - transportation mode = link layer protocol
 - travel agent = routing algorithm

5: DataLink Layer 5a-5

Link Layer Services

- Framing, link access:
 - encapsulate datagram into frame, adding header, trailer
 - channel access if shared medium
 - 'physical addresses' used in frame headers to identify source, dest
 - different from IP address!
- Reliable delivery between adjacent nodes
 - we learned how to do this already (chapter 3)!
 - seldom used on low bit error link (fiber, some twisted pair)
 - wireless links: high error rates
 - Q: why both link-level and end-end reliability?

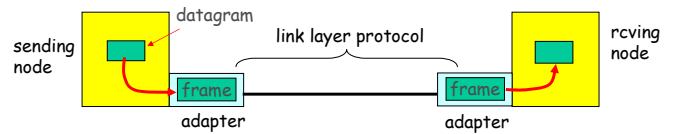
5: DataLink Layer 5a-6

Link Layer Services (more)

- ❑ **Flow Control:**
 - pacing between adjacent sending and receiving nodes
- ❑ **Error Detection:**
 - errors caused by signal attenuation, noise.
 - receiver detects presence of errors:
 - signals sender for retransmission or drops frame
- ❑ **Error Correction:**
 - receiver identifies *and corrects* bit error(s) without resorting to retransmission
- ❑ **Half-duplex and full-duplex**
 - with half duplex, nodes at both ends of link can transmit, but not at same time

5: DataLink Layer 5a-7

Adaptors Communicating



- ❑ link layer implemented in "adaptor" (aka NIC)
 - Ethernet card, PCMCIA card, 802.11 card
- ❑ sending side:
 - encapsulates datagram in a frame
 - adds error checking bits, rdt, flow control, etc.
- ❑ receiving side
 - looks for errors, rdt, flow control, etc
 - extracts datagram, passes to rcving node
- ❑ adapter is semi-autonomous
- ❑ link & physical layers

5: DataLink Layer 5a-8

Chapter 5 outline

- ❑ 5.1 Introduction and services
- ❑ 5.2 Error detection and correction
- ❑ 5.3 Multiple access protocols
- ❑ 5.4 LAN addresses and ARP
- ❑ 5.5 Ethernet
- ❑ 5.6 Hubs, bridges, and switches
- ❑ 5.7 Wireless links and LANs
- ❑ 5.8 PPP
- ❑ 5.9 ATM
- ❑ 5.10 Frame Relay

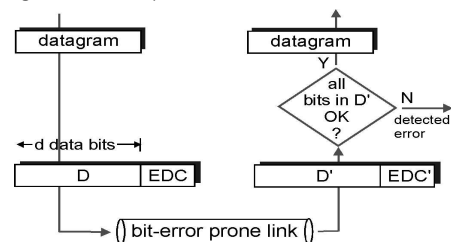
5: DataLink Layer 5a-9

Error Detection

EDC= Error Detection and Correction bits (redundancy)

D = Data protected by error checking, may include header fields

- Error detection not 100% reliable!
 - protocol may miss some errors, but rarely
 - larger EDC field yields better detection and correction

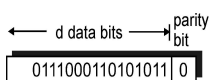


5: DataLink Layer 5a-10

Parity Checking

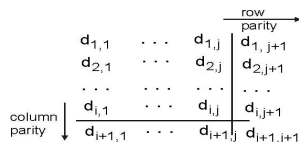
Single Bit Parity:

Detect single bit errors



Two Dimensional Bit Parity:

Detect and correct single bit errors



101011 101011
 111100 101011
 011101 011101
 001010 001010
 no errors parity error
 correctable single bit error

5: DataLink Layer 5a-11

Internet checksum

Goal: 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:

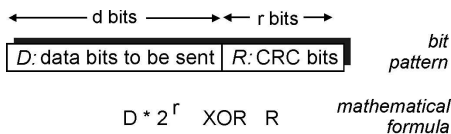
- ❑ compute checksum of received segment
- ❑ check if computed checksum equals checksum field value:
 - NO - error detected
 - YES - no error detected. *But maybe errors nonetheless?* More later

5: DataLink Layer 5a-12

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
 - $\langle D, R \rangle$ exactly divisible by G (modulo 2)
 - receiver knows G , divides $\langle D, R \rangle$ by G . If non-zero remainder: error detected!
 - can detect all burst errors less than $r+1$ bits
- widely used in practice (ATM, HDCL)

Teoria dell'Informazione
/Teoria dei codici vi
spiegherà le motivazioni
matematiche !!!



5: DataLink Layer 5a-13

Chapter 5 outline

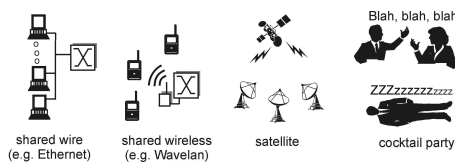
- 5.1 Introduction and services
- 5.2 Error detection and correction
- 5.3 Multiple access protocols
- 5.4 LAN addresses and ARP
- 5.5 Ethernet
- 5.6 Hubs, bridges, and switches
- 5.7 Wireless links and LANs
- 5.8 PPP
- 5.9 ATM
- 5.10 Frame Relay

5: DataLink Layer 5a-14

Multiple Access Links and Protocols

Two types of "links":

- point-to-point
 - PPP for dial-up access
 - point-to-point link between Ethernet switch and host
- broadcast (shared wire or medium)
 - traditional Ethernet
 - upstream HFC
 - 802.11 wireless LAN



5: DataLink Layer 5a-15

Multiple Access protocols

- single shared broadcast channel
- two or more simultaneous transmissions by nodes: interference
 - only one node can send successfully at a time
- multiple access protocol
- distributed algorithm that determines how nodes share channel, i.e., determine when node can transmit
- communication about channel sharing must use channel itself!
- what to look for in multiple access protocols:

5: DataLink Layer 5a-16

Ideal Multiple Access Protocol

Broadcast channel of rate R bps

1. When one node wants to transmit, it can send at rate R .
2. When M nodes want to transmit, each can send at average rate R/M
3. Fully decentralized:
 - no special node to coordinate transmissions
 - no synchronization of clocks, slots
4. Simple

5: DataLink Layer 5a-17

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"
 - tightly coordinate shared access to avoid collisions

5: DataLink Layer 5a-18

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

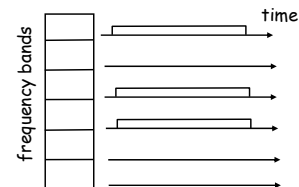


5: DataLink Layer 5a-19

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



5: DataLink Layer 5a-20

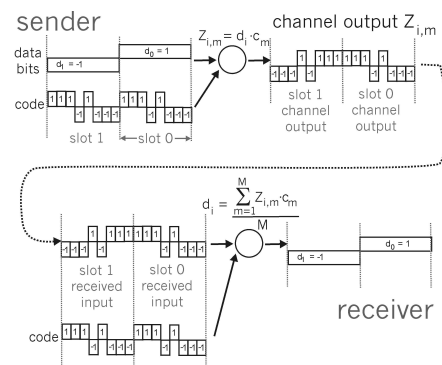
Channel Partitioning (CDMA)

CDMA (Code Division Multiple Access)

- unique "code" assigned to each user; i.e., code set partitioning
- used mostly in wireless broadcast channels (cellular, satellite, etc)
- all users share same frequency, but each user has own "chipping" sequence (i.e., code) to encode data
- M chips = 1 bit time. Ex. Of chipping sequence: 00011011. To send a '1' 00011011 to send a '0' the complement of the chipping sequence 11100100
- encoded signal = (original data) X (chipping sequence)
- decoding: inner-product of encoded signal and chipping sequence
- allows multiple users to "coexist" and transmit simultaneously with minimal interference (if codes are "orthogonal")

5: DataLink Layer 5a-21

CDMA Encode/Decode



Bipolar notation used for pedagogical purposes: binary 0 being -1 and binary 1 being +1

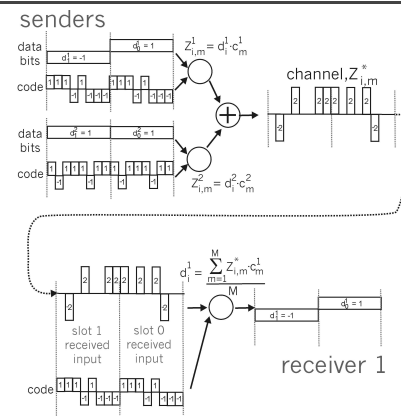
5: DataLink Layer 5a-22

Orthogonal codes properties

- Orthogonal codes: given two sequences S and T , $S \cdot T = 1/m \sum S_i T_i$ is 0
- If $S \cdot T = 0$ also $S \cdot \bar{T} = 0$
- $S \cdot S = 1$
- $S \cdot \bar{S} = -1$
- If multiple stations trasnmits with orthogonal codes it is enough to compute $S \cdot C$ with S received signal and C source chipping sequence to retrieve what trasmitted form the source. Why? Magic?
- Derives from orthogonal codes
- $S \cdot C = (A+B+C) \cdot C = A \cdot C + B \cdot C + C \cdot C = 0 + 0 + 1$

5: DataLink Layer 5a-23

CDMA: two-sender interference



5: DataLink Layer 5a-24

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 → "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

5: DataLink Layer 5a-25

Slotted ALOHA

Assumptions

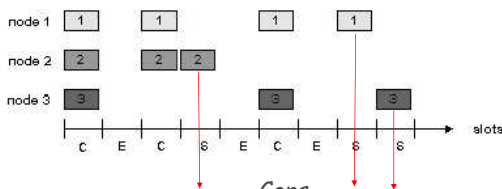
- all frames same size
- time is divided into equal size slots, time to transmit 1 frame
- nodes start to transmit frames only at beginning of slots
- nodes are synchronized
- if 2 or more nodes transmit in slot, all nodes detect collision

Operation

- when node obtains fresh frame, it transmits in next slot
- no collision, node can send new frame in next slot
- if collision, node retransmits frame in each subsequent slot with prob. p until success

5: DataLink Layer 5a-26

Slotted ALOHA



Pros

- single active node can continuously transmit at full rate of channel
- highly decentralized: only slots in nodes need to be in sync
- simple

Cons

- collisions, wasting slots
- idle slots
- nodes may be able to detect collision in less than time to transmit packet

5: DataLink Layer 5a-27

Slotted Aloha efficiency

Efficiency is the long-run fraction of successful slots when there's many nodes, each with many frames to send

- Suppose N nodes with many frames to send, each transmits in slot with probability p
- prob that 1st node has success in a slot = $p(1-p)^{N-1}$
- prob that any node has a success = $Np(1-p)^{N-1}$

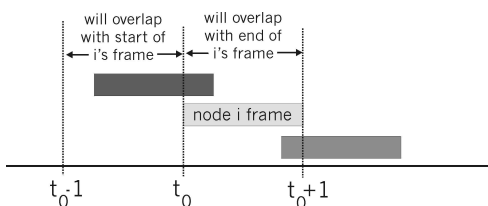
- For max efficiency with N nodes, find p^* that maximizes $Np(1-p)^{N-1}$
- For many nodes, take limit of $Np^*(1-p^*)^{N-1}$ as N goes to infinity, gives $1/e = .37$

At best: channel used for useful transmissions 37% of time!

5: DataLink Layer 5a-28

Pure (unslotted) ALOHA

- unslotted Aloha: simpler, no synchronization
- when frame first arrives
 - transmit immediately
- collision probability increases:
 - frame sent at t_0 collides with other frames sent in $[t_0-1, t_0+1]$



5: DataLink Layer 5a-29

Pure Aloha efficiency

$P(\text{success by given node}) = P(\text{node transmits}) \cdot$

$$\begin{aligned}
 & P(\text{no other node transmits in } [p_0-1, p_0]) \cdot \\
 & P(\text{no other node transmits in } [p_0, p_0+1]) \\
 & = p \cdot (1-p)^{N-1} \cdot (1-p)^{N-1} \\
 & = p \cdot (1-p)^{2(N-1)}
 \end{aligned}$$

... choosing optimum p and then letting $n \rightarrow \infty$...

Even worse ! $= 1/(2e) = .18$

5: DataLink Layer 5a-30

CSMA (Carrier Sense Multiple Access)

CSMA: listen before transmit:

- If channel sensed idle: transmit entire frame
- If channel sensed busy, defer transmission

- Human analogy: don't interrupt others!

5: DataLink Layer 5a-31

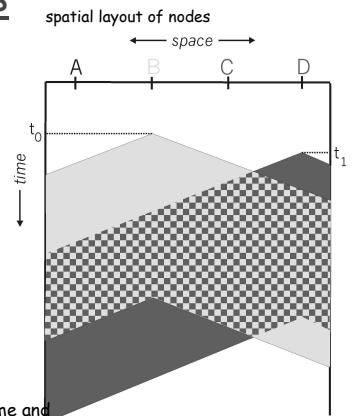
CSMA collisions

collisions *can* still occur:
propagation delay means
two nodes may not hear
each other's transmission

collision:
entire packet transmission
time wasted

note:
role of distance & propagation
delay in determining collision
probability

What if collision? Wait for some time and
try again



5: DataLink Layer 5a-32

Exponential back-off

- Choose the waiting time randomly in an interval
- We would like to wait for a short time if not many colliding nodes
- Need to choose among a longer interval if many potentially colliding nodes (to avoid collisions when retransmitting)

5: DataLink Layer 5a-33

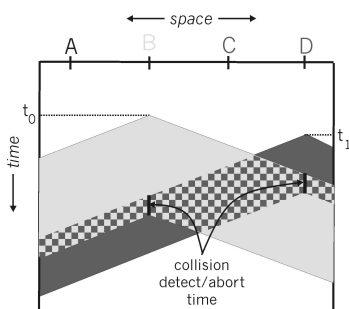
CSMA/CD (Collision Detection)

CSMA/CD: carrier sensing, deferral as in CSMA

- collisions *detected* within short time
- colliding transmissions aborted, reducing channel wastage
- collision detection:
 - easy in wired LANs: measure signal strengths, compare transmitted, received signals
 - difficult in wireless LANs: receiver shut off while transmitting
- human analogy: the polite conversationalist

5: DataLink Layer 5a-34

CSMA/CD collision detection



5: DataLink Layer 5a-35

5: DataLink Layer 5a-36

"Taking Turns" MAC protocols

channel partitioning MAC protocols:

- share channel efficiently and fairly at high load
- inefficient at low load: delay in channel access, 1/N bandwidth allocated even if only 1 active node!

Random access MAC protocols

- efficient at low load: single node can fully utilize channel
- high load: collision overhead

"taking turns" protocols

look for best of both worlds!

5: DataLink Layer 5a-37

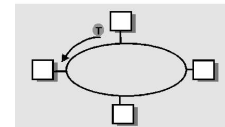
"Taking Turns" MAC protocols

Polling:

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

Token passing:

- control **token** passed from one node to next sequentially.
- token message
- concerns:
 - token overhead
 - latency
 - single point of failure (token)



5: DataLink Layer 5a-38

Summary of MAC protocols

- What do you do with a shared media?
 - Channel Partitioning, by time, frequency or code
 - Time Division, Code Division, Frequency Division
 - Random partitioning (dynamic),
 - ALOHA, S-ALOHA, CSMA, CSMA/CD
 - carrier sensing: easy in some technologies (wire), hard in others (wireless)
 - CSMA/CD used in Ethernet
 - Taking Turns
 - polling from a central site, token passing

5: DataLink Layer 5a-39

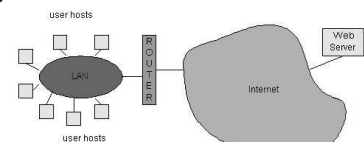
LAN technologies

Data link layer so far:

- services, error detection/correction, multiple access

Next: LAN technologies

- addressing
- Ethernet
- hubs, bridges, switches
- 802.11



5: DataLink Layer 5a-40

LAN Addresses and ARP

32-bit IP address:

- network-layer address
- used to get datagram to destination IP network (recall IP network definition)

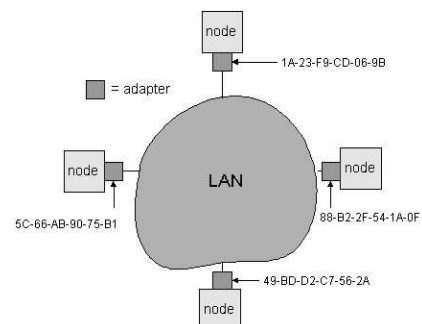
LAN (or MAC or physical or Ethernet) address:

- 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

5: DataLink Layer 5a-41

LAN Addresses and ARP

Each adapter on LAN has unique LAN address



5: DataLink Layer 5a-42

LAN Address (more)

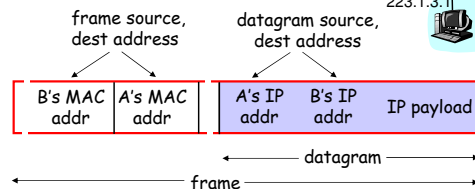
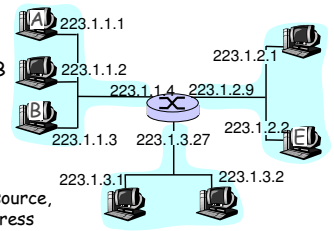
- MAC address allocation administered by IEEE
- manufacturer buys portion of MAC address space (to assure uniqueness)
- Analogy:
 - (a) MAC address: like Social Security Number
 - (b) IP address: like postal address
- MAC flat address => portability
 - can move LAN card from one LAN to another
- IP hierarchical address NOT portable
 - depends on IP network to which node is attached

5: DataLink Layer 5a-43

Recall earlier routing discussion

Starting at A, given IP datagram addressed to B:

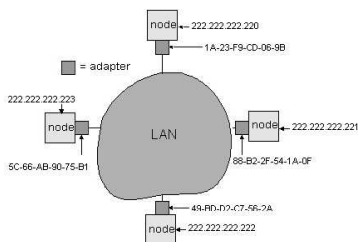
- look up net. address of B, find B on same net. as A
- link layer send datagram to B inside link-layer frame



5: DataLink Layer 5a-44

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)

5: DataLink Layer 5a-45

ARP protocol

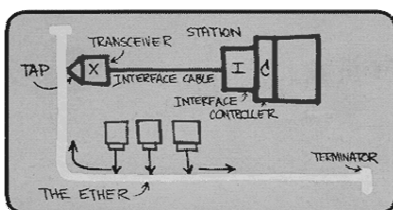
- A wants to send datagram to B, and A knows B's IP address.
- Suppose B's MAC address is not in A's ARP table.
- A broadcasts ARP query packet, containing B's IP address
 - all machines on LAN receive ARP query
- B receives ARP packet, replies to A with its (B's) MAC address
 - frame sent to A's MAC 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
 - USED to save ARP messages: if a receive an ARP message I cache all the informations associated to it
- ARP is "plug-and-play":
 - nodes create their ARP tables without intervention from net administrator

5: DataLink Layer 5a-46

Ethernet

"dominant" LAN technology:

- cheap \$20 for 100Mbs!
- first widely used LAN technology
- Simpler, cheaper than token LANs and ATM
- Kept up with speed race: 10, 100, 1000 Mbps

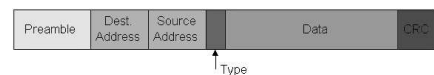


Metcalfe's Ethernet sketch

5: DataLink Layer 5a-47

Ethernet Frame Structure

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



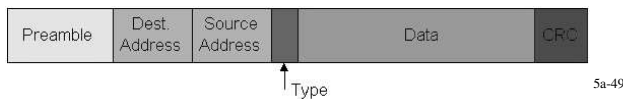
Preamble:

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

5: DataLink Layer 5a-48

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



5: DataLink Layer 5a-50

Unreliable, connectionless service

- Connectionless: No handshaking between sending and receiving adapter.
- Unreliable: receiving adapter doesn't send acks or nacks to sending adapter
 - stream of datagrams passed to network layer can have gaps
 - gaps will be filled if app is using TCP
 - otherwise, app will see the gaps

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

5: DataLink Layer 5a-51

Ethernet CSMA/CD algorithm

1. Adaptor gets datagram from and 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 !
4. 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 K*512 bit times and returns to Step 2

5: DataLink Layer 5a-52

Ethernet's CSMA/CD (more)

Jam Signal: make sure all other transmitters are aware of collision; 48 bits;
Bit time: .1 microsec for 10 Mbps Ethernet ;
for K=1023, wait time is about 50 msec

Exponential Backoff:

- *Goal*: adapt retransmission attempts to estimated current load
 - heavy load: random wait will be longer
- first collision: choose K from {0,1}; delay is K x 512 bit transmission times
- after second collision: choose K from {0,1,2,3}...
- after ten collisions, choose K from {0,1,2,3,4,...,1023}

See/interact with Java applet on AWL Web site: highly recommended !

5: DataLink Layer 5a-53

CSMA/CD efficiency

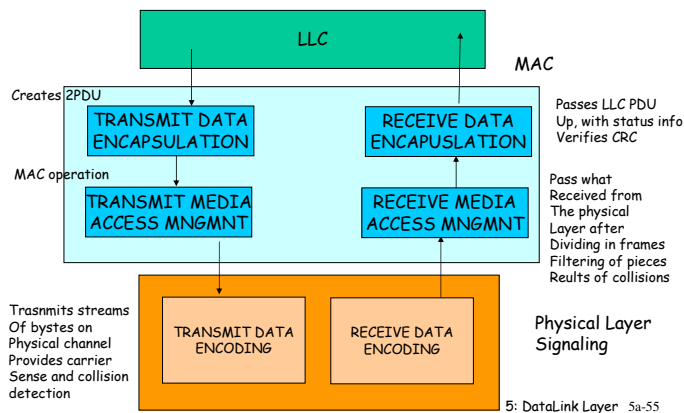
- T_{prop} = max prop between 2 nodes in LAN
- t_{trans} = time to transmit max-size frame

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

- Efficiency goes to 1 as t_{prop} goes to 0
- Goes to 1 as t_{trans} goes to infinity
- Much better than ALOHA, but still decentralized, simple, and cheap

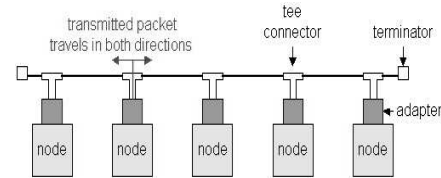
5: DataLink Layer 5a-54

Ethernet MAC sublayer module



Ethernet Technologies: 10Base2

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

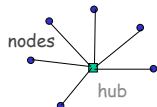


- repeaters used to connect up to multiple segments
- repeater repeats bits it hears on one interface to its other interfaces: physical layer device only!
- has become a legacy technology

5: DataLink Layer 5a-56

10BaseT and 100BaseT

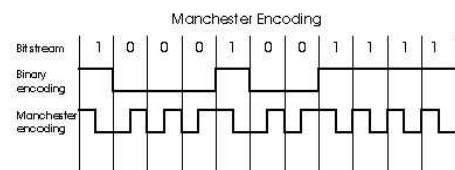
- 10/100 Mbps rate; latter called "fast ethernet"
- T stands for Twisted Pair
- Nodes connect to a hub: "star topology"; 100 m max distance between nodes and hub



- Hubs are essentially physical-layer repeaters:
 - bits coming in one link go out all other links
 - no frame buffering
 - no CSMA/CD at hub: adapters detect collisions
 - provides net management functionality

5: DataLink Layer 5a-57

Manchester encoding



- Used in 10BaseT, 10Base2
- Each bit has a transition
- Allows clocks in sending and receiving nodes to synchronize to each other
 - no need for a centralized, global clock among nodes!
- Hey, this is physical-layer stuff!

5: DataLink Layer 5a-58

Ethernet: some numbers..

- Slo time 512 bit times
- Interframegap 9.6 micros
- Number of times max for retransmitting a frame 16
- Backoff limit ($2^{\text{backoff limit}}$ indicates max length of the backoff interval): 10
- Jam size: 32 bits
- Max frame size: 1518 bytes
- Min frame size 64 bytes (512 bits)
- Address size: 48 bits

5: DataLink Layer 5a-59

5: DataLink Layer 5a-60

Chapter 5 outline

- 5.1 Introduction and services
- 5.2 Error detection and correction
- 5.3 Multiple access protocols
- 5.4 LAN addresses and ARP
- 5.5 Ethernet
- 5.6 Hubs, bridges, and switches
- 5.7 Wireless links and LANs
- 5.8 PPP
- 5.9 ATM
- 5.10 Frame Relay

5: DataLink Layer 5a-61

Interconnecting LAN segments

- Hubs
- Bridges
- Switches
 - Remark: switches are essentially multi-port bridges.
 - What we say about bridges also holds for switches!

5: DataLink Layer 5a-62

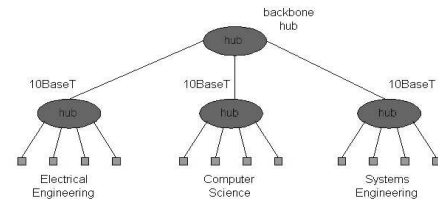
Why LAN interconnection: Some examples

- Different departments, administratively autonomous, may have different LANs, possibly using different technologies
- A LAN segment has limited length AND limited capacity

5: DataLink Layer 5a-63

Interconnecting with hubs

- Devices with relay and some management capabilities
- Backbone hub interconnects LAN segments
- Extends max distance between nodes
- But individual segment collision domains become one large collision domain
 - if a node in CS and a node EE transmit at same time: collision
- Can't interconnect 10BaseT & 100BaseT



5: DataLink Layer 5a-64

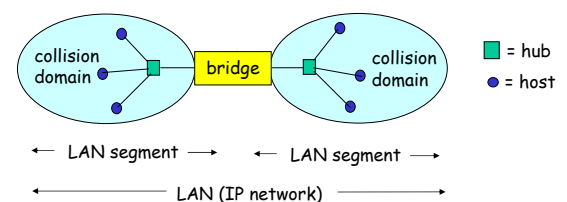
Bridges

- Link layer device
 - stores and forwards Ethernet frames
 - examines frame header and selectively forwards frame based on MAC dest address
 - when frame is to be forwarded on segment, uses CSMA/CD to access segment
- transparent
 - hosts are unaware of presence of bridges
- plug-and-play, self-learning
 - bridges do not need to be configured

5: DataLink Layer 5a-65

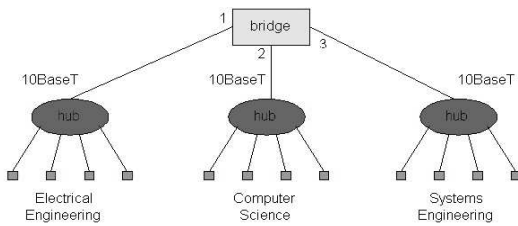
Bridges: traffic isolation

- Bridge installation breaks LAN into LAN segments
- bridges filter packets:
 - same-LAN-segment frames not usually forwarded onto other LAN segments
 - segments become separate collision domains



5: DataLink Layer 5a-66

Forwarding



How do determine to which LAN segment to forward frame?

- Looks like a routing problem...

5: DataLink Layer 5a-67

Self learning

- A bridge has a bridge table
- entry in bridge table:
 - (Node LAN Address, Bridge Interface, Time Stamp)
 - stale entries in table dropped (TTL can be 60 min)
- bridges *learn* (backward learning) which hosts can be reached through which interfaces
 - when frame received, bridge "learns" location of sender: incoming LAN segment
 - records sender/location pair in bridge table

5: DataLink Layer 5a-68

Filtering/Forwarding

When bridge receives a frame:

```

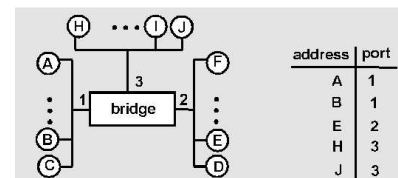
index bridge table using MAC dest address
if entry found for destination
then{
    if dest on segment from which frame arrived
        then drop the frame (filtering)
    else forward the frame on interface indicated
}
else flood
    
```

forward on all but the interface on which the frame arrived

5: DataLink Layer 5a-69

Bridge example

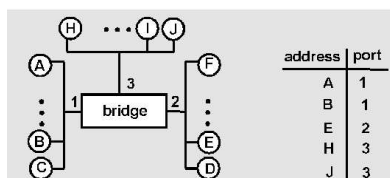
Suppose C sends frame to D and D replies back with frame to C.



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

5: DataLink Layer 5a-70

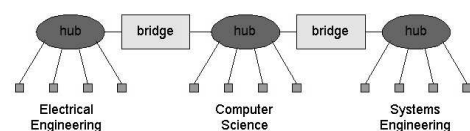
Bridge Learning: example



- D generates frame for C, sends
- bridge receives frame
 - notes in bridge table that D is on interface 2
 - bridge knows C is on interface 1, so *selectively* forwards frame to interface 1

5: DataLink Layer 5a-71

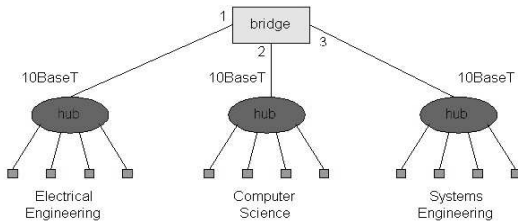
Interconnection without backbone



- Not recommended for two reasons:
 - single point of failure at Computer Science hub
 - all traffic between EE and SE must path over CS segment

5: DataLink Layer 5a-72

Backbone configuration

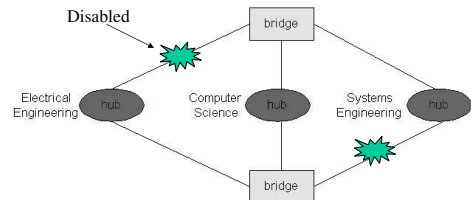


Recommended !

5: DataLink Layer 5a-73

Bridges Spanning Tree

- ❑ for increased reliability, desirable to have redundant, alternative paths from source to dest
- ❑ with multiple paths, cycles result - bridges may multiply and forward frame forever
- ❑ solution: organize bridges in a spanning tree by disabling subset of interfaces



5: DataLink Layer 5a-74

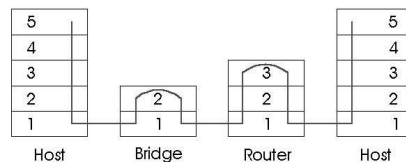
Some bridge features

- ❑ Isolates collision domains resulting in higher total max throughput
- ❑ limitless number of nodes and geographical coverage
- ❑ Can connect different Ethernet types
- ❑ Transparent ("plug-and-play"): no configuration necessary

5: DataLink Layer 5a-75

Bridges vs. Routers

- ❑ both store-and-forward devices
 - routers: network layer devices (examine network layer headers)
 - bridges are link layer devices
- ❑ routers maintain routing tables, implement routing algorithms
- ❑ bridges maintain bridge tables, implement filtering, learning and spanning tree algorithms



5: DataLink Layer 5a-76

Routers vs. Bridges

Bridges + and -

- + Bridge operation is simpler requiring less packet processing
- + Bridge tables are self learning
- All traffic confined to spanning tree, even when alternative bandwidth is available
- Bridges do not offer protection from broadcast storms

5: DataLink Layer 5a-77

Routers vs. Bridges

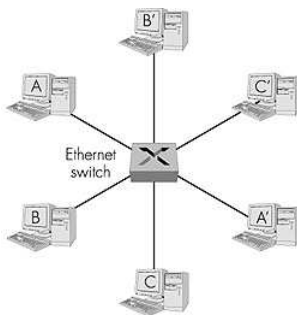
Routers + and -

- + arbitrary topologies can be supported, cycling is limited by TTL counters (and good routing protocols)
- + provide protection against broadcast storms
- require IP address configuration (not plug and play)
- require higher packet processing
- ❑ bridges do well in small (few hundred hosts) while routers used in large networks (thousands of hosts)

5: DataLink Layer 5a-78

Ethernet Switches

- ❑ Essentially a multi-interface bridge
- ❑ layer 2 (frame) forwarding, filtering using LAN addresses
- ❑ large number of interfaces
- ❑ often: individual hosts, star-connected into switch
 - Ethernet, but no collisions!
 - Switching: A-to-A' and B-to-B' simultaneously, no collisions (in the case of hosts directly connected)



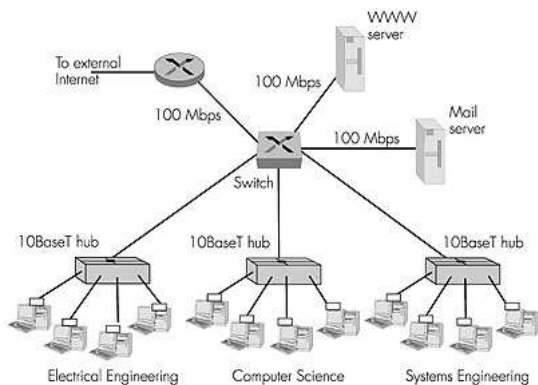
5: DataLink Layer 5a-79

Ethernet Switches

- ❑ cut-through switching: frame forwarded from input to output port without awaiting for assembly of entire frame
 - slight reduction in latency
- ❑ combinations of shared/dedicated, 10/100/1000 Mbps interfaces

5: DataLink Layer 5a-80

Not an atypical LAN (IP network)



5: DataLink Layer 5a-81

Summary comparison

	<u>hubs</u>	<u>bridges</u>	<u>routers</u>	<u>switches</u>
traffic isolation	no	yes	yes	yes
plug & play	yes	yes	no	yes
optimal routing	no	no	yes	no
cut through	yes	no	no	yes

5: DataLink Layer 5a-82

Chapter 5 outline

- ❑ 5.1 Introduction and services
- ❑ 5.2 Error detection and correction
- ❑ 5.3 Multiple access protocols
- ❑ 5.4 LAN addresses and ARP
- ❑ 5.5 Ethernet
- ❑ 5.6 Hubs, bridges, and switches
- ❑ 5.7 Wireless links and LANs
- ❑ 5.8 PPP
- ❑ 5.9 ATM
- ❑ 5.10 Frame Relay

5: DataLink Layer 5a-83

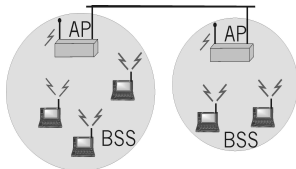
IEEE 802.11 Wireless LAN

- ❑ 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 base-station and ad-hoc network versions

5: DataLink Layer 5a-84

Base station approach

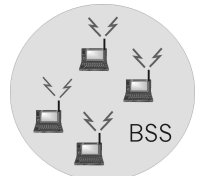
- ❑ Wireless host communicates with a base station
 - base station = access point (AP)
- ❑ Basic Service Set (BSS) (a.k.a. "cell") contains:
 - wireless hosts
 - access point (AP): base station
- ❑ BSS's combined to form distribution system (DS)



5: DataLink Layer 5a-85

Ad Hoc Network approach

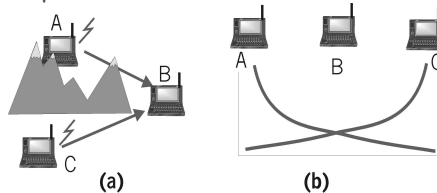
- ❑ No AP (i.e., base station)
- ❑ wireless hosts communicate with each other
 - to get packet from wireless host A to B may need to route through wireless hosts X,Y,Z
- ❑ Applications:
 - "laptop" meeting in conference room, car
 - interconnection of "personal" devices
 - battlefield
- ❑ IETF MANET (Mobile Ad hoc Networks) working group



5: DataLink Layer 5a-86

IEEE 802.11: multiple access

- ❑ Collision if 2 or more nodes transmit at same time
- ❑ CSMA makes sense:
 - get all the bandwidth if you're the only one transmitting
 - shouldn't cause a collision if you sense another transmission
- ❑ Collision detection has problems (send and receive simultaneously not allowed) and doesn't work: hidden terminal problem



5: DataLink Layer 5a-87

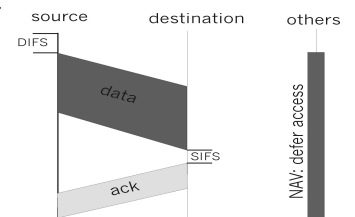
IEEE 802.11 MAC Protocol: CSMA/CA

802.11 CSMA: sender

- if sense channel idle for **DIFS** (Distributed Interframe Space) sec. then transmit entire frame (no collision detection)
- if sense channel busy then binary backoff (waits until channel sensed idled + random interval selected according to binary backoff rules)

802.11 CSMA receiver

- if received OK return ACK after **SIFS** (Short InterFrame Spacing) (ACK is needed due to hidden terminal problem)



5: DataLink Layer 5a-88

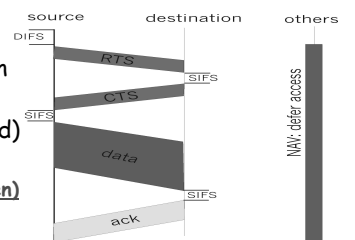
Collision avoidance mechanisms

- ❑ Problem:
 - two nodes, hidden from each other, transmit complete frames to base station
 - wasted bandwidth for long duration !
- ❑ Solution:
 - small reservation packets
 - nodes track reservation interval with internal "network allocation vector" (NAV)

5: DataLink Layer 5a-89

Collision Avoidance: RTS-CTS exchange

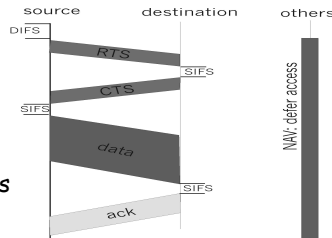
- ❑ sender transmits short RTS (request to send) packet: indicates duration of transmission
- ❑ receiver replies with short CTS (clear to send) packet
 - notifying (possibly hidden) nodes
- ❑ hidden nodes will not transmit for specified duration: NAV



5: DataLink Layer 5a-90

Collision Avoidance: RTS-CTS exchange

- RTS and CTS short:
 - collisions less likely, of shorter duration
 - end result similar to collision detection
- IEEE 802.11 allows:
 - CSMA
 - CSMA/CA: reservations
 - polling from AP



5: DataLink Layer 5a-91

A word about Bluetooth

- Low-power, small radius, wireless networking technology
 - 10-100 meters
- omnidirectional
 - not line-of-sight infrared
- Interconnects gadgets
- 2.4-2.5 GHz unlicensed radio band
- up to 721 kbps
- Interference from wireless LANs, digital cordless phones, microwave ovens:
 - frequency hopping helps
- MAC protocol supports:
 - error correction
 - ARQ
- Each node has a 12-bit address

5: DataLink Layer 5a-92

Chapter 5 outline

- 5.1 Introduction and services
- 5.2 Error detection and correction
- 5.3 Multiple access protocols
- 5.4 LAN addresses and ARP
- 5.5 Ethernet
- 5.6 Hubs, bridges, and switches
- 5.7 Wireless links and LANs
- 5.8 PPP
- 5.9 ATM
- 5.10 Frame Relay

5: DataLink Layer 5a-93

5: DataLink Layer 5a-94

Point to Point Data Link Control

- one sender, one receiver, one link: easier than broadcast link:
 - no Media Access Control
 - no need for explicit MAC addressing
 - e.g., dialup link, ISDN line
- popular point-to-point DLC protocols:
 - PPP (point-to-point protocol)
 - HDLC: High level data link control (Data link used to be considered "high layer" in protocol stack!)

5: DataLink Layer 5a-95

PPP Design Requirements [RFC 1557]

- packet framing: encapsulation of network-layer datagram in data link frame
 - carry network layer data of any network layer protocol (not just IP) *at same time*
 - ability to demultiplex upwards
- bit transparency: must carry any bit pattern in the data field
- error detection (no correction)
- connection liveness: detect, signal link failure to network layer
- network layer address negotiation: endpoint can learn/configure each other's network address

5: DataLink Layer 5a-96

PPP non-requirements

- ❑ no error correction/recovery
- ❑ no flow control
- ❑ out of order delivery OK
- ❑ no need to support multipoint links (e.g., polling)

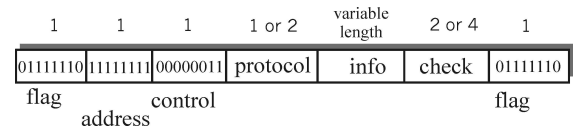
Error recovery, flow control, data re-ordering
all relegated to higher layers!

5: DataLink Layer 5a-97

PPP Data Frame

- ❑ Flag: delimiter (framing)
- ❑ Address: does nothing (only one option)
- ❑ Control: does nothing; in the future possible multiple control fields
- ❑ Protocol: upper layer protocol to which frame delivered (eg, PPP-LCP, IP, IPCP, etc)

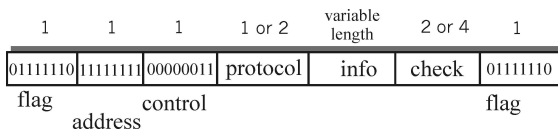
Can avoid sending



5: DataLink Layer 5a-98

PPP Data Frame

- ❑ info: upper layer data being carried
- ❑ check: cyclic redundancy check for error detection



5: DataLink Layer 5a-99

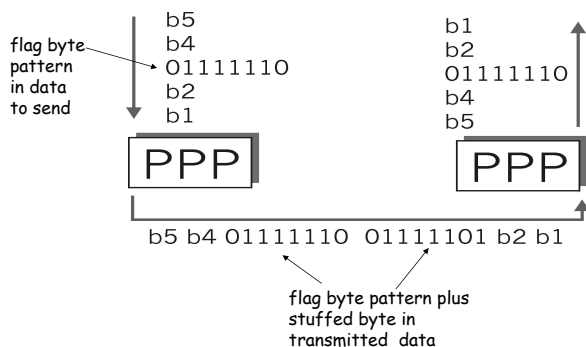
Byte Stuffing

- ❑ "data transparency" requirement: data field must be allowed to include flag pattern <01111110> (used as flag for sake of framing)
 - Q: is received <01111110> data or flag?

- ❑ Sender: adds ("stuffs") extra < 01111110> byte after each < 01111110> data byte
- ❑ Receiver:
 - two 01111110 bytes in a row: discard first byte, continue data reception
 - single 01111110: flag byte

5: DataLink Layer 5a-100

Byte Stuffing

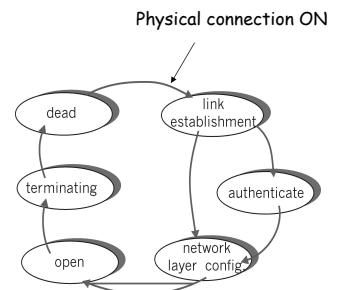


5: DataLink Layer 5a-101

PPP Data Control Protocol

Before exchanging network-layer data, data link peers must

- ❑ configure PPP link (max. frame length, authentication)
- ❑ learn/configure network layer information
 - for IP: carry IP Control Protocol (IPCP) msgs (protocol field: 8021) to configure/learn IP address



5: DataLink Layer 5a-102

Chapter 5 outline

- 5.1 Introduction and services
- 5.2 Error detection and correction
- 5.3 Multiple access protocols
- 5.4 LAN addresses and ARP
- 5.5 Ethernet
- 5.6 Hubs, bridges, and switches
- 5.7 Wireless links and LANs
- 5.8 PPP
- 5.9 ATM
- 5.10 Frame Relay

5: DataLink Layer 5a-103

5: DataLink Layer 5a-104

Frame Relay

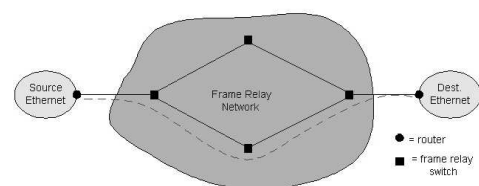
Like ATM:

- wide area network technologies
- Virtual-circuit oriented
- origins in telephony world
- can be used to carry IP datagrams
 - can thus be viewed as link layers by IP protocol

5: DataLink Layer 5a-105

Frame Relay

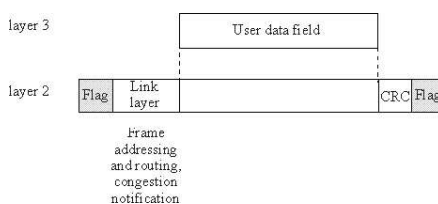
- Designed in late '80s, widely deployed in the '90s
- Frame relay service:
 - no error control
 - end-to-end congestion control



5: DataLink Layer 5a-106

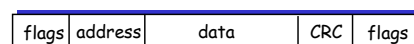
Frame Relay (more)

- Designed to **interconnect** corporate customer LANs
 - typically permanent VC's: "**pipe**" carrying aggregate traffic between two routers
 - switched VC's: as in ATM
- corporate customer **leases** FR service from public Frame Relay network (eg, Sprint, ATT)



5: DataLink Layer 5a-107

Frame Relay (more)



- Flag bits, 01111110, delimit frame
- address:
 - 10 bit VC ID field
 - 3 congestion control bits
 - FECN: forward explicit congestion notification (frame experienced congestion on path)
 - BECN: congestion on reverse path
 - DE: discard eligibility

5: DataLink Layer 5a-108

Frame Relay -VC Rate Control

- Committed Information Rate (CIR)
 - defined, "guaranteed" for each VC
 - negotiated at VC set up time
 - customer pays based on CIR
- DE bit: Discard Eligibility bit
 - Edge FR switch measures traffic rate for each VC; marks DE bit
 - DE = 0: high priority, rate compliant frame; deliver at "all costs"
 - DE = 1: low priority, eligible for congestion discard

5: DataLink Layer 5a-109

Frame Relay - CIR & Frame Marking

- **Access Rate:** rate **R** of the access link between **source router** (customer) and **edge FR switch** (provider); $64\text{Kbps} < R < 1,544\text{Kbps}$
- Typically, **many VCs** (one per destination router) multiplexed on the same access trunk; each VC has own **CIR**
- Edge FR switch **measures** traffic rate for each VC; it **marks** (ie DE = 1) frames which **exceed CIR** (these may be later dropped)
- Internet's more recent **differentiated service** uses similar ideas

5: DataLink Layer 5a-110

Chapter 5 outline

- 5.1 Introduction and services
- 5.2 Error detection and correction
- 5.3 Multiple access protocols
- 5.4 LAN addresses and ARP
- 5.5 Ethernet
- 5.6 Hubs, bridges, and switches
- 5.7 Wireless links and LANs
- 5.8 PPP
- 5.9 ATM
- 5.10 Frame Relay

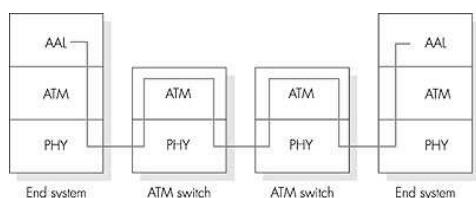
5: DataLink Layer 5a-111

Asynchronous Transfer Mode: ATM

- **1990's/00 standard for high-speed** (155Mbps to 622 Mbps and higher) *Broadband Integrated Service Digital Network* architecture
- **Goal:** *integrated, end-end transport of carry voice, video, data*
 - meeting timing/QoS requirements of voice, video (versus Internet best-effort model)
 - "next generation" telephony: technical roots in telephone world
 - packet-switching (fixed length packets, called "cells") using virtual circuits

5: DataLink Layer 5a-112

ATM architecture



- adaptation layer: only at edge of ATM network
 - data segmentation/reassembly
 - roughly analogous to Internet transport layer
- ATM layer: "network" layer
 - cell switching, routing
- physical layer

5: DataLink Layer 5a-113

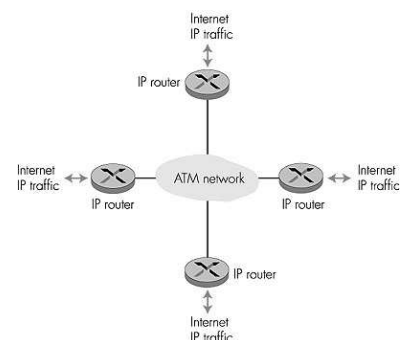
ATM: network or link layer?

Vision: end-to-end transport: "ATM from desktop to desktop"

- ATM is a network technology

Reality: used to connect IP backbone routers

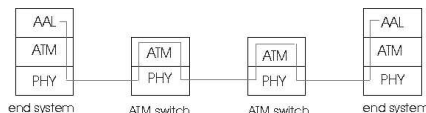
- "IP over ATM"
- ATM as switched link layer, connecting IP routers



5: DataLink Layer 5a-114

ATM Adaptation Layer (AAL)

- **ATM Adaptation Layer (AAL):** "adapts" upper layers (IP or native ATM applications) to ATM layer below
- AAL present **only in end systems**, not in switches
- AAL layer segment (header/trailer fields, data) fragmented across multiple ATM cells
 - analogy: TCP segment in many IP packets

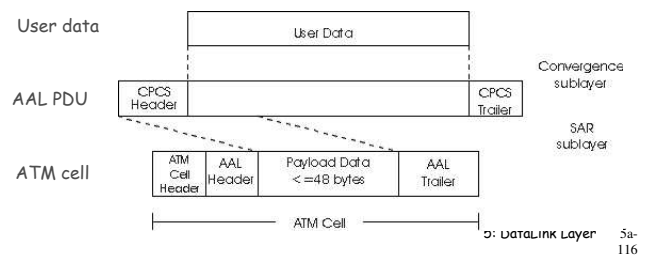


5: DataLink Layer 5a-115

ATM Adaptation Layer (AAL) [more]

Different versions of AAL layers, depending on ATM service class:

- AAL1: for CBR (Constant Bit Rate) services, e.g. circuit emulation
- AAL2: for VBR (Variable Bit Rate) services, e.g., MPEG video
- AAL5: for data (eg, IP datagrams)



5: DataLink Layer 5a-116

AAL5 - Simple And Efficient AL (SEAL)

- **AAL5: low overhead AAL** used to carry IP datagrams
 - 4 byte cyclic redundancy check
 - PAD ensures payload multiple of 48bytes
 - large AAL5 data unit to be fragmented into 48-byte ATM cells

CPCS-PDU payload	PAD	Length	CRC
0-65535	0-47	2	4

5: DataLink Layer 5a-117

ATM Layer

Service: transport cells across ATM network

- analogous to IP network layer
- very different services than IP network layer

Network Architecture	Service Model	Guarantees ?				Congestion feedback
		Bandwidth	Loss	Order	Timing	
Internet	best effort	none	no	no	no	no (inferred via loss)
ATM	CBR	constant rate	yes	yes	yes	no congestion
ATM	VBR	guaranteed rate	yes	yes	yes	no congestion
ATM	ABR	guaranteed minimum	no	yes	no	yes
ATM	UBR	none	no	yes	no	no

5: DataLink Layer 5a-118

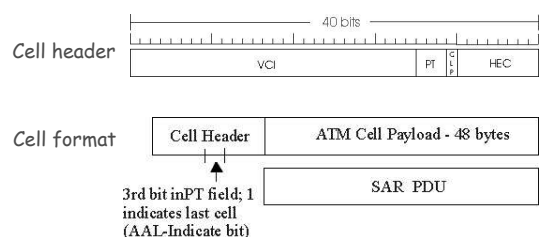
ATM Layer: Virtual Circuits

- **VC transport:** cells carried on VC from source to dest
 - call setup, teardown for each call *before* data can flow
 - each packet carries VC identifier (not destination ID)
 - *every* switch on source-dest path maintain "state" for each passing connection
 - link, switch resources (bandwidth, buffers) may be *allocated* to VC: to get circuit-like perf.
- **Permanent VCs (PVCs)**
 - long lasting connections
 - typically: "permanent" route between to IP routers
- **Switched VCs (SVC):**
 - dynamically set up on per-call basis

5: DataLink Layer 5a-119

ATM Layer: ATM cell

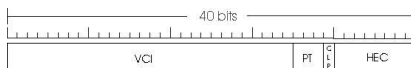
- 5-byte ATM cell header
- 48-byte payload
 - Why?: small payload -> short cell-creation delay for digitized voice
 - halfway between 32 and 64 (compromise!)



5: DataLink Layer 5a-120

ATM cell header

- **VCI:** virtual channel ID
 - will *change* from link to link thru net
- **PT:** Payload type (e.g. RM cell versus data cell)
- **CLP:** Cell Loss Priority bit
 - CLP = 1 implies low priority cell, can be discarded if congestion
- **HEC:** Header Error Checksum
 - cyclic redundancy check



5: DataLink Layer 5a-121

ATM Physical Layer

Two pieces (sublayers) of physical layer:

- Transmission Convergence Sublayer (TCS): adapts ATM layer above to PMD sublayer below
- Physical Medium Dependent: depends on physical medium being used

TCS Functions:

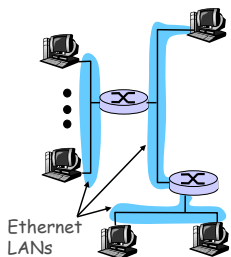
- Header **checksum** generation: 8 bits CRC
- **Cell delineation**
- With "unstructured" PMD sublayer, transmission of **idle cells** when no data cells to send

5: DataLink Layer 5a-122

IP-Over-ATM

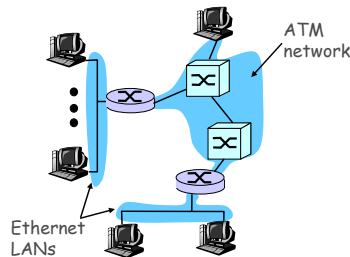
Classic IP only

- 3 "networks" (e.g., LAN segments)
- MAC (802.3) and IP addresses



IP over ATM

- replace "network" (e.g., LAN segment) with ATM network
- ATM addresses, IP addresses

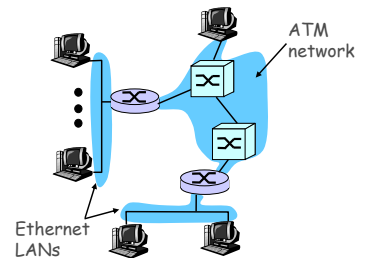


5: DataLink Layer 5a-123

IP-Over-ATM

Issues:

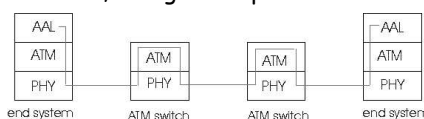
- IP datagrams into ATM AAL5 PDUs
- from IP addresses to ATM addresses
 - just like IP addresses to 802.3 MAC addresses!



5: DataLink Layer 5a-124

Datagram Journey in IP-over-ATM Network

- at Source Host:
 - IP layer maps between IP, ATM dest address (using ARP)
 - passes datagram to AAL5
 - AAL5 encapsulates data, segments cells, passes to ATM layer
- ATM network: moves cell along VC to destination
- at Destination Host:
 - AAL5 reassembles cells into original datagram
 - if CRC OK, datagram is passed to IP



5: DataLink Layer 5a-125

Chapter 5: Summary

- principles behind data link layer services:
 - error detection, correction
 - sharing a broadcast channel: multiple access
 - link layer addressing, ARP
- link layer technologies: Ethernet, hubs, bridges, switches, IEEE 802.11 LANs, PPP, ATM, Frame Relay
- journey down the protocol stack now **OVER!**

5: DataLink Layer 5a-126