Introduction to Computer Networks

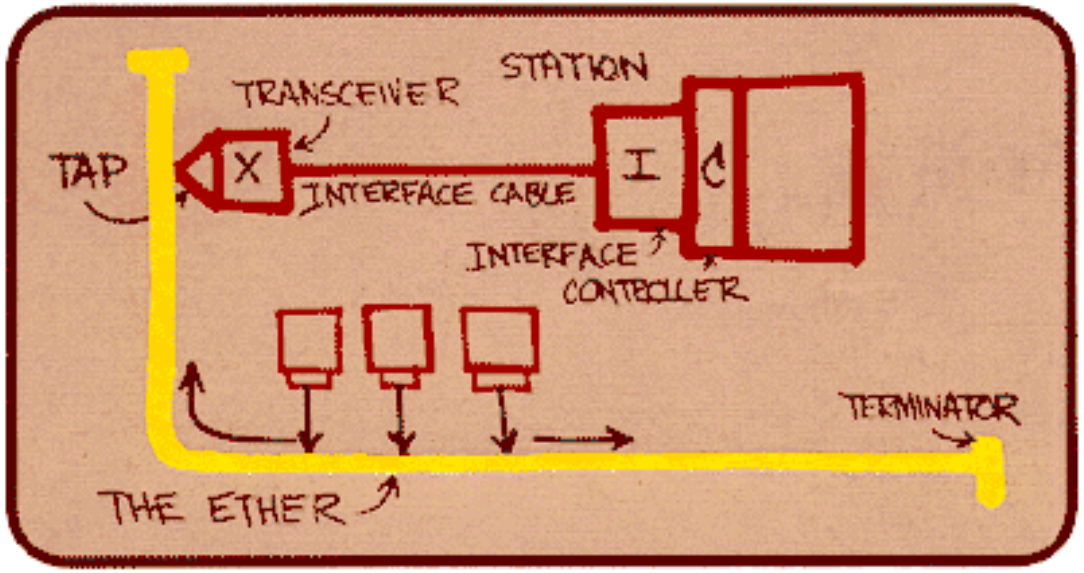
IEEE 802.3 Ethernet

Outline

- Introduction
- Ethernet Topologies
- Ethernet Frame Format
- Ethernet MAC Protocol -- CSMA/CD
- 802.3 Ethernet Standards

Ethernet

- Most successful local area networking technology of last 30 years.
- First widely used LAN technology
- kept up with speed race: 10 Mbps – 100 Gbps



Metcalfe's Ethernet sketch

Ethernet

- Developed in the mid-1970s by researchers at the Xerox Palo Alto Research Centers (PARC).
- DEC and Intel joined Xerox to define a 10-Mbps Ethernet standard in 1978.
- This standard formed the basis for IEEE standard 802.3
- More recently 802.3 has been extended to include
  - 100-Mbps version called Fast Ethernet,
  - 1000-Mbps version called Gigabit Ethernet,
  - 10 Gigabit Ethernet, and also
  - 100 Gigabit Ethernet

Ethernet: Unreliable, Connectionless

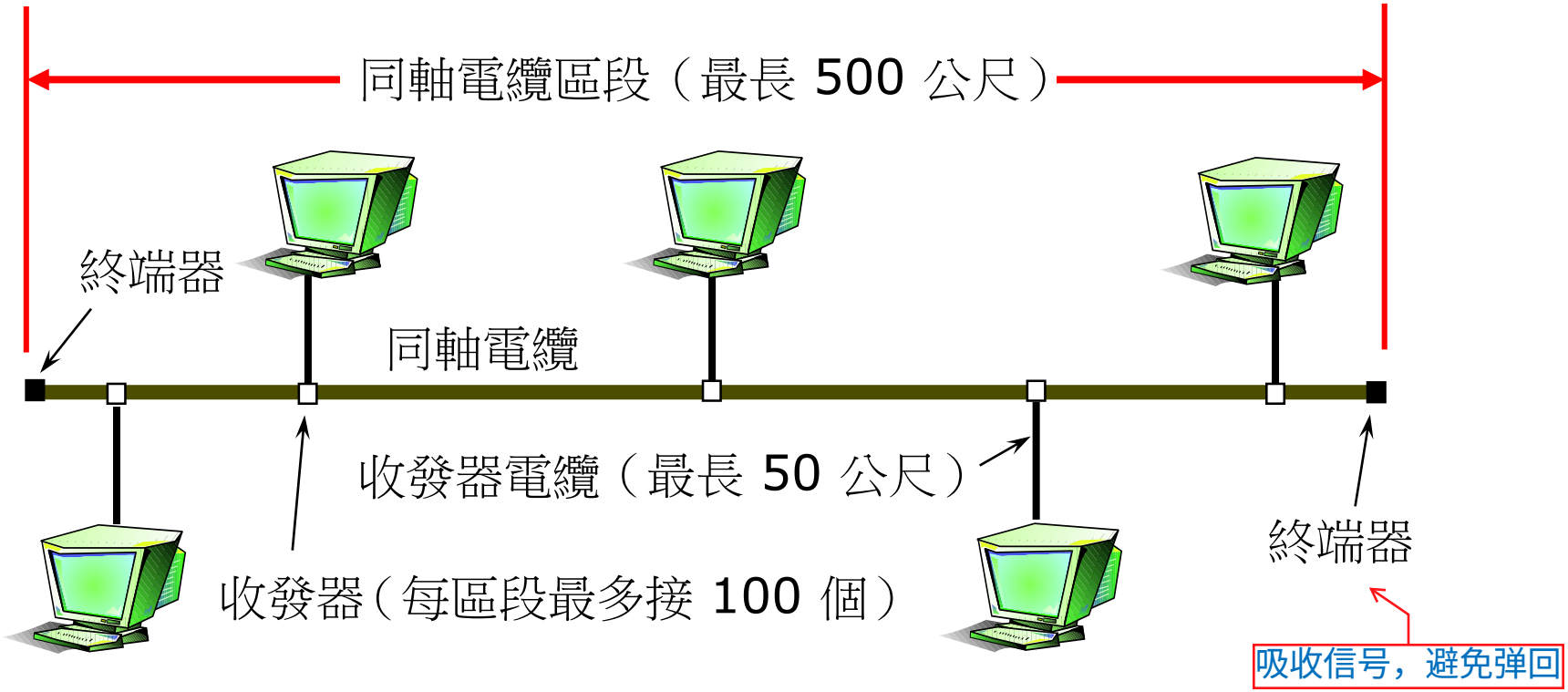
- **Connectionless:** No handshaking between sending and receiving NICs
- **Unreliable:** receiving NIC doesn't send ACKs or NACKs to sending NIC
- Ethernet's MAC protocol: **Carrier Sense Multiple Access with Collision Detection (CSMA/CD)**

Outline

- Introduction
- Ethernet Topologies
- Ethernet Frame Format
- Ethernet MAC Protocol -- CSMA/CD
- 802.3 Ethernet Standards

Bus Topology

- **Bus topology popular through mid 90s**
  - **all nodes in same collision domain (can collide with each other)**



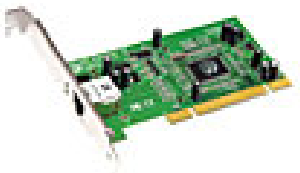
Ethernet (10Base5)

网段，最长500m

- An Ethernet segment is implemented on a coaxial cable of up to 500 m.
- Hosts connect to an Ethernet segment by tapping into it.

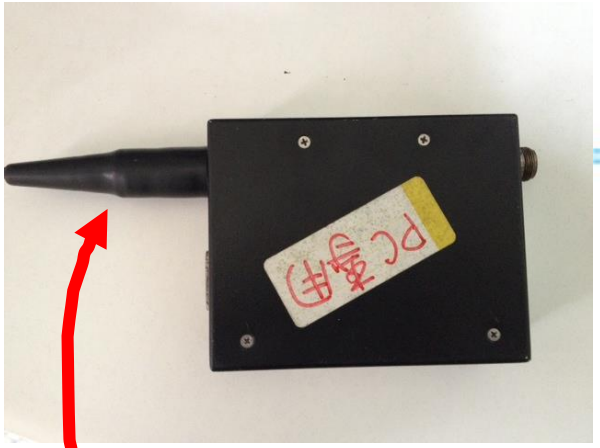
既可传送也可接收

- A transceiver (a small device directly attached to the tap) detects when the line is idle and drives signal when the host is transmitting.
- The transceiver also receives incoming signal.
- The transceiver is connected to an Ethernet adaptor which is plugged into the host. But now most are built in into the computers.
- The protocol is implemented on the adaptor.

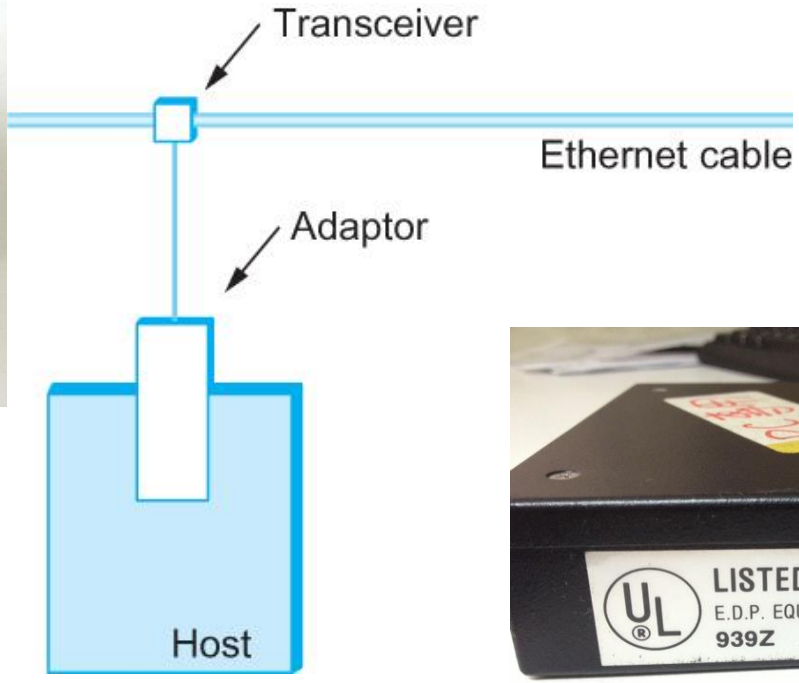


Ethernet (10Base5)

10Mbps，Base技术只有一个channel，最长500m

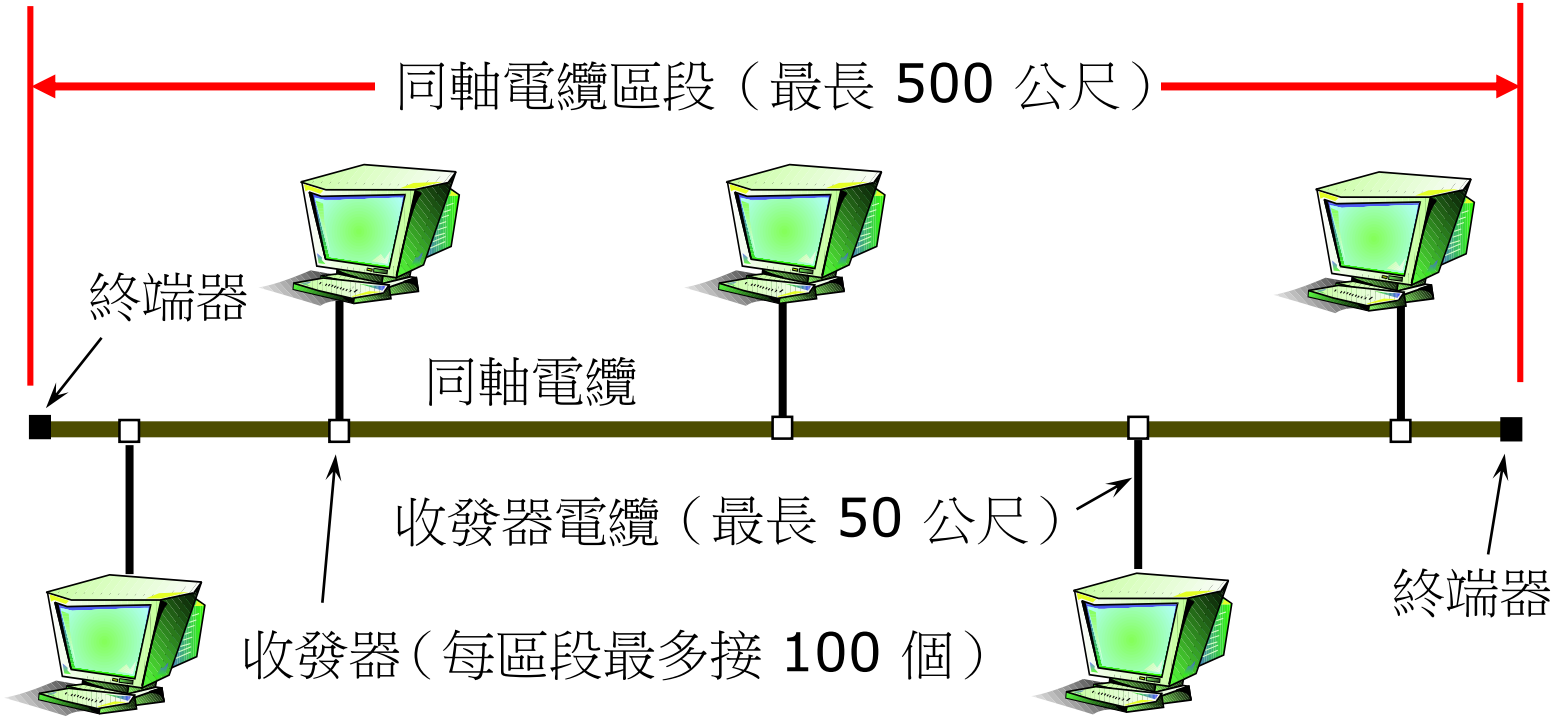


Terminator

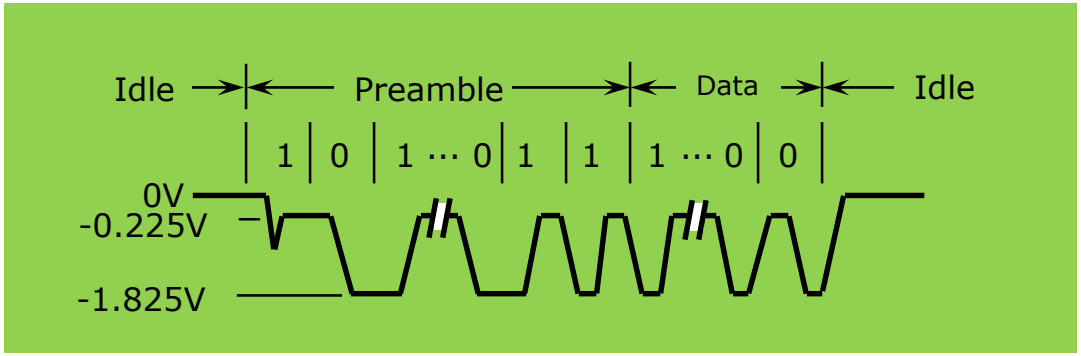


Ethernet transceiver, adaptor, and terminator

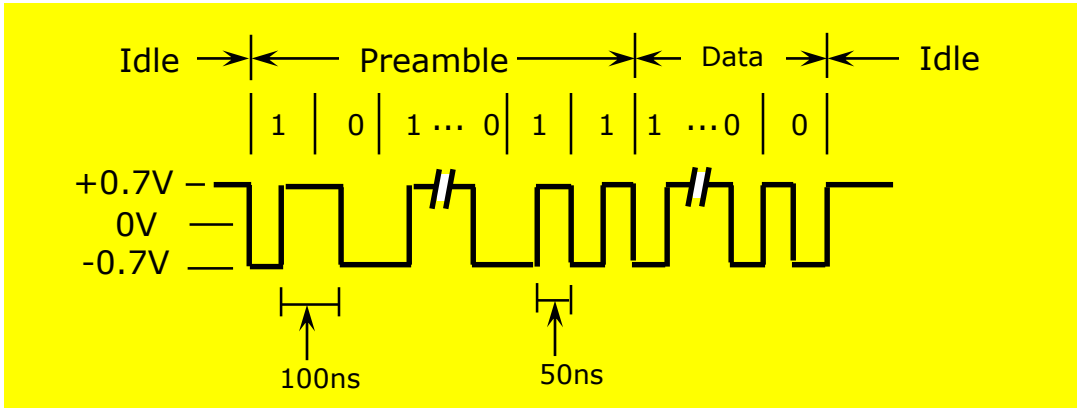
Network Configuration Example 1 (Single segment)



Cable Signaling (Manchester Encoding)



Coaxial Cable



Transceiver Cable

- Each bit has a transition
- Allows clocks in sending and receiving nodes to synchronize to each other

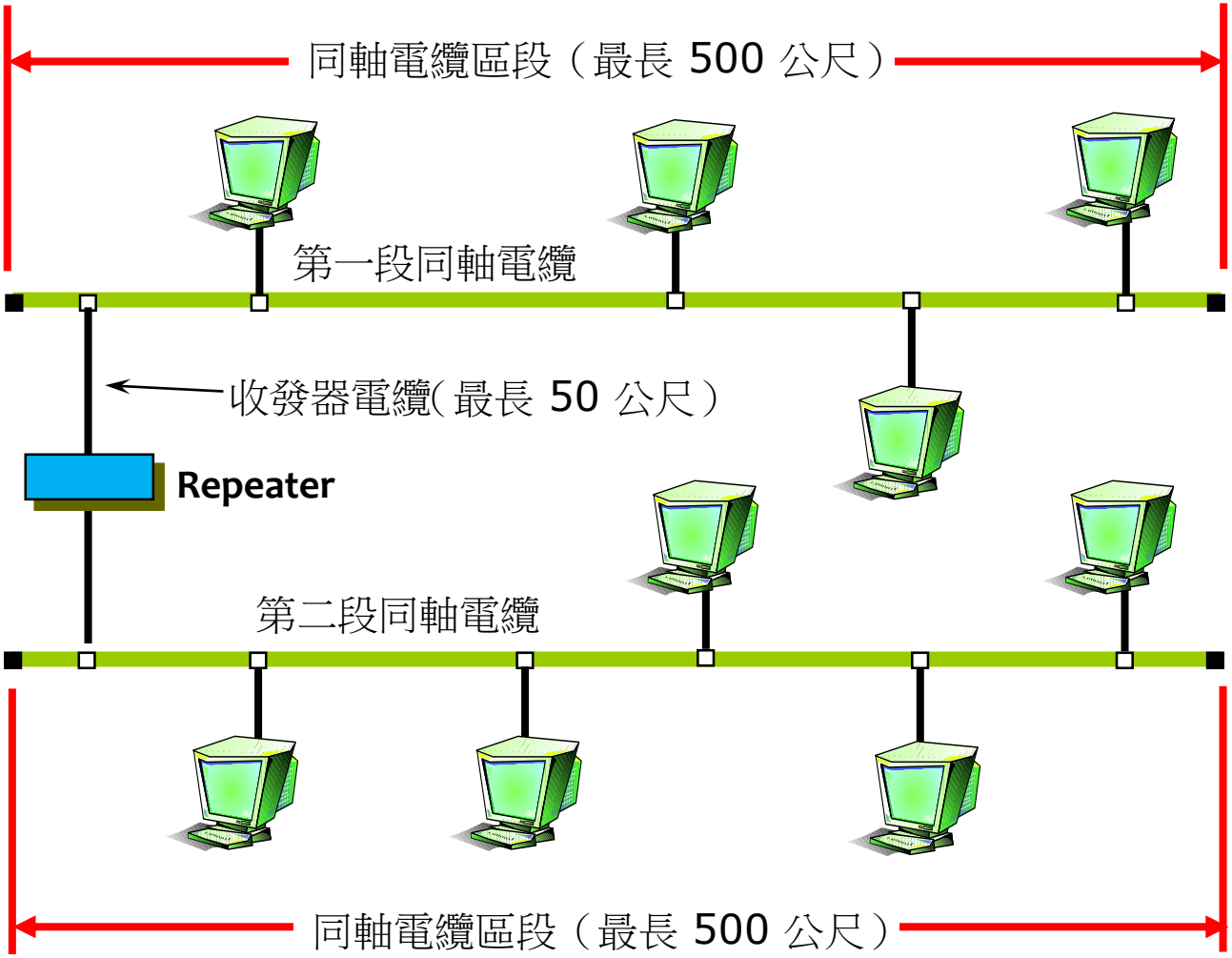
Ethernet (10Base5)

- Multiple Ethernet segments can be joined together by *repeaters.*
- A *repeater* is a device that forwards digital signals.
- No more than four repeaters may be positioned between any pair of hosts.
  - An Ethernet has a total reach of only 2500 m.

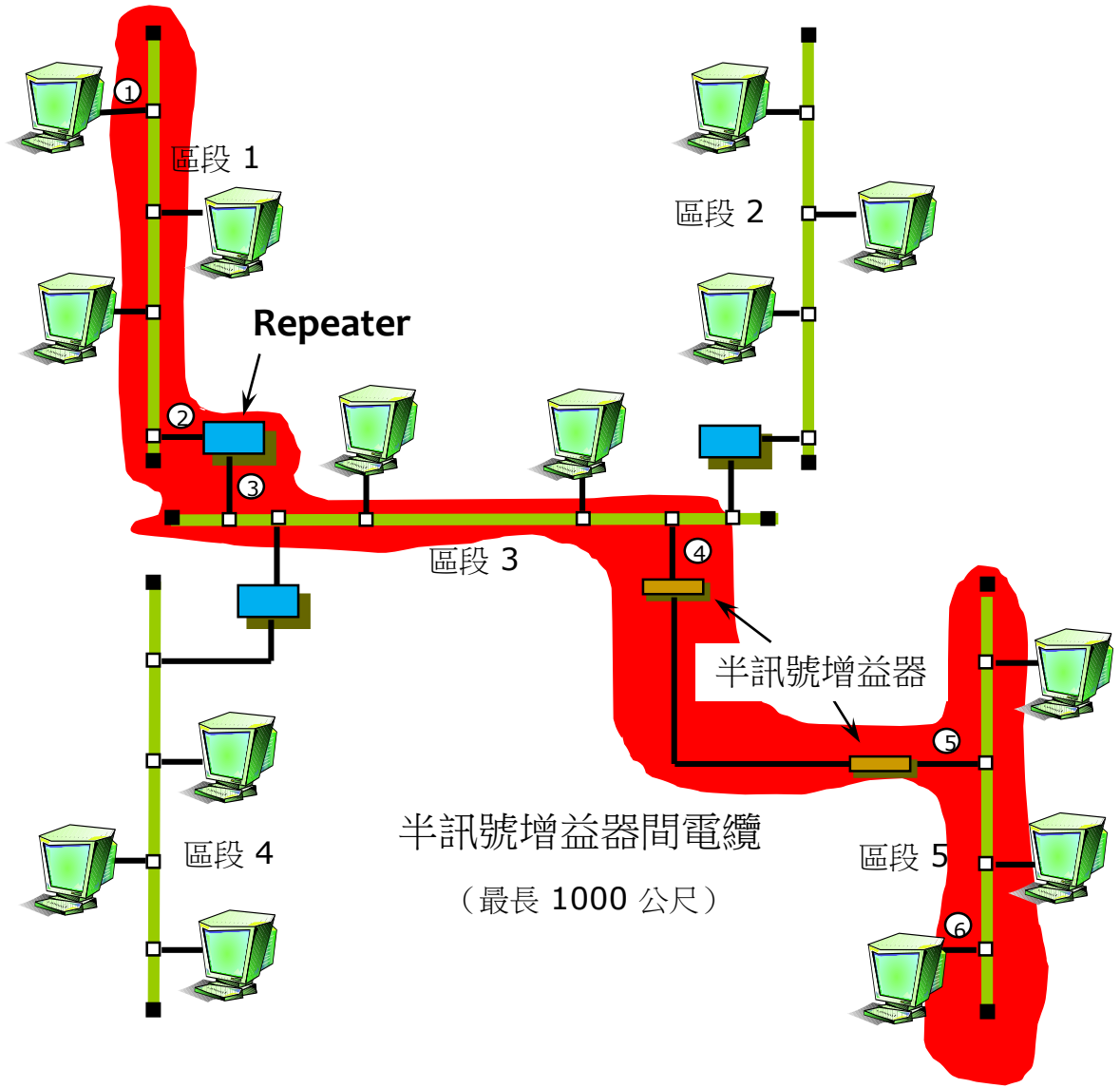
第一层设备，只负责转送讯号而不管MAC地址

将衰减的信号恢复成方波

Network Configuration Example 2 (Two segments)



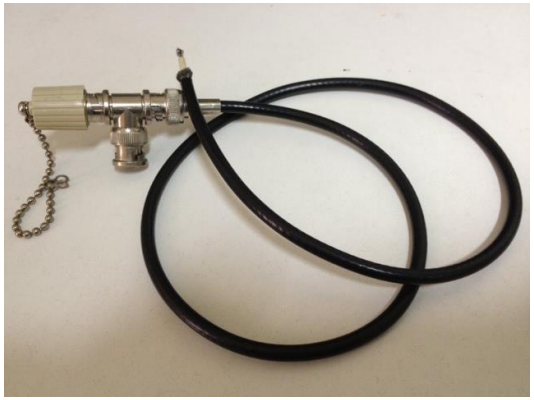
Network Configuration Example 3 (Five segments, maximum)



Ethernet (10Base2)

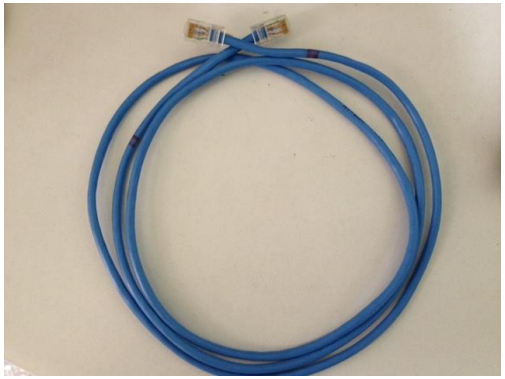
- **New Technologies in Ethernet**
  - Instead of using coax cable, an Ethernet can be constructed from a thinner cable known as 10Base2 (the original was 10Base5)

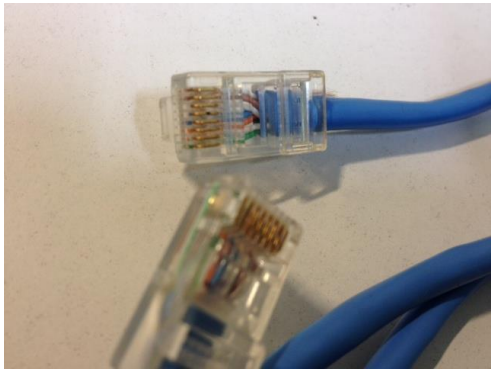
    更轻便，传输距离也更短
    - 10 means the network operates at 10 Mbps
    - Base means the cable is used in a baseband system
    - 2 means that a given segment can be no longer than 200 m



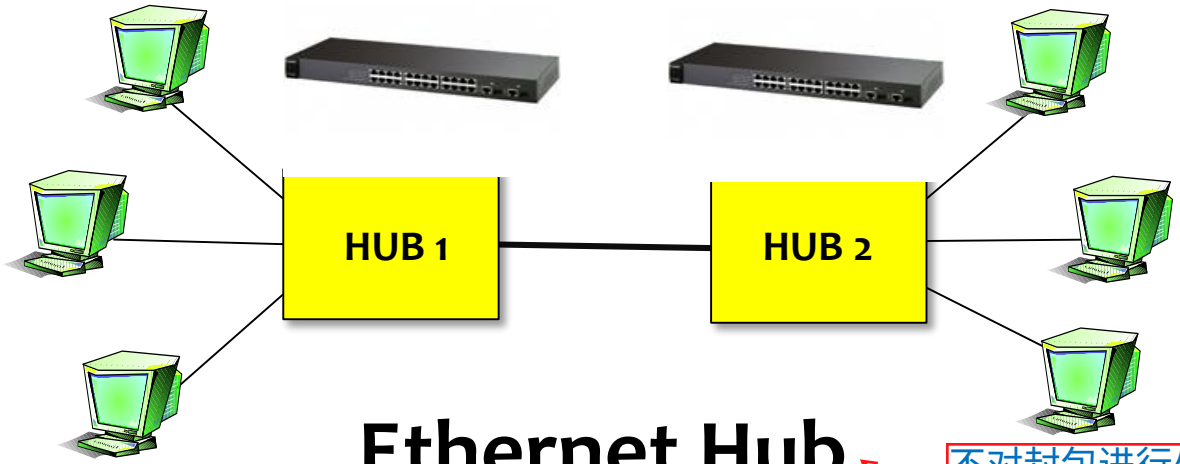
Ethernet (10BaseT)

- New Technologies in Ethernet
  - Another cable technology is 10BaseT
    - T stands for twisted pair 双绞线
    - Limited to 100 m in length
  - With 10BaseT, the common configuration is to have several point to point segments coming out of a multiway repeater, called *Hub* 管理收纳



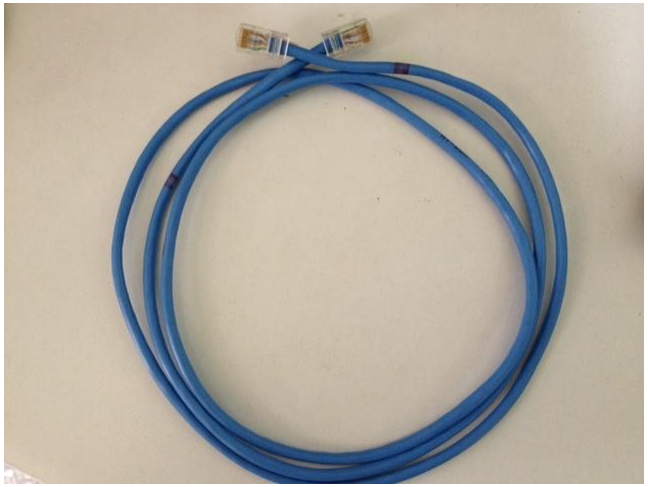


Ethernet



Ethernet Hub

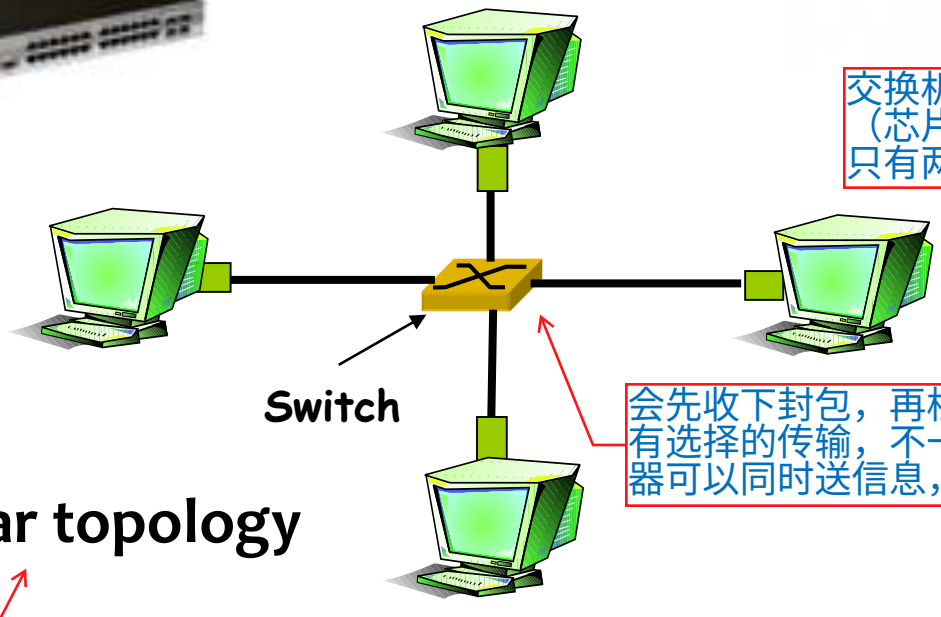
不对封包进行处理，收到后即广播出去，相当于Repeater，是第一层设备。如果多个机器同时送会造成冲突



Star Topology

- **Today: Star topology prevails**
  - active *switch* in center
  - each "spoke" runs a (separate) Ethernet protocol (nodes do not collide with each other)

交换机的每个接口都内置网卡（芯片），相当于每个网段上只有两个机器：主机和交换机



会先收下封包，再根据封包内容进行处理，有选择的传输，不一定会广播，所以多台机器可以同时送信息，对应的网卡会进行处理

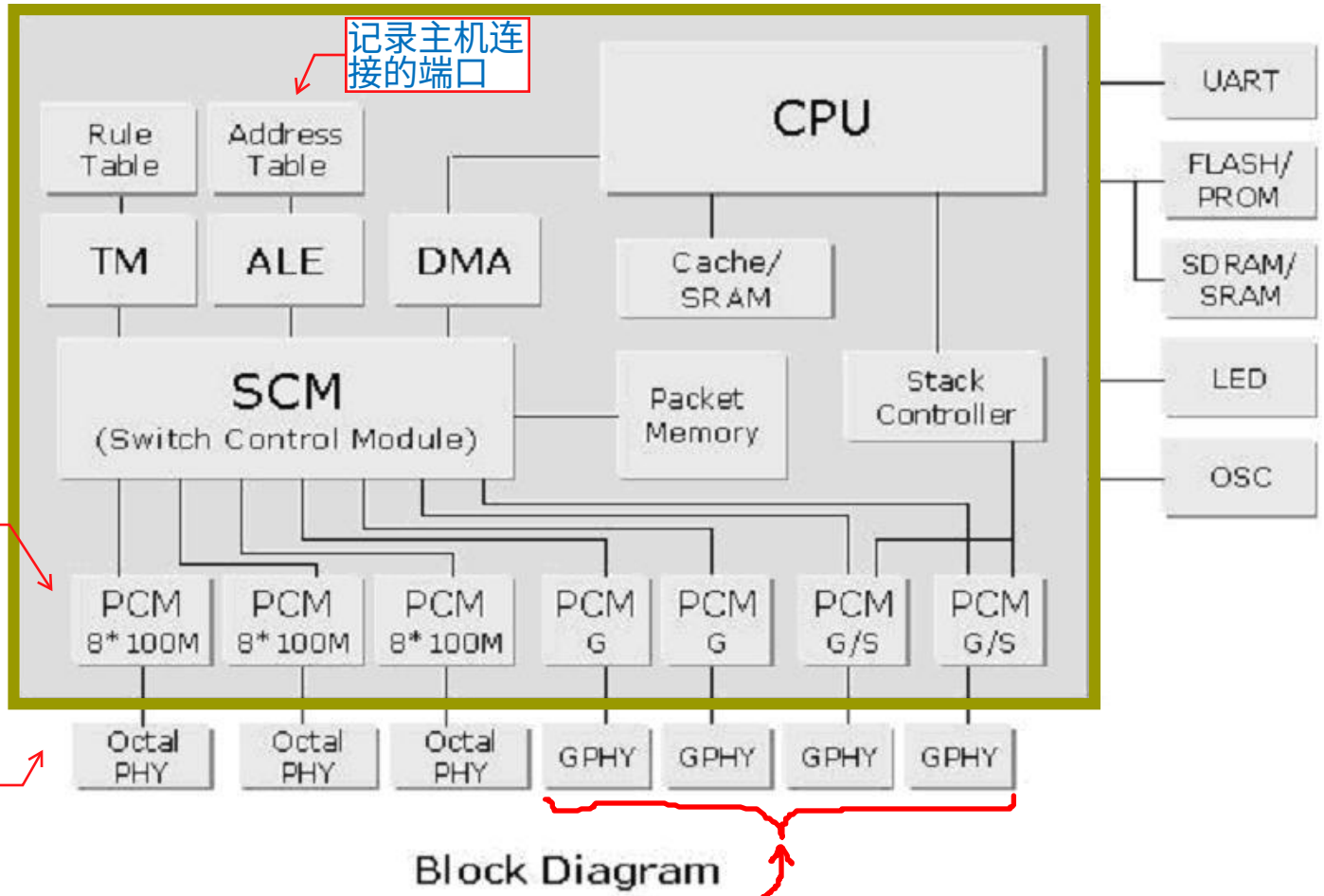
Star topology

星型结构

Ethernet-Switch

- To speed up the transmission rate of Ethernet Hub without changing the interface cards on stations.
- Ether-Switch Architecture
- Each Ethernet port can have a transmission simultaneously.

Ethernet Switch ASIC example



Acute Leo AQ6628 24+4 Ethernet Switch ASIC

每个封包进来，同时查表然后同时交换出去->速度很快而且所有工作都在芯片完成；能够以封包进来的速度完成处理并转送出去->while speed 线速

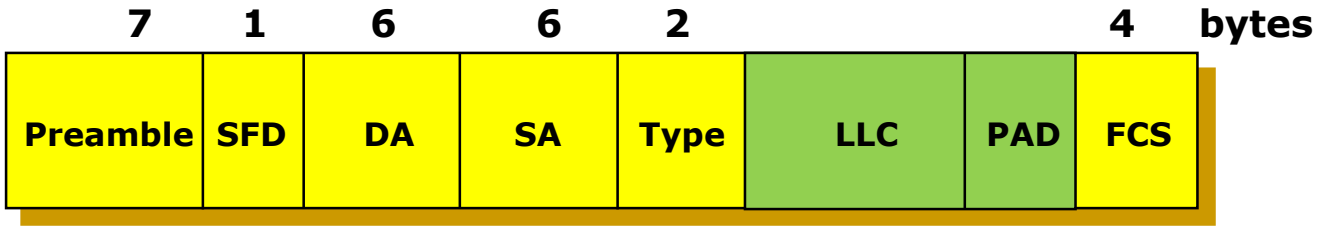
Outline

- Introduction
- Ethernet Topologies
- Ethernet Frame Format
- Ethernet MAC Protocol -- CSMA/CD
- 802.3 Ethernet Standards

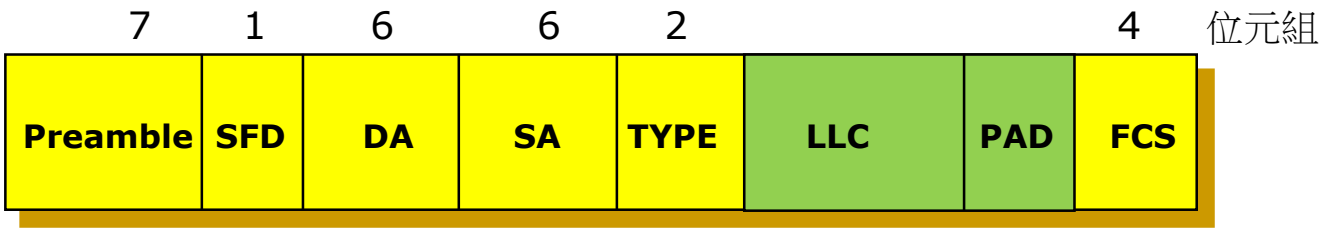
Ethernet Frame Format

Frame format

- Preamble (64bits): allows the receiver to synchronize with the signal (sequence of alternating 0s and 1s).
- Source and Destination MAC Addresses (48bits each).
  网卡的ID具有全球唯一性
- Packet type (16bits): acts as demux key to identify the higher level protocol.
- Data (up to 1500 bytes)
  出于公平
  - Minimally a frame must contain at least 46 bytes of data.
  - Frame must be long enough to detect collision.
- FCS: CRC (32bit)



Ethernet Frame Format



- Preamble: (101010...1010) for Synchronization
- SFD: Start Frame Delimiter (10101011)
- DA: Destination MAC Address
- SA: Source MAC Address
- Packet type (16bits): acts as demux key to identify the higher level protocol.
- LLC-Frame: Up to 1500 bytes
- PAD: Padding when LLC-Frame < 46 bytes
- FCS: Frame Check Sequence (CRC-32)
- MAC-frame size -- from DA to FCS

Cyclical Redundancy Check 循环冗余检验

內容至少46byte+首尾18byte

  - Min 64 bytes to distinguish from collision
  - Max 1518 bytes to prevent dominating bandwidth

內容最多1500byte，首尾18byte不包括Preamble&SFD

Ethernet Addresses

- Each host on an Ethernet (in fact, every Ethernet host in the world) has a unique Ethernet Address.
- The address belongs to the adaptor, not the host.
  - It is usually burnt into ROM.

地址是属于网卡的，而非主机，不插网卡的主机是没有地址的

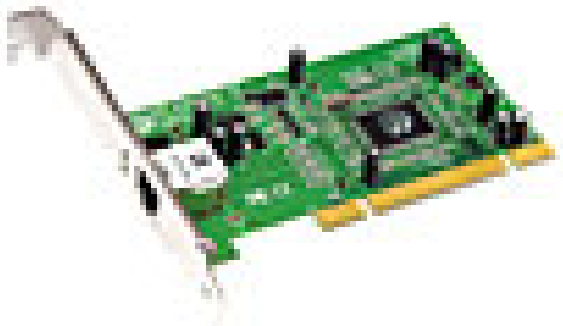
- Ethernet addresses are typically printed in a human readable format
  - As a sequence of six numbers separated by colons.
  - Each number corresponds to 1 byte of the 6 byte address and is given by a pair of hexadecimal digits, one for each of the 4-bit nibbles in the byte
  - Leading 0s are dropped.

十六进制

  - For example, 8:0:2b:e4:b1:2 is
    - 00001000 00000000 00101011 11100100 10110001 00000010

Ethernet Addresses

- To ensure that every adaptor gets a unique address, each manufacturer of Ethernet devices is allocated a different prefix that must be prepended to the address on every adaptor they build
  - AMD has been assigned the 24bit prefix 8:0:20



Ethernet Addresses

- Each frame transmitted on an Ethernet is received by every adaptor connected to that Ethernet.
- Each adaptor recognizes those frames addressed to its address and passes only those frames on to the host.
- In addition to *unicast* address, an Ethernet address consisting of all 1s is treated as a *broadcast* address.
  - All adaptors pass frames addressed to the *broadcast* address up to the host.
- Similarly, an address that has the first bit set to 1 but is not the *broadcast* address is called a *multicast* address.
  - A given host can program its adaptor to accept some set of *multicast* addresses.

Ethernet Addresses

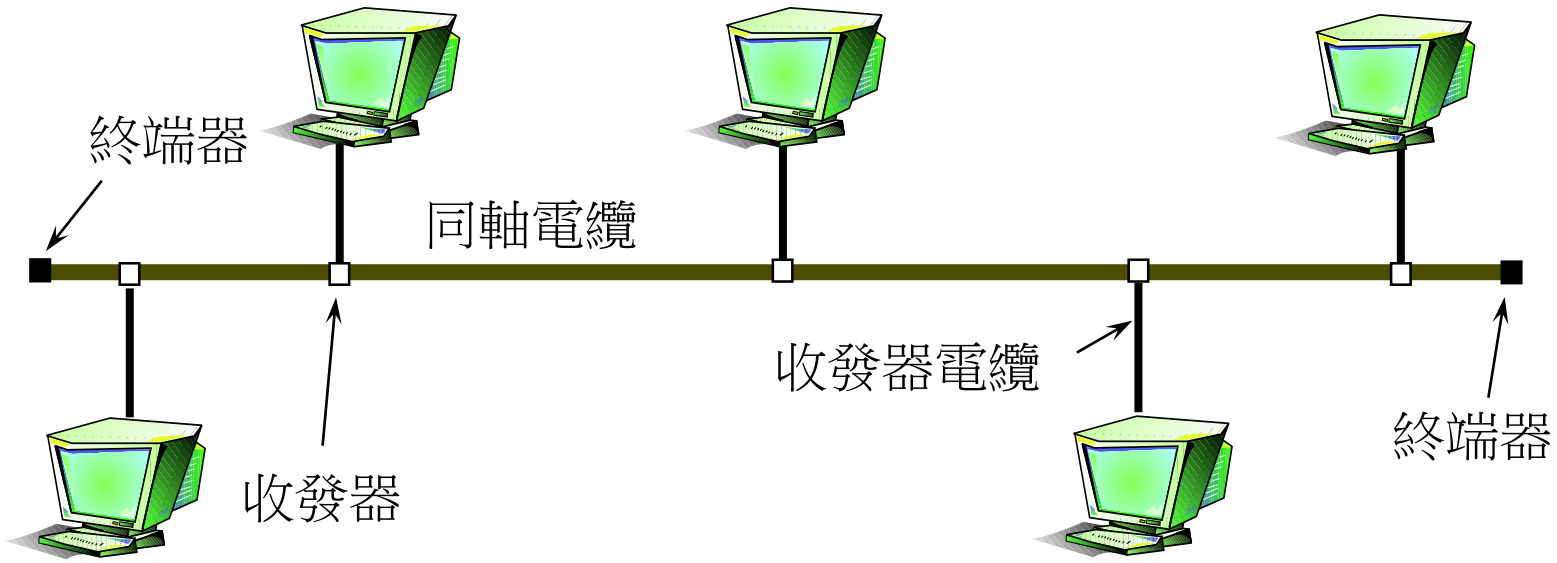
- To summarize, an Ethernet adaptor receives all frames and accepts
  - Frames addressed to its own address
  - Frames addressed to the broadcast address
  - Frames addressed to a multicast address if it has been instructed

Outline

- Introduction
- Ethernet Topologies
- Ethernet Frame Format
- Ethernet MAC Protocol -- CSMA/CD
- 802.3 Ethernet Standards

Ethernet MAC protocol

- Any signal placed on the Ethernet by a host is broadcast over the entire network
  - Signal is propagated in both directions.
  - Repeaters forward the signal on all outgoing segments.
  - Terminators attached to the end of each segment absorb the signal.



CSMA (Carrier Sense Multiple Access)

CSMA: listen before transmit:

If channel sensed idle: transmit entire frame

- If channel sensed busy, defer transmission

CSMA collisions

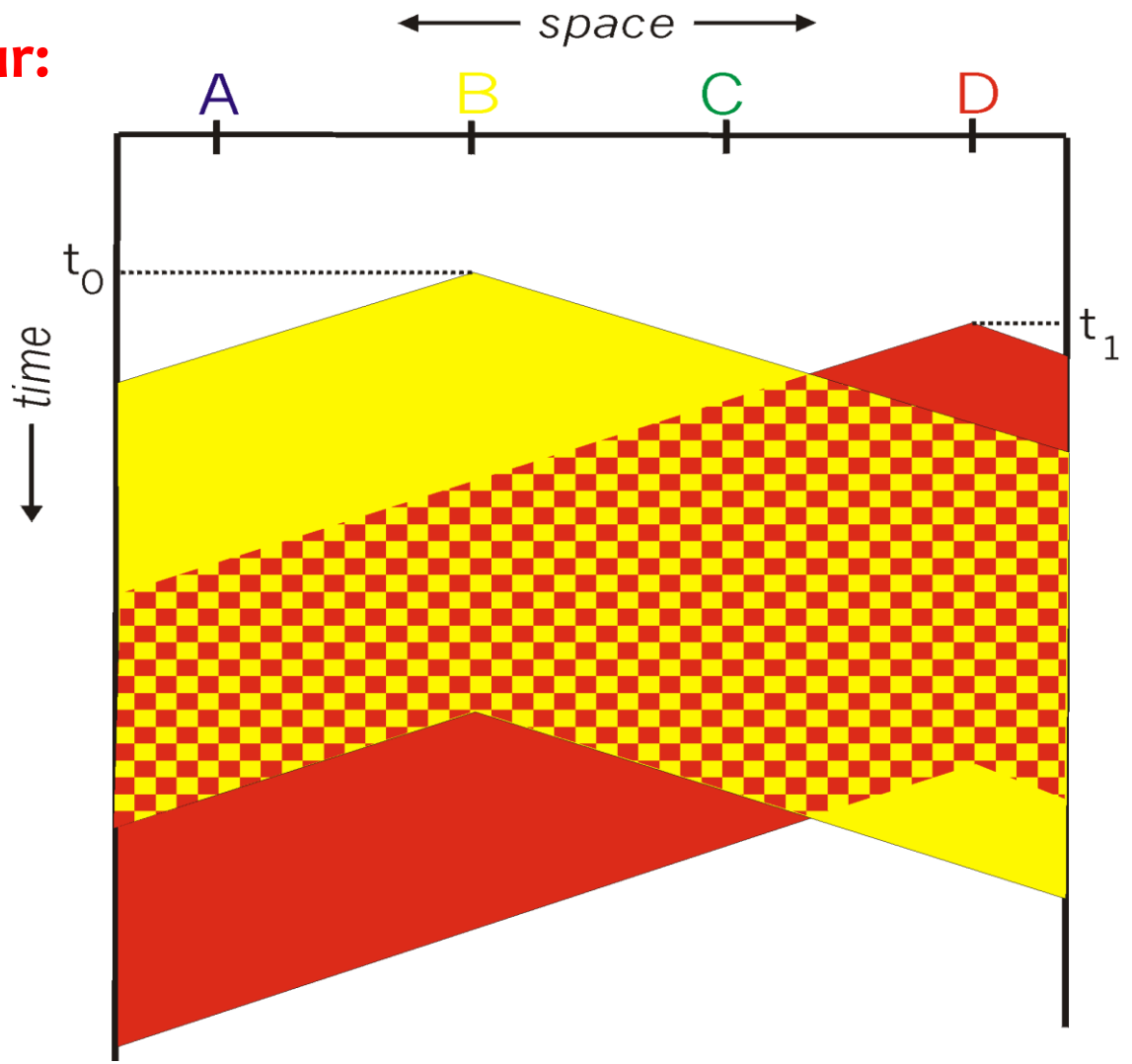
collisions can still occur:

propagation delay means two nodes may not hear each other's transmission

collision:

entire packet transmission time wasted

发送以后继续听，发现冲撞
不作反应仍然将封包送完



CSMA/CD (Collision Detection)

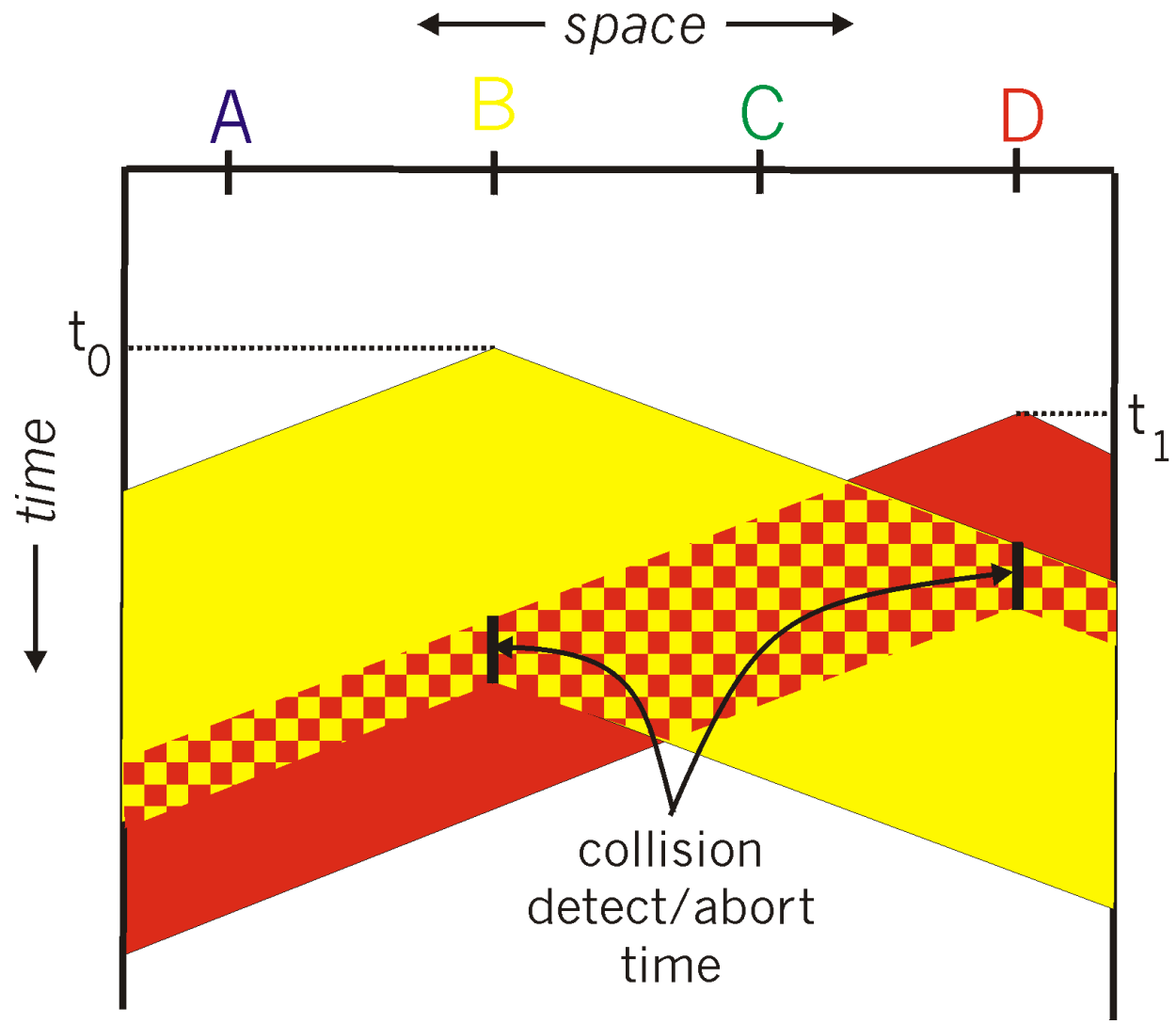
CSMA/CD: carrier sensing, deferral as in CSMA

- collisions *detected* within short time
- colliding transmissions aborted, reducing channel wastage

检测到封包冲撞即停止传送

- Collision detection:
  - Measure signal strengths, compare transmitted, received signals

CSMA/CD collision detection



CSMA/CD

- Carrier Sense Multiple Access with Collision Detection (CSMA/CD).
  - A set of nodes send and receive frames over a shared link.
  - Carrier sense means that all nodes can distinguish between an idle and a busy link.
  - Collision detection means that a node listens as it transmits and can therefore detect when a frame it is transmitting has collided with a frame transmitted by another node.

CSMA/CD

- When the adaptor has a frame to send and the line is idle, it transmits the frame immediately.
- When the adaptor has a frame to send and the line is busy, it waits for the line to go idle and then transmits immediately.
- The Ethernet is said to be 1-persistent protocol because an adaptor with a frame to send transmits with probability 1 whenever a busy line goes idle.

监听到封包传送结束后发起传输的可能性

CSMA/CD

- Since there is no centralized control it is possible for two (or more) adaptors to begin transmitting at the same time,
  - Either because both found the line to be idle,
  - Or, both had been waiting for a busy line to become idle.
- When this happens, the two (or more) frames are said to be *collide* on the network.

CSMA/CD

- Since Ethernet supports collision detection, each sender is able to determine that a collision is in progress.
- At the moment an adaptor detects that its frame is colliding with another, it first makes sure to transmit a 32-bit jamming sequence and then stops transmission.

确保所有工作站都知道
这个封包已经被冲撞

  - Thus, a transmitter will minimally send 96 bits in the case of collision
    - 64-bit preamble + 32-bit jamming sequence

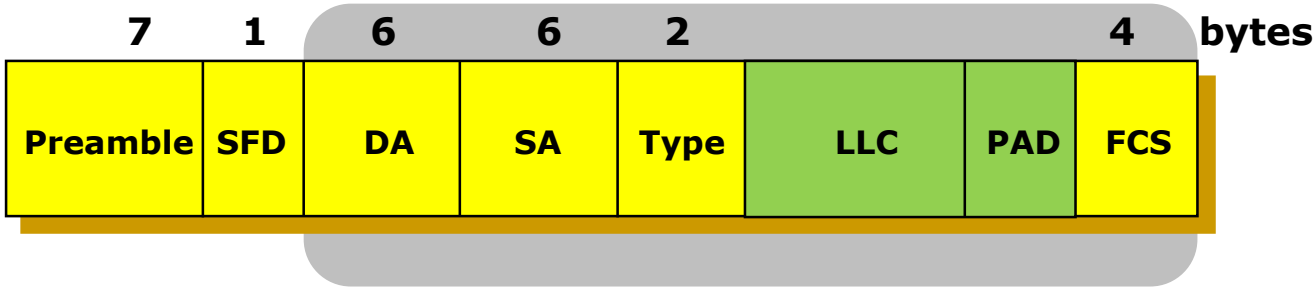
CSMA/CD

- One way that an adaptor will send only 96 bits (called a *runt frame*) is if the two hosts are close to each other.
- In case the two hosts are farther apart, they would have had to transmit longer, and thus send more bits, before detecting the collision.

Collision Window

- The worst case scenario happens when the two hosts are at opposite ends of the Ethernet.
- To know for sure that the frame its just sent did not collide with another frame, the transmitter may need to send as many as 512 bits.
  - Every Ethernet frame must be at least 512 bits (64 bytes) long.
    - 14 bytes of header + 46 bytes of data + 4 bytes of CRC

若封包短于此长度，则不足够确认信号到网络另一端再回来是否发生了冲撞

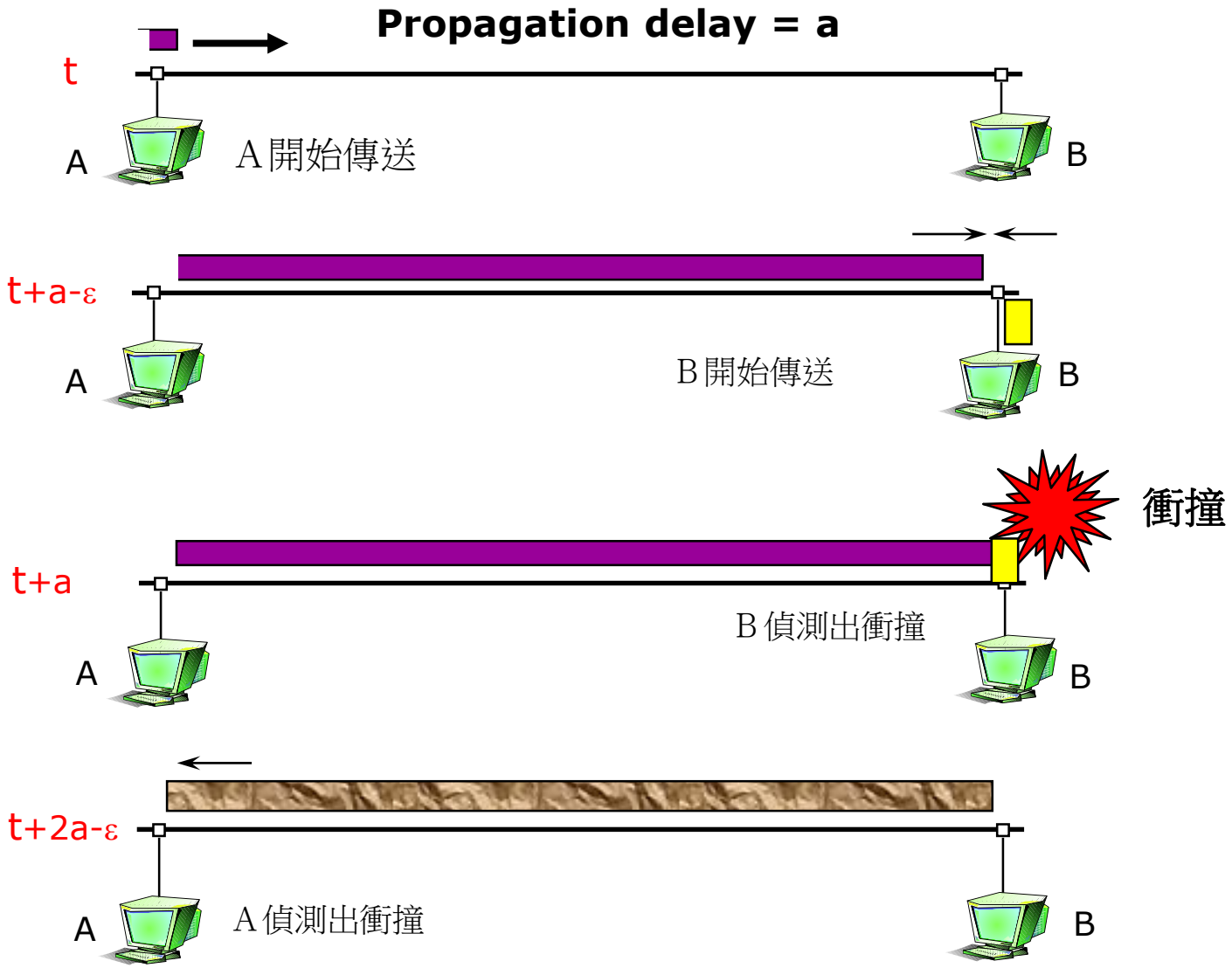


Collision Window

- Why 512 bits (64 bytes) ?
  - Why is its length limited to 2500 m?
- Collision Window = round-trip delay (2a)

- The farther apart two nodes are, the longer it takes for a frame sent by one to reach the other, and the network is vulnerable to collision during this time

Collision Detection Window for CSMA/CD (=2a)



Collision Window

- A begins transmitting a frame at time t
- a denotes the one link latency
- The first bit of A's frame arrives at B at time $t + a$
- Suppose an instant before host A's frame arrives, host B begins to transmit its own frame
- B's frame will immediately collide with A's frame and this collision will be detected by host B
- Host B will send the 32-bit jamming sequence
- Host A will not know that the collision occurred until B's frame reaches it, which will happen at $t + 2a$
- Host A must continue to transmit until this time in order to detect the collision
  - Host A must transmit for $2a$ to be sure that it detects all possible collisions

Collision Window

- Consider that a maximally configured Ethernet is 2500 m long, and there may be up to four repeaters between any two hosts, the round trip delay has been determined to be 51.2 μs
  - Which on 10 Mbps Ethernet corresponds to 512 bits
  - 10 Mbps x 51.2 μs = 512 bits ← 64byte
- The other way to look at this situation,
  - We need to limit the Ethernet's maximum latency to a fairly small value (51.2 μs) for the access algorithm to work
    - Hence the maximum length for the Ethernet is on the order of 2500 m.

Exponential Backoff Algorithm

- Once an adaptor has detected a collision, and stopped its transmission, it waits a certain amount of time and tries again.
- Each time the adaptor tries to transmit but fails, it doubles the amount of time it waits before trying again.
- This strategy of doubling the delay interval between each retransmission attempt is known as *Exponential Backoff*.

Exponential Backoff Algorithm

- The adaptor first delays either 0 or 51.2 μs, selected at random.
- If this effort fails, it then waits 0, 51.2, 102.4, 153.6 μs (selected randomly) before trying again;
  - This is k * 51.2 for $k = 0, 1, 2, 3$
- After the third collision, it waits k * 51.2 for $k = 0 \ldots 2^3 - 1$ (again selected at random).
- In general, the algorithm randomly selects a k between 0 and $2^n - 1$ and waits for k * 51.2 μs, where n is the number of collisions experienced so far.

最多冲撞16次，16次都冲撞就放弃。但2^16太大了，控制等待时间最长为2^10-1，即冲撞10次以后就都在0-1023之间挑选等待时间

CSMA/CD Protocol

- Carrier Sense before transmission
- Carrier Sense while transmission
- Collision: Two or more stations transmitting simultaneously
- Backoff: Random delay after collision
- Deference: Defers transmission if channel is sensed busy
- Collision Window (Slot time): Round-trip propagation delay time plus some carrier sense time. In IEEE 802.3, this value is defined to be 51.2 us.

CSMA/CD Collision Handling

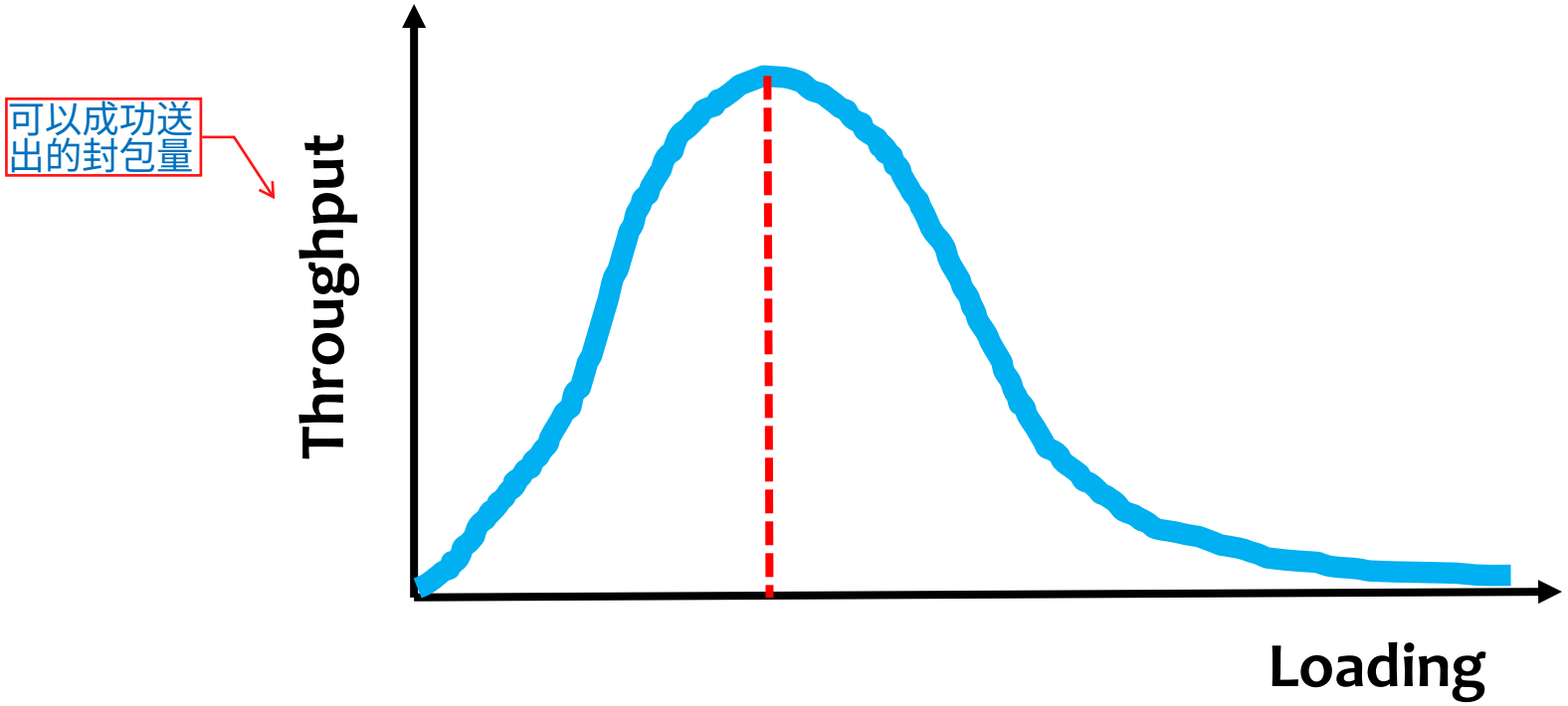
- Collision Signal is generated by Physical layer.
- Jam signal (collision enforcement): To make sure that all stations involved in the collision will detect collision. A pattern of 32 bits.
- Collision backoff and retransmission method (Truncated Binary Exponential Backoff Algorithm, BEBA):
  - n : number of collisions experienced (n <= 16)
  - k : Min (n,10) -- Truncation
  - r : Random delay time (unit: slot time), 0 <= r < 2^k

CSMA/CD Collision Handling

- Slot time = 51.2 us.
- Disadvantage of BEBA:
  - Last-in-First-out effect: Stations with no or few collisions will have a better chance to transmit before stations that have waited longer.

Ethernet Performance

- Ethernets work best under lightly loaded conditions.
- Under heavy loads, too much of the network's capacity is wasted by collisions.

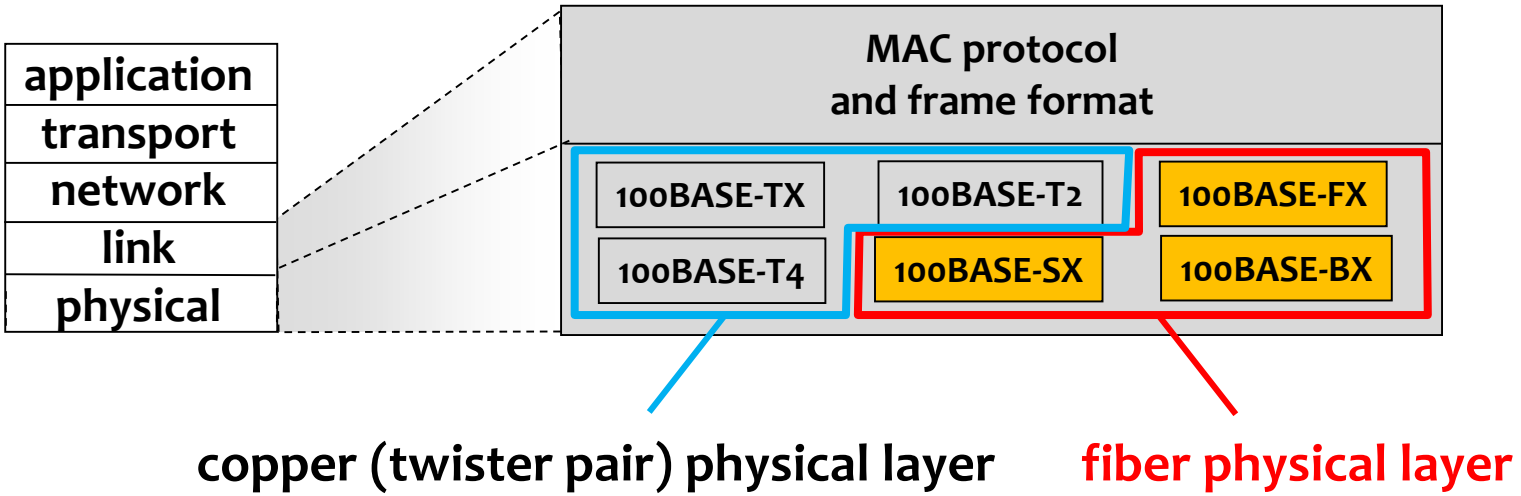


Outline

- Introduction
- Ethernet Topologies
- Ethernet Frame Format
- Ethernet MAC Protocol -- CSMA/CD
- 802.3 Ethernet Standards

802.3 Ethernet Standards: Link & Physical Layers

- *Many* different Ethernet standards
  - common MAC protocol (CSMA/CD) and frame format
  - different speeds: 2 Mbps, 10 Mbps, 100 Mbps, 1Gbps, 10Gbps, 100Gbps
  - different physical layer media: fiber, cable



Summary

- MAC Protocol -- CSMA/CD
- Connection less, unreliable transmission
- Topology from Bus to Star (switches)
- Half-duplex transmission in Bus topology
  - Work best under lightly loaded conditions
  - Too much collision under heavy load
- Full-duplex transmission in Switch topology (point-to-point)
  - No more collisions !!
  - Excellent performance (wired speed)

努力送出，冲撞16次以内都会重送。但还有一种不可靠情况是，没有发生冲撞，但受到干扰，sender不知道，只要没有冲撞就不会重送，但receiver接收到后通过CRC检查发现封包错误，就会丢弃。所以不可靠

线路专用，还是跑CSMA/CD，但不会侦测到冲撞