Computer Networks

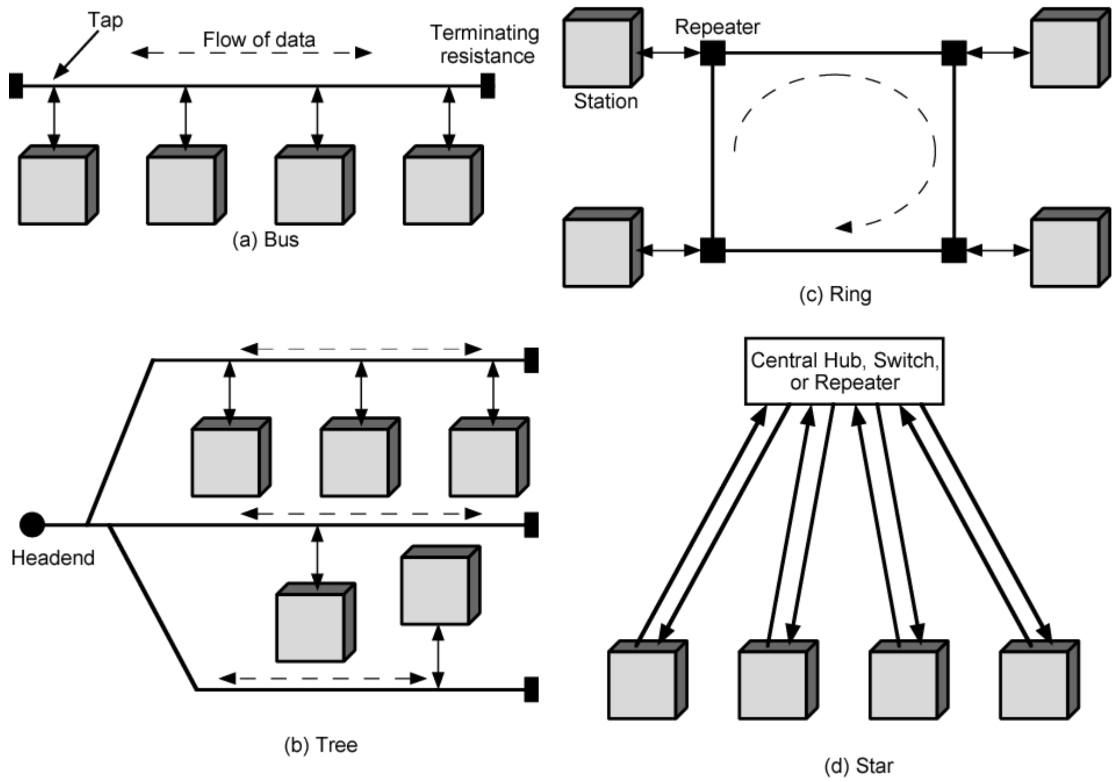
Wenzhong Li
Nanjing University
Fall 2014

Chapter 2. Direct Link Networks

- Link Service and Framing
- Error Detection and Reliable Transmission
- HDLC, PPP, and SONET
- Token Ring
- Ethernet
- Bridges and Layer-2 switch
- Wireless Networks
- Network Performance

Ethernet

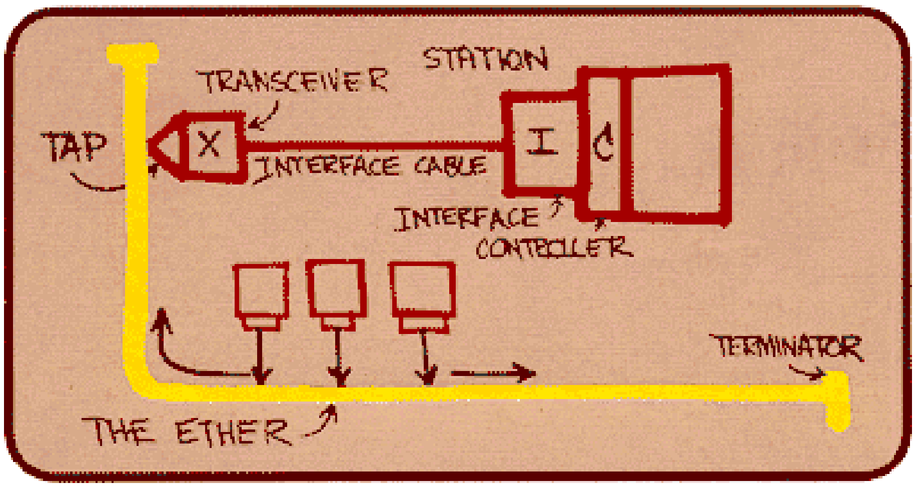
Different Topologies of LAN



Ethernet

"Dominant" wired LAN technology:

- First widely used LAN technology
- Simpler, cheaper than token LANs and ATM
- Keep up with speed race: 10 Mbps ~ 10 Gbps



Metcalfe's Ethernet sketch

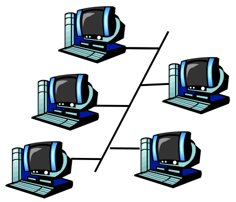
Ethernet

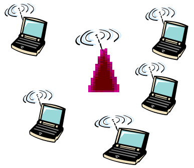
- Multiple access protocols
- CSMA/CD
- IEEE 802.3
- High-Speed Ethernet

Multiple access protocols

Multiple Access Links

- two types of “links”:
  - point-to-point
    - PPP for dial-up access
    - point-to-point link between Ethernet switch, host
  - *broadcast (shared wire or medium)*
    - old-fashioned Ethernet
    - 802.11 wireless LAN





- Properties of Multiple Access Links
  - Single shared broadcast channel
  - Two or more simultaneous transmissions by nodes: interference
  - Collision: node receives two or more signals at the same time

Multiple Access Protocols

- An ideal multiple access protocol

given a broadcast channel of rate R bps, we want

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

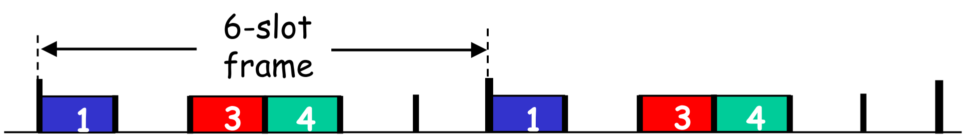
Handling Multiple Access

- Multiple access control (MAC)
  - Determine when node can transmit on shared media
- Three classes:
- Channel Partitioning
  - Divide channel into smaller “pieces” (time slots, frequency, code)
  - Allocate piece to node for exclusive use
- Taking turns
  - Nodes take turns to transmit, nodes with more to send can take longer turns
- Random Access
  - Channel not divided, allow collisions
  - Coordinate or recover from collisions

Channel Partitioning with TDMA

- TDMA: time division multiple access
- Access to channel in "slots and rounds"
- Each station gets fixed length slot (packet trans time) in each round
- Unused slots go idle

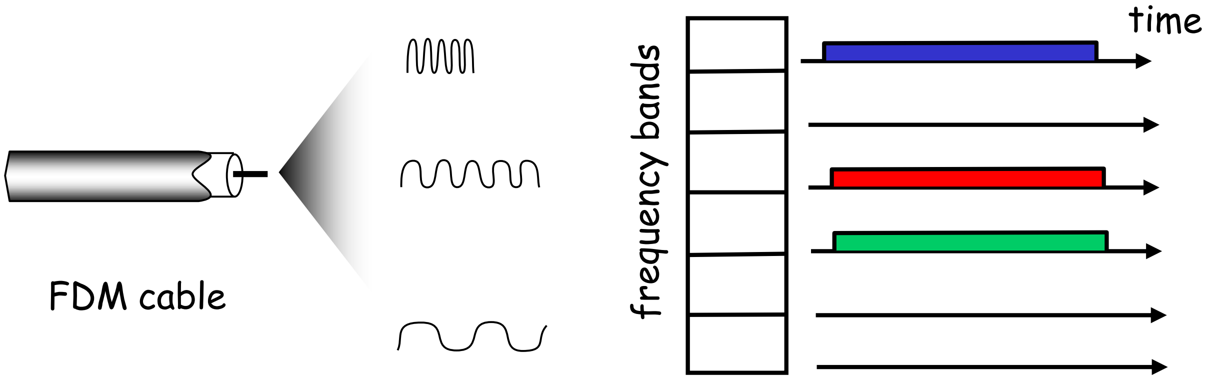
- Example: a 6-station LAN, 1,3,4 have packets, slots 2,5,6 idle



Channel Partitioning with 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: a 6-station LAN, 1,3,4 have packets, frequency bands 2,5,6 idle



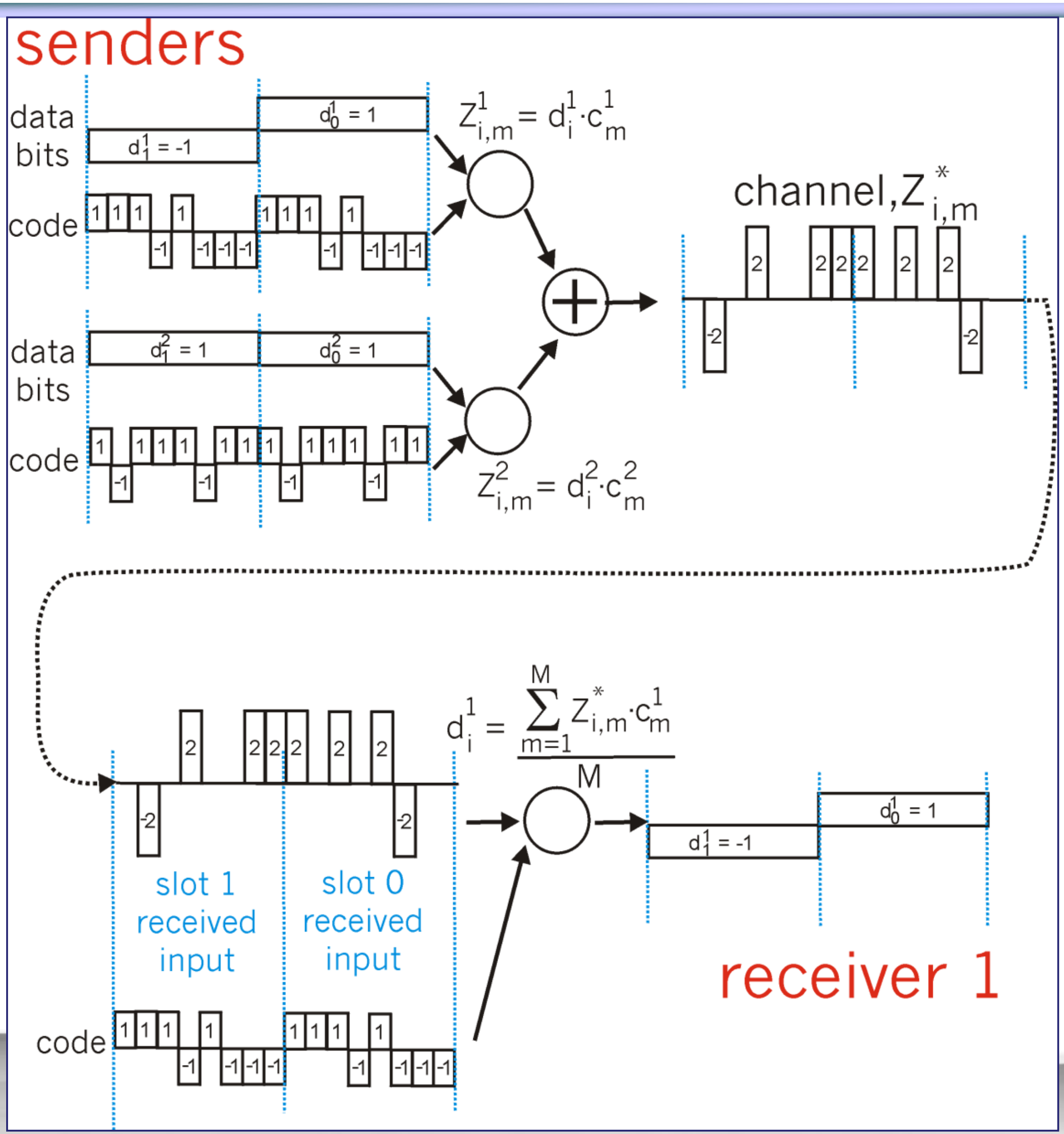
Channel Partitioning with CDMA

- CDMA: Code Division Multiple Access
  - Used in wireless broadcast channels (cellular, satellite, etc)
- All nodes share same frequency, but each node has own "chipping" sequence (i.e., code set) to encode data

- Encoded signal = (original data) $\times$ (chipping sequence)
- Decoding = inner-product of encoded signal and chipping sequence

- If codes are "orthogonal"
  - Multiple nodes can transmit simultaneously with minimal interference

CDMA: Example

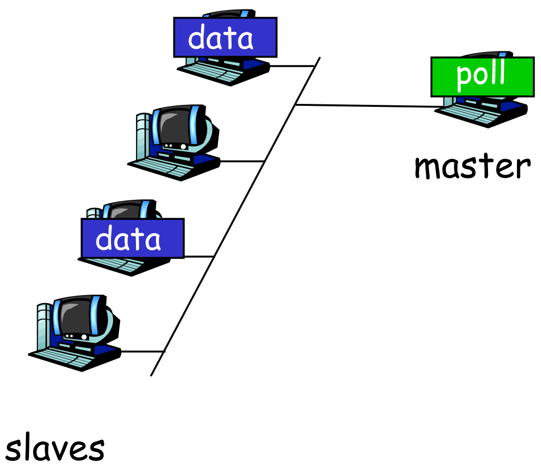


Taking Turns

Polling:
- Master node “invites” slave nodes to transmit in turn
- Typically used with “dumb” slave devices

Concerns:
- Polling overhead
- Latency
- Single point of failure (master)

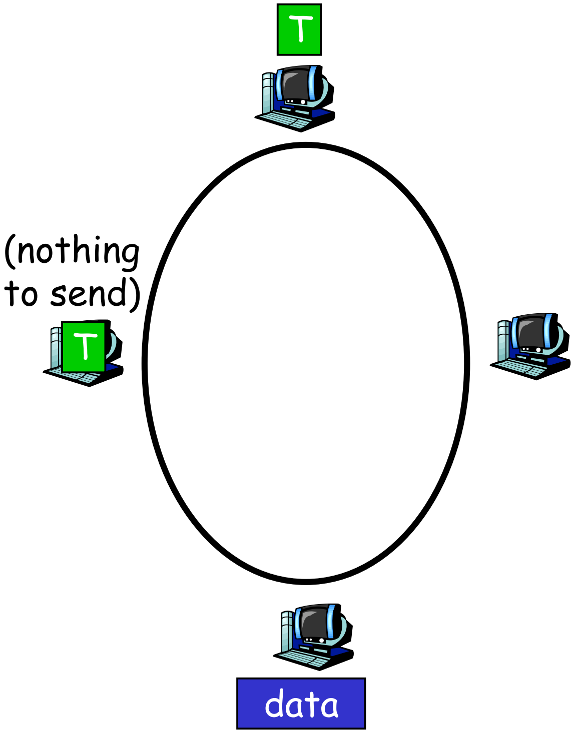
- e.g. Bluetooth



Taking Turns

Token passing:

- Control token passed from one node to next sequentially
- Token to message



Concerns:

- Token overhead
- Latency
- Single point of failure (token)

- IBM Token Ring, FDDI

Random Access Protocols

- When node has packet to send
  - Transmit at full channel data rate R
  - No priori coordination among nodes
- Two or more transmitting nodes → collision

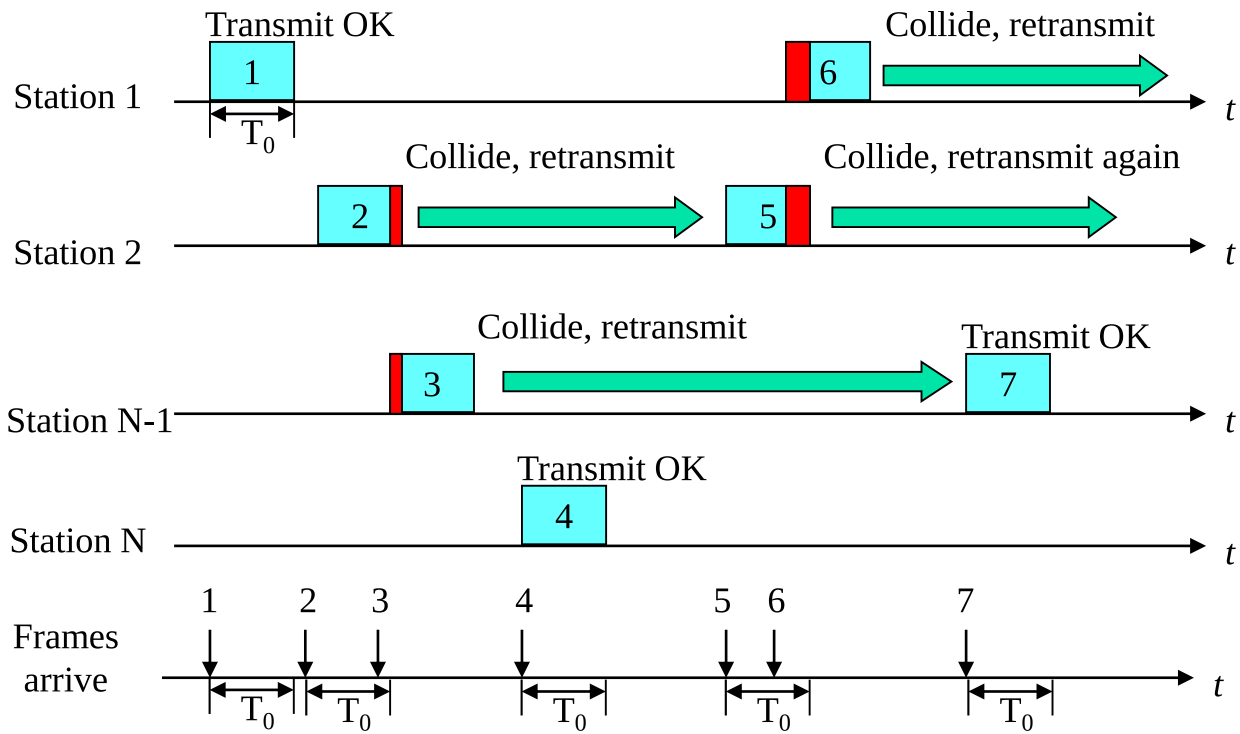
- Random access MAC protocol specifies:
  - How to detect / avoid collisions
  - How to recover from collisions (e.g. via delayed retransmissions)

- Examples of random access MAC protocols:
  - ALOHA, Slotted ALOHA
  - CSMA, CSMA/CD, CSMA/CA

ALOHA

- Additive Link On-line HAwaii system
  - Developed for Packet Radio networks by Hawaii University
- Sender
  - When station has frame, it sends
  - If *ACK*, fine.
    - If not, retransmit with probability p, and wait with probability (1-p)
  - If no *ACK* after repeated transmissions, give up
- Receiver
  - Use frame check sequence (as in *HDLC*)
  - If frame OK and address matches receiver, send *ACK*
- Frame may be damaged by noise or collision
  - Another station transmitting at the same time
  - Any overlap of frames causes collision

Illustration of ALOHA



ALOHA Efficiency

- Collision occurs when frames overlap
- Successful transmission probability for node i, assuming N node

P(success) = P(node i transmits) $\cdot$ P(no other node transmits in $[t_0-1, t_0]$

$\cdot$ P(no other node transmits in $[t_0, t_0+1]$

$= p \cdot (1-p)^{N-1} \cdot (1-p)^{N-1} = p \cdot (1-p)^{2(N-1)}$

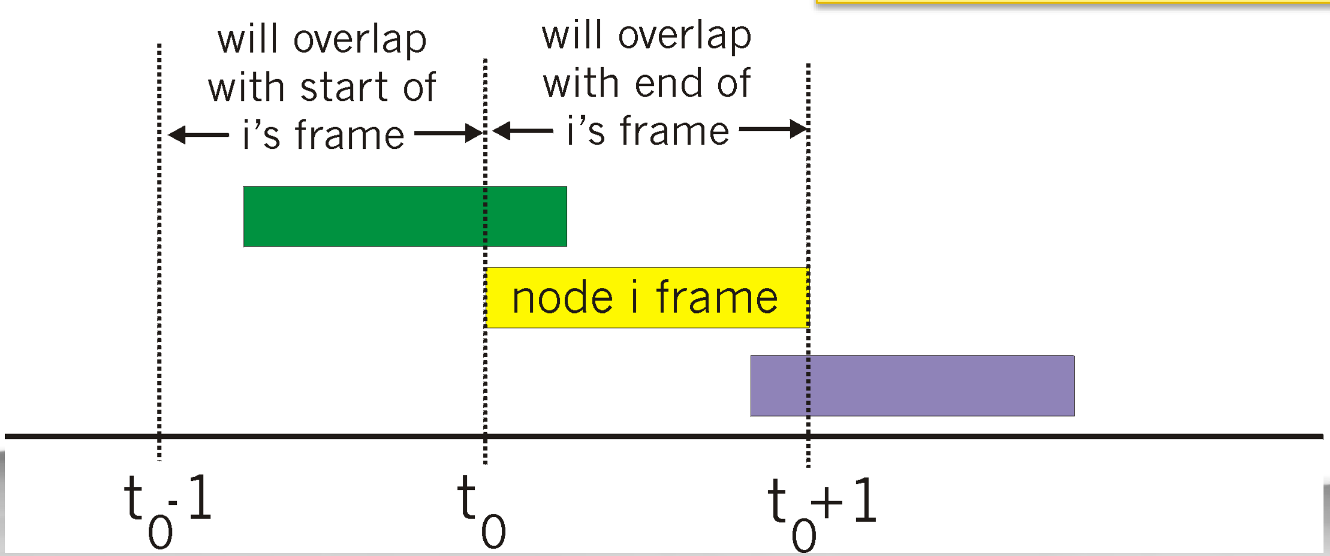
P(any success)=N·P(success) $= \mathbf{N}p \cdot (1-p)^{2(N-1)}$

… choosing optimum p and then letting N -> ∞

P(any success) $\approx 1/(2e) = 0.18$

Max utilization is 18%

efficiency: long-run fraction of successful slots (many nodes, all with many frames to send)



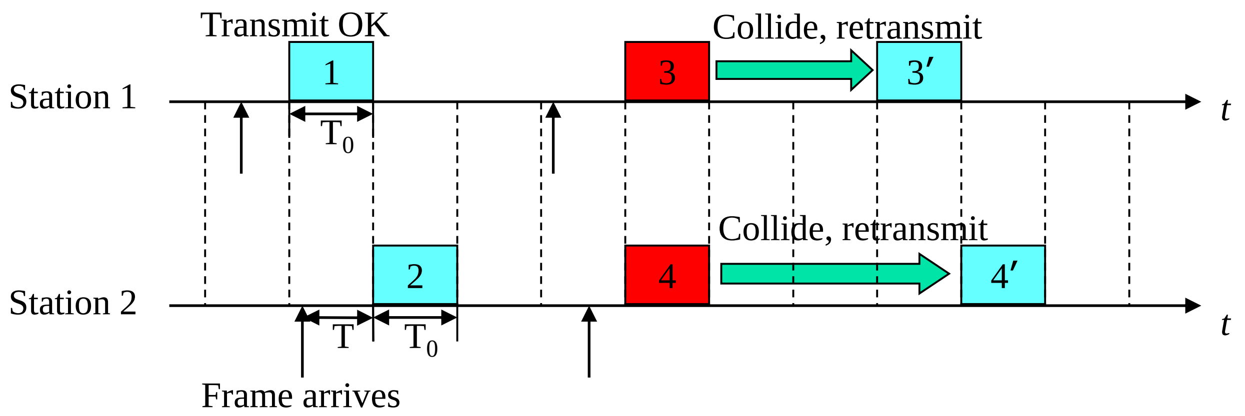
Slotted ALOHA

- All frames have same size
- Time in uniform slots equal to frame transmission time (T_0)
- Nodes are synchronized (need central clock or other sync mechanism)
- Transmission begins at slot boundary
- Frames either miss or overlap totally

operation:

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

Illustration of Slotted-ALOHA



- Node retransmits frame in each subsequent slot with prob. p until success

Slotted ALOHA: Efficiency

- *Suppose: N* nodes with many frames to send, each transmits in slot with probability *p*
- Prob (given node has success in a slot) = $p(1-p)^{N-1}$
- Prob (*any* node has a success) = $Np(1-p)^{N-1}$
- Max efficiency: find p* that maximizes $Np(1-p)^{N-1}$
- When n -> ∞

  max efficiency ≈ 1/e = 0.37
- Max utilization is 37%

CSMA

CSMA

- CSMA (carrier sense multiple access,载波侦听多路访问)
- Suppose
  - Propagation time is much less than transmission time
  - All stations know that a transmission has started almost immediately
- Method
  - listen before transmit:
  - Sender listen for clear medium (carrier sense), if medium idle, transmit
  - If channel sensed busy, wait reasonable time (round trip plus *ACK* contention)
  - If no *ACK* then retransmit
- Longer frame and shorter propagation gives better utilization

Nonpersistent CSMA (非持续CSMA)

- Station wishing to transmit listens
  1. If medium is idle, transmit; otherwise, go to 2
  2. If busy, wait amount of random time (delay) and repeat 1
- Random delays reduces probability of collisions
  - Two stations waiting will take different time to begin transmission
- Capacity is wasted, since medium will remain idle following end of transmission
  - Even if one or more stations waiting
- Nonpersistent stations are deferential

1-persistent CSMA

- To avoid idle channel time, 1-persistent protocol used
- Station wishing to transmit listens
  1. If medium idle, transmit; otherwise, go to step 2
  2. If medium busy, listen until idle; then transmit immediately
- 1-persistent stations selfish
  - If two or more stations waiting, collision guaranteed

p-Persistent CSMA

- Try making compromise
  - Attempts to reduce collisions like Nonpersistent
  - And reduce idle time like 1-persistent
- Rules
  1. If medium idle, transmit with probability p, and delay one time unit with probability (1–p)
     - Typically, time unit = maximum propagation delay
  2. If medium busy, listen until idle and repeat step 1
  3. If transmission is delayed one time unit, repeat step 1
- What is an effective value of p?

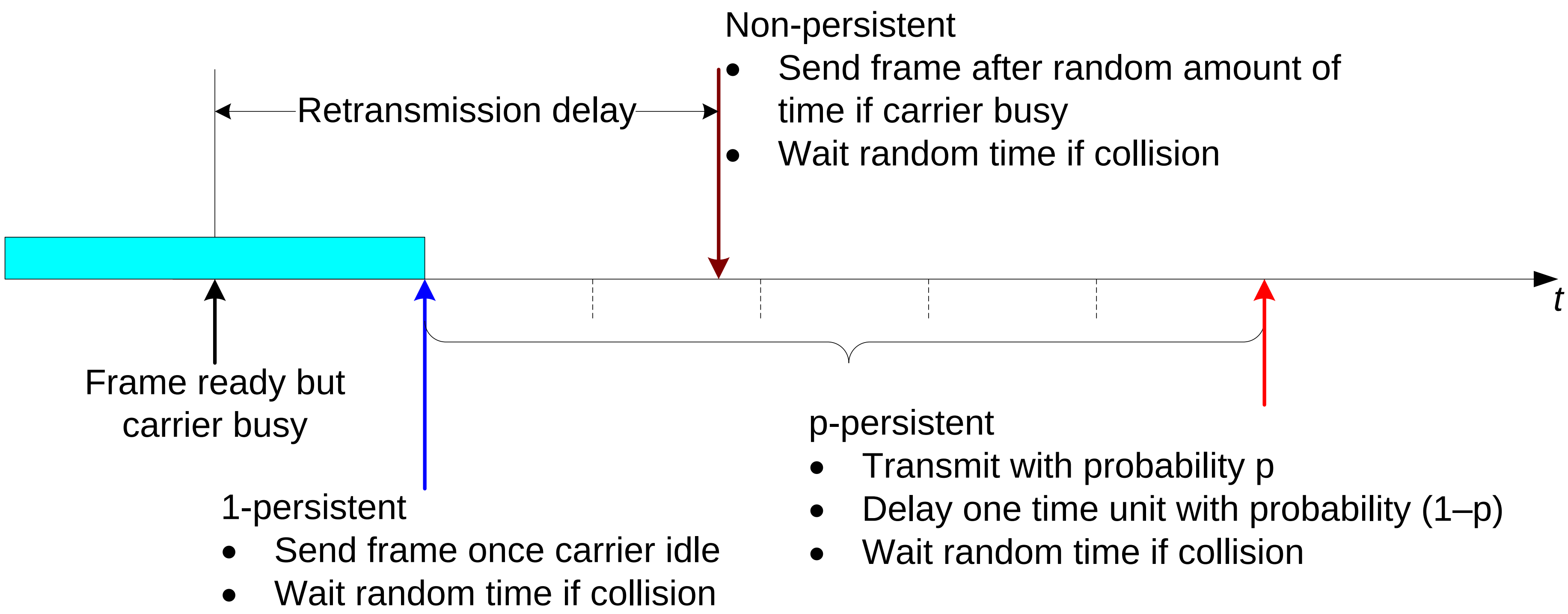
Value of p

- Objective: avoid instability under heavy load
- Suppose: N stations waiting to send
  - The best value of p in theory is 1/N

- If heavy load expected, p small
- However, as p made smaller, stations wait longer

- In general, this gives very long delays

Different Types of CSMA



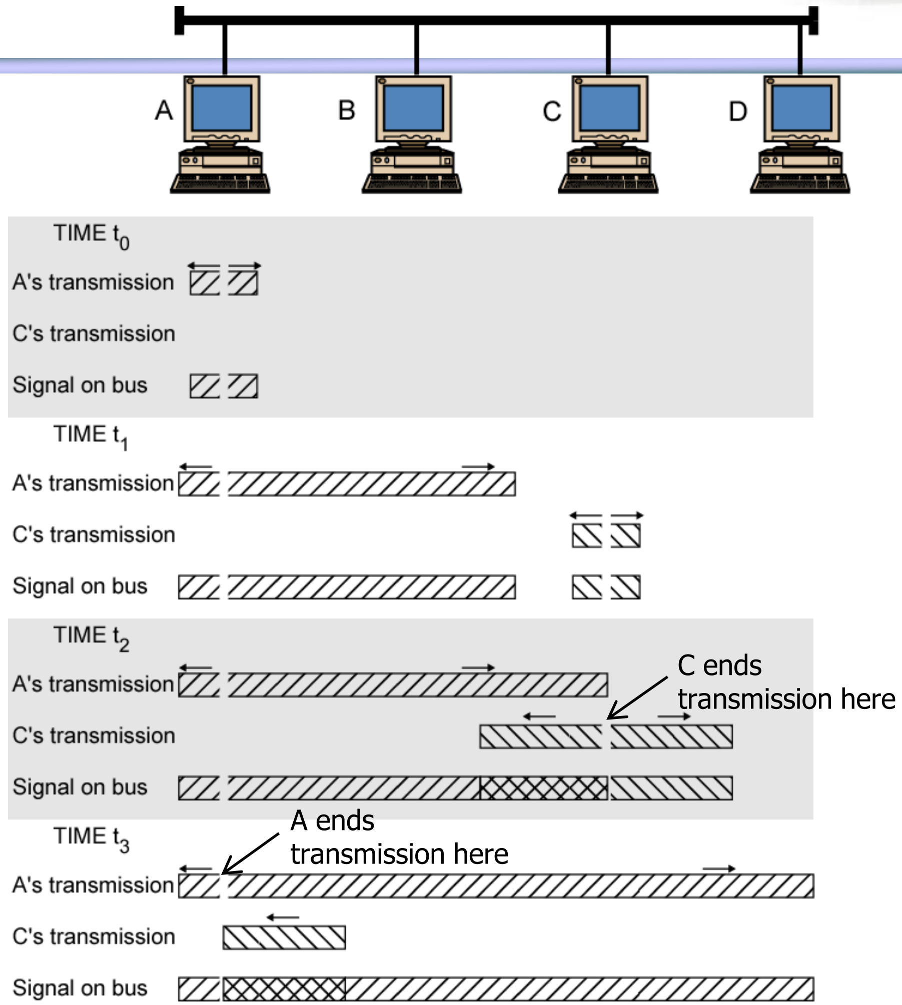
CSMA/CD

- With CSMA, collision occupies medium for duration of transmission
  - Colliding transmissions aborted once detected
- Stations listen whilst transmitting

1. If medium idle, transmit; otherwise, step 2
2. If busy, listen for idle, then transmit immediately
3. If collision detected, send jam signal then abort
4. After jam, wait random time then start from step 1

Jam信号是连续几个字节全1的信号，旨在加强冲突，使其他设备易于检测

CSMA/CD Operation



The Persistence Algorithm

- IEEE 802.3 uses 1-persistent
  - Both non-persistent and p-persistent have performance problems
- Collision handling for 1-persistent
  - Wasted time due to collisions is short
  - With random backoff, unlikely to collide on next tries
  - Binary exponential backoff used

Binary Exponential Backoff

- Attempt to transmit repeatedly if repeated collisions
  - First 10 attempts, mean value of random delay doubled
  - Value then remains same for 6 further attempts
  - After 16 unsuccessful attempts, station gives up and reports error
- 1-persistent algorithm with binary exponential backoff efficient over wide range of loads
  - Low loads, 1-persistence guarantees efficiency
  - High loads, at least as stable as other techniques
- Backoff algorithm gives last-in, first-out effect
  - Stations with few collisions transmit first

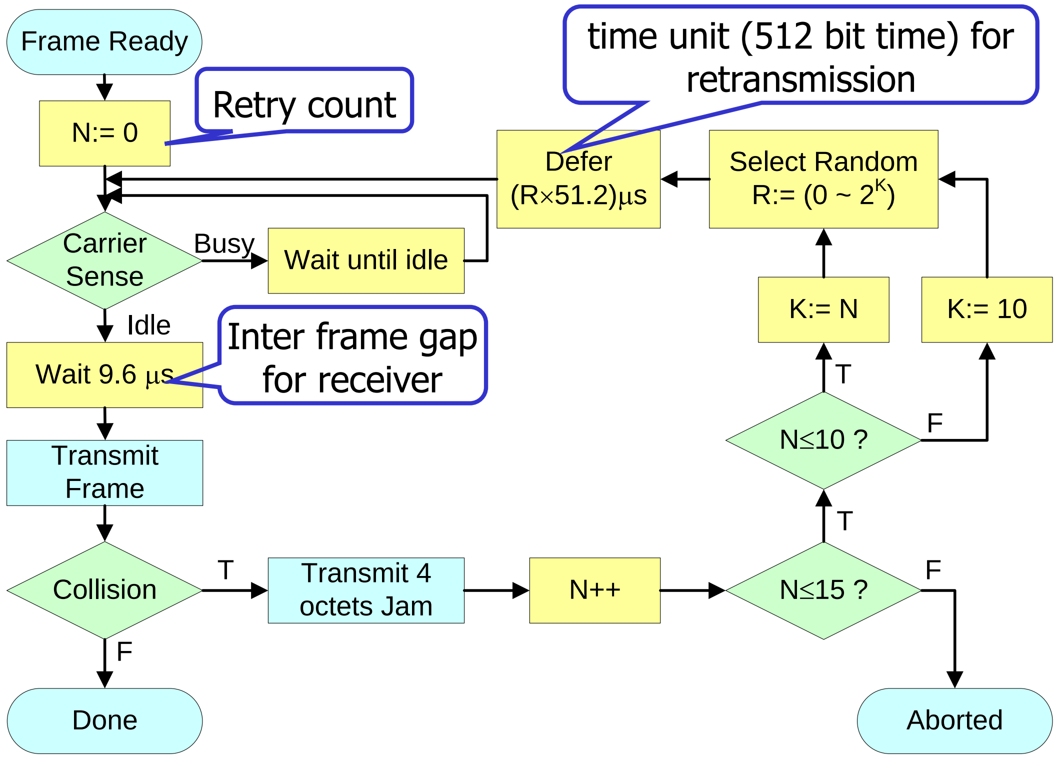
Collision Detection

- On baseband bus, collision produces much higher signal voltage than signal
  - Collision detected if cable signal greater than single station signal
- Signal attenuated over distance
  - Jam needed
  - Limit distance to 500m (10Base5) or 200m (10Base2)
- For twisted pair (star-topology) activity on more than one port is collision
  - Special collision presence signal

Ethernet CSMA/CD algorithm

1. NIC receives datagram from network layer, creates frame
2. If NIC senses channel idle, starts frame transmission. If NIC senses channel busy, waits until channel idle, then transmits.
3. If NIC transmits entire frame without detecting another transmission, NIC is done with frame !
4. If NIC detects another transmission while transmitting, aborts and sends jam signal
5. After aborting, NIC enters binary (exponential) backoff:
   - after mth collision, NIC chooses K at random from {0,1,2, …, 2m-1}. NIC waits K·512 bit times, returns to Step 2
   - longer backoff interval with more collisions

IEEE 802.3 Transmission Algorithm



CSMA/CD efficiency

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

- efficiency goes to 1
  - as t_{prop} goes to 0
  - as t_{trans} goes to infinity
- better performance than ALOHA: and simple, cheap, decentralized!

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

IEEE 802.3

Ethernet Frame Format

Ethernet v2 MAC frames （DEC，Intel，Xerox，1980，简称DIXv2）

Preamble	DA	SA	Type	Data	FCS
8 octets	6	6	2	46~1500	4

IEEE 802.3 MAC frames

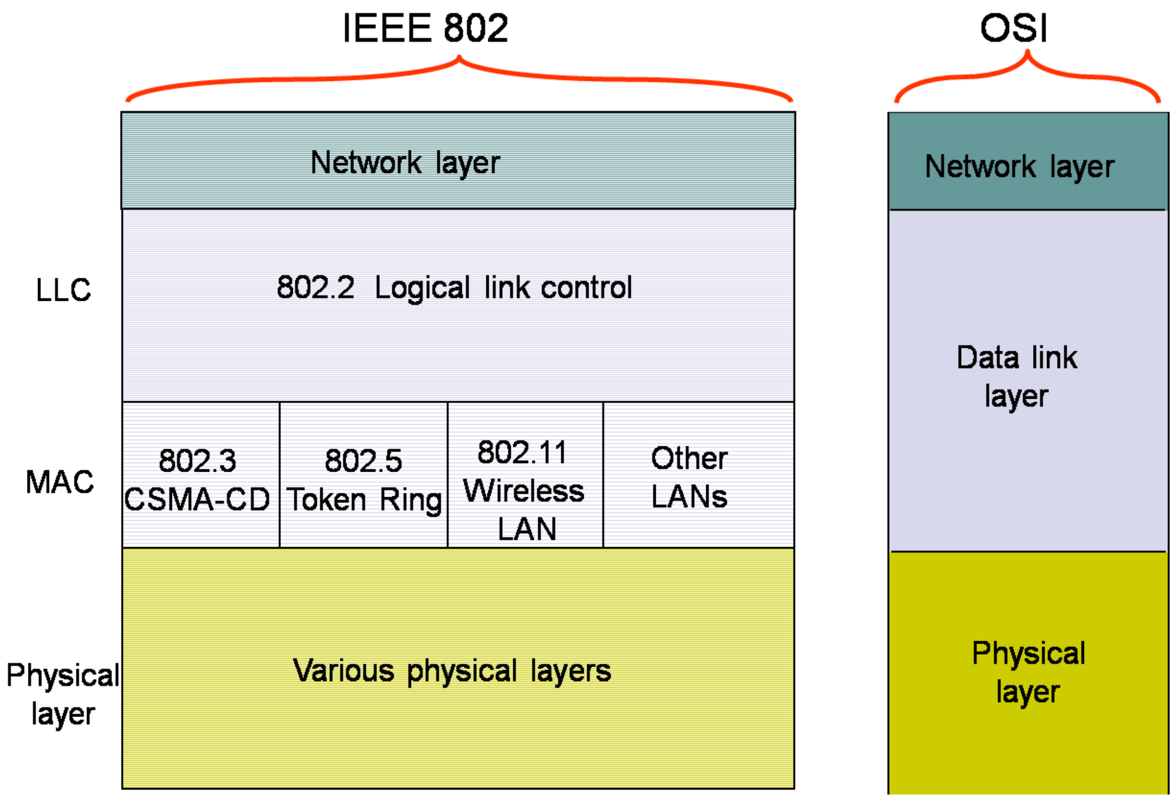
46 to 1500 octets

7 octets	1	6	6	2	≥0	≥0	4
Preamble	SFD	DA	SA	Length	LLC Data	Pad	FCS

SFD = Start of frame delimiter
DA = Destination address
SA = Source address
FCS = Frame check sequence

- Preamble
  - 7 octets(八位组) with pattern 10101010 followed by one octet with pattern 10101011
  - Used to synchronize receiver, sender clock rates
- SFD：帧起始定界符
- Addresses (DA, SA)
  - 6 octets MAC addresses
  - e.g. 08:00:2b:e4:b1:02 stands for 00001000 00000000 00101011 11100100 10110001 00000010
- Type/Length
  - Indicates higher layer protocol, IP, Novell IPX, AppleTalk
  - Or length of LLC data (in IEEE 802.3)
  - 区别：IEEE 802.3没有分配1536 （十进制）以下的数为协议类型代码，故值>=1536时，为类型字段，是dixv2帧。如果从源地址之后的2个字节小于1536，则是长度字段，为IEEE802.3帧。
- FCS
  - CRC checked at receiver

IEEE 802.3



IEEE 802
OSI
Network layer
LLC
802.2 Logical link control
MAC
802.3
CSMA-CD
802.5
Token Ring
802.11
Wireless
LAN
Other
LANs
Physical
layer
Various physical layers
Network layer
Data link
layer
Physical
layer

IEEE 802 – Logical Link Control

Ethernet: Unreliable and connectionless

- No handshaking between sending and receiving NICs
- Receiver doesn't send ACK or NAK to sender

Logical Link Control

- Handle logical links between the stations
- Flow and error control

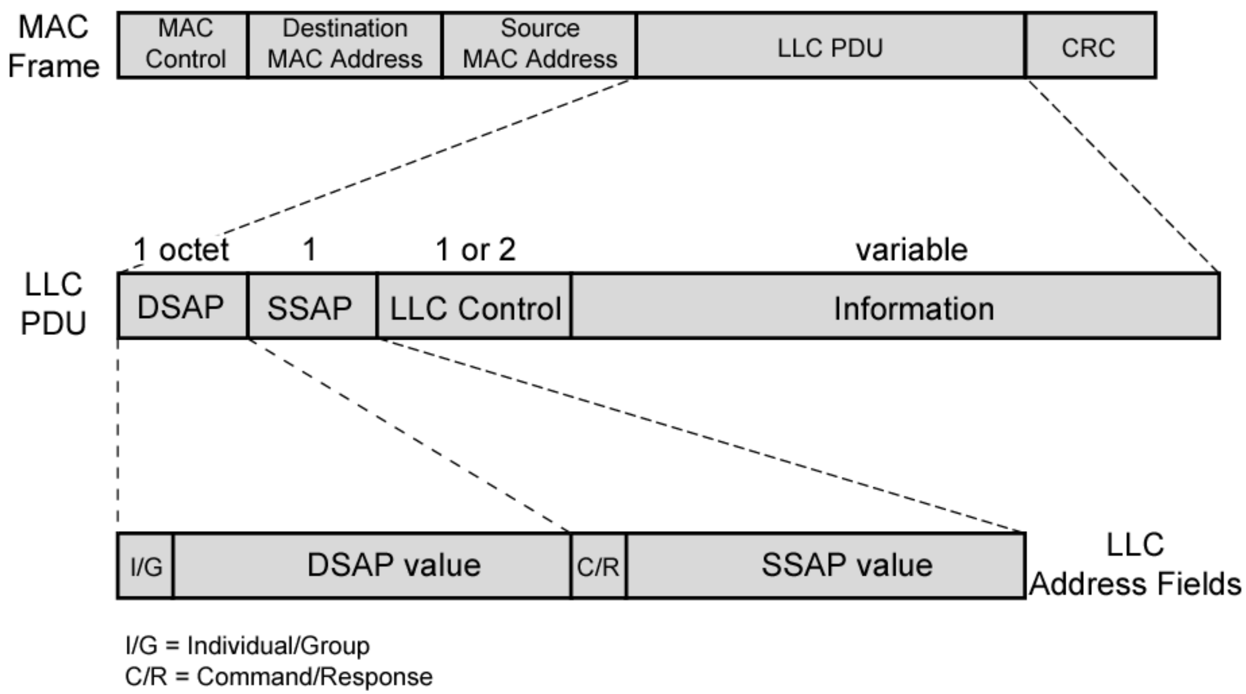
- Rare used on Ethernet, but on WiFi and Token Ring

LLC Services

Based on HDLC, 3 defined:

- Unacknowledged connectionless service
  - Nothing added
- Acknowledged connectionless service
  - Add ACK and NAK, stop-and-wait
- Connection mode service
  - HDLC in Asynchronous balanced mode

LLC PDU and MAC Frame



DSAP: destination service access point
SSAP: Source service access point

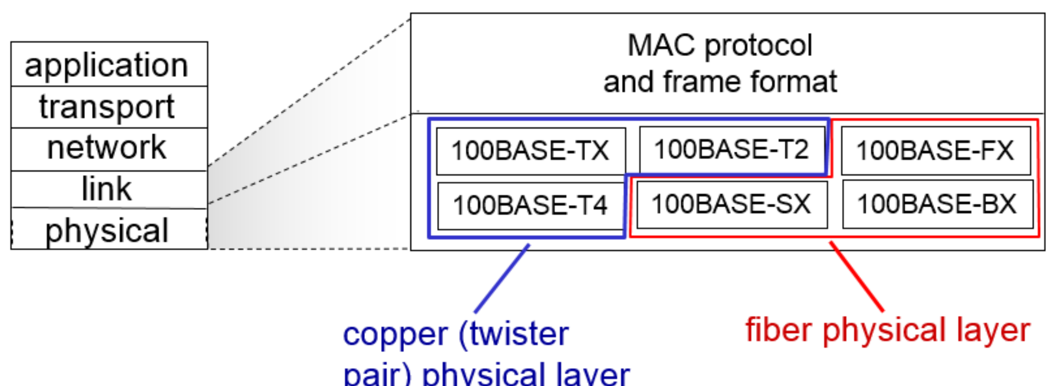
802.3 Physical Layer

- *many* different Ethernet standards
  - common MAC protocol and frame format
  - different speeds: 2 Mbps, 10 Mbps, 100 Mbps, 1Gbps, 10G bps
  - different physical layer media: fiber, cable

speed

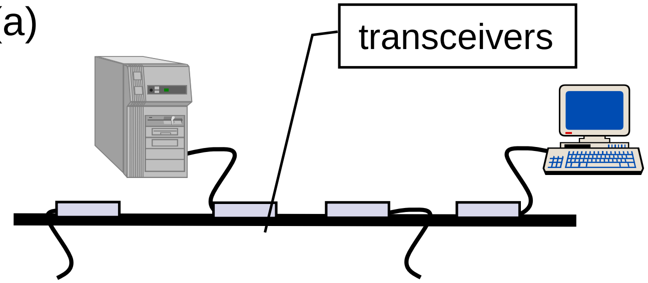
medium

	10base5	10base2	10baseT	10baseFX
Medium	Thick coax	Thin coax	Twisted pair	Optical fiber
Max. Segment Length	500 m	200 m	100 m	2 km
Topology	Bus	Bus	Star	Point-to-point link



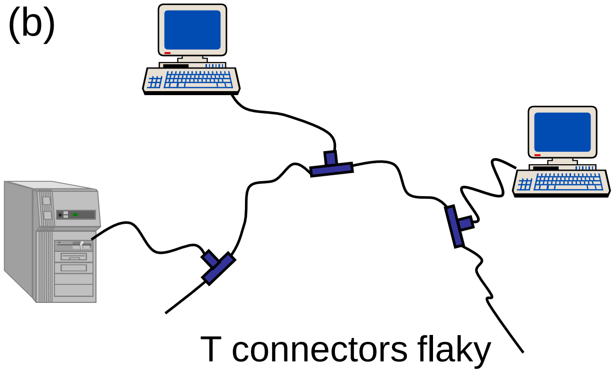
802.3 Physical Layer

(a)



Thick Coax: Stiff, hard to work with

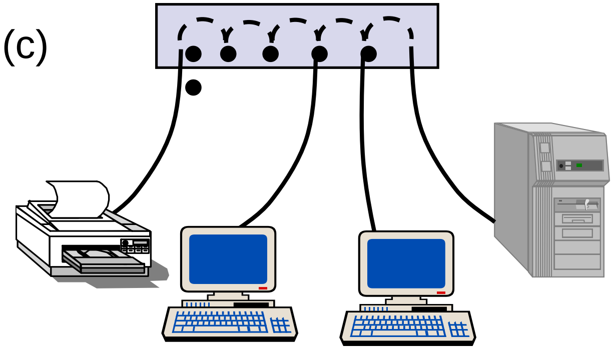
(b)



T connectors flaky

Single collision domain

(c)



Twisted Pair Cheap Easy to work with Reliable

High-Speed Ethernet

High-Speed Ethernet

- Still use IEEE 802.3 MAC protocol and frame format
- Star topology

- Speedup by modern wiring and signaling techniques
  - 100Mbps Fast Ethernet
  - Gigabit Ethernet
  - 10Gbps Ethernet

100Mbps Fast Ethernet

- 100BASE-TX
  - 2 pairs of twisted-pair cable (STP and Category 5 UTP)
  - MLT-3 signaling scheme
- 100BASE-FX
  - 2 optical fiber cables
  - 4B/5B-NRZI code group stream
- 100BASE-T4
  - 4 pairs of Cat. 3 or Cat. 5 UTP
  - Data stream split into 3 separate streams
  - Ternary signaling scheme (8B6T)

100Mbps Fast Ethernet

	100BASE-TX		100BASE-FX	100BASE-T4
Medium	2 pair, STP	2 pair, cat 5 UTP	2 optical fibre	4 pair, cat 3,4,5 UTP
Signaling	MLT-3	MLT-3	4B5B, NRZI	8B6T, NRZ
Transmission Rate	100 Mbps			
Topology	Star			
Max Length (m)	100	100	200	100
Network Span (m)	200	200	400	200

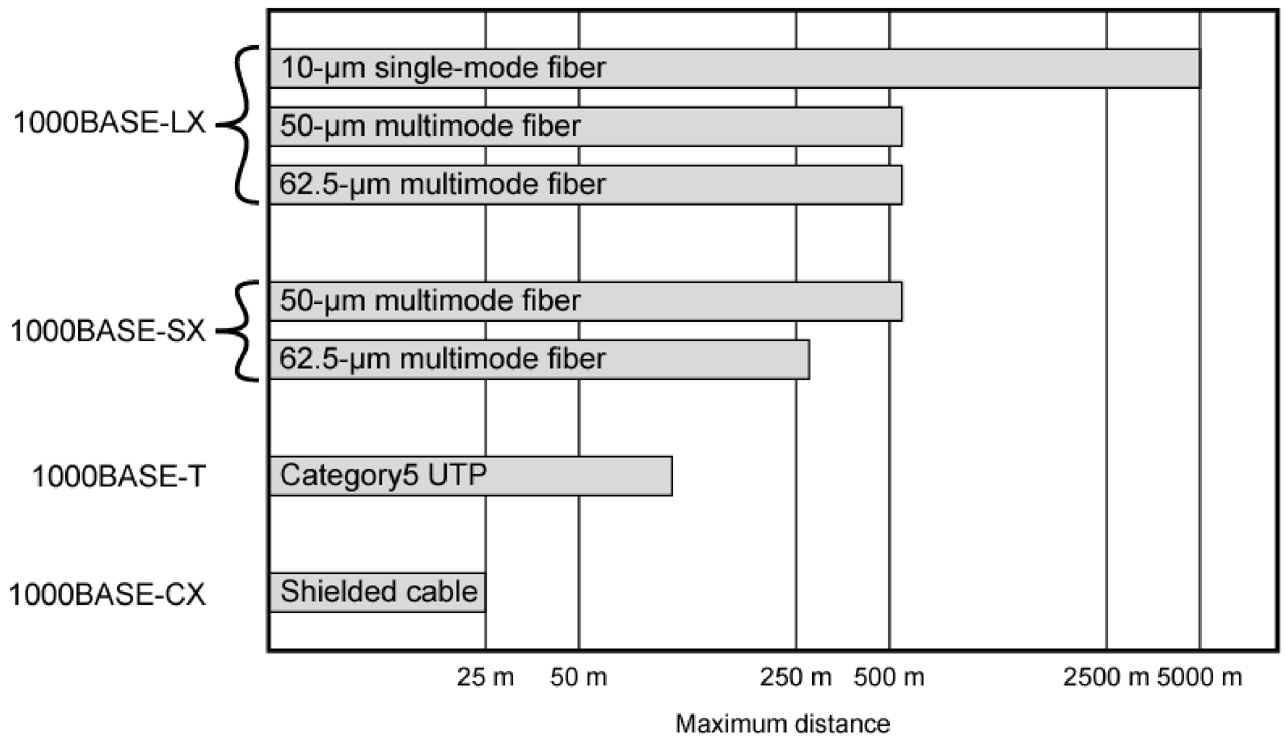
Gigabit Ethernet

- Carrier extension
  - At least 4096 bit-times long
- Frame bursting
  - Treat multiple small frames as a large one
- Support Half Duplex and Full Duplex Operation
  - Half Duplex
  - Full Duplex
    - Use switch
    - Each station constitutes separate collision domain, thus no need CSMA/CD

Gigabit Physical Layer

- 1000Base-SX
  - Short wavelength, multi mode fiber
- 1000Base-LX
  - Long wavelength, multi or single mode fiber
- 1000Base-CX
  - Copper jumpers <25m, shielded twisted pair
- 1000Base-T
  - 4 pairs, cat 5 UTP
- Signaling
  - Fiber: 8B/10B; UTP: 4D-PAM5

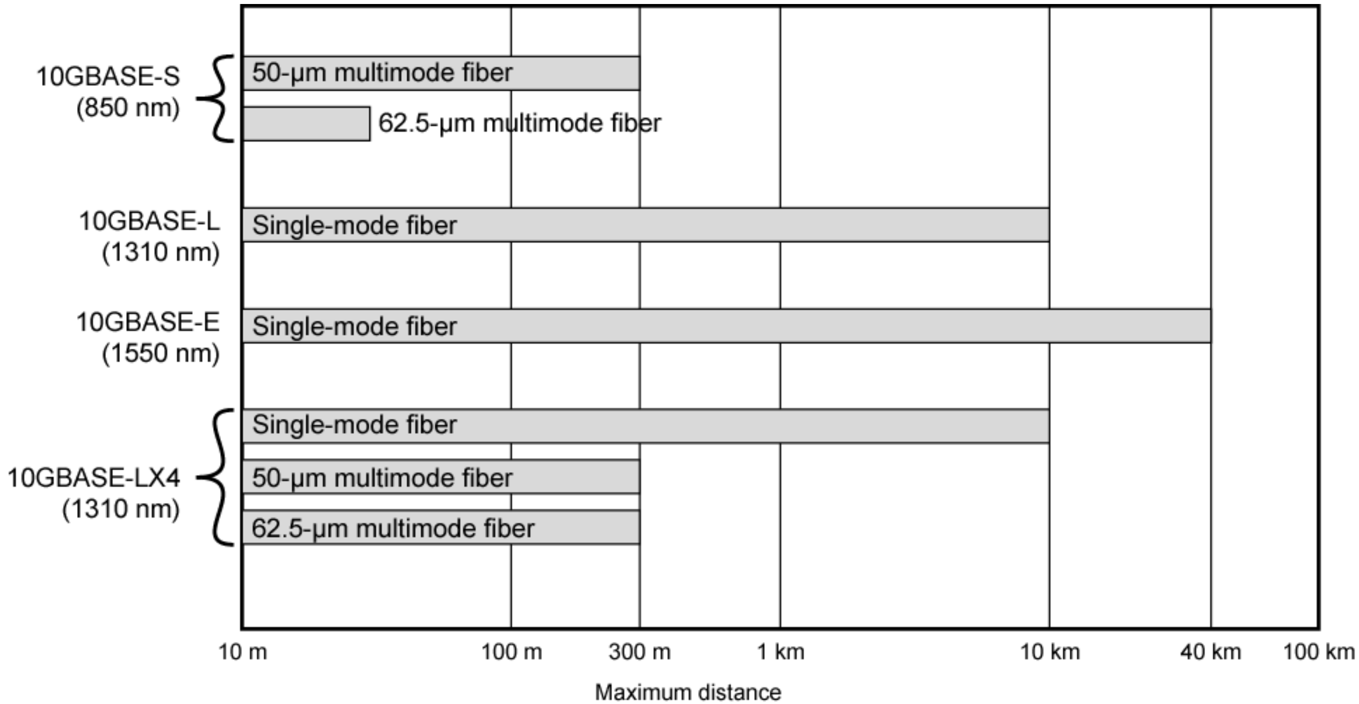
Gigabit Distance Options



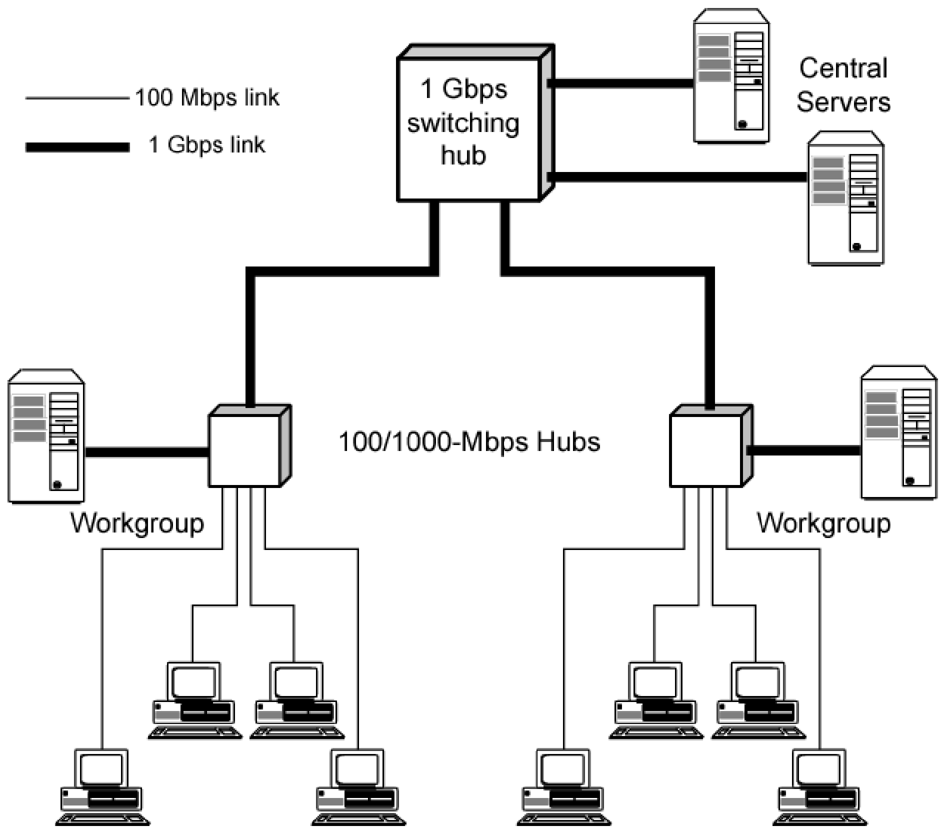
10Gbps Ethernet

- Full-duplex mode only
- 10GBASE-S (short)
  - 850 nm on multi-mode fiber, Up to 300 m
- 10GBASE-L (long)
  - 1310 nm on single-mode fiber, Up to 10 km
- 10GBASE-E (extended)
  - 1550 nm on single-mode fiber, Up to 40 km
- 10GBASE-LX4
  - 1310 nm on single-mode or multimode fiber, Up to 10 km
  - Wavelength-division multiplexing (WDM) bit stream across 4 light waves

10Gbps Ethernet Distance Options



Gigabit Ethernet Configuration



Summary

- Multiple access protocols
  - 信道切分，轮流访问，随机访问
- CSMA/CD
  - Nonpersistent，1-persistent，p-persistent
  - CSMA/CD原理，算法
- IEEE 802.3
  - 以太网帧格式
- High-Speed Ethernet

Homework

- 书本16章习题：16.2, 16.3, 16.6, 16.9, 16.10, 16.14