



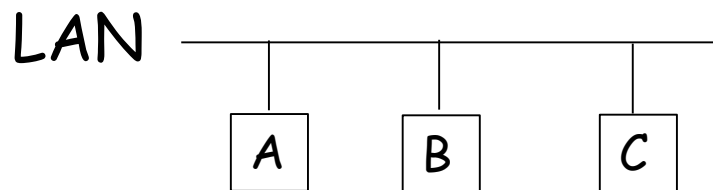
第四单元 介质访问控制子层

- 概述
- 以太网
- 透明网桥算法
- 生成树算法
- 虚拟局域网
- 交换机
- 令牌环网



概述

- 多路访问链路（广播链路）采用共享介质连接所有站点。发送站点通过广播方式发送数据并占用整个共享介质的带宽。由于每个站点只需要一条线接入网络就可以访问所有站点，这种网络一般安装简单，价格便宜。局域网(Local Area Network, LAN)都是使用这种链路。

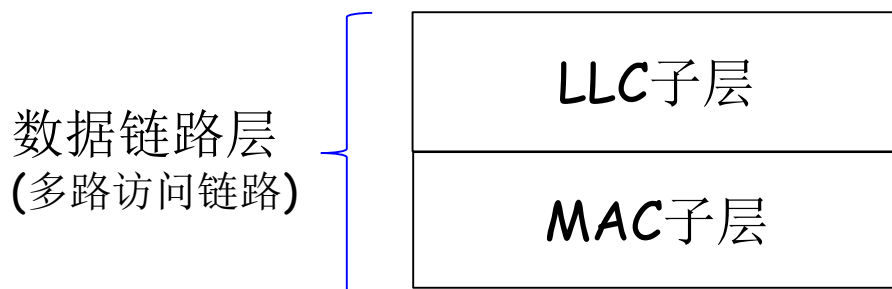


- 在多路访问链路中多个站点同时发送数据，则会产生冲突。这种问题是点到点链路没有的，因此，需要重新考虑数据链路层的功能设计。

如何解决冲突问题？

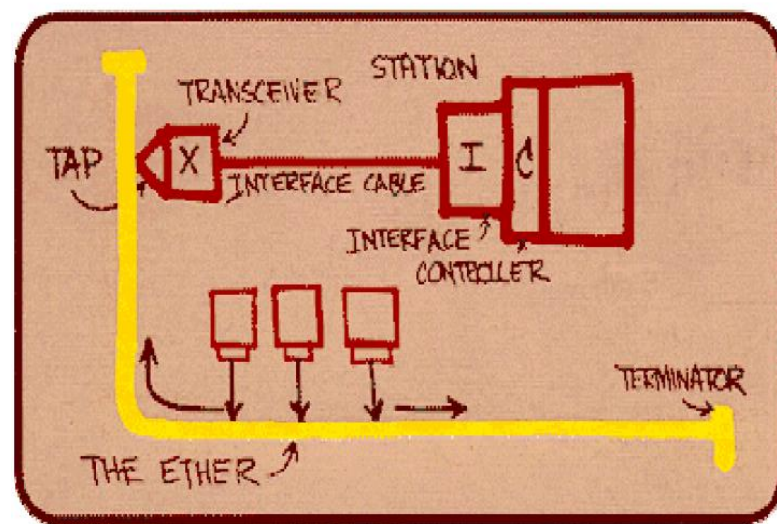
Random Access Protocol
Take Turns

- ❑ **OSI**把为解决冲突问题而专门在数据链路层划分出的一个子层，就是介质访问控制子层(**Media Access Control, MAC子层**)，其功能是控制和协调所有站点对共享介质的访问，以避免或减少冲突。
- ❑ 因为**MAC**子层不提供可靠的数据传输，所以，在**MAC**子层之上又定义了一个子层，逻辑链路控制子层(**Logic Link Control, LLC**)，用来为上层协议提供服务：
 - (1) **LLC1**提供无确认无连接服务；
 - (2) **LLC2**提供有确认面向连接的服务；
 - (3) **LLC3**提供有确认无连接的服务。



以太网

- ❑ 1970年, 夏威夷大学 ALOHA, 1971年, slotted ALOHA (无线)
- ❑ 1973年, Xerox公司Palo Alto研究中心(PARC)的Metcalfe成功研制了以太网(2.94Mbps, 无源基带), 1975年7月 Metcalfe和Boggs发表了关于以太网的论文, 并申请了专利。
- ❑ 1980年, DEC, Intel, Xerox提出了DIXv1 (10Mbps), 1982年, DIXv2。
- ❑ 1985年, IEEE提出802.3标准(10Mbps), 即以太网协议。3COM网, Novell网。
- ❑ 1995年发布了100Mbps的快速以太网标准IEEE 802.3u。
- ❑ 1998年, 发布了1000Mbps的吉比特以太网标准IEEE 802.3ab。

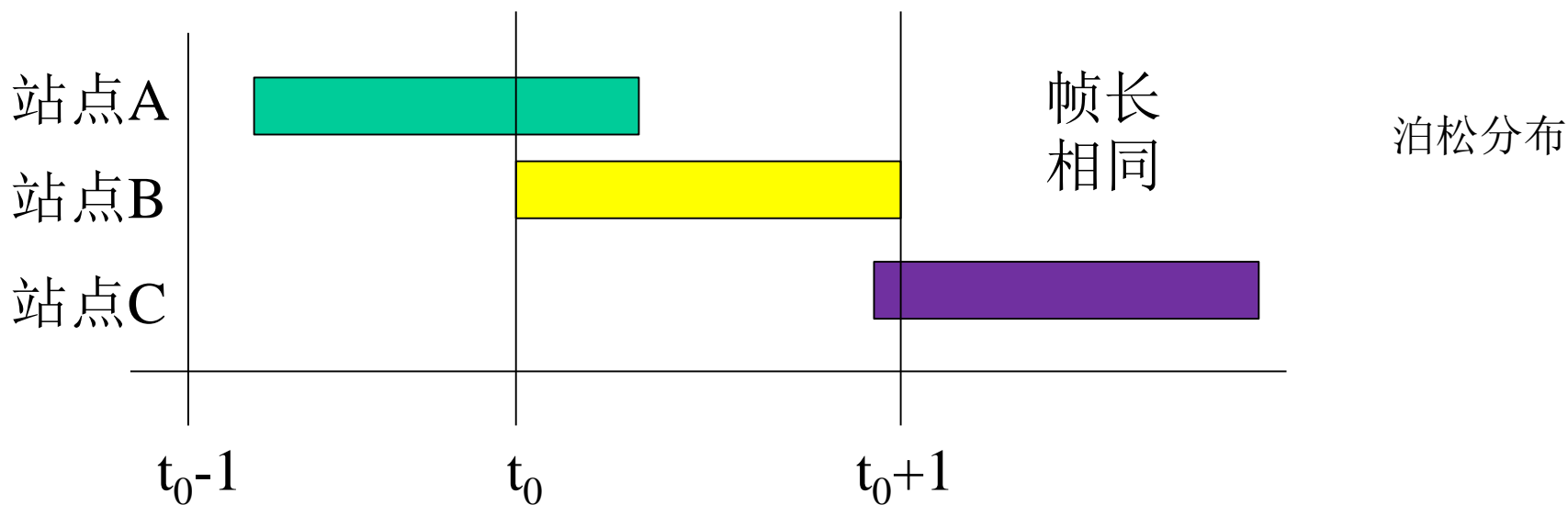


Bob Metcalfe的以太网设计图

IEEE(Institute of Electrical and Electronics Engineers)为电气和电子工程师协会, 念成I triple E (三个E)

纯ALOHA

想发送就发送，超时未收到确认则认为发生了冲突。



假设一个站点在任何时刻的发送概率是 p ，如果有 N 个站点有很多帧要发送，则最大效率 E 为：

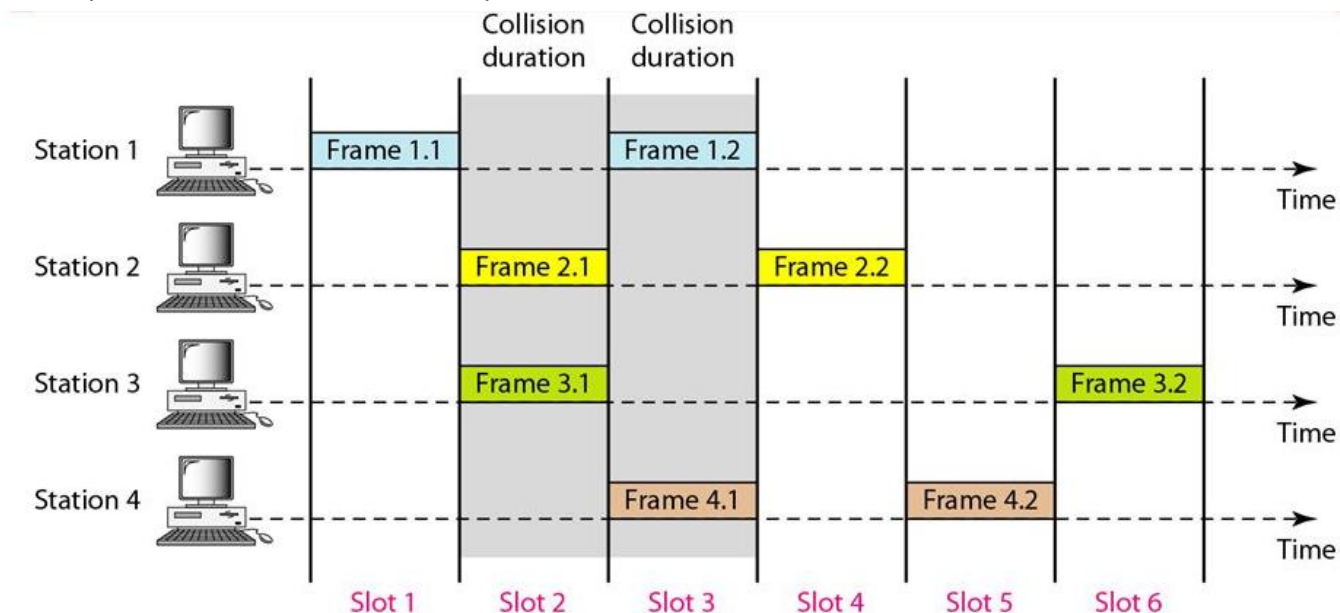
$$N * p * (1-p)^{N-1} * (1-p)^{N-1} \quad \text{最大 } E \approx 1/2e \approx 18.4\%$$

该站点在 $[t_0, t_0+1]$ 发送的概率 其它站点在 $[t_0-1, t_0]$ 不发送的概率 其它站点在 $[t_0, t_0+1]$ 不发送的概率

当 p 为令表达式最大， $N \rightarrow \infty$ 时

分槽ALOHA

把时间分为长度相同的时槽(slot)，每个站点只在时槽开始时发送。信道空，立即以概率 p 发送，以概率 $1-p$ 延迟一个时间槽；信道忙，延迟一个时间槽。



如果 N 个站点要发送帧，它们在每个时槽发送的概率为 p ，则最大效率 E 为

$$N * p * (1-p)^{N-1} \quad \text{最大 } E \approx 1/e \approx 36.8\%$$

当 p 为令表达式最大， $N \rightarrow \infty$ 时

CSMA

(Carrier Sense Multiple Access)

发送前先监听信道

信道空，立即发送；信道忙，持续监听，转前面。

1-persistent CSMA
(以太网)

信道空，发送；信道忙，延迟一段随机长度的时间，转前面。

non-persistent CSMA

信道空，立即以概率 p 发送，以概率 $1-p$ 延迟一个时间槽；
信道忙，延迟一个时间槽，转前面。

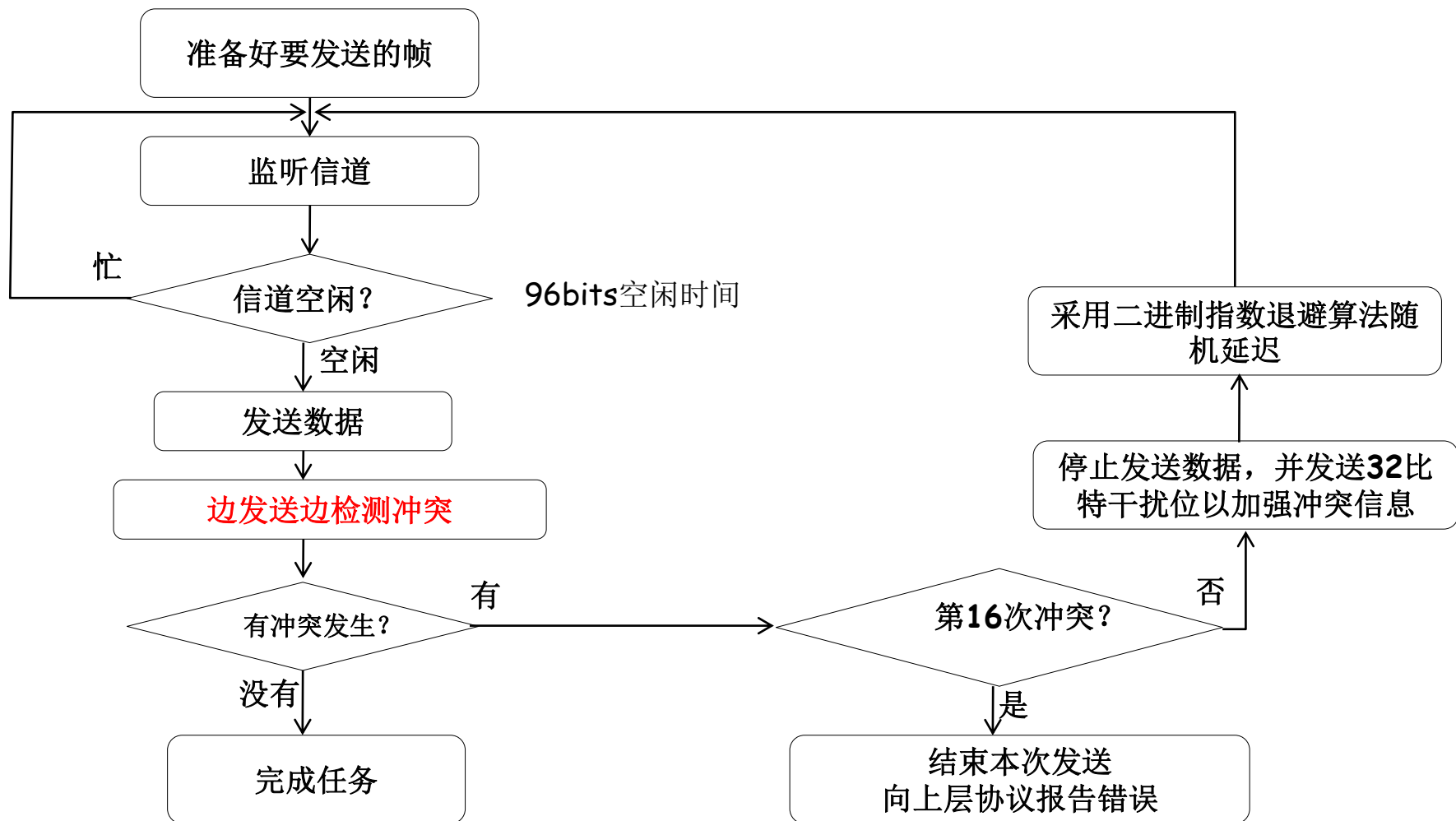
p-persistent CSMA
(分槽 ALOHA)

以太网的MAC层协议

□ 发送帧的方法--CSMA/CD(Carrier Sense Multiple Access With Collision Detection)协议:

- (1) 发送数据帧之前先监听信道。如果信道空闲，立即发送。如果信道忙，则持续监听，直到信道空闲，立即发送。
- (2) 边发送边检测冲突。如果发送完毕都没有检测到冲突，则发送成功。
- (3) 如果发送时检测到冲突，则停止发送，在发送32位干扰位(jamming signal)以加强冲突信号后采用二进制指数退避算法随机延迟一段时间，然后转(1)。

❑ CSMA/CD协议的流程图



❑ 二进制指数退避算法(binary exponential backoff)

第一次冲突 从0个或1个时间片中随机选择一个进行延迟

第二次冲突 从0, 1, 2, 3个时间片中随机选择一个

.

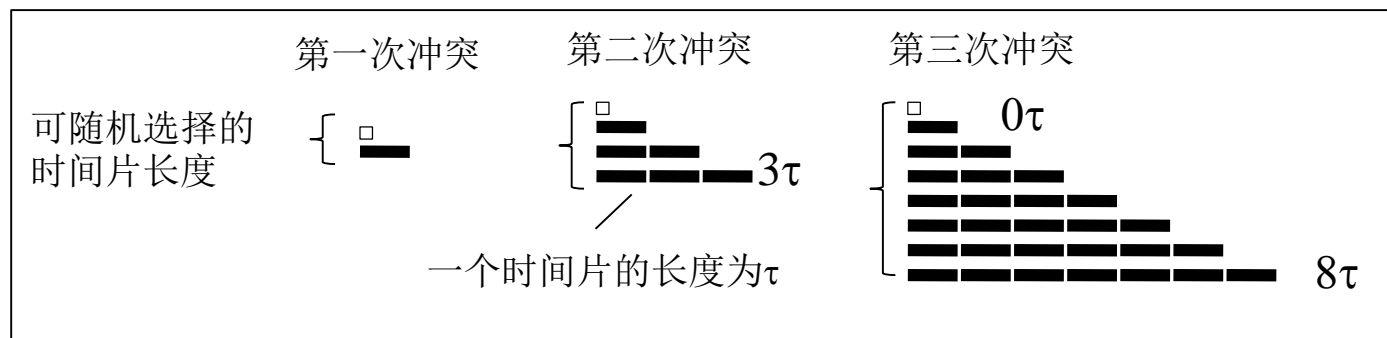
.

.

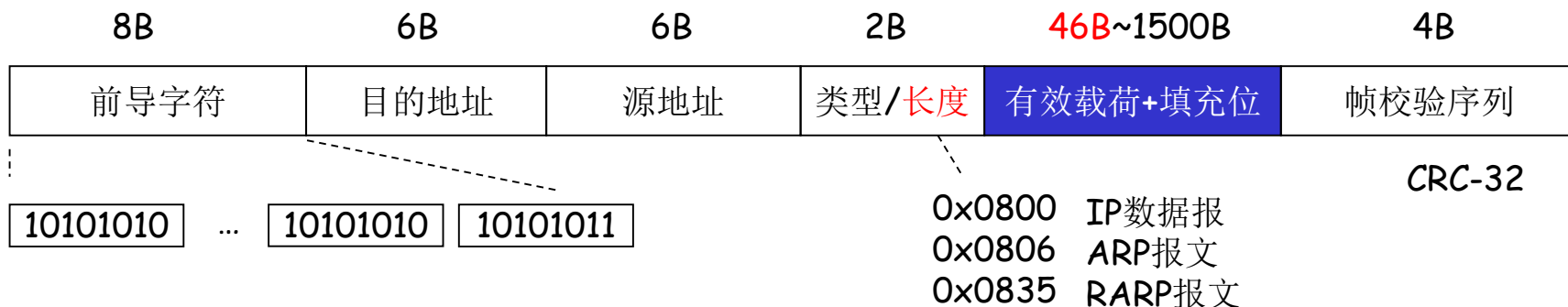
第 i 次冲突 从0, 1, ..., $2^j - 1$ 个时间片随机选择一个. $i < 16$. $j = \min(i, 10)$ 。

前十次冲突后可选时间片数量每次加倍, 后五次冲突后可选时间片数量不变, 所以也称为**截止式(truncated)二进制指数退避算法**。

其中, 时间片 τ 的长度为**512**比特的时间, **10Mbps**的以太网为**51.2 μs** 。



802.3的MAC帧格式



802.3的MAC帧格式

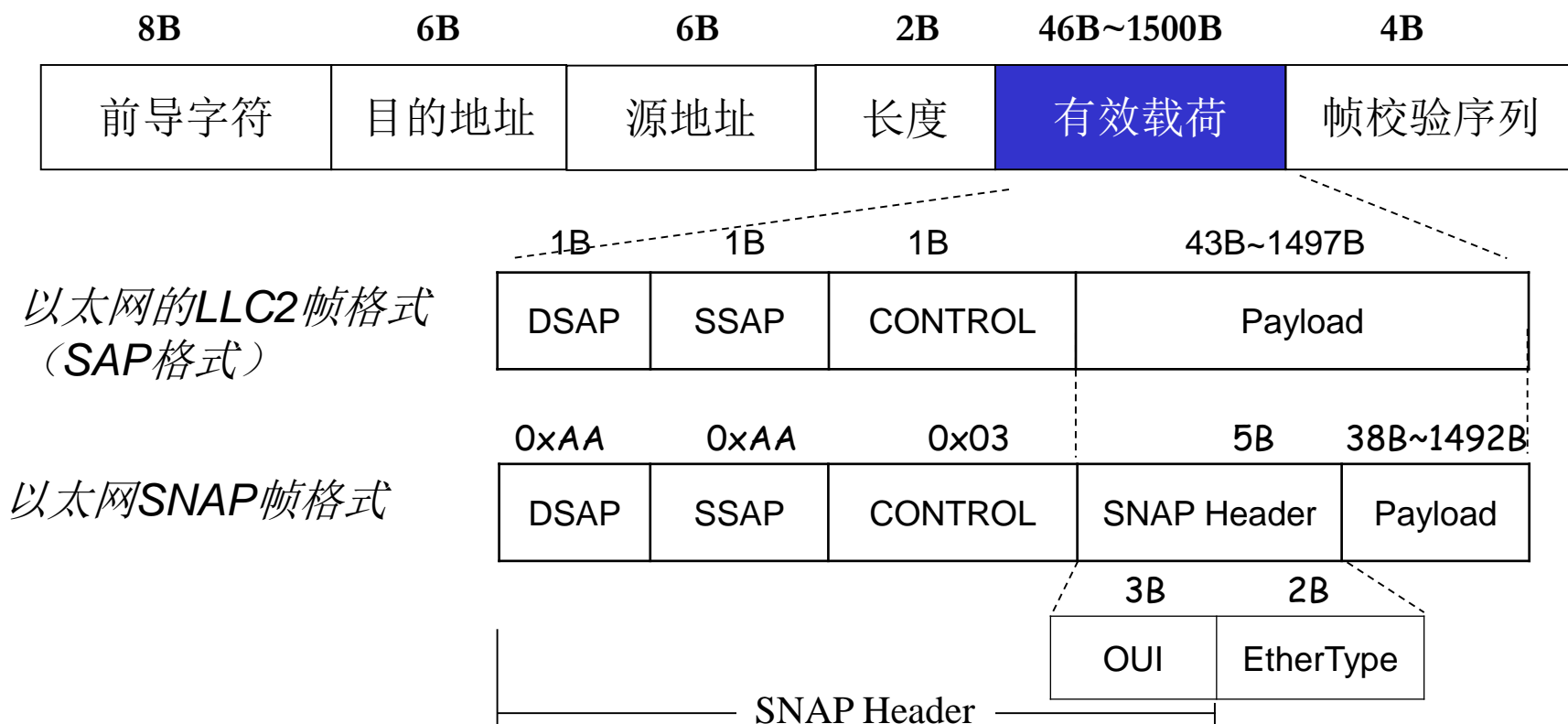
- 前导字符(Preamble): 同步字符(7B)和起始定界符(Start of Frame Delimiter)(1B)。
- 有效载荷(Payload): 用户数据。不足46字节时加入填充字节(任何字节)至46字节。
- 类型/长度字段(Type/Length): 指明上层协议(>1500)或有效载荷的长度(≤1500)。
- 帧校验序列(Frame Check Sequence): 对目的地址、源地址、类型/长度和有效载荷(加填充位)字段进行CRC-32校验。

[类型/长度字段]

- | | |
|-----------------|-----------------------|
| DIXv2帧 | ——> 类型。≥1536 (0x0600) |
| 802.3帧(原) | ——> 长度。(SNAP格式) |
| 802.3帧(1997年修订) | ——> 类型/长度。 |

用于
长度
后协
议?

● 802.3 帧为长度字段时的扩展格式



- DSAP, SSAP: 目的SAP和源SAP, 用来指明Payload的协议类型。例如: 0x06和0xAA指明LLC2帧的payload为IP协议和SNAP协议。
- Control: 与HDLC的Control字段功能相同。0x03表示无编号帧。
- OUI: 000000表示是通用以太网类型, 否则为厂商号。
- EtherType: 以太类型, 其值与含义和前面类型字段完全相同。
- SNAP: 子网访问协议(Subnetwork Access Protocol)。802.11采用了SNAP协议。

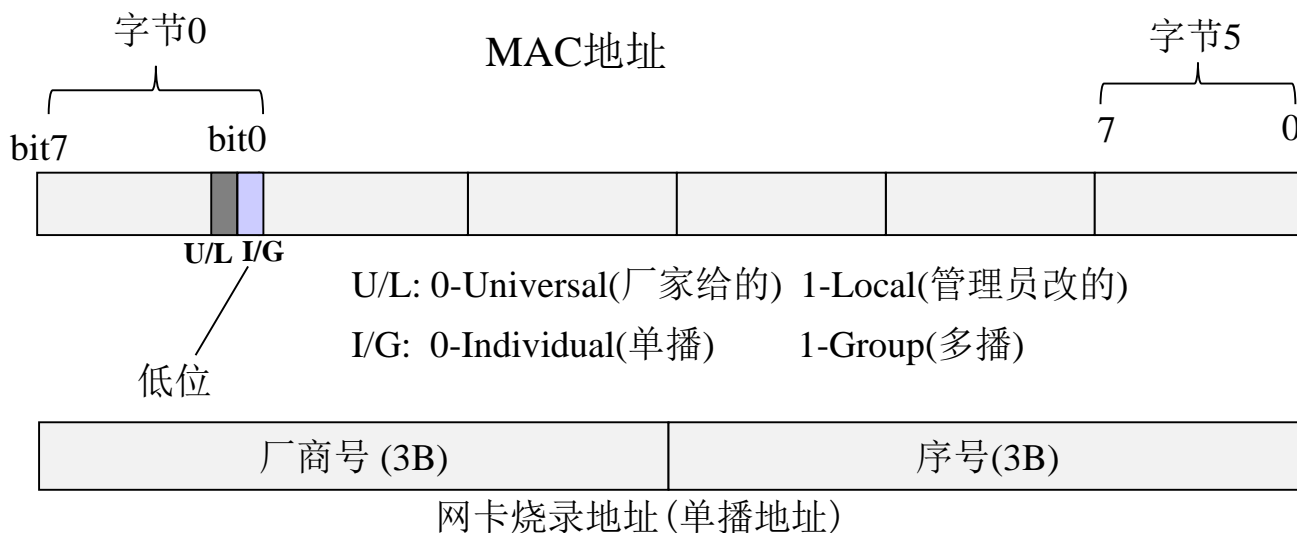
- 源地址和目标地址(6B)

- 源地址一般为发送者的单播地址。目标地址可以是接收者的单播地址，也可以是多播地址和广播地址。

- **单播地址**：全球唯一。每个网卡(或接口)一个，最高字节的最低有效位为0。如：06-01-02-01-2C-4B。也称为网卡地址，烧录地址(Burned-In-Address, BIA)，MAC地址，硬件地址，物理地址。

多播地址的字节0的第0位为1，并且地址非全1。如：01-00-5E-20-01-4B。

广播地址的48位全为1。



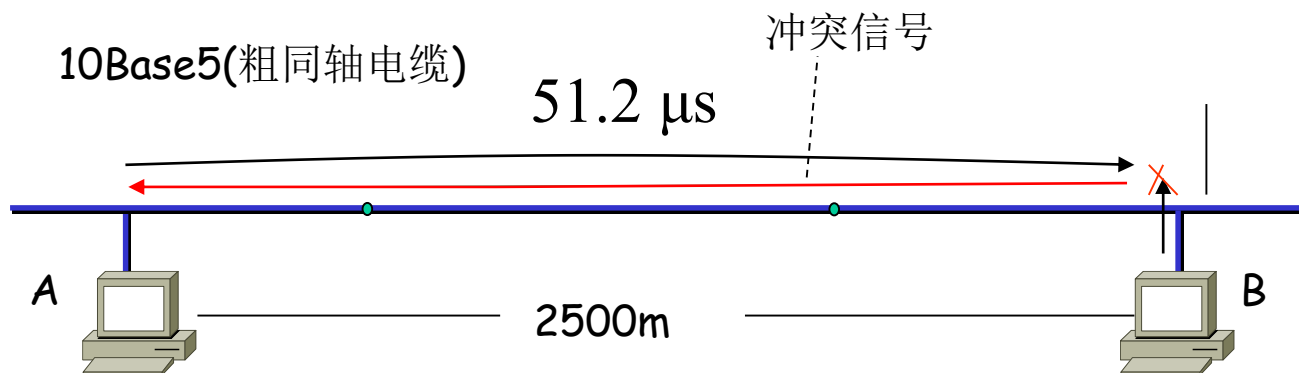
IEEE为每个厂商分配的唯一**厂商号**(Organization Unique Identifier, OUI)。例如，00-00-0C和00-AA-00分别为思科和Intel的厂商号。厂商再为其生产的每个网卡或接口分配**序号**。

从字节0开始发送，每个字节从低位(bit0)开始发送(802.3和802.4),802.5和802.6从高位(bit7)开始发送。

接收帧的方法

- (1) 以太网站点(网卡)会缓存所有的帧
- (2) 如果缓存的帧有错 (长度错误, **CRC**错等), 则丢弃它。
- (3) 如果缓存的帧的目的地址为**单播地址**并且与接收该帧的网卡的**MAC**地址一致, 则接收它。如果目的地址为**多播地址**并且为网卡预设的多播地址之一, 或者为**广播地址**, 也接收它。其它情况则丢弃它。
- (4) 如果把网卡设置为**混杂模式**则会接收所有无错的帧。

最短帧问题



- 以太网(10M bps)相距最远的两个站点（上图站点A和B)之间的 信号往返时间为 $51.2\ \mu\text{s}$ 。
- 假如站点A发送的数据在快到达站点B时与其发送的数据冲突，因为发送站点只在发送时才检测冲突，为了检测到返回的冲突信号，则要求站点A此时还在发送，故帧长至少为512b(64B)。
- 64B也称为争用窗口(**contention window**)长度。

CSMA/CD 的效率分析

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

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

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

以太网(802.3)的物理层(1)

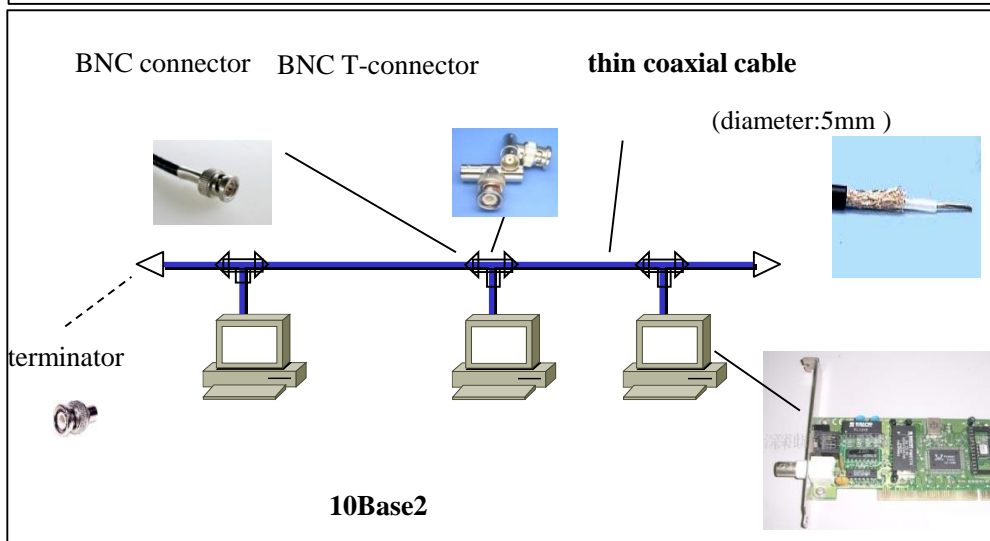
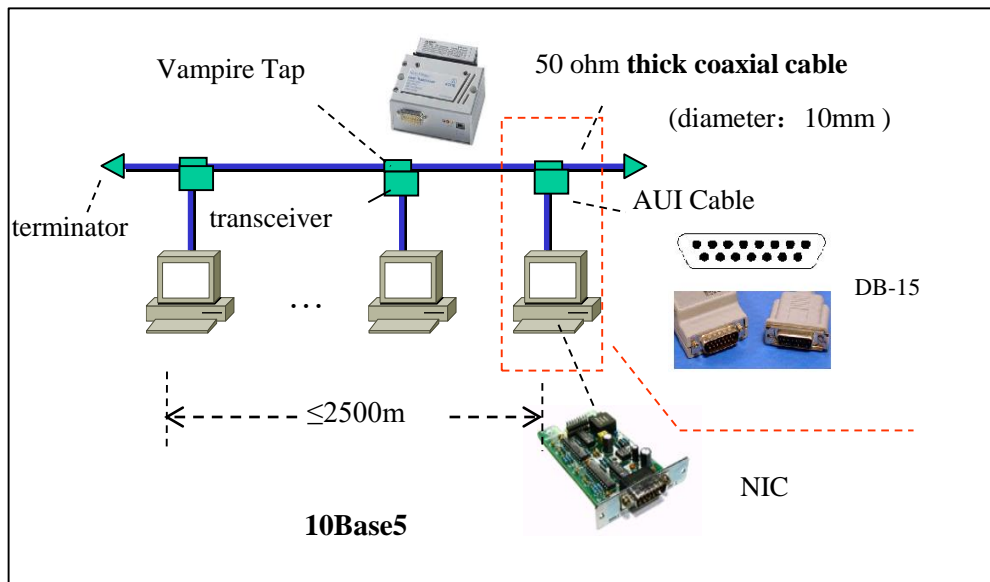
□ 物理层规范

名称	传输技术	传输介质	网段最大距离	每段最多节点数	IEEE标准及发布时间
10Base5	基带	50欧姆粗同轴电缆	500m	100	802.3, 1980
10Base2	基带	50欧姆细同轴电缆	185m	30	802.3a, 1985
10BaseT	基带	Cat-3,Cat-5 UTP	100m	1024	802.3i, 1990
10BaseF	基带	光纤, 850nm波长	2000m	1024	802.3j, 1993

- **传输方法:** 均使用异步传输, 即信道空闲时以太网设备不任何发送信号。
- **编码方法:** 采用曼彻斯特编码。
- **命名规则:** 10BaseT ---- 10表示10Mbps, Base表示基带传输, T表示双绞线; 10base2的2表示最大距离200m。

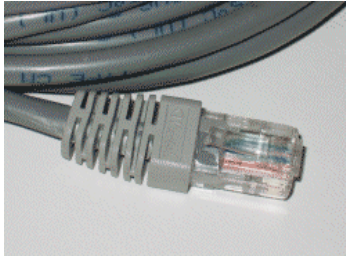
早期的以太网为10Broad36

以太网(802.3)的物理层(2)



以太网(802.3)的物理层(3)

10BaseT



RJ-45

Pin Position

8
7
6
5
4
3
2
1

线对 2
线对 3
线对 1
线对 4

引脚
1 = 绿/白
2 = 绿
3 = 橙/白
4 = 蓝/白
5 = 蓝
6 = 橙/白
7 = 棕/白
8 = 棕

T568A

T568B

引脚
1 = 橙/白
2 = 橙
3 = 绿/白
4 = 蓝
5 = 蓝/白
6 = 绿
7 = 棕/白
8 = 棕

TIA/EIA-T568-A 和 TIA/EIA-T568-B 引脚线

UTP

1,2 3,6

直通连接法

Hub/Switch

UTP

1,2 3,6

交叉线连接法

hub

hub

T568A: 1(TD+), 2(TD-), 3(RD+), 6(RD-)

T568B: 1(RD+), 2(RD-), 3(TD+), 6(TD-)

接线标准	1	2	3	4	5	6	7	8
568A	白绿	绿	白橙	蓝	白蓝	橙	白棕	棕
568B	白橙	橙	白绿	蓝	白蓝	绿	白棕	棕

该种较常用


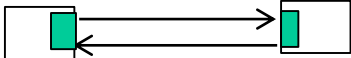
ST connector

SC connector

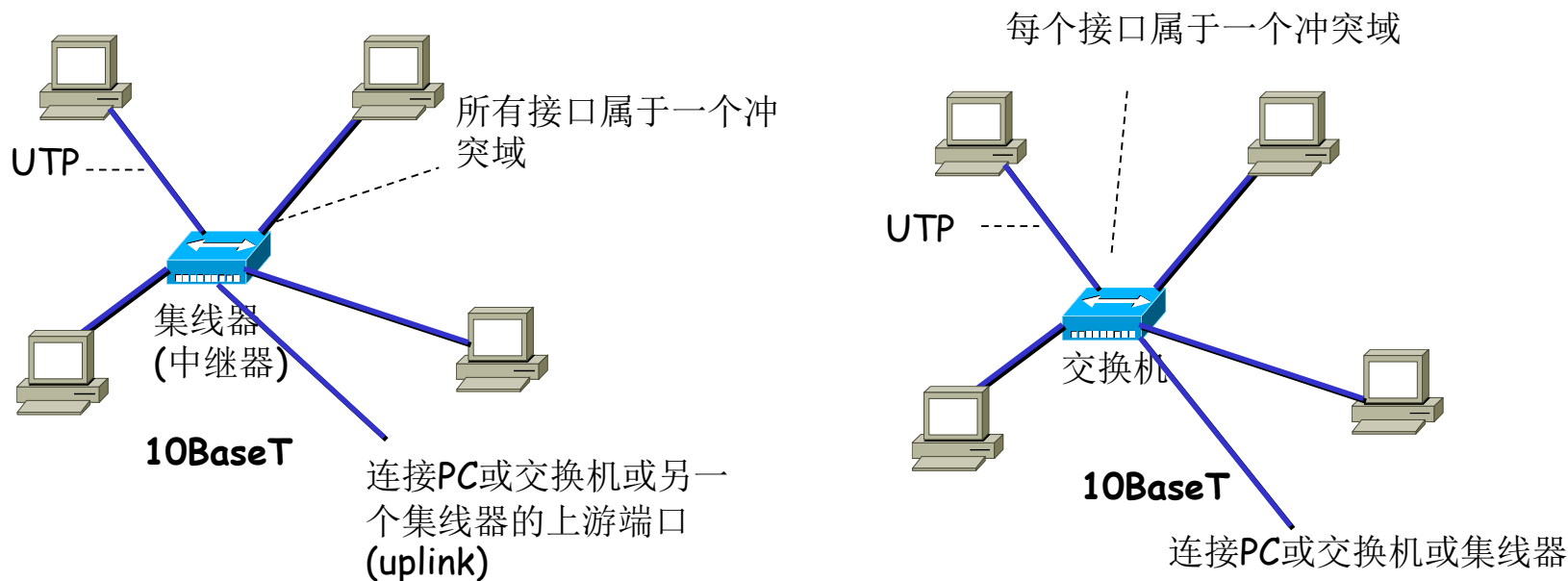
10BaseF

SC connector

ST connector

以太网(802.3)的物理层(4)



- 如果通过两个接口同时发送数据会产生冲突，则这两个接口属于同一个冲突域 (collision domain)。一个广播帧可以到达的所有接口属于同一个广播域。
- 属于同一个冲突域的以太网部分称为网段(segment)。

快速以太网概述

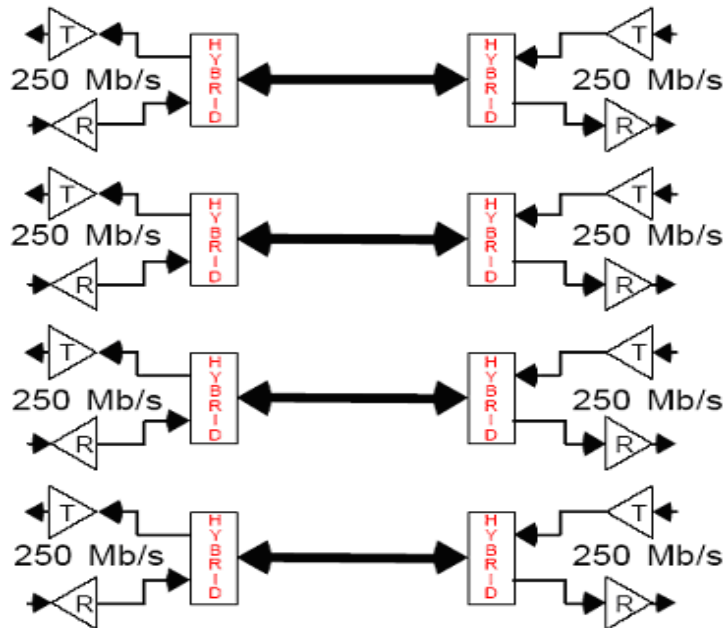
- 1995年，IEEE发布了100Mbps的快速以太网的标准IEEE 802.3u。
- 与802.3相比，快速以太网只是把传输速率提高到100Mbps, 其它均保持不变：
 - 1) **MAC**子层的协议不变：**CSMA/CD**协议不变，帧格式不变。
 - 2) 最大距离改为100m (10base5的最大距离为2500m)。
 - 3) 帧间空隙隔依然为96b，即0.96 μ s。

	电缆	接头	编码方案	距离	传输方向
100Base-TX	Cat 5/Cat 5e UTP(两对) 100-ohm STP(两对)	RJ-45/DB-9	4B/5B MLT-3	100m	全双工
100Base-T4	Cat 3/Cat 4/Cat 5 UTP(四对) (思科不支持)	RJ-45	8B/6T 一种3级编码	100m	半双工
100Base-FX	2个多模光纤	SC或ST	4B/5B NRZI	2000m	全双工

千兆以太网

(802.3ab : 1000Base-T)

- 使用5类或以上UTP的4对双绞线。每对线每个方向125M波特率和250Mbps。采用了复杂的编码方式8B1Q4和4D-PAM5，并利用混波电路将发送信号和接收信号耦合在同一线缆中传输，使其互不干扰，从而实现了双向传输。
- 可以采用全双工(交换机)或半双工(集线器)传输。

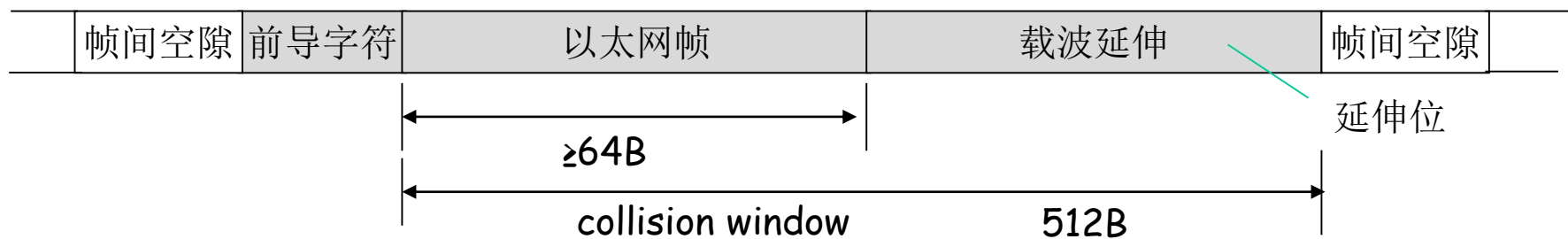


编码和调制方法:

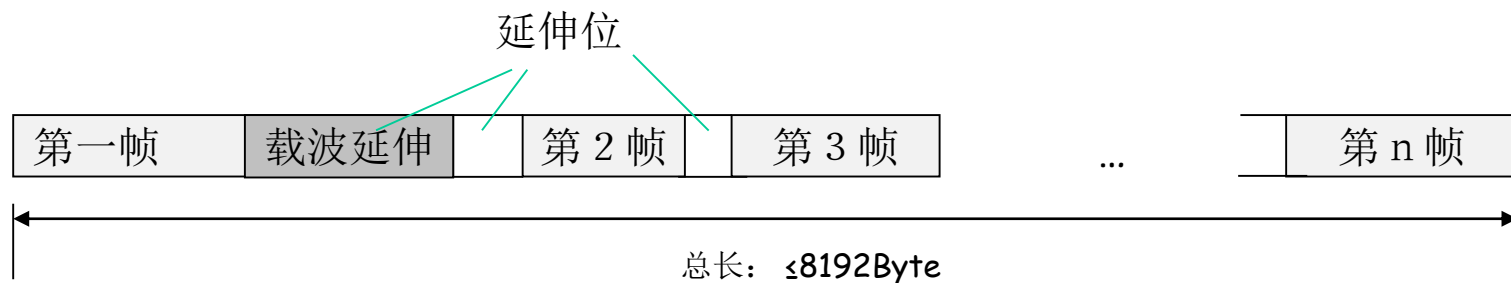
8B1Q4(8-bit 1-Quinary Quarter)
4D-PAM5(4-dimensional 5-level
Pulse Amplitude Modulation)。

- 和快速以太网一样，802.3ab除了把传输速率提高1000Mbps，其它不变。由于1000Base-T的速度提高了100倍，故半双工的冲突域也要减少为1/100，即25m。这个距离很难接受，怎么办？可以采用载波延伸技术延长距离到200m。

(1) 载波延伸(carrier extension): 通过附加填充字节，令帧长至少为512B，网络半径可以延长到200m。填充字节称为延伸位(extension bits)。



(2) 帧突发(frame bursting): 发送很多短帧时采用载波延伸浪费很大。可以采用帧突发进行改进，即一次传送多个短帧。第一帧小于512B时进行载波延伸，后面的帧直接发出，不用加载波延伸。每一帧之间有一个小的间隔，填入延伸位。



万兆以太网

(10 Gigabit Ethernet)

□ 特性

- (1) 保持帧格式不变
- (2) 光纤或双绞线, 全双工
- (3) 无冲突, 不使用CSMA/CD算法

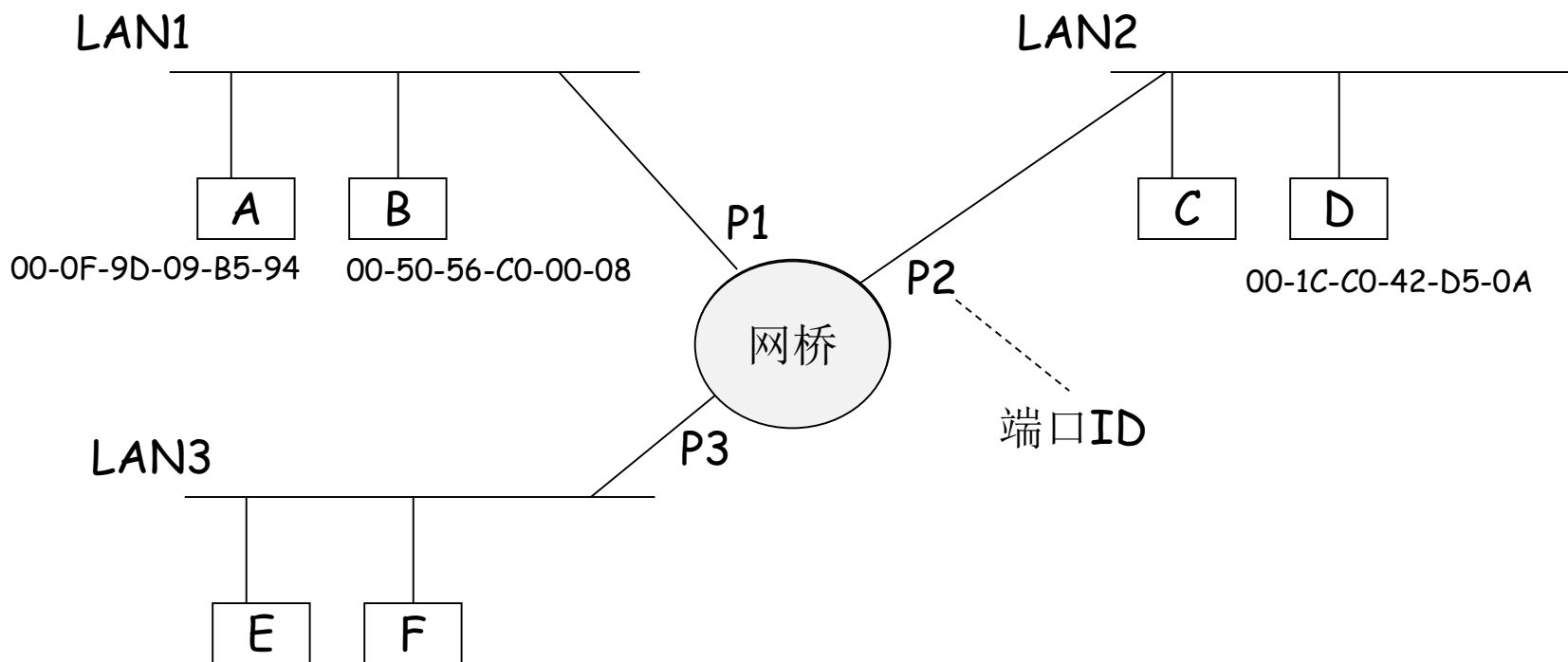
□ 标准

- 802.3ae-2002 (fiber 10GBase-SR, -LR, -ER and -LX4 PMDs)
802.3ak-2004 (10GBase-CX4 copper twin-ax InfiniBand type cable),
802.3an-2006 (10GBASE-T copper twisted pair)
802.3ap-2007 (copper backplane 10GBase-KR and -KX4 PMDs) and
802.3aq-2006 (fiber 10GBase-LRM PMD with enhanced equalization).
PMDs-physical medium-dependent sublayers
- The 802.3ae-2002 and 802.3ak-2004 amendments were consolidated into the IEEE 802.3-2005 standard. IEEE 802.3-2005 and the other amendments were consolidated into IEEE Std 802.3-2008.

http://en.wikipedia.org/wiki/10-gigabit_Ethernet

透明网桥 (1)

□ 扩展局域网



- 用网桥(bridge)连接若干局域网(LAN)可以建造一个更大的局域网, 称为桥接的局域网(bridged LAN) 或 扩展局域网(extended LAN)。
- 原来的局域网就成为该扩展局域网的一部分, 称为该扩展局域网的一个网段(Segment)。

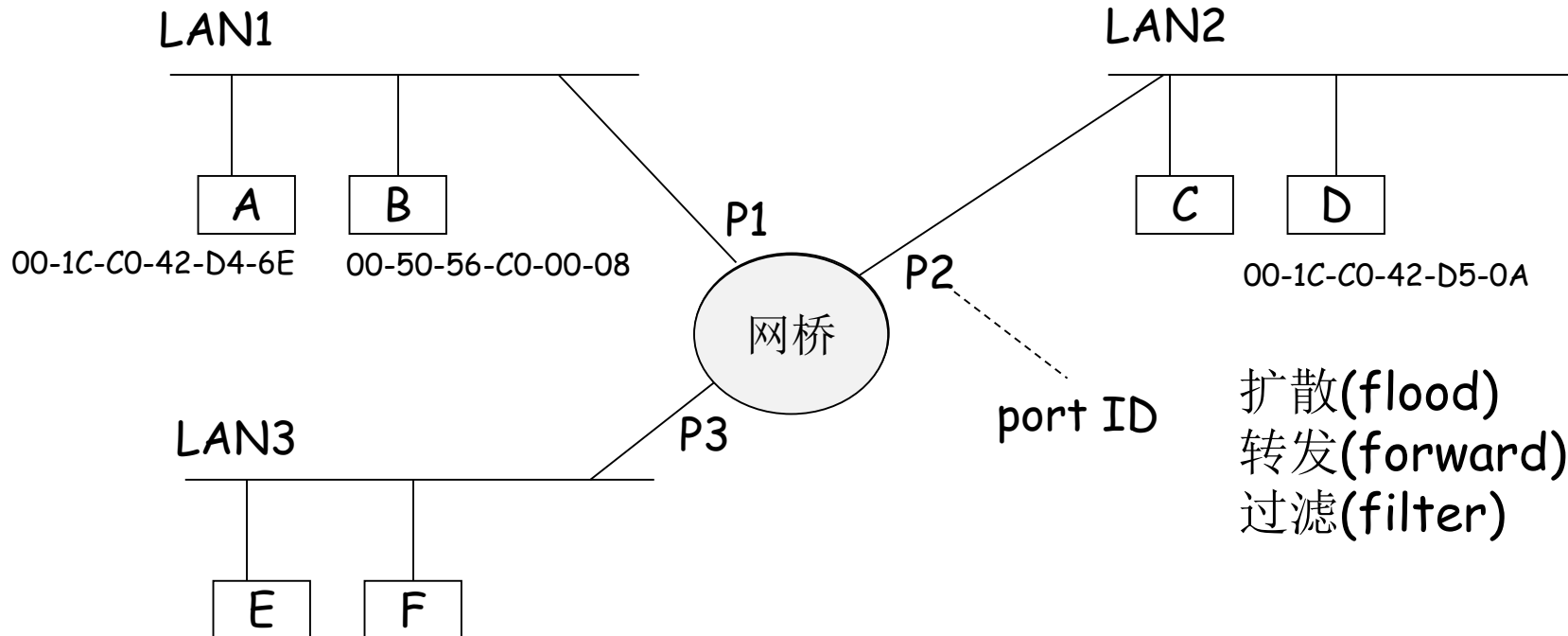
透明网桥 (2)

□ MAC地址表

Time-To-Live(生存期)

目的MAC 地址	转发端口	TTL(sec)
00-50-56-C0-00-08	P1	60
00-1C-C0-42-D5-0A	P2	120

端口(port)



透明网桥 (3)

当网桥收到一个单播帧，它会用该帧的目的地址查询**MAC**地址表

1. 如果没有查到，则扩散(**flood**)该帧。
2. 如果查到，则看查到的端口是否为收到该帧的端口，
如果是，则丢弃该帧(**filter**)，
否则，把该帧从查到的端口发送出去(**forward**)。

当网桥收到一个多播或广播帧，它会直接扩散(**flood**)该帧。

扩散(flood)就是网桥把收到的帧转发到除了该帧的接收端口之外的所有其它端口。

IEEE 802.1D 透明网桥和生成树算法

透明网桥(4)

□ 自学习

- **MAC**地址表初始为空。
- 网桥从端口接收所有的帧，并把接收到的帧的源地址和接收端口记录到**MAC**地址表中：如果该源地址在**MAC**地址表中不存在，则增加一个新记录，并启动超时定时器；如果存在，则更新接口并重置超时定时器。
- 网桥会自动删除超时的记录。

为什么叫透明网桥(Transparent Bridges)?

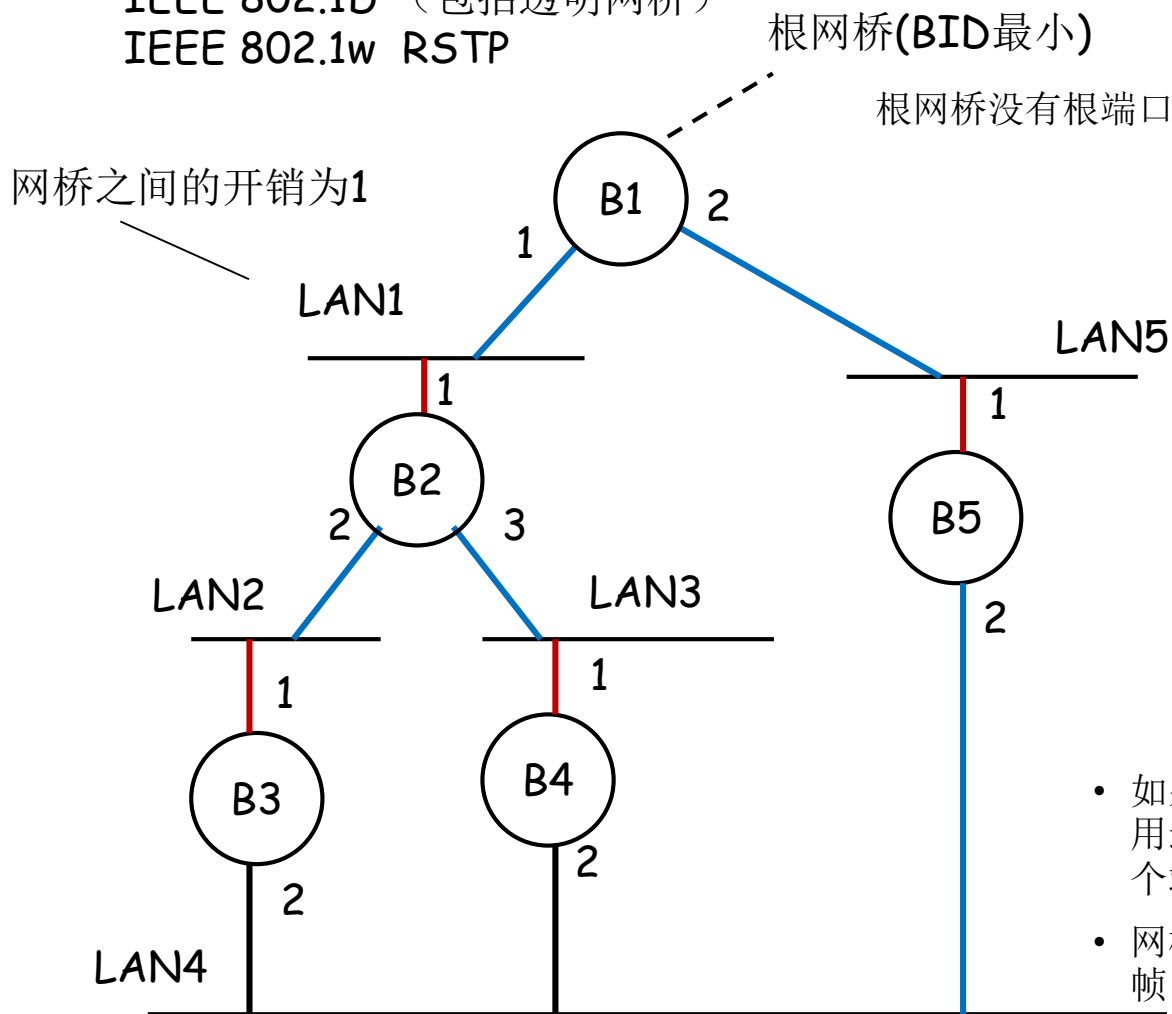
生成树协议

(Spanning Tree Protocol)

IEEE 802.1D (包括透明网桥)

IEEE 802.1w RSTP

配置消息BPDU (Bridge PDU, 以太网的多播帧):
<当前根BID, 到根的距离, 发送BID, 发送端口>



网桥ID(Bridge ID, BID): 优先权(2B)和网桥的MAC地址(6B), 这里用网桥下标表示。

端口PID(Port ID, PID): 优先权(1B)和序号(1B), 这里直接用序号表示。

— 根端口(网桥上离根最近的端口)

— 指定网桥 (网段上离根最近的网桥)。指定端口: 指定网桥上与网段连接的端口

— 阻塞端口(网桥上既非根端口又非指定端口)

- 如果到根网桥的最短路径有多条, 可以采用最短路径上的下一个网桥ID以及下一个端口ID用于打破平衡 (取更小的)。
- 网桥只在根端口和指定端口之间转发数据帧。

STP链路(接口)权重

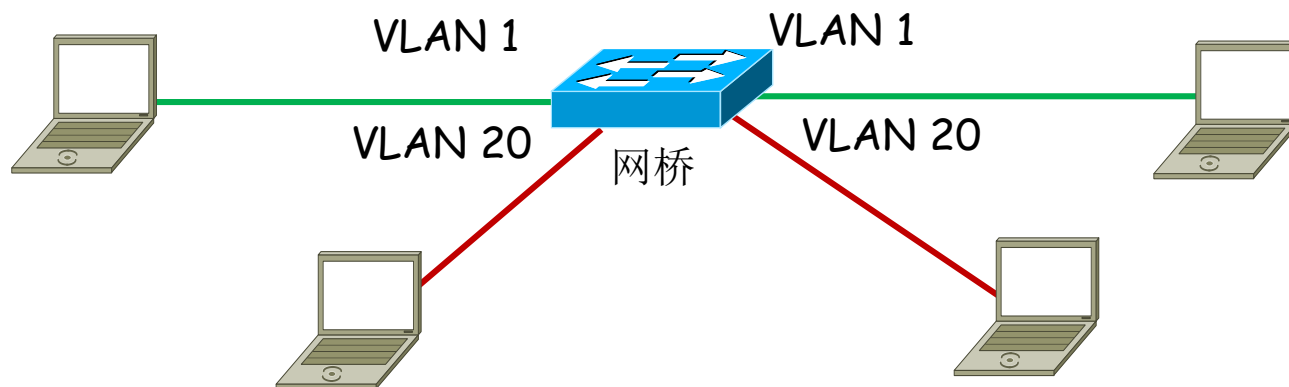
Port Speed	Cost Value	Range
10Gbit/s	2	1~65535
1Gbit/s	4	1~65535
100Mbit/s	19	1~65535
10Mbit/s	100	1~65535

the Short Method
(以前使用)

Port Speed	Recommended Value	Recommended Range
10Gbit/s	2000	200~20000
1Gbit/s	20000	2000~200000
100Mbit/s	200000	20000~2000000
10Mbit/s	2000000	200000~20000000

the Long Method
(现在常用)

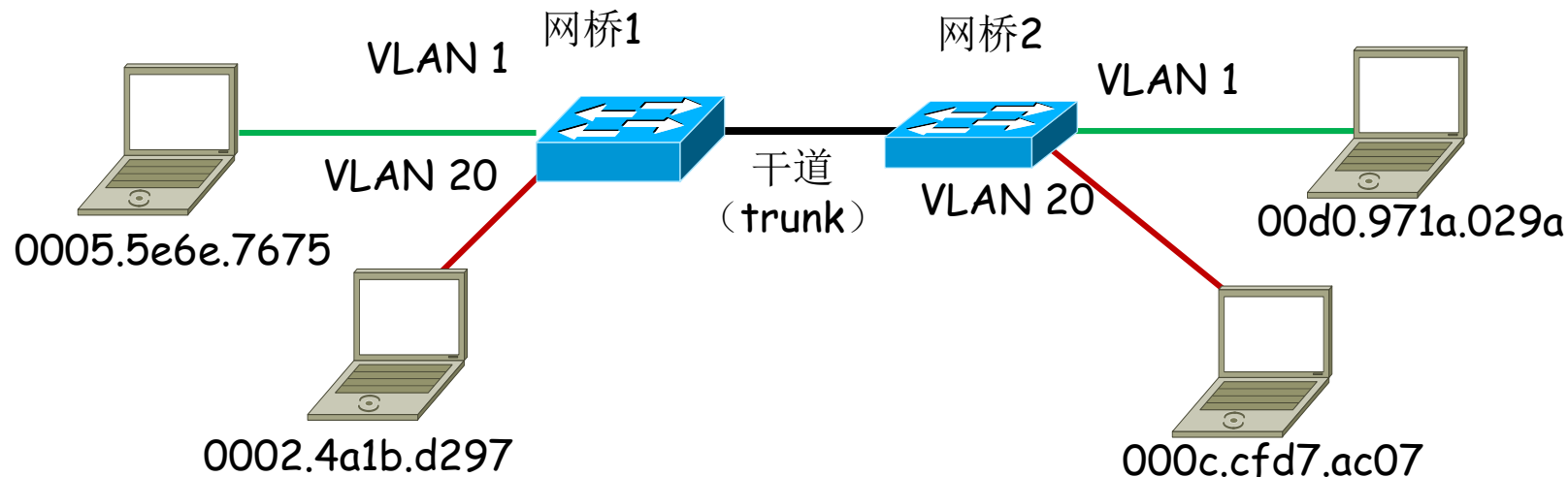
虚拟局域网(1)



如果网桥只在具有相同颜色的端口(Port) 之间转发帧，就会把原来的局域网分割成多个相互隔离的局域网，称为虚拟局域网(Virtual LAN, VLAN)。

所谓的颜色其实就是**VLAN ID**，是由管理员为每个端口配置的一个标识。具有相同的**VLAN ID**的端口处于同一个**VLAN**，端口的默认VLAN为**VLAN 1**。

虚拟局域网(2)

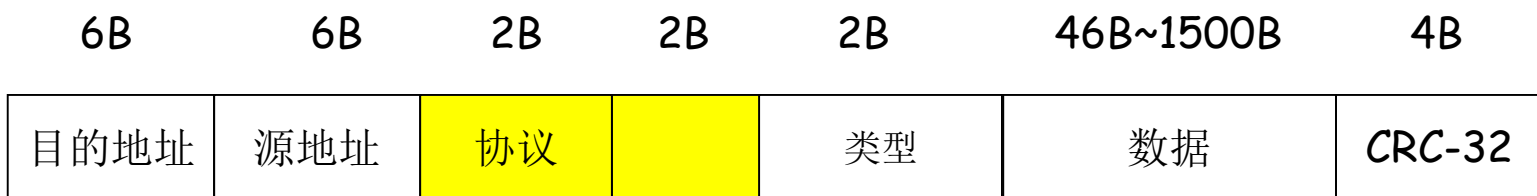


一个VLAN的帧只能转发到属于同一个VLAN的端口或者干道端口。只有发往干道端口的帧才需要加上VLAN ID。

从干道收到的帧中如果没有VLAN ID，则认为是本征VLAN(Native VLAN)，默认为VLAN 1。发往干道的Native VLAN的帧不加VLAN ID。

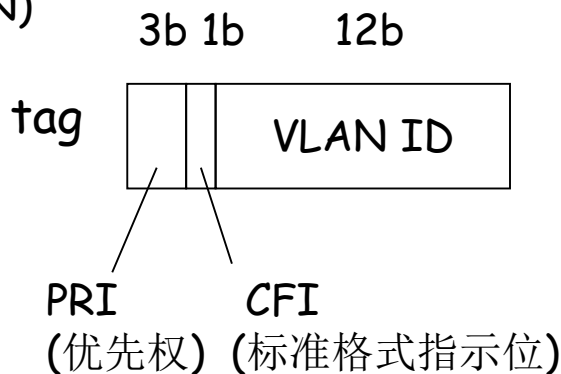
虚拟局域网(3)

IEEE 802.1Q 的帧



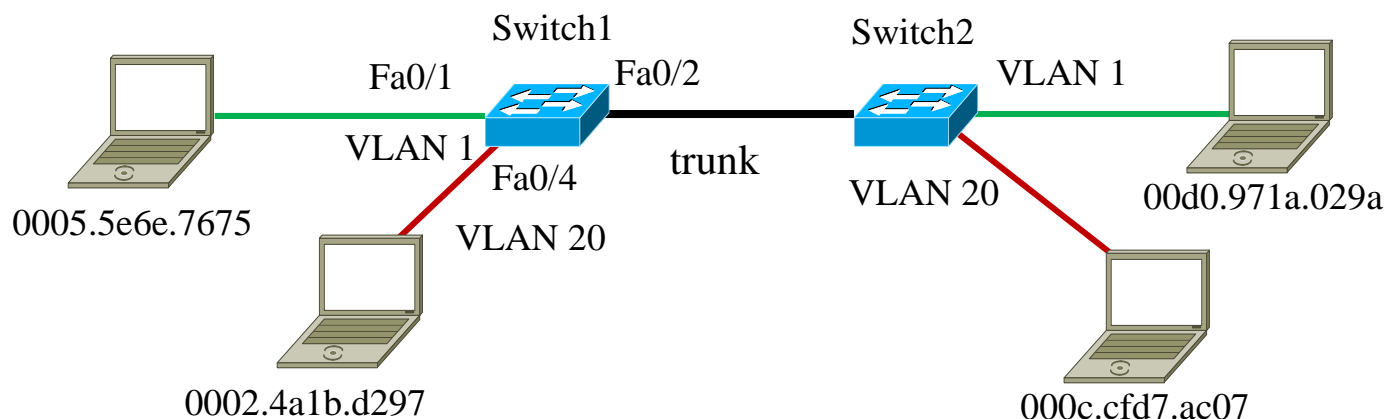
0x8100(VLAN)

以太网的帧加入VLAN ID



虚拟局域网(4)

□ VLAN 举例



Switch1#show mac-address-table

Vlan	Mac Address	Type	Ports
----	-----	-----	-----
1	0005.5e6e.7675	DYNAMIC	Fa0/1
1	00d0.971a.029a	DYNAMIC	Fa0/2
20	0002.4a1b.d297	DYNAMIC	Fa0/4
20	000c.cfd7.ac07	DYNAMIC	Fa0/2

* 左表中的DYNAMIC 是指通过学习得到的
而不是手工配置的

虚拟局域网(5)

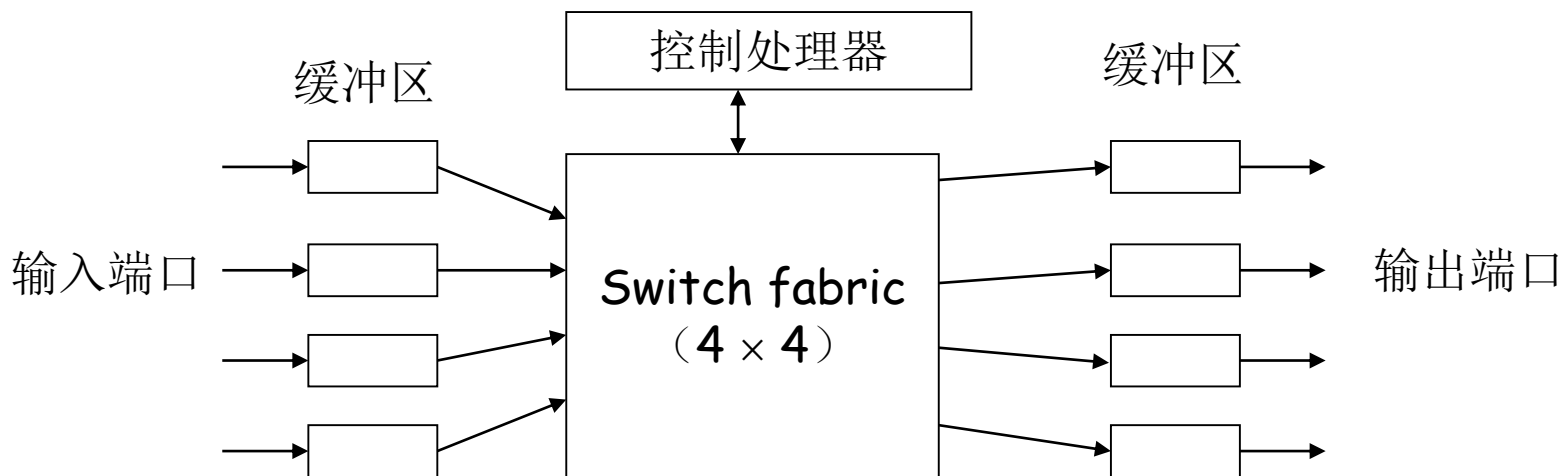
□ CST, PVST+ and MSTP

- IEEE 802.1Q中定义了由所有VLAN共享一棵树的公共生成树 (Common Spanning Tree, CST)。
- 具有思科专利的PVST(Per-VLAN Spanning Tree)协议为每个VLAN配置一颗生成树。由于 PVST 只能用于思科自己的VLAN标准ISL，思科又定义了同时可用于IEEE 802.1Q的PVST+标准。思科的设备现在默认使用PVST+。
- 多生成树MSTP (Multiple Spanning Tree Protocol)起初单独由IEEE 802.1s定义，后来并入IEEE 802.1Q-2005。它是RSTP的一个扩展，并可以把VLAN分组，每个VLAN组使用一颗生成树。
- BID: PRIORITY(4b)+VLAN ID(12b)+MAC addr(6B)
PRIORITY(4b):0,4096,... ,32768(default),..., 61440

交换机 (1)

□ 概述

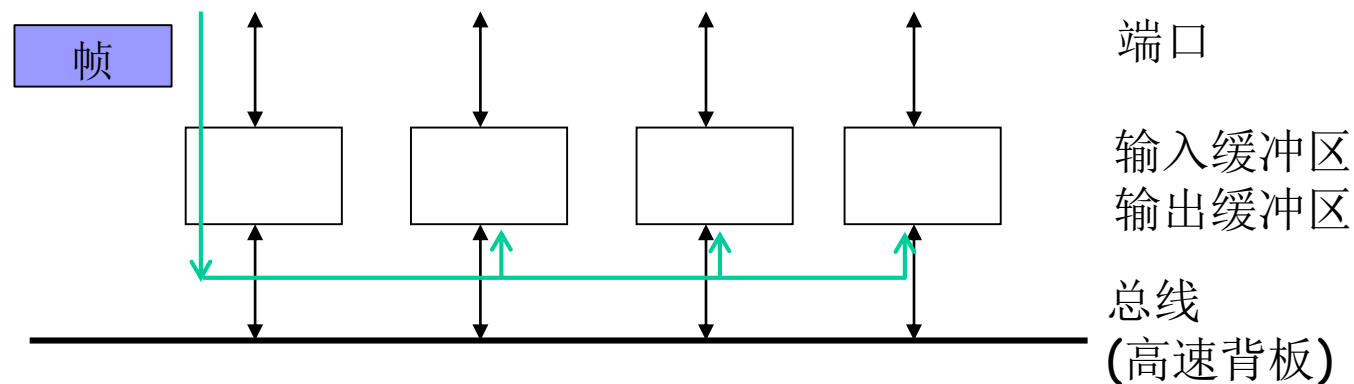
- 交换机(switch)是一个把多个网段连接起来的设备，也称为多端口网桥。



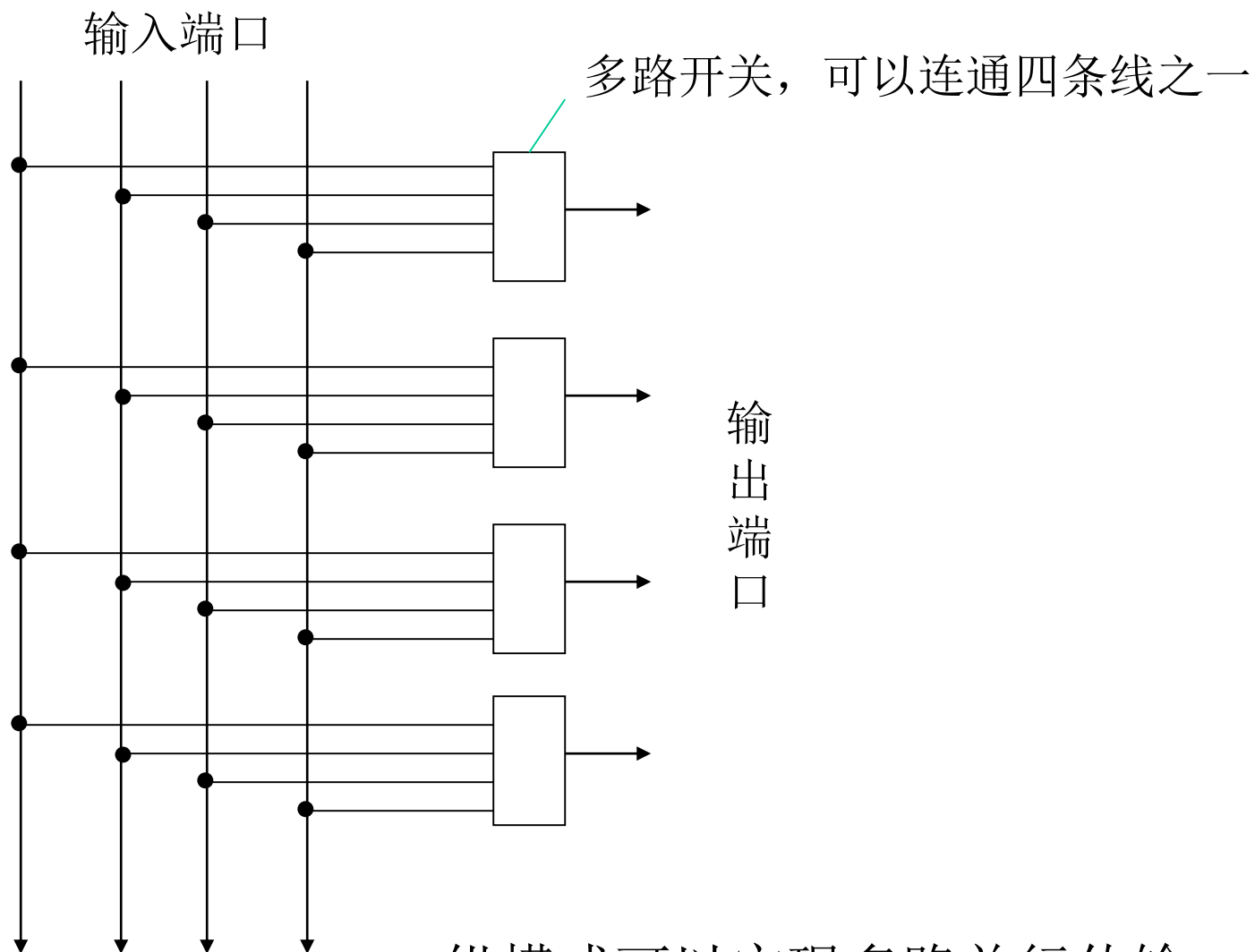
交换机(2)

□ 交换结构(Fabrics)

- 共享总线式交换机



- 纵横式 (crossbar)



纵横式可以实现多路并行传输。

交换机(3)

□ 转发方法

- **存储转发(Store and forward):** 交换机接收整个帧后转发它。大部分交换机都采用这种转发模式。
- **直通(Cut through):** 交换机不用收到整个帧而是收到帧的硬件地址后立即转发它。如果输出(outgoing port)忙, 则会转为存储转发。
- **无碎片(Fragment free):** 交换机不用收到整个帧而是收到帧的前64个字节(冲突窗口)后立即转发它。
- **适应性交换(Adaptive switching):** 自动在上面三种方式进行选择。

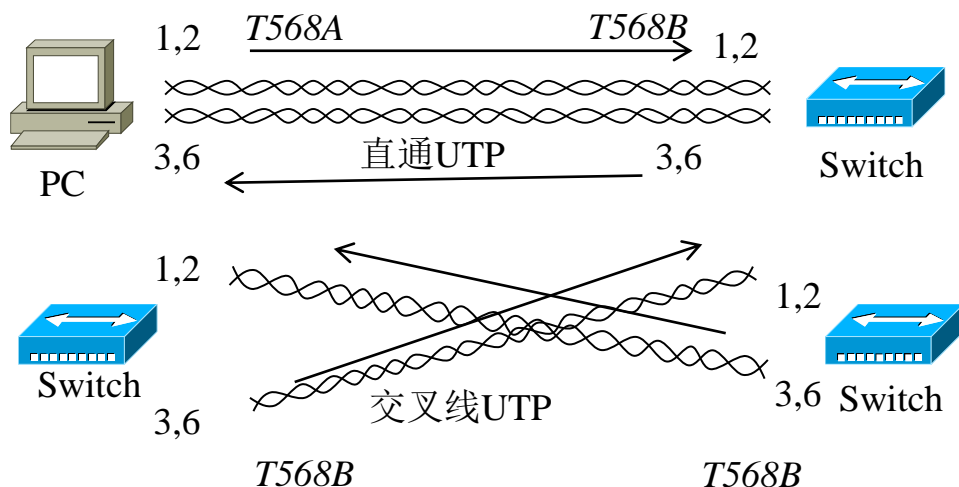
交换机(4)

□ 全双工模式

交换机可以工作在全双工模式下，因为没有冲突，CSMA/CD算法可以被关闭。

□ 自动翻转(Auto-MDIX)

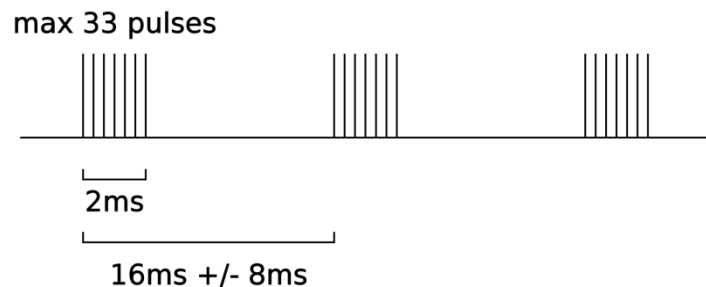
大部分交换机可以自动选择连接方式：交叉线或直通线



交换机(5)

□ 自适应(Autonegotiation)

两个站点周期性使用快速链路脉冲(fast link pulse,FLP)选择
10M/100M/1000M bps 自适应



快速以太网

属性(P):

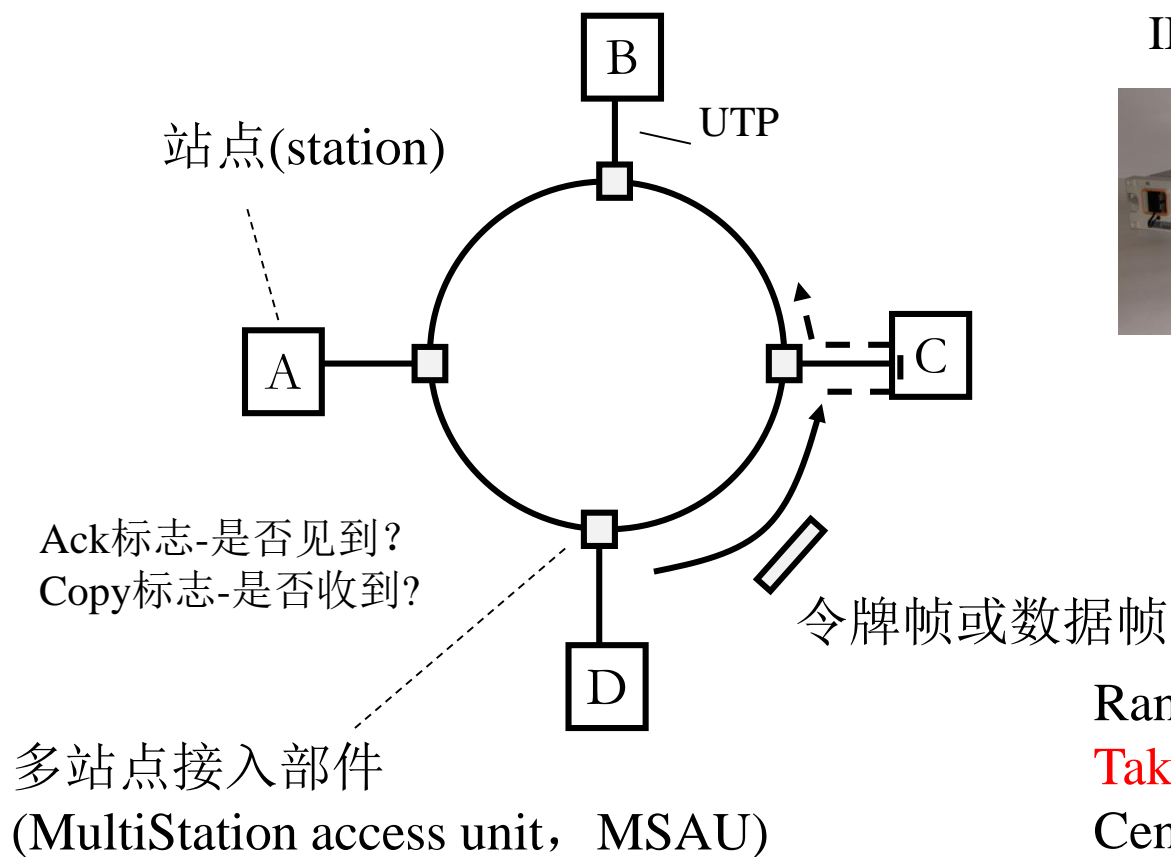
- 大量发送卸载 V2 (IPv4)
- 大量发送卸载 V2 (IPv6)
- 关闭唤醒
- 唤醒幻数据包
- 唤醒模式匹配
- 极大帧
- 接收方缩放
- 接收缓冲区
- 流控制
- 每秒最大 IRQ
- 能效以太网
- 速度和双工
- 网络地址
- 中断裁决

值(V):

- 自动协商
- 10 Mbps 半双工
- 10 Mbps 全双工
- 100 Mbps 半双工
- 100 Mbps 全双工
- 自动协商

令牌环网(1)

令牌环网(Token Ring)是一个通过在站点之间传递令牌防止冲突并且具有优先权的星形LAN，其标准为IEEE 802.5。现在有千兆令牌环网。



IBM 8228 MSAU



监控站、无主帧

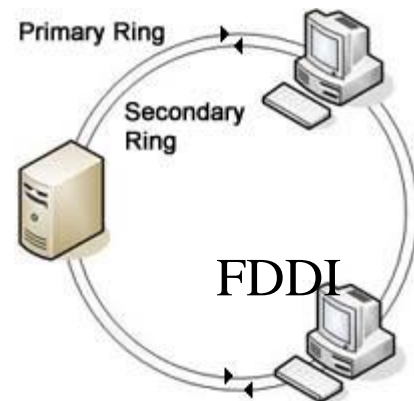
Random Access Protocol
Take Turns Protocol
Centralized Protocol

令牌环网(2)

数据传送过程:

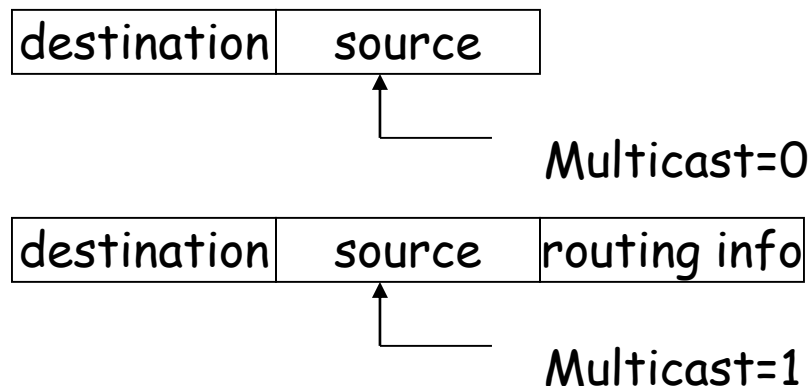
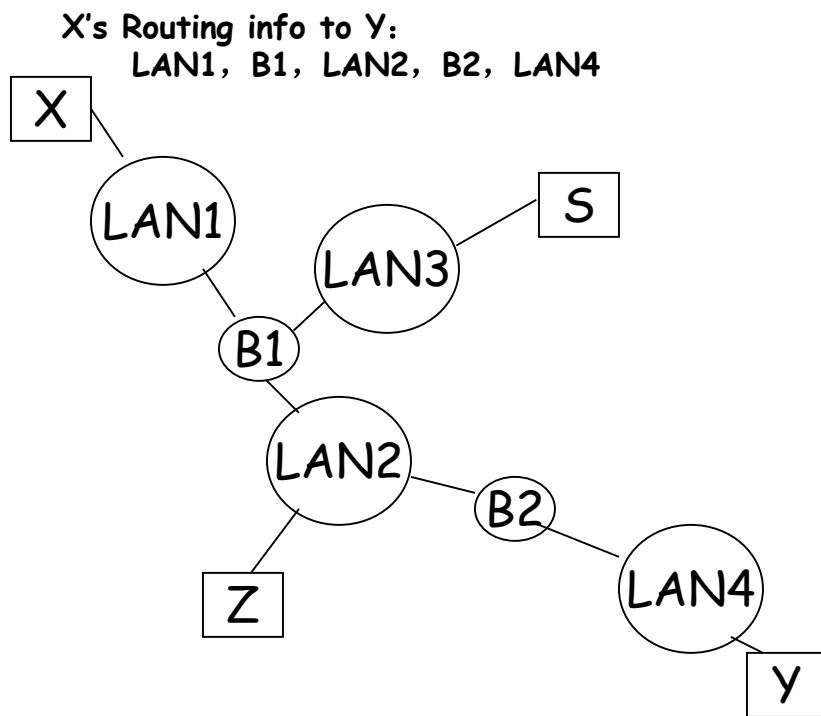
- 令牌(帧)绕环而行。
- 只有截获令牌的站点才可以发送数据帧。
- 发送的数据帧通过所有的活动站点。
- 目的站点拷贝数据帧。
- 只有发送方移除数据帧。
- 当没有数据帧要发送或者持有时间到，当前的发送站点要释放令牌。被释放的令牌继续绕环而行。

光纤分布式数据接口(Fiber Distributed Data Interface, FDDI)是另一种采用了令牌环的局域网，是一种100 Mbps的光纤局域网。

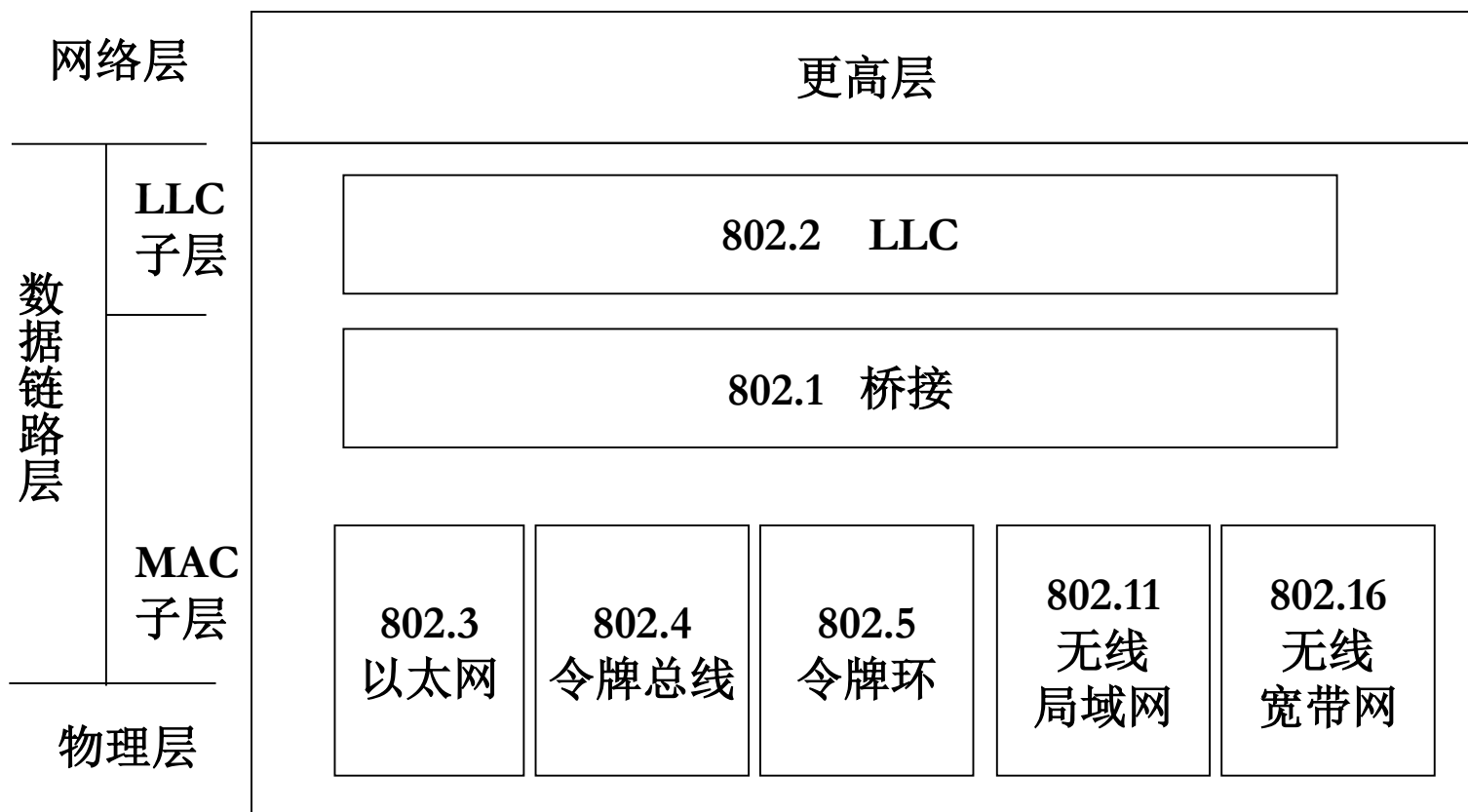


源路由网桥

- ❑ 源路由桥接算法是由**IBM**开发的用于令牌环网的协议。
- ❑ 为了兼容普通交换机，源路由网桥交换机也必须实现透明网桥的功能。



IEEE 802系列标准



802.1D-透明网桥和生成树
802.1Q-VLAN

LLC1提供无确认无连接服务
LLC2提供有确认面向连接的服务
LLC3提供有确认无连接的服务

总结

- ❑ 概述：介质访问控制层的功能
- ❑ 以太网：**CSMA/CD**算法，二进制指数退避算法
10M/100M/1000M bps 以太网
- ❑ 透明网桥算法：
- ❑ 生成树算法(STP)
- ❑ **VLAN**
- ❑ 交换机
- ❑ 令牌环网
- ❑ **IEEE 802**标准