

# 第五章 MAC子层进阶- 以太网

授课教师：崔 勇

清华大学



计 算 机 网 络  
教 案 社 区

致谢社区成员

袁华 华南理工大学

曹轶臻 中国传媒大学

赵婧如 西安邮电大学

徐敬东 南开大学



# 回顾：介质访问子层基础



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 信道分配问题

- 常见的局域网拓扑均共享信道：可能多个站点同时请求占用信道！
- 解决：介质的多路访问控制

## ➤ 随机访问协议：努力减少冲突

- Aloha：任性->时隙
- CSMA：先听后发；不坚持听，不一定发

## ➤ 受控访问协议：避免冲突

- 位图；二进制倒数计数；令牌环

## ➤ 有限竞争：自适应树搜索协议

## ➤ 以太网的共享访问基础

- 经典以太网：指数后退的1-持续CSMA/CD
- 最小帧长：取决于以太网上最远距离的RTT & 链路带宽

**减少冲突是关键**

越来越优雅，越来越谦让！

(1-持续=>P-持续)

持续侦听，概率p发送

**避免冲突：提前排位配置**



# 思考

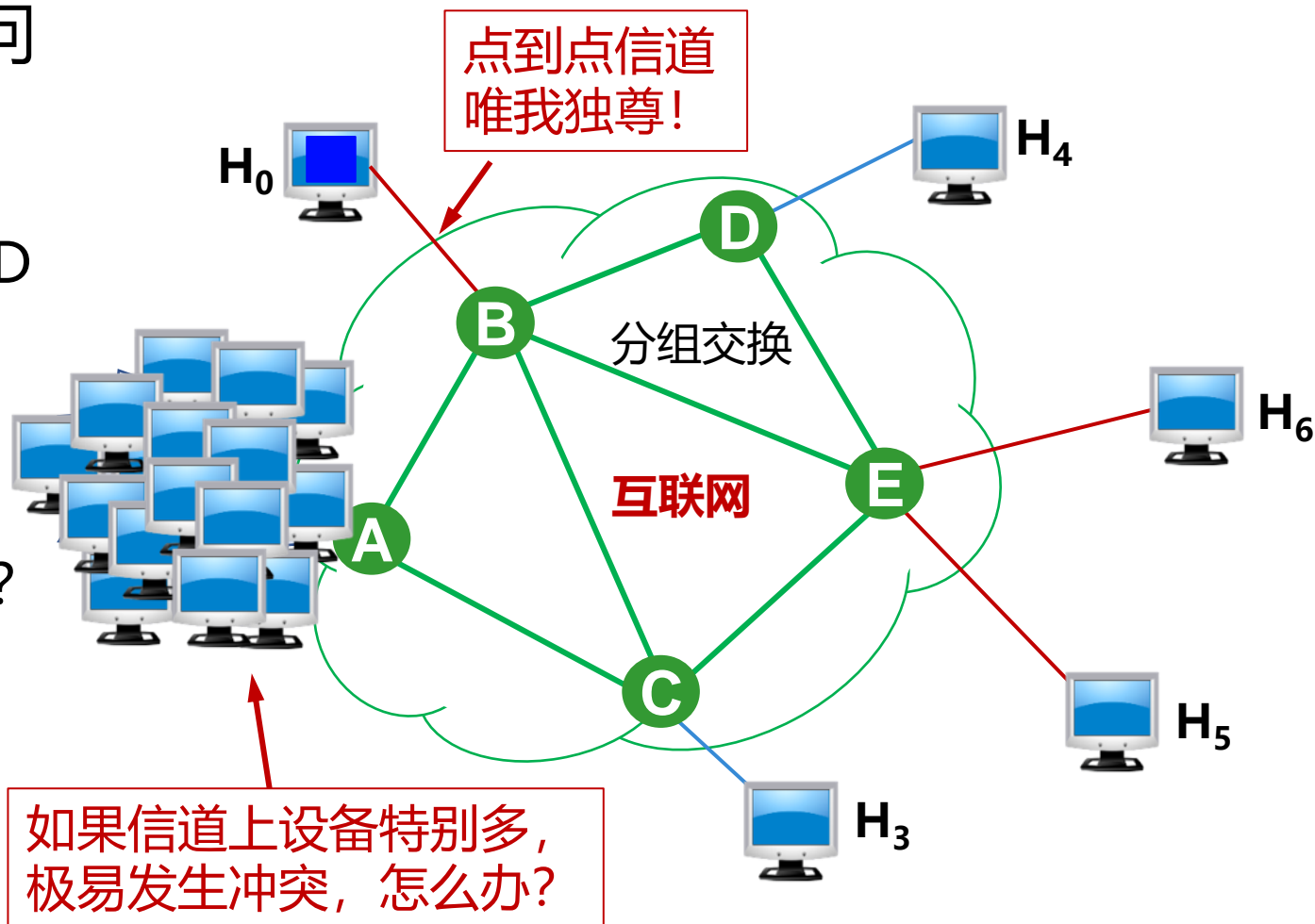


清华大学  
Tsinghua University



计算机网络教案社区

- 理论基础：共享信道访问
  - 从Aloha到多种CSMA
  - 受控访问：配置太烦人
  - 以太网：1-持续CSMA/CD
- 实践：经典以太网
  - 谁给谁？帧结构.....
- 新问题：大规模局域网
  - 设备多冲突严重，怎么办？
  - 信道效率100%就满意吗？
- 思路：用“域”隔离
  - 共享信道设备数量少一点
  - 互联互通怎么办？





# 主要内容与学习目标



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 主要内容

- 经典以太网
- 链路层交换原理
- 生成树协议
- 链路层交换机
- 虚拟局域网

## ➤ 学习目标

- 理解经典以太网的设计思路和帧结构
- 掌握链路层交换原理
- 掌握生成树协议与MAC表维护
- 掌握交换机的三种工作模式
- 了解虚拟局域网的使用场景和实现方法



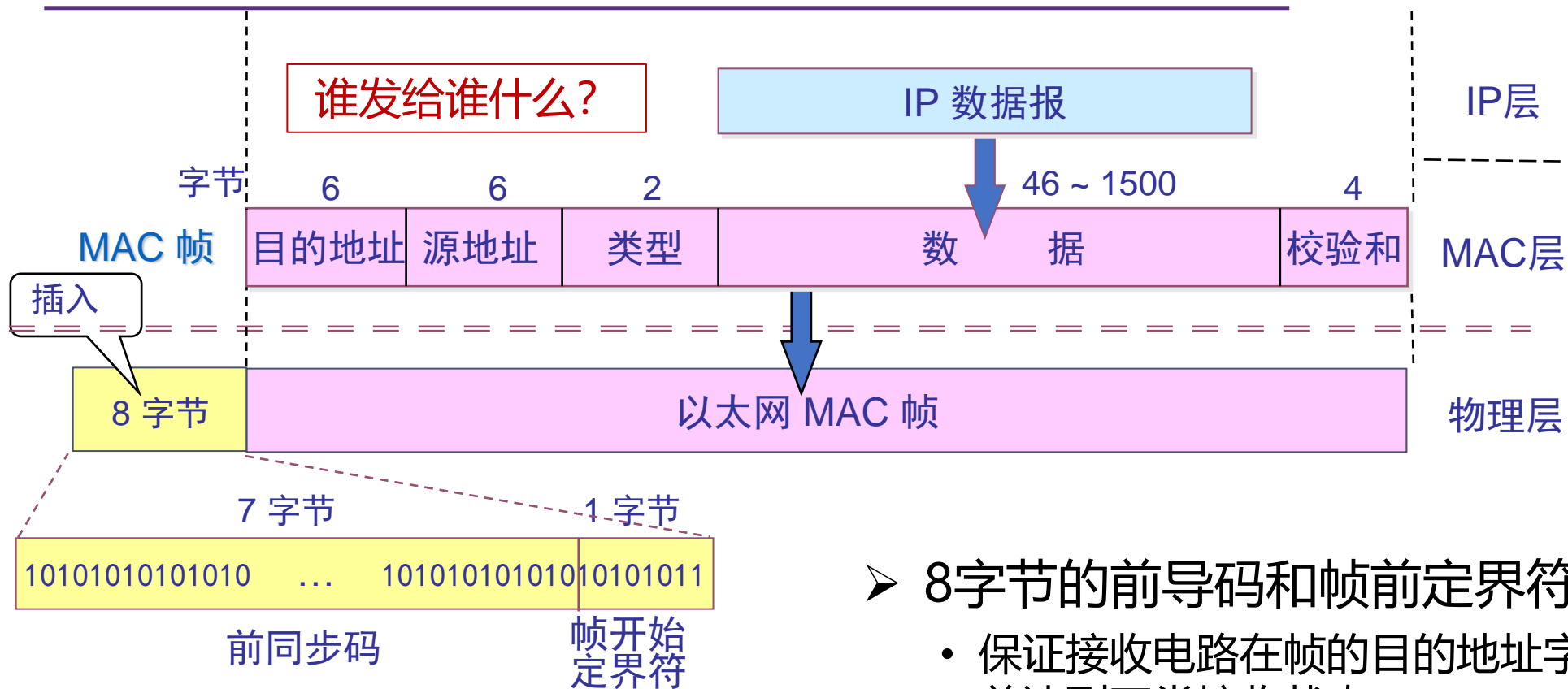
# 经典以太网 - MAC子层协议



清华大学  
Tsinghua University



计算机网络教案社区



## ➤ 8字节的前导码和帧前定界符

- 保证接收电路在帧的目的地址字段到来之前达到正常接收状态
- 不需要保留提交上层，也不计入帧头长度



# 经典以太网 - MAC子层协议



- 硬件地址又称为物理地址，或 MAC 地址
- MAC帧中的源地址和目的地址长度均为6字节

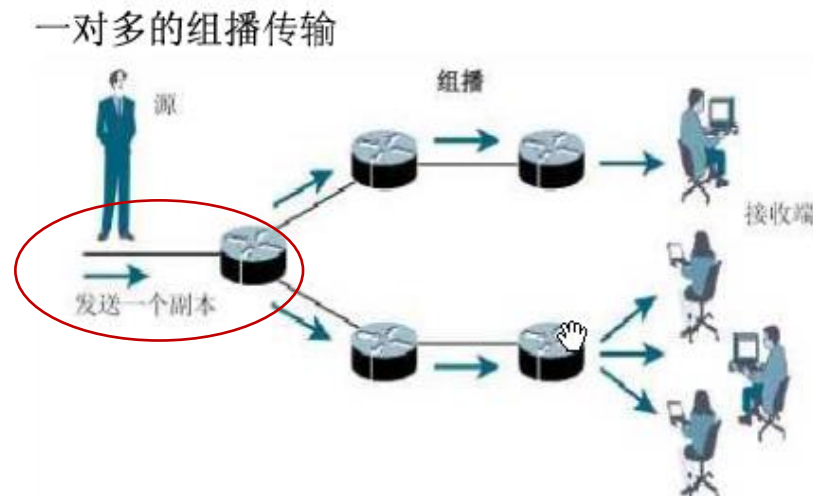
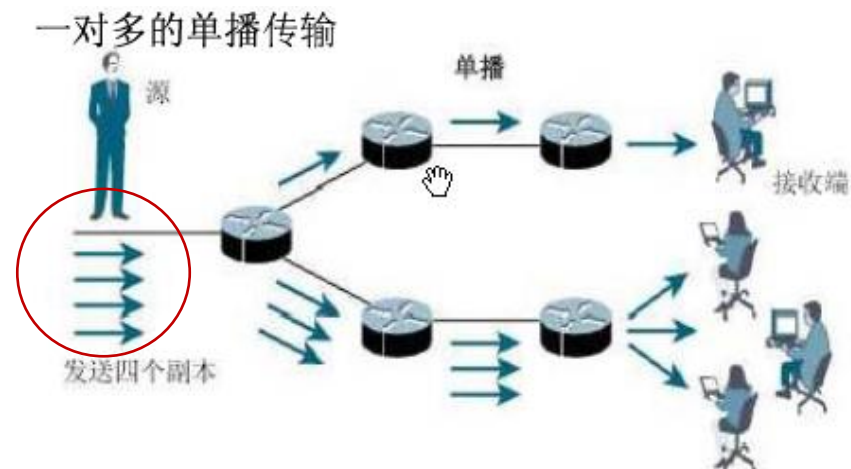
6	6	2	46 ~ 1500	4
目的地址	源地址	类型	数 据	校验和

MAC层支持单播、广播和组播？

单播 (unicast) : 5C-26-0A-7E-4E-4C

广播 (broadcast) : FF-FF-FF-FF-FF-FF

组播 (multicast) : 01-00-5E-00-00-00





# 经典以太网 - MAC子层协议



清华大学  
Tsinghua University



计算机网络教案社区

- 硬件地址又称为物理地址，或 MAC 地址
- MAC帧中的源地址和目的地址长度均为6字节

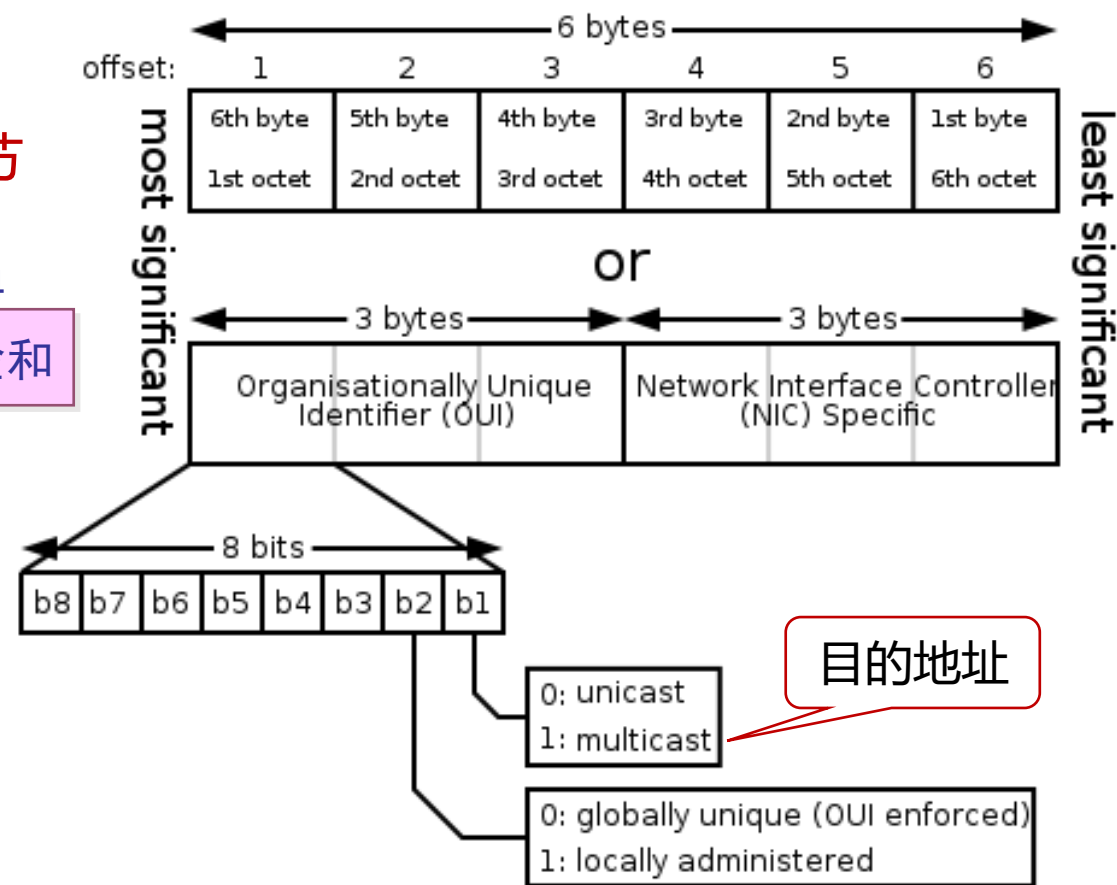


MAC层如何区分单播广播和组播?

单播 (unicast) : 5C-26-0A-7E-4E-4C

广播 (broadcast) : FF-FF-FF-FF-FF-FF

组播 (multicast) : 01-00-5E-00-00-00





# 经典以太网 - MAC子层协议



清华大学  
Tsinghua University



计算机网络教案社区

- OUI (Organizationally Unique Identifier)
  - IEEE Registration Authority是负责注册和管理组织唯一标识符 (OUI) 的管理机构
- 在Windows上使用`ipconfig /all`命令查看MAC地址 (Linux: `ifconfig`)

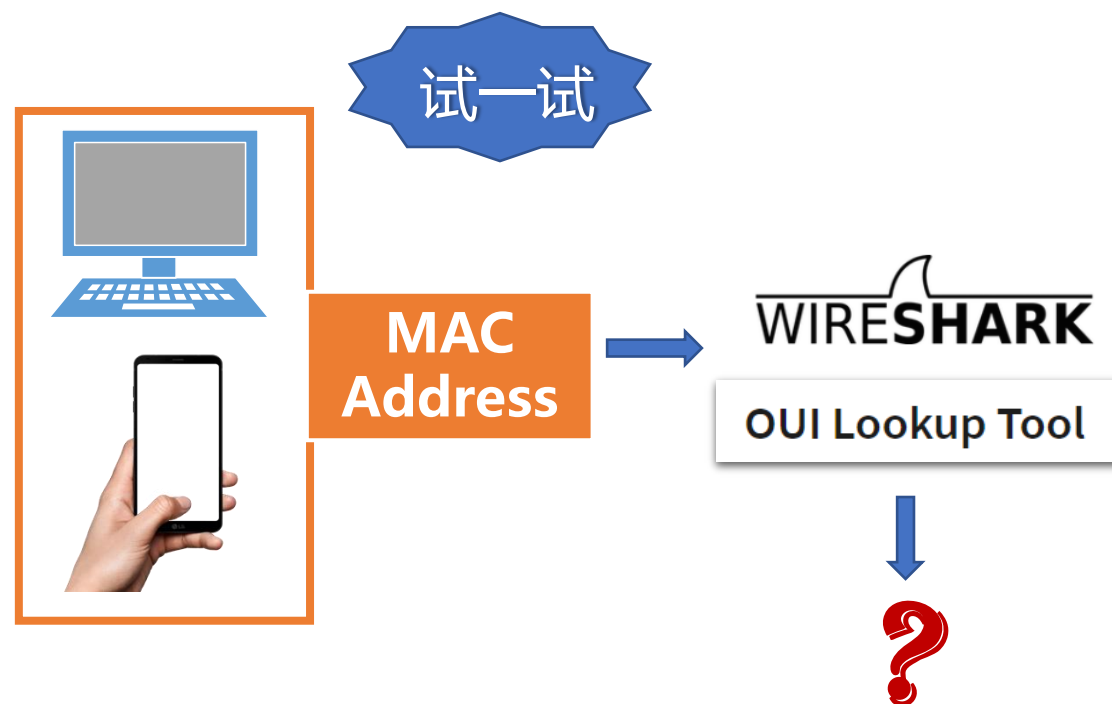
```
Administrator: 命令提示符
C:\>ipconfig /all

Windows IP Configuration

Host Name . . . . . : DESKTOP-HOME
Primary Dns Suffix . . . . . :
Node Type . . . . . : Hybrid
IP Routing Enabled. . . . . : No
WINS Proxy Enabled. . . . . : No

Ethernet adapter 以太网:

Connection-specific DNS Suffix  . : 
Description . . . . . : Total(R) Ethernet Connection (2) I218-V
Physical Address. . . . . : 1C-87-2C-72-86-BC
Dhcp Enabled. . . . . : Yes
Autoconfiguration Enabled . . . . : Yes
Link-local IPv6 Address . . . . . : fe80::80b7:4e6a:7e36:9220%16(Preferred)
IPv4 Address. . . . . : 10.0.0.21(Preferred)
Subnet Mask . . . . . : 255.255.255.0
Lease Obtained. . . . . : 2021年2月11日 13:33:27
Lease Expires . . . . . : 2021年2月16日 10:27:57
Default Gateway . . . . . : 10.0.0.1
DHCP Server . . . . . : 10.0.0.1
DHCPv6 IAID . . . . . : 52201260
DHCPv6 Client DUID. . . . . : 00-01-00-01-10-DA-2A-02-1C-87-2C-72-86-BC
DNS Servers . . . . . : 10.0.0.1
NetBIOS over Tcpip. . . . . : Enabled
```







# 经典以太网 - MAC子层协议



清华大学  
Tsinghua University



计算机网络教案社区

MAC  
帧格式

6	6	2	46 ~ 1500	4
目的地址	源地址	类型	数 据	校验和

目的地址	源地址	Length	数 据	校验和
------	-----	--------	-----	-----

如何判定帧长度?

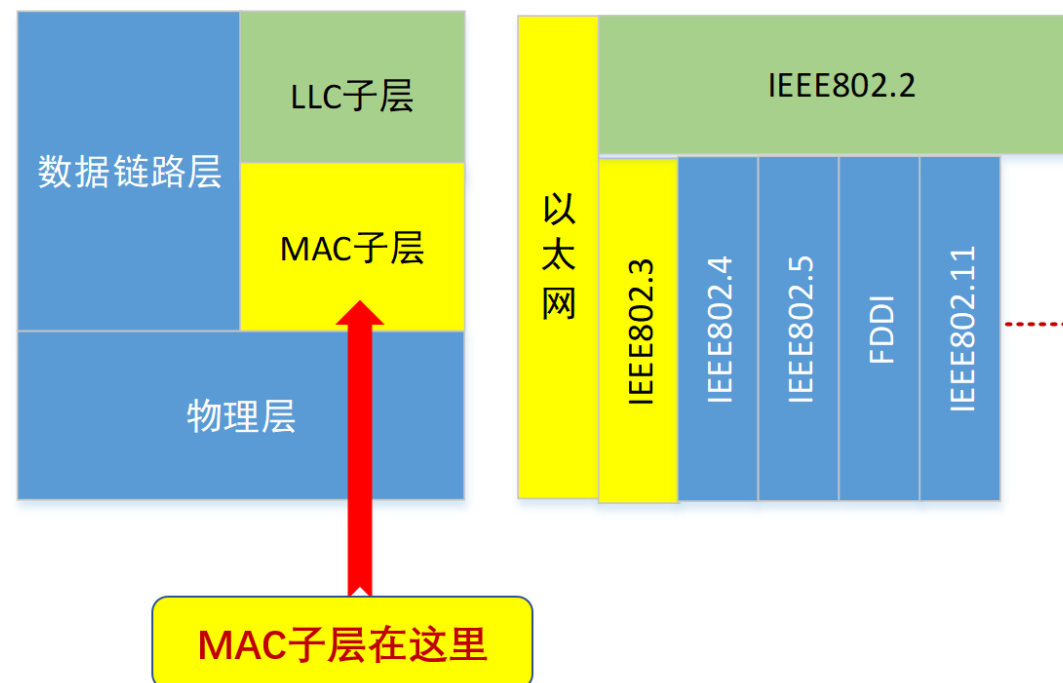
(a) DIX Ethernet V2

(b) IEEE 802.3

如何判定上层类型?

以太网类型字段

IPv4: 0x0800  
ARP: 0x0806  
IPv6: 0x86DD  
PPP: 0x880B  
PPPoE: 0x8864  
...





# 经典以太网 - MAC子层协议

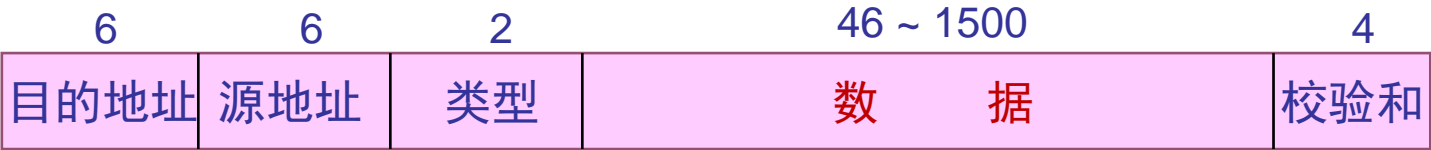
➤ 数据字段

- 46 ~ 1500字节

- 最小帧长 = 46+18 = 64B

- 最大帧长 = 1500+18 = 1518B （最大传输单元MTU：1500B）

➤ 数据字段不足46字节，需要填充整数字节（Padding）至46字节，以保证以太网MAC帧不小于64字节



```
> Frame 25: 60 bytes on wire (480 bits), 60 bytes captured (480 bits) on interface 0
v Ethernet II, Src: b0:7f:b9:ff:70:aa, Dst: 1c:87:2c:72:86:bc
  > Destination: 1c:87:2c:72:86:bc
  > Source: b0:7f:b9:ff:70:aa
  Type: ARP (0x0806)
  Padding: 0000000000000000000000000000000000000000
> Address Resolution Protocol (reply)
```

0000	1c 87 2c 72 86 bc b0 7f b9 ff 70 aa 08 06 00 01	...,r.... ..p....
0010	08 00 06 04 00 02 b0 7f b9 ff 70 aa 0a 00 00 01	..... ..p....
0020	1c 87 2c 72 86 bc 0a 00 00 16 00 00 00 00 00 00	...,r.... ..
0030	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00	.....



# 经典以太网 - MAC子层协议



清华大学  
Tsinghua University



计算机网络教案社区



## ➤ 校验和

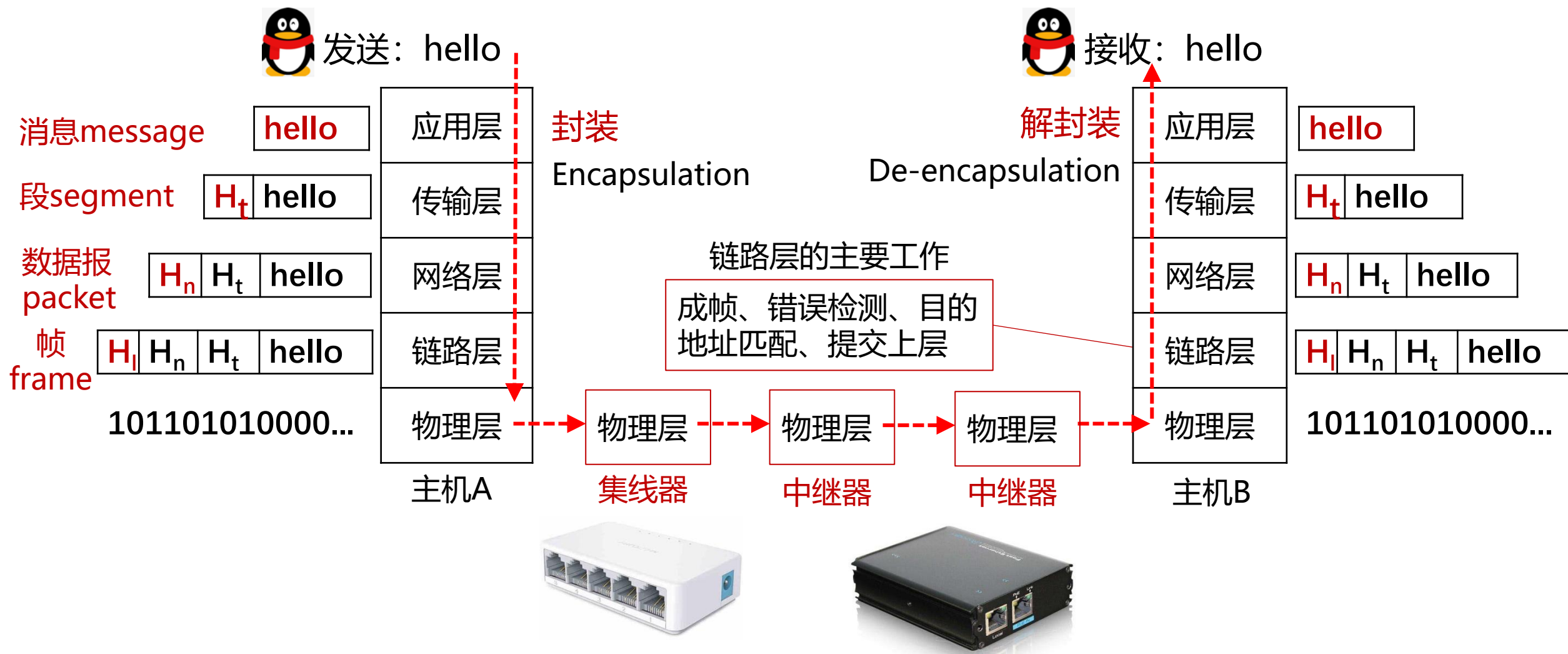
- **FCS**, Frame Check Sequence
- 使用**CRC32**计算除了校验和以外的其他字段

## ➤ 对于检查出的**无效 MAC 帧**就简单地丢弃，以太网不负责重传丢弃的帧

- 用收到的帧检验序列 FCS 查出有差错
- 数据字段的长度不在 46 ~ 1500 字节之间
- 数据字段的长度与长度字段的值不一致
- 帧的长度不是整数个字节



# 回顾：网络体系结构分层模型





# 以太网的发展



清华大学  
Tsinghua University



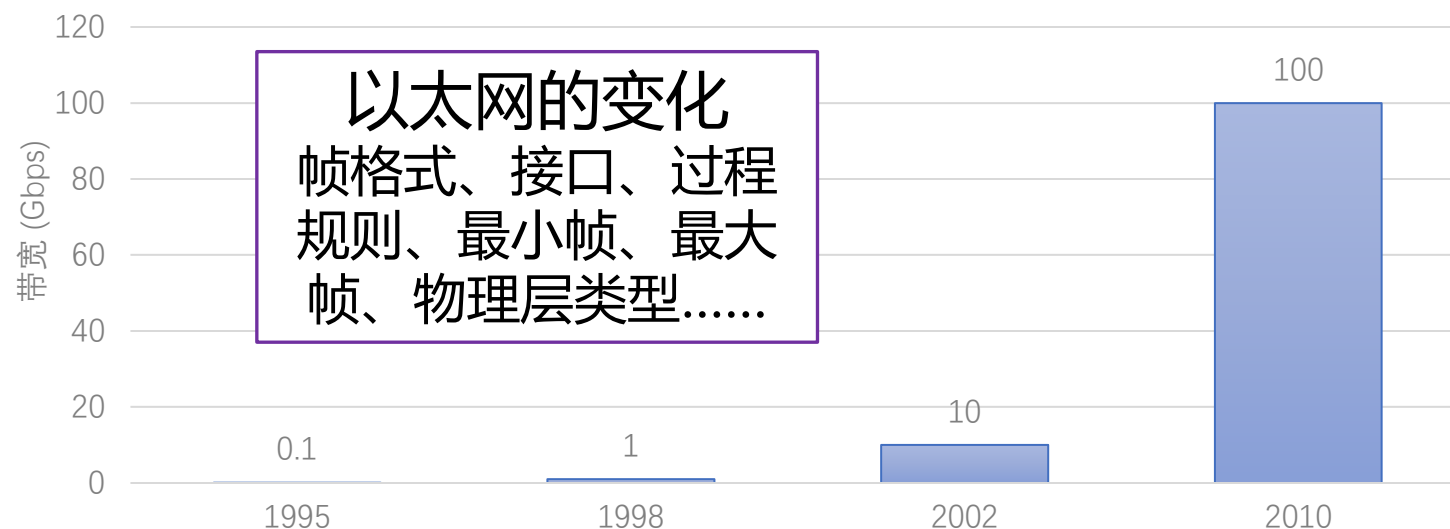
计算机网络教案社区

快速以太网  
IEEE 802.3u  
100Mbps

千兆以太网  
IEEE 802.3ab  
1Gbps

万兆以太网  
IEEE 802.3ae  
10Gbps

百G以太网  
802.3ba等  
40-100Gbps



不断增长的带宽

以太网的主战场在哪里？

灵活性

简单性

兼容性

廉价

易维护

易扩展

可靠性



# 以太网的未来

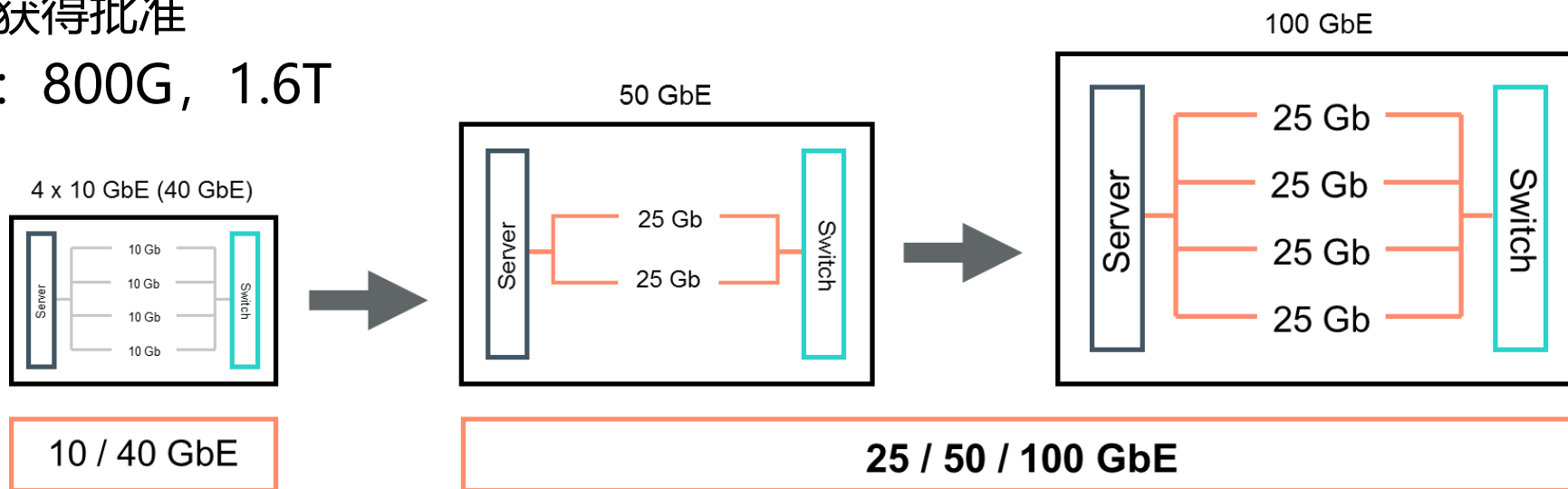


清华大学  
Tsinghua University



计算机网络教案社区

- 25/50G和第二代100G以太网
  - 2014年发布25G以太网标准IEEE 802.3by，弥补了10G以太网的低带宽和40G以太网的高成本缺陷
  - 25G以太网采用了25Gb/s单通道技术，可基于4个25Gbps光纤通道实现100G传输
- 400GbE和200GbE
  - 2017年，由IEEE P802.3bs使用与100GbE大致相似的技术开发的400GbE和200GbE标准获得批准
- 未来：800G，1.6T





# 主要内容与学习目标



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 主要内容

- 经典以太网
- 链路层交换原理
- 生成树协议
- 链路层交换机
- 虚拟局域网

## ➤ 学习目标

- 了解经典以太网的帧结构
- 掌握链路层交换原理
- 掌握生成树协议与MAC表维护
- 掌握交换机的三种工作模式
- 了解虚拟局域网的使用场景和实现方法



# 数据链路层交换原理

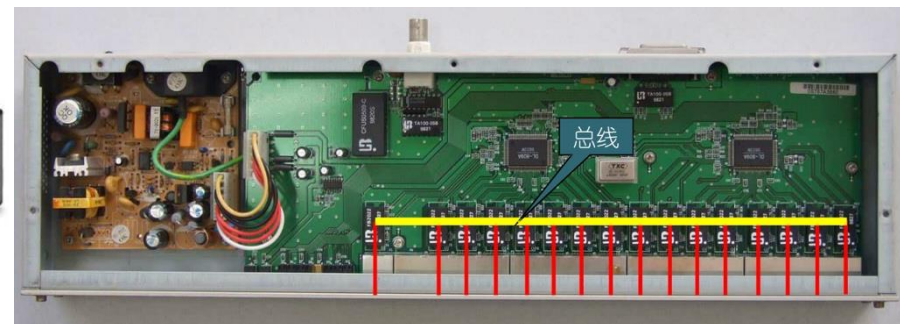


清华大学  
Tsinghua University



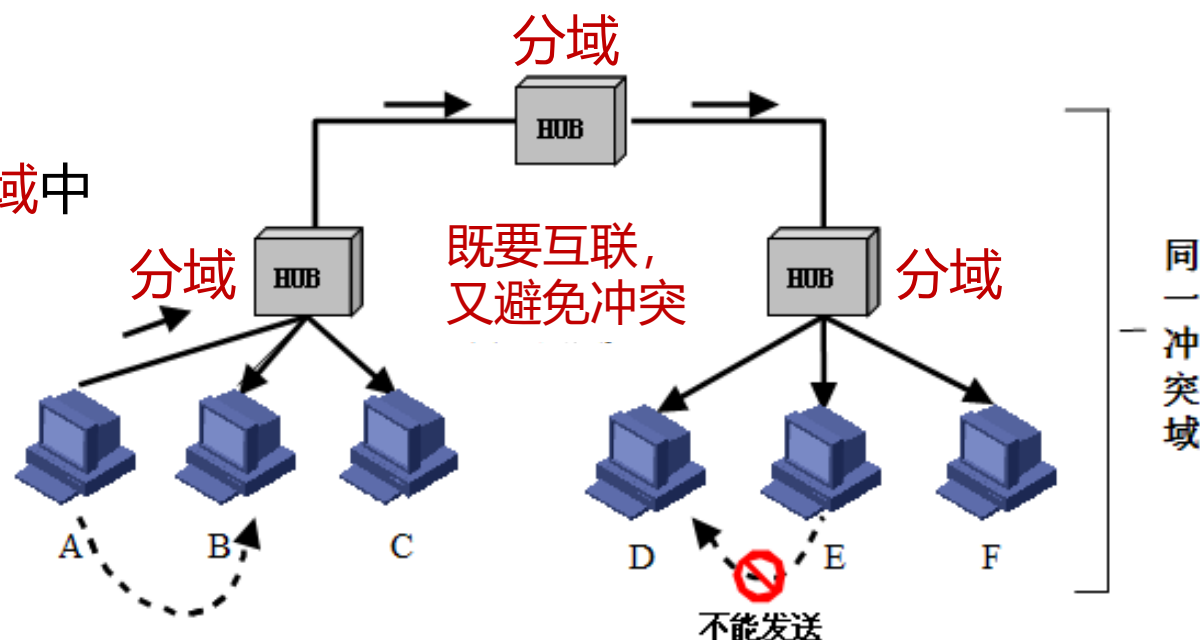
计算机网络教案社区

- 使用**集线器 (HUB)** 组建以太网
  - Hub所有端口内部都是连通的
  - 使同一根总线
  - 和Repeater一样，也是物理层设备
- 使用Hub扩展以太网
  - 集线器不能增加容量
  - 用集线器组成更大的局域网都在一个**冲突域**中
  - Hub级连：限制了网络的可扩展性



面对可扩展性，优雅和谦让无能为力，怎么办？

谁发给谁？中间设备如何协助和优化？







# 数据链路层交换原理



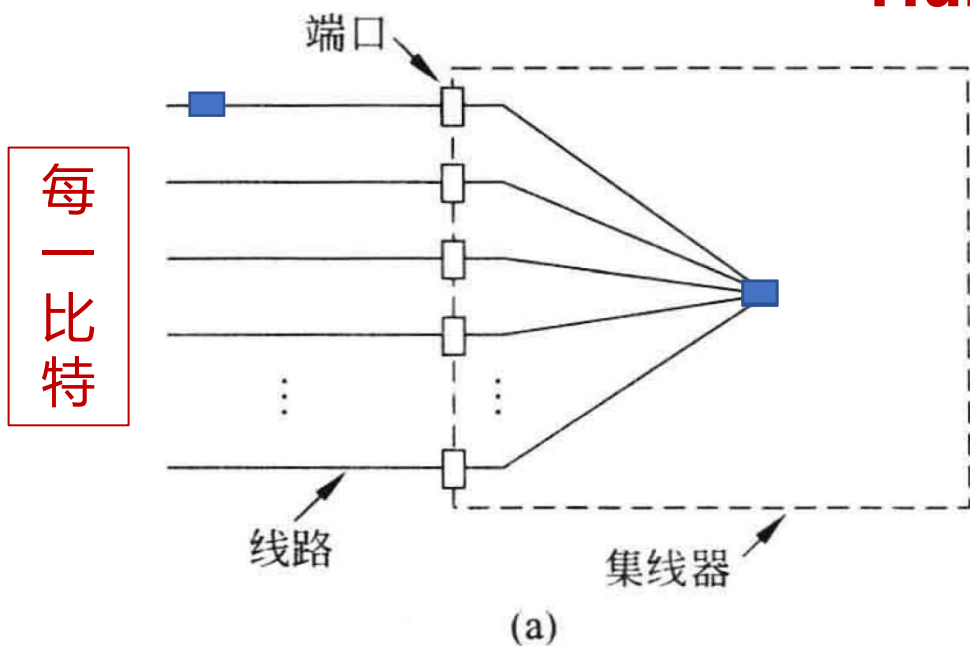
清华大学  
Tsinghua University



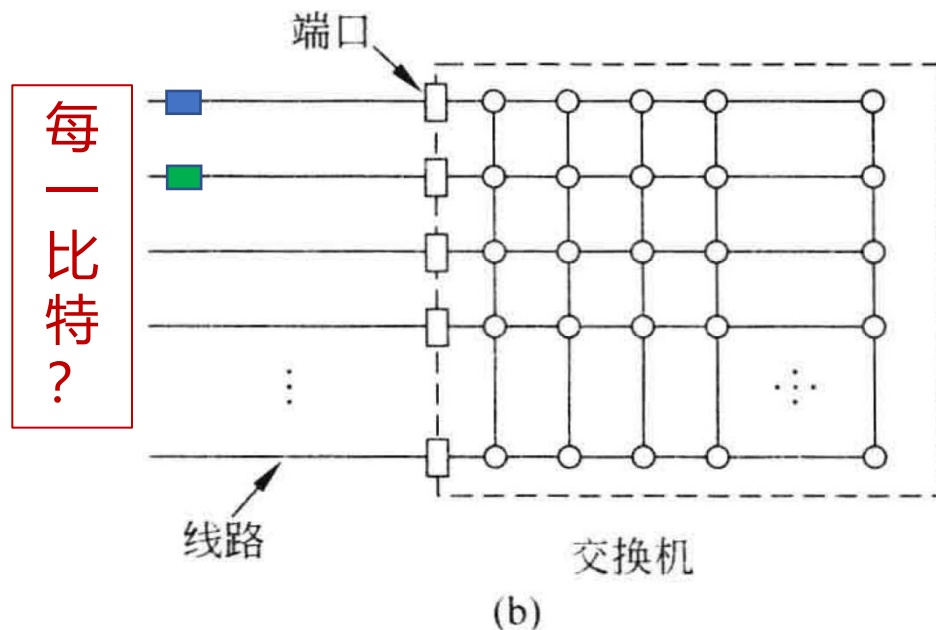
计算机网络教案社区

## Hub vs Switch

**WINNER**



- 内部连接所有线缆，逻辑上等同于**单根总线**的经典以太网
- 所有站都位于**同一个冲突域**，必须使用CSMA/CD协议



- 内部通过**高速背板**连接所有端口
- 每个端口都有独立的冲突域，在**全双工**模式下端口可以同时收发，则不需要CSMA/CD
- 可以实现**并行**传输



# 数据链路层交换原理

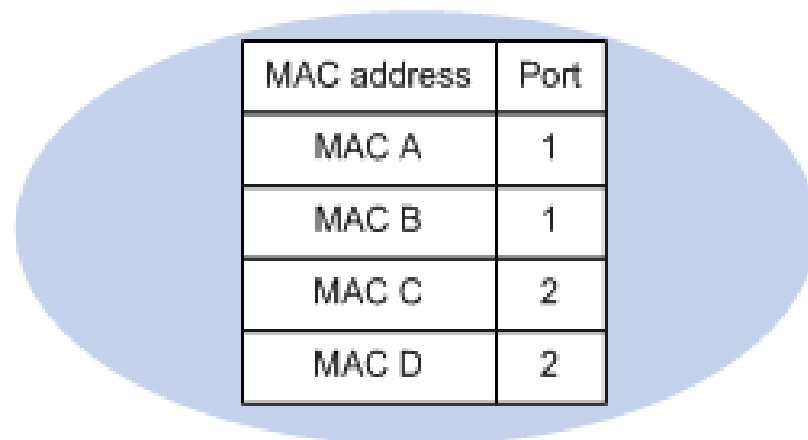


清华大学  
Tsinghua University

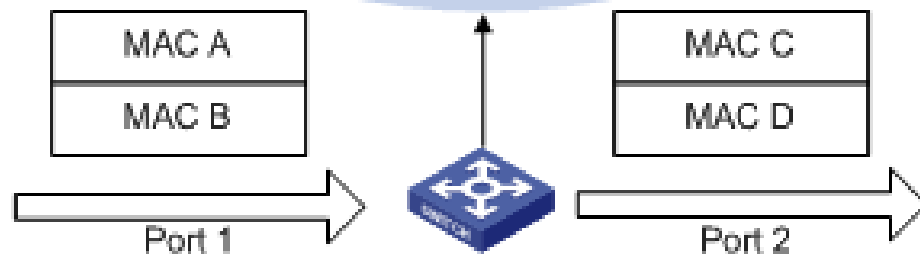


计算机网络教案社区

- 交换式以太网的核心是**交换机 (Switch)**
  - 工作在数据链路层，检查MAC 帧的**目的地址**对收到的帧进行转发
  - 交换机通过高速背板把帧传送到目标端口



- 混杂模式 (promiscuous mode)
  - Hacker
  - 网络分析



目标与理想



# 数据链路层交换原理



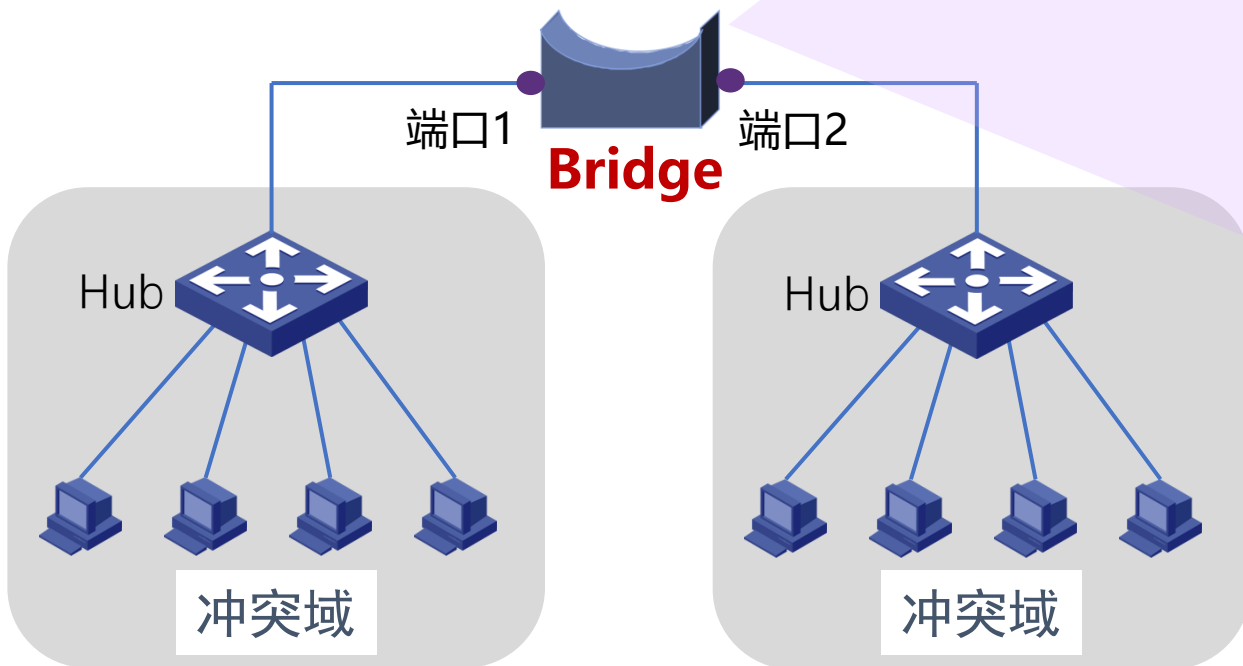
清华大学  
Tsinghua University



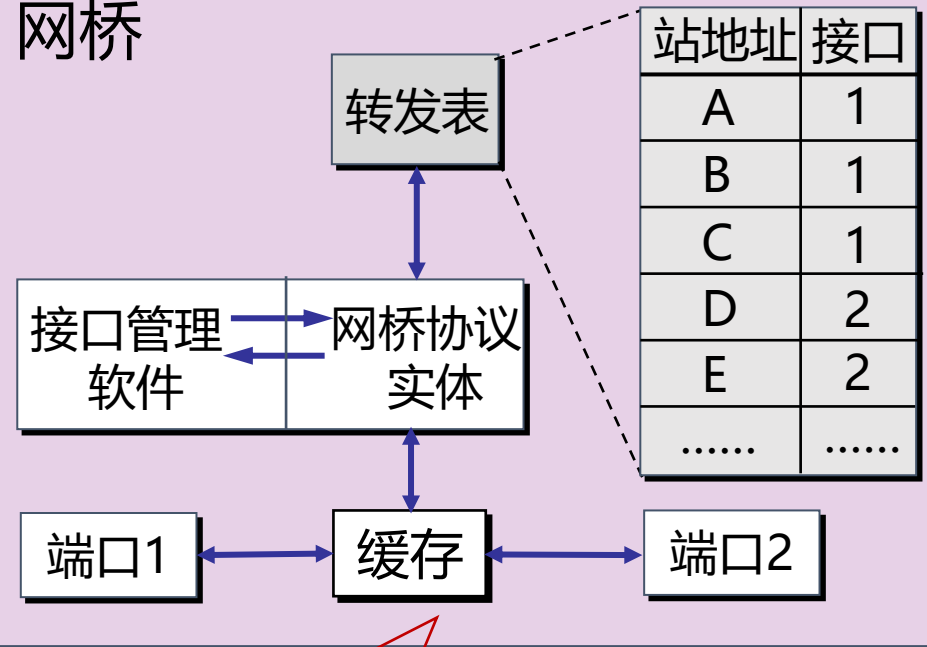
计算机网络教案社区

## ➤ 数据链路层设备扩充网络

- 网桥或交换机
- 分隔了冲突域
- 连接不同类型局域网



## 网桥



需要存储转发吗？

如何通过存储转发隔离冲突域？



# 数据链路层交换原理

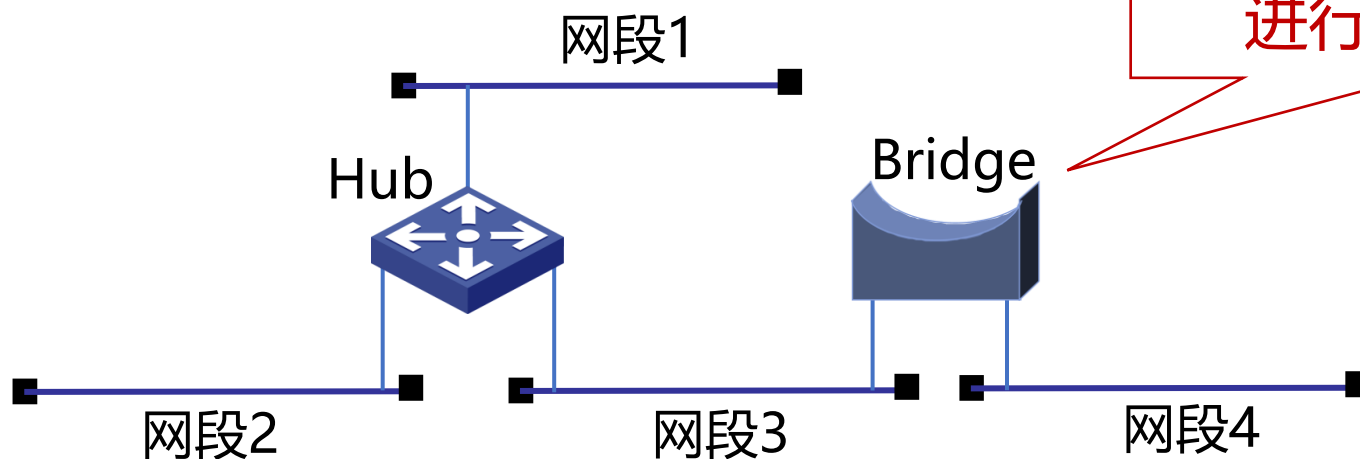


清华大学  
Tsinghua University



计算机网络教案社区

➤ 练习：图中有几个冲突域？





# 数据链路层交换原理



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 什么是理想的网桥？

- 即插即用，网络无需任何配置
- 提升网络效率，避免不必要的转发
- 网络中的站点无需感知网桥的存在与否
- 网桥本身也应尽量避免复杂的配置

如何实现透明网桥？

无配置且  
避免不必要的转发

自动学习？





# 数据链路层交换原理



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ Forwarding (转发)

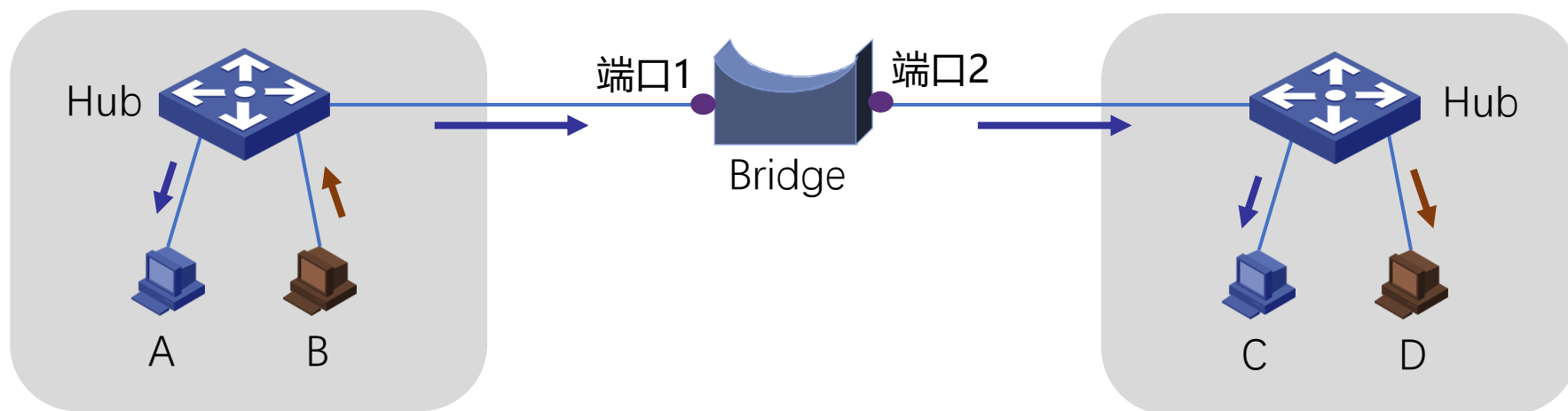
B → D 发出数据帧

逆向学习源地址  
并根据**目的地址**  
查询MAC地址表

MAC地址表完善时

MAC地址表	
MAC地址	端口
MAC_A	1
MAC_B	1
MAC_C	2
MAC_D	2

找到匹配项!  
从对应端口2转发出去





# 数据链路层交换原理



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ Filtering (过滤)

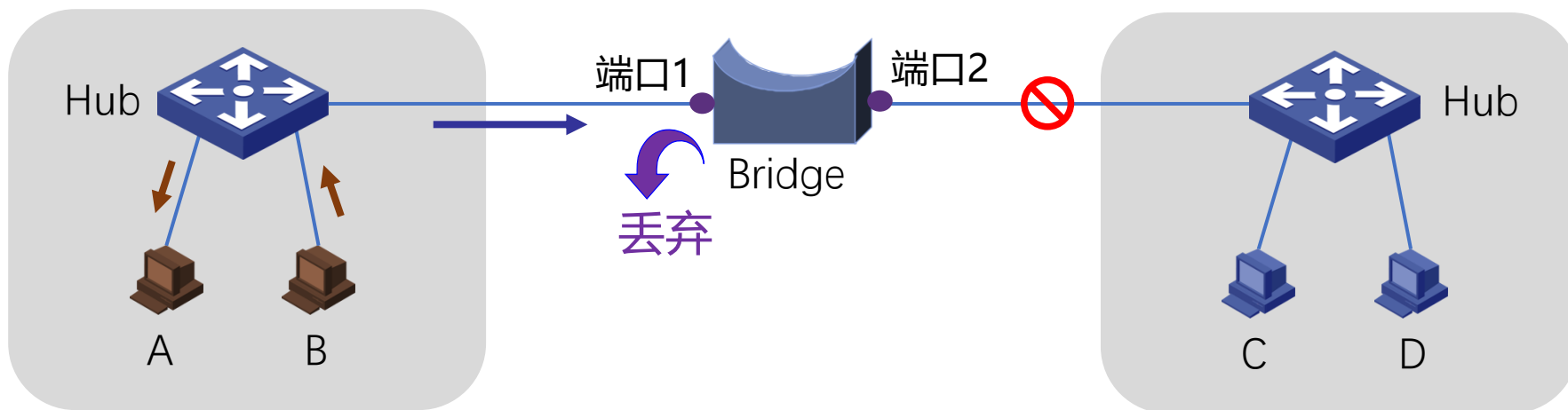
B → A 发出数据帧

逆向学习源地址  
并根据**目的地址**  
查询MAC地址表

MAC地址表完善时

MAC地址表	
MAC地址	端口
MAC_A	1
MAC_B	1
MAC_C	2
MAC_D	2

找到匹配项!  
入境口=出境口, 丢弃!





# 数据链路层交换原理



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ Flooding (泛洪)

A → B 发出数据帧

找不到匹配表项怎么办?

从所有端口 (除了入境口) 发送出去

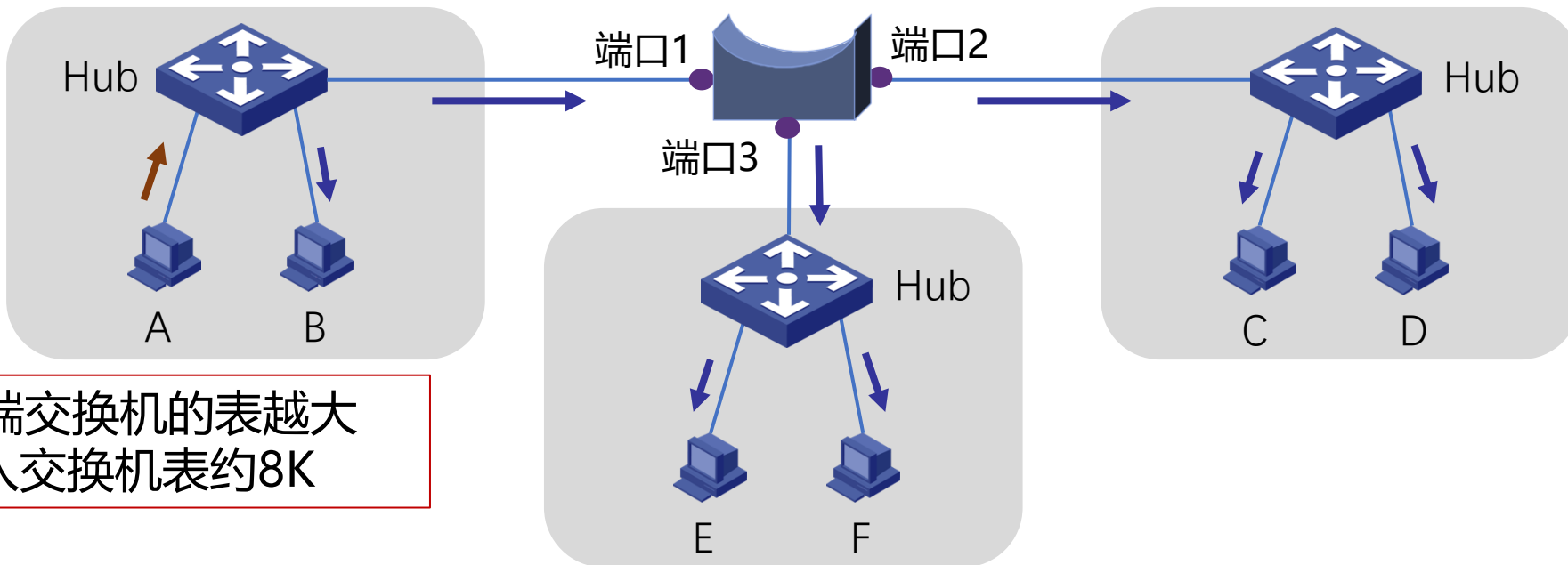
数据被发送到无关网段☹

MAC地址表不完善时

MAC地址表	
MAC地址	端口
MAC_A	1

MAC地址表会存满而溢出吗?  
是不是存在安全隐患?

要想学好  
请先学坏



越高端交换机的表越大  
接入交换机表约8K





# 数据链路层交换原理



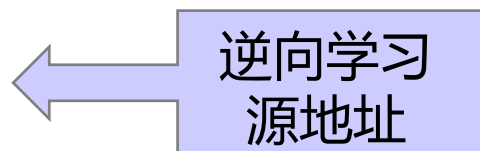
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ MAC地址表的构建-逆向学习源地址

A→B发出数据帧



MAC地址表	
MAC地址	端口
MAC_A	1

记录帧到达时间

设定老化时间（默认300s）

当老化时间到期时，该表项会被清除

Why?





# 数据链路层交换原理



清华大学  
Tsinghua University



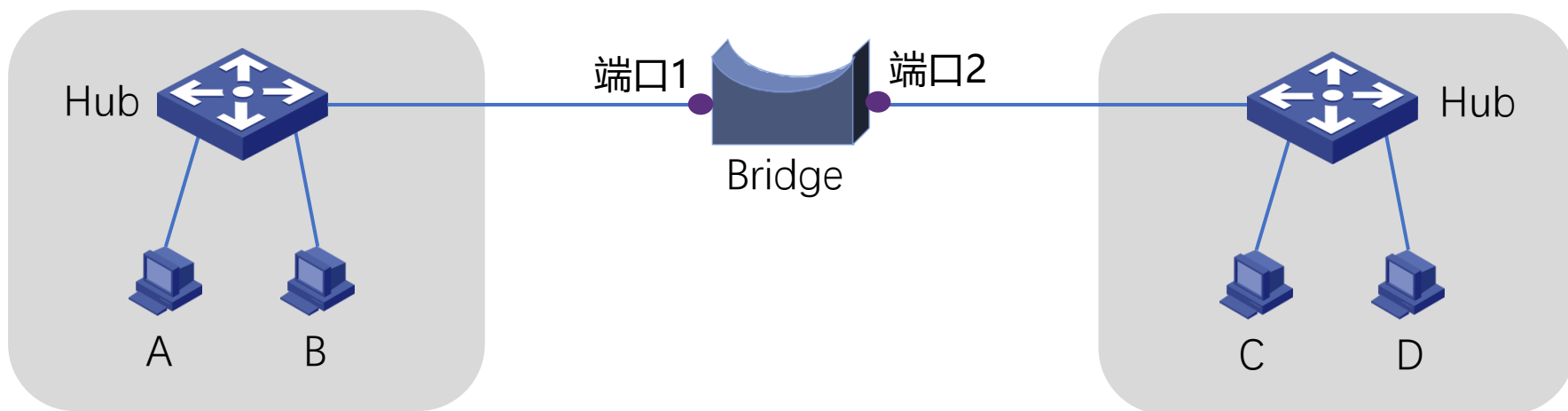
计算机网络教案社区

## ➤ 发送帧的站MAC地址被学习

网桥怎样学习到B\C\D的  
MAC地址?

MAC地址表	
MAC地址	端口
MAC_A	1
MAC_B	1
MAC_C	2
MAC_D	2

主机向外发送数据时  
其MAC地址就会被学习





# 数据链路层交换原理



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 网桥更新MAC地址表的过程

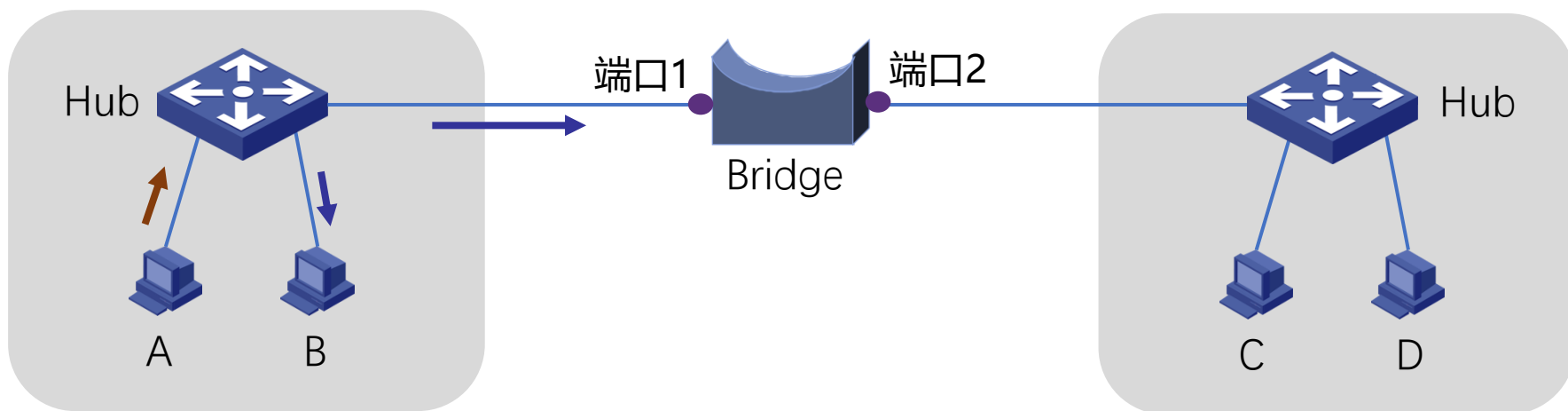
再次：A发出数据帧



MAC地址表	
MAC地址	端口
MAC_A	1
MAC_B	1
MAC_C	2
MAC_D	2

网桥发现MAC\_A已在表中!

更新该表项的帧达到时间，  
重置老化时间





# 数据链路层交换原理

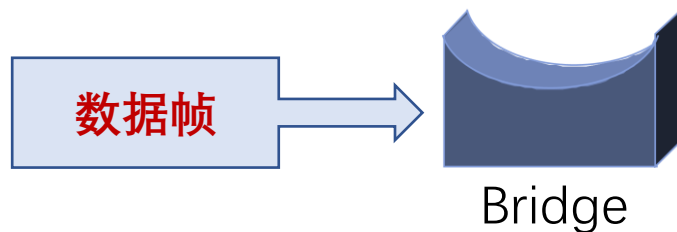


清华大学  
Tsinghua University



计算机网络教案社区

- 网桥通过逆向学习构建MAC地址表
- 网桥对于入境帧处理：forwarding、filtering、flooding



存储+转发  
不仅仅是转发

- Flooding（泛洪）
  - 两种目的地址的帧，需要泛洪：
    - 广播帧：目的地址为FF-FF-FF-FF-FF-FF的数据帧
    - 未知单播帧：目的地址不在MAC地址转发表中的单播数据帧



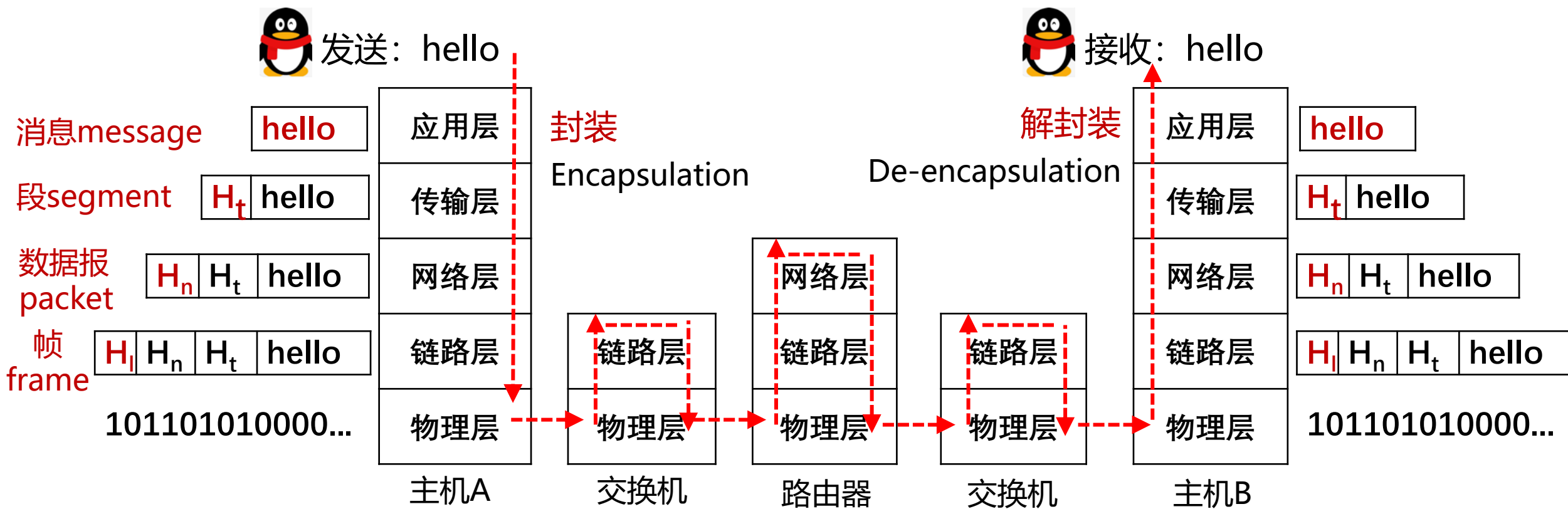
# 数据链路层交换原理



清华大学  
Tsinghua University



计算机网络教案社区



链路层的主要工作

成帧、错误检测、目的地址匹配、提交上层

**MAC表匹配和转发**



# 数据链路层交换原理



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 透明网桥工作原理（小结）

### • 逆向学习

- ✓ 增加表项：帧的源地址对应的项不在表中
- ✓ 删除表项：老化时间到期
- ✓ 更新表项：帧的源地址在表中，更新时间戳

### • 入帧转发过程（三选一），查目的地址是否在MAC地址表中

- ✓ 如果有，且入境端口  $\neq$  出境端口，则从对应的出境端口转发帧
- ✓ 如果有，且入境端口  $=$  出境端口，则丢弃帧（即过滤帧）
- ✓ 如果没有，则向除入境端口以外的其它所有端口泛洪帧



# 主要内容与学习目标



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 主要内容

- 经典以太网
- 链路层交换原理
- 生成树协议
- 链路层交换机
- 虚拟局域网

## ➤ 学习目标

- 了解经典以太网的帧结构
- 掌握链路层交换原理
- 掌握生成树协议与MAC表维护
- 掌握交换机的三种工作模式
- 了解虚拟局域网的使用场景和实现方法



# 生成树协议

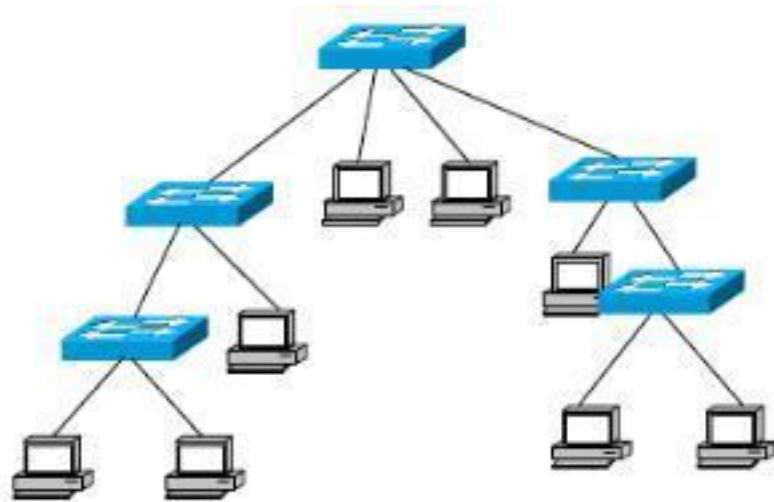


清华大学  
Tsinghua University



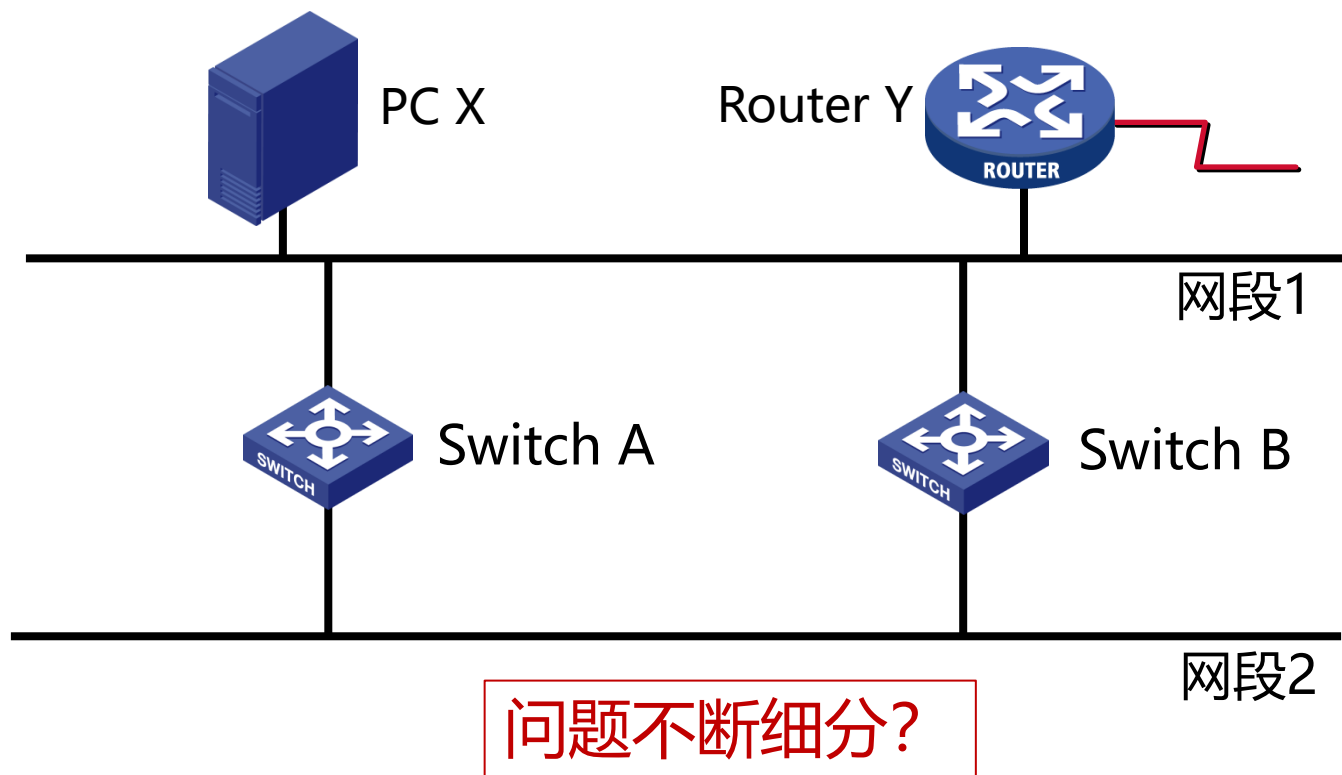
计算机网络教案社区

如何用多台交换机  
连接成大网络？



树型拓扑结构  
优缺点分析？

- 可靠的网络：需要冗余拓扑
- 可能产生的问题：导致物理环路







# 生成树协议



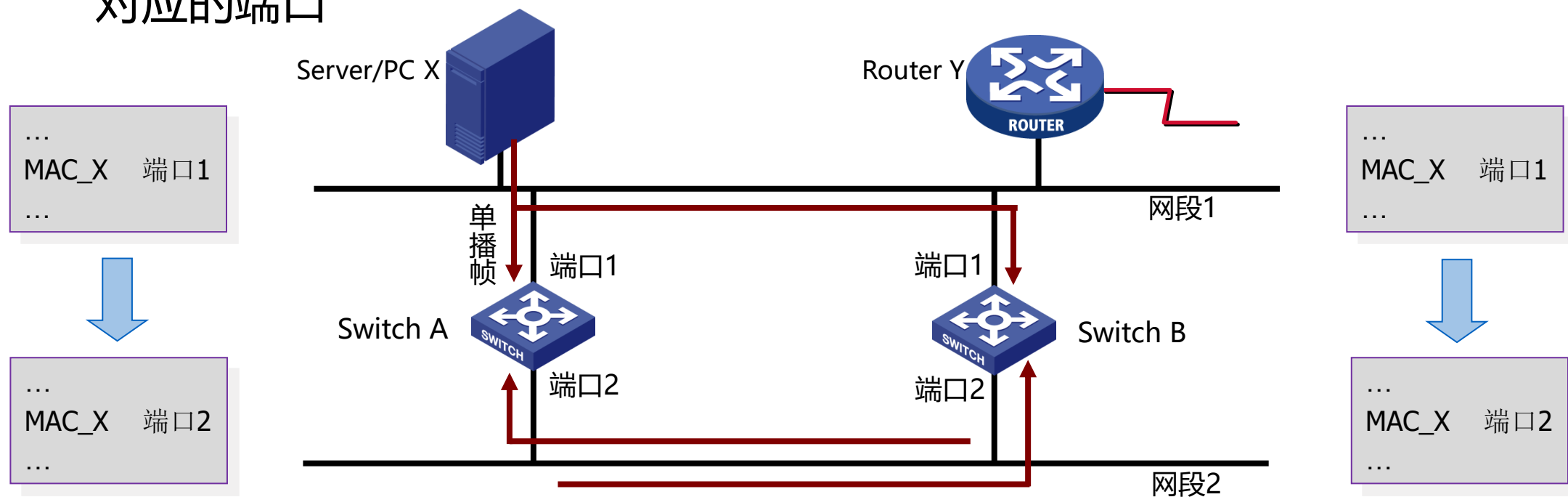
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 物理环路引发的**问题1**：MAC地址表不稳定

- 当一个帧的多个副本到达不同端口时，交换机会不断修改同一MAC地址对应的端口





# 生成树协议



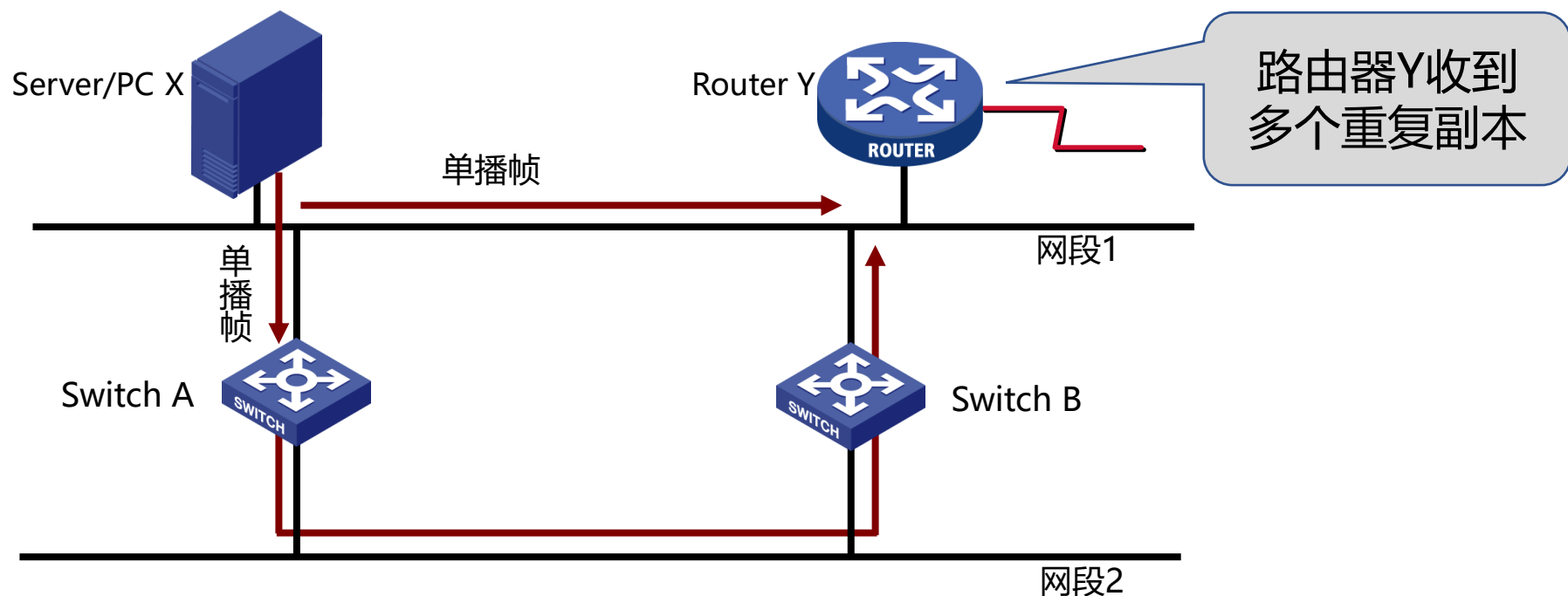
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 物理环路引发的**问题2**：重复帧

- X发送到环路的单播帧，造成目的设备Y收到重复的帧



假设所有交换机的MAC地址表中均没有路由器Y的MAC地址



# 生成树协议



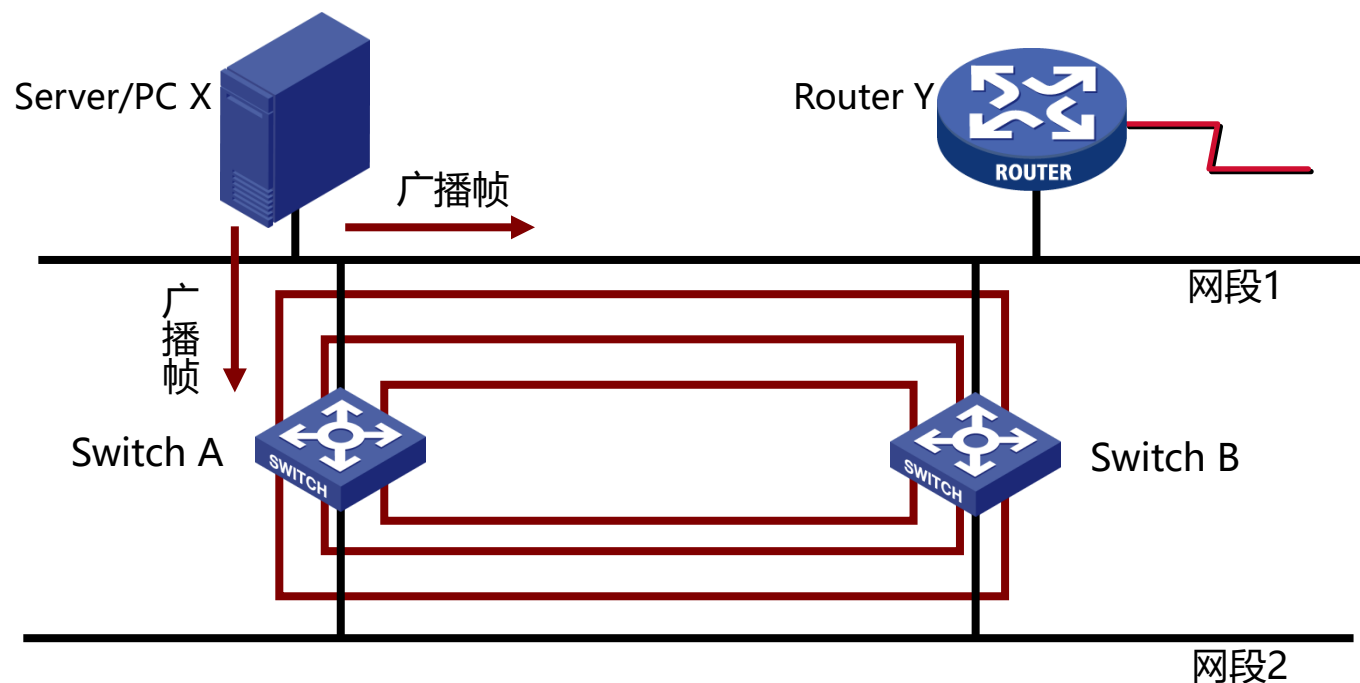
清华大学  
Tsinghua University



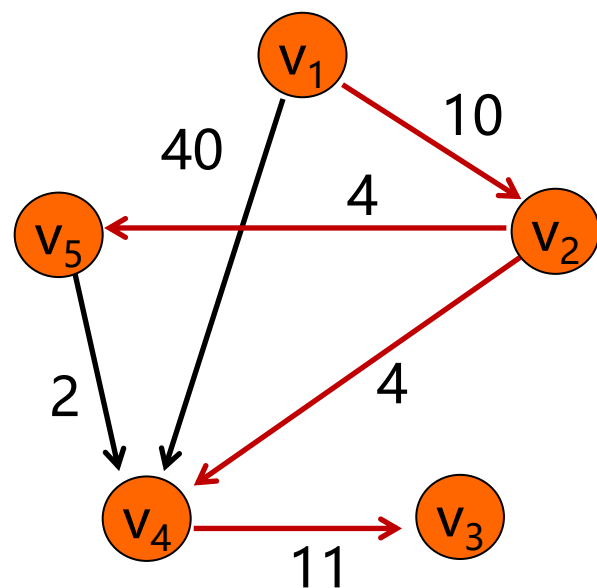
计算机网络教案社区

## ➤ 物理环路引发的**问题3**：广播风暴

- 交换机（网桥）在物理环路上无休止地泛洪广播流量，无限循环，迅速消耗网络资源



物理环路的解决思路？



自动构造生成树/支撑树



# 生成树协议



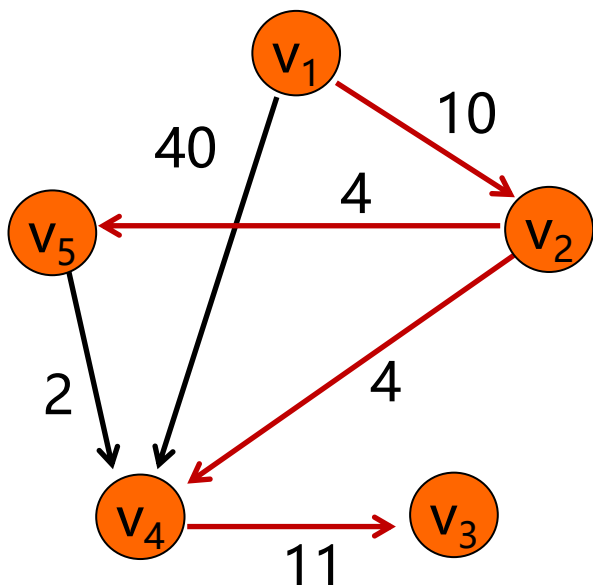
清华大学  
Tsinghua University



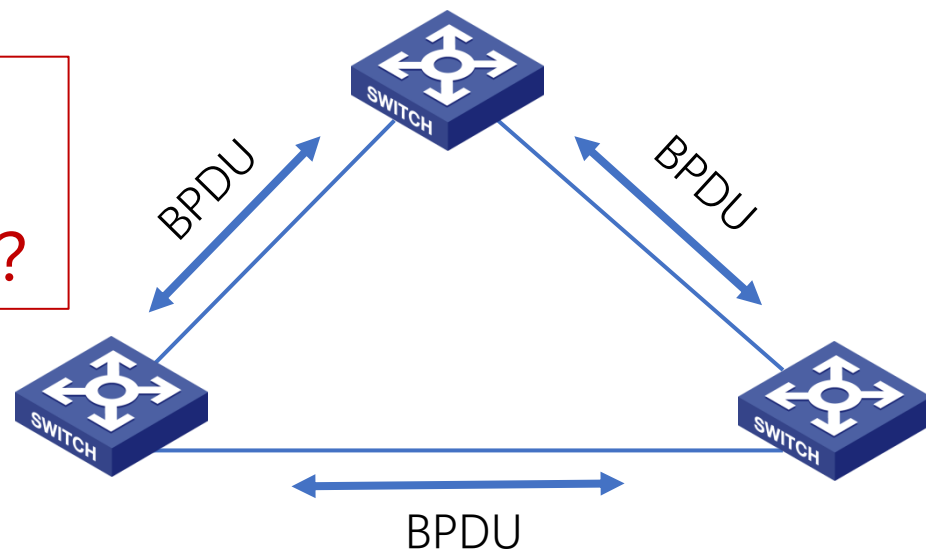
计算机网络教案社区

## ➤ 怎么得到一棵无环的生成树呢？

- 参与的交换机（网桥）：**收发**桥协议数据单元BPDU
- 选举产生**根桥**、**根端口**、**指定端口**，形成生成树



- 采用分布式&无配置方法
- 哪些端口工作？





# 生成树协议



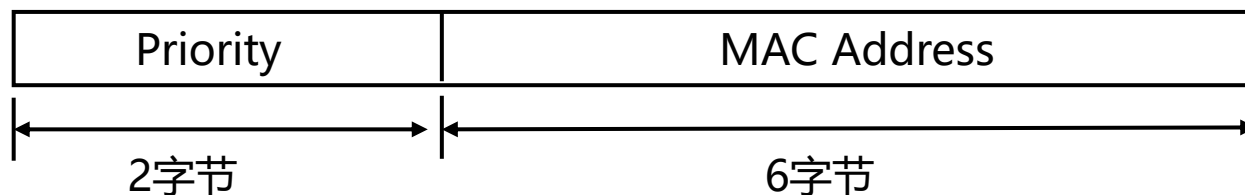
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 桥协议数据单元BPDU包含的四个关键信息

- **根桥ID** (Root ID) : 被选为根的桥ID
  - 桥ID共8字节, 由2字节的优先级和6字节的MAC地址组成的



- **根路径开销** (Root Path Cost) : 到根桥的最小路径开销
- **指定桥ID** (Designated Bridge ID) : 生成和转发BPDU的桥ID
- **指定端口ID** (Designated Port ID) : 发送BPDU的端口ID



# 生成树协议



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 生成树的三个选举过程

(1) 选举**根桥**(Root Bridge)

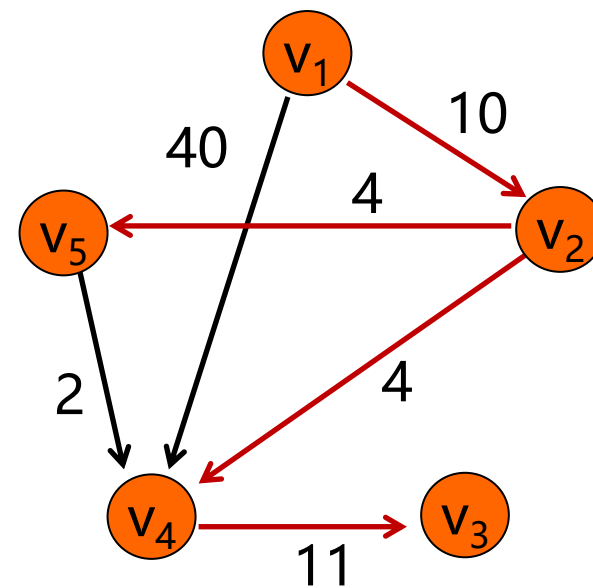
整个局域网的根 (树根)

(2) 为每个非根桥选出一个**根端口**  
(Root Port)

该非根桥使用选出的这个根端口去根桥

(3) 为每个网段确定一个**指定端口**  
(Designated Port)

即**代表端口**，代表这个网段向其他网段转发数据





# 生成树协议



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 生成树的选举过程1：选举根桥

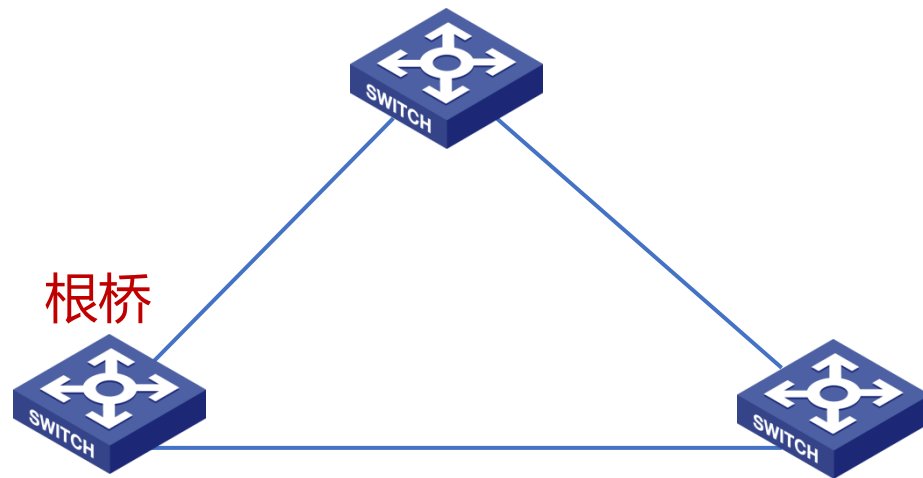
- 同一广播域中的所有交换机均参与选举
- 在给定广播域内只有一个根桥，其它均为非根桥
- 根桥的所有端口都处在转发状态

可以接收和发送数据帧

如何选举根桥？

- 首先比较优先级，**优先级**数值最小的交换机胜出成为根桥
- 如果优先级数值相等，**MAC地址**最小的交换机成为根桥

Bridge ID  
优先级： 32769  
MAC地址： 00-0A-00-11-11-11



Bridge ID  
优先级： 24577  
MAC地址：  
00-0A-00-33-33-33

Bridge ID  
优先级： 32769  
MAC地址：  
00-0A-00-22-22-22



# 生成树协议

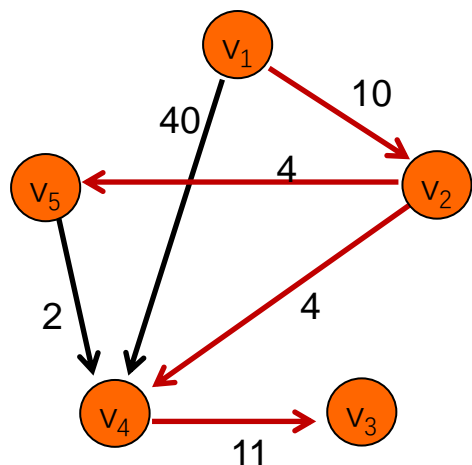


清华大学  
Tsinghua University

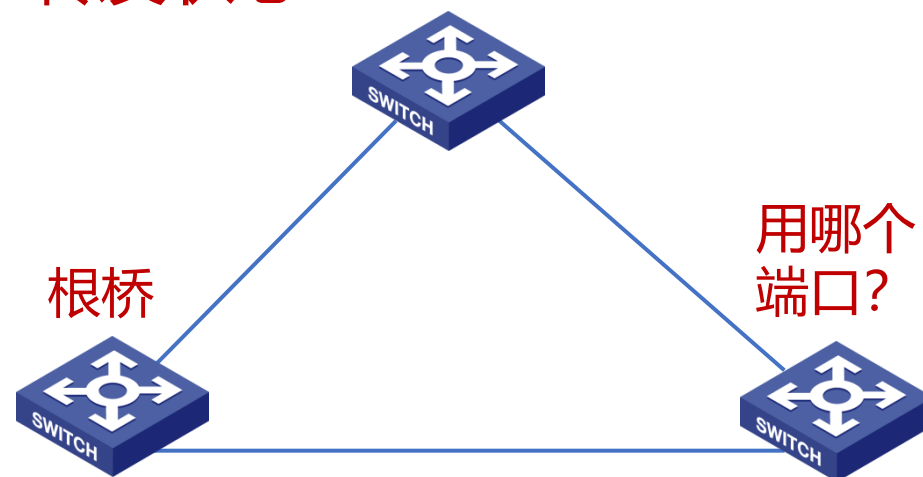


计算机网络教案社区

- 生成树的选举过程2：为每个非根桥选出一个根端口
- 每个非根桥，通过**比较**其每个端口到根桥的**根路径开销**，选出根端口
  - 具有最小根路径开销的端口被选作根端口
  - 如果多个端口的根路径**开销相同**，则**端口ID最小**的端口被选作根端口
  - 非根桥只能有一个根端口，根端口处于**转发状态**



如何定义  
路径开销  
???







# 生成树协议



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 什么是根路径开销？

- 端口（链路）开销值由IEEE定义（如下表），也可通过手工配置改变
- 根桥的根路径开销为0
- 非根桥的根路径开销为到根桥的路径上所有端口（链路）开销之和

速率值	开销 (IEEE802.1D-1998)
10Mbps	100
100Mbps	19
1Gbps	4
10Gbps	2
> 10Gbps	1



# 生成树协议

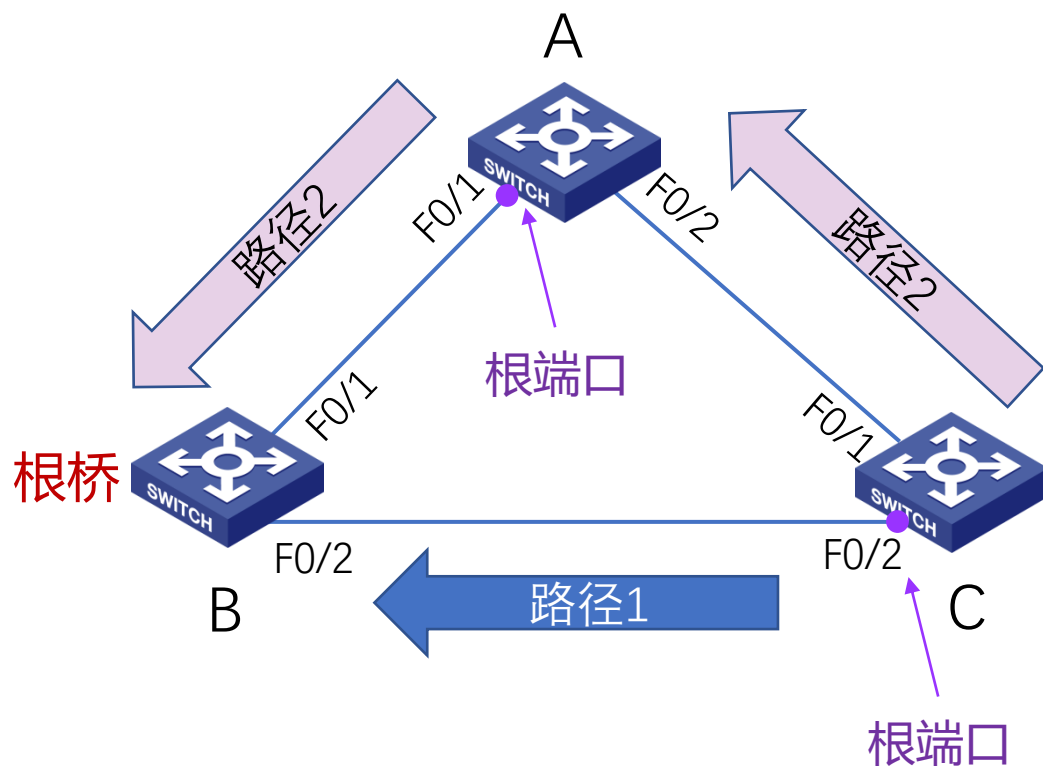


清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 根端口选举实例



图中交换机端口速率均为100Mb/s

即每跳链路开销均为19

C到达根桥有两条路径：

- 路径1开销为19
- 路径2开销为19+19=38

路径1开销较小！

F0/2相比于F0/1到达根桥的根路径开销更小，因此F0/2为C的根端口



# 生成树协议



清华大学  
Tsinghua University



计算机网络教案社区

- 生成树的选举过程3：为每个网段确定一个指定端口
  - 对于每一个网段，在所有连接到它的交换机（网桥）端口中进行选择
  - 一个具有最小根路径开销的端口，作为该网段的指定端口
  - 指定端口处于转发状态，负责该网段的数据转发
  - 连接该网段的其他端口，若既不是指定端口，也不是根端口，则阻塞



# 生成树协议

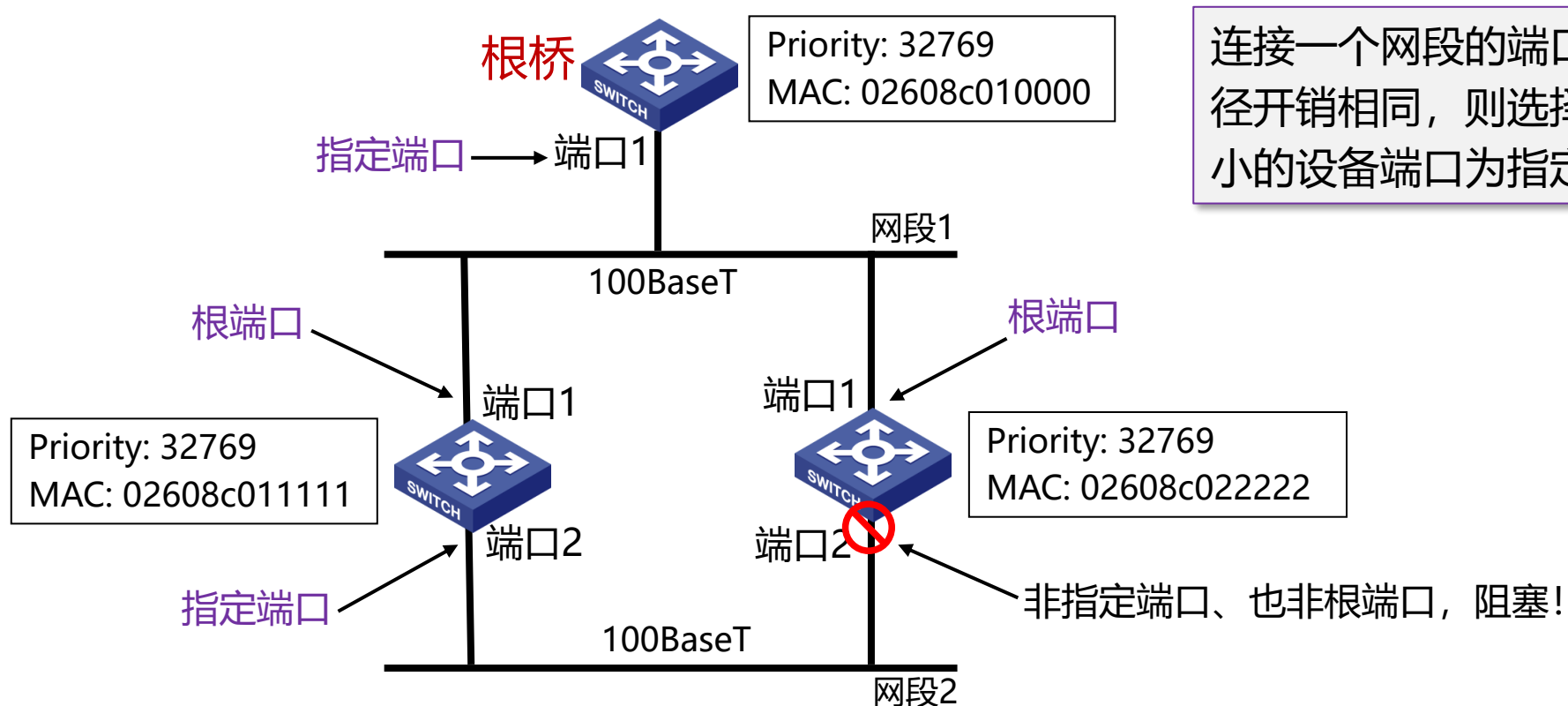


清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 指定端口选举实例





# 生成树协议



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 端口角色与端口状态

- 经过上述构造生成树的三个过程，端口角色便确定了

端口角色	英文名称	端口状态
指定端口	Designated port	Forwarding
根端口	Root port	Forwarding
非指定端口/根端口 (通常称为备用端口或冗余端口)	Alternate port	Blocking

判断题：根桥的所有端口都是指定端口吗？ 正确



# 生成树协议



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 端口角色与端口状态

- 事实上，802.1D标准给出了五种端口状态

端口角色	端口状态	端口行为
未启用STP功能的端口	Disabled	不收发BPDU报文，接收或转发数据
<b>非</b> 指定端口或根端口	Blocking	接收但不发送BPDU，不接收或转发数据
--	Listening	接收并发送BPDU，不接收或转发数据
--	Learning	接收并发送BPDU，不接收或转发数据
指定端口或根端口	Forwarding	接收并发送BPDU，接收并转发数据

如何适应网络的动态变化？



# 生成树协议



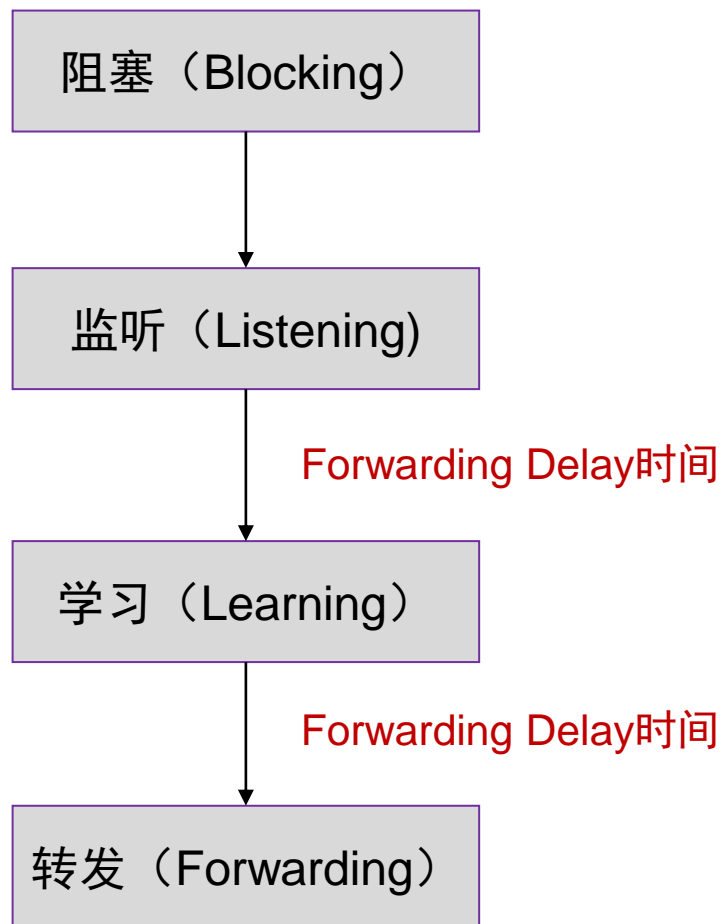
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 端口状态的迁移

- 端口角色确定为指定端口或根端口后，从Blocking状态经Listening和Learning才能到Forwarding状态
- 默认的Forwarding Delay时间是**15秒**
- 能保证当网络的拓扑发生改变时，新的配置信息能够传遍整个网络，从而避免由于网络为收敛而造成临时环路

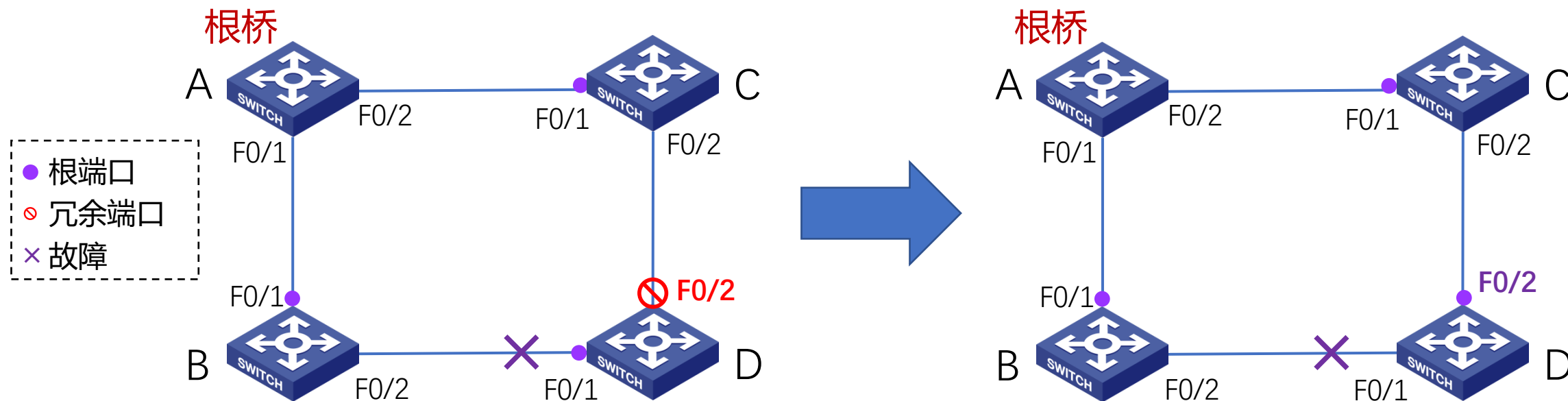




# 生成树协议

## ➤ 生成树的某“枝”断掉了，怎么办？

- 当由交换机（网桥）或链路故障导致网络拓扑改变时，重新构造生成树



交换机D的端口F0/2从Blocking状态到Forwarding状态至少要经过两倍Forwarding Delay时间，导致网络的连通性至少要30秒之后才能恢复





# 生成树协议

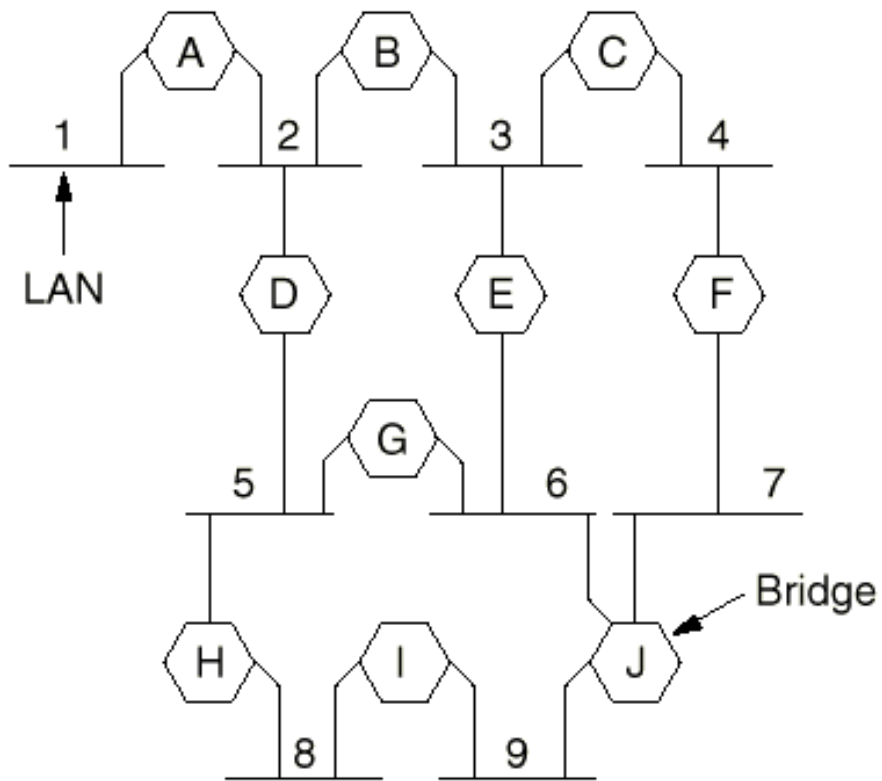
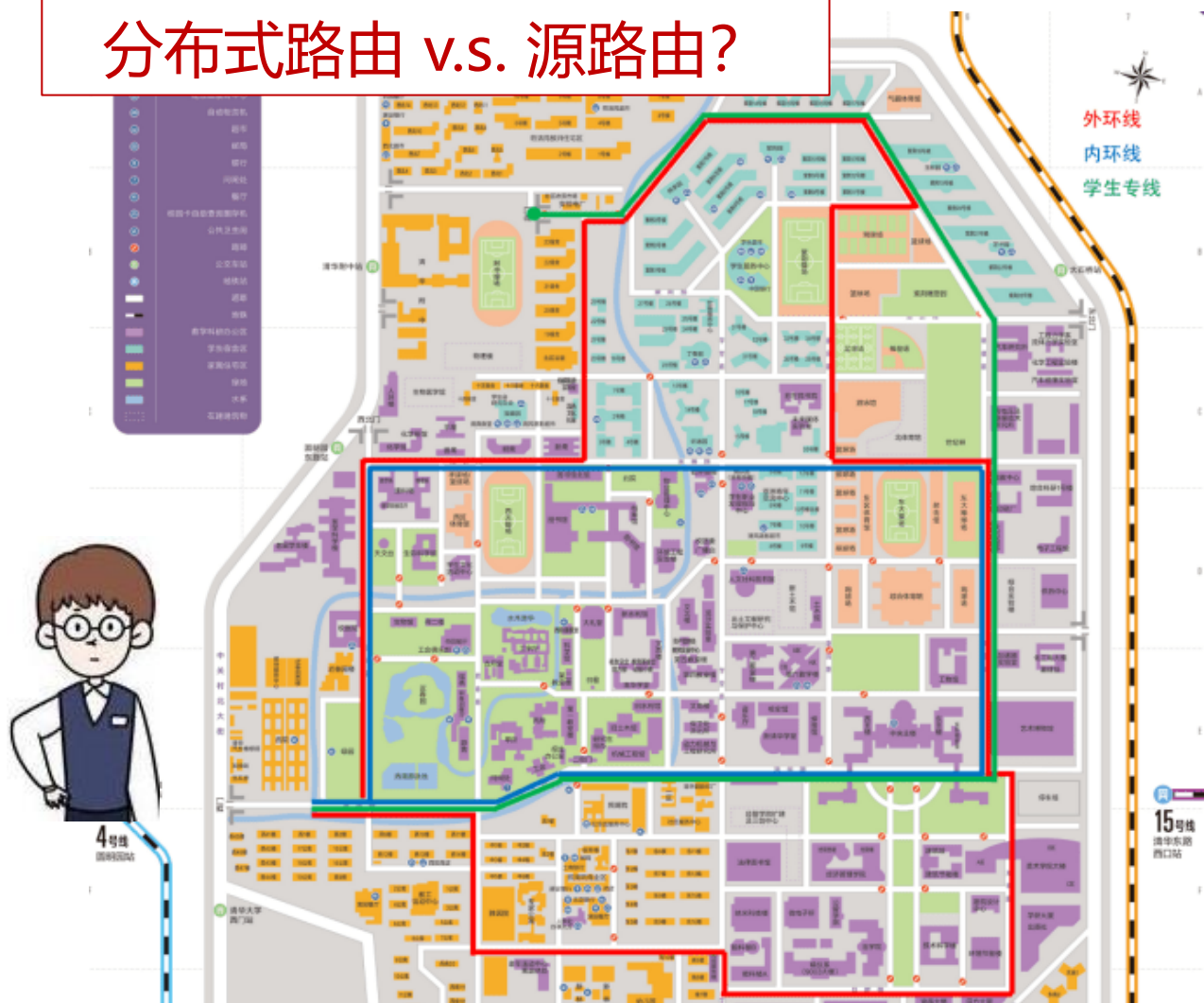


清华大学  
Tsinghua University



计算机网络教案社区

分布式路由 v.s. 源路由?



每个网桥独立维护MAC表，  
基于MAC表匹配独立转发



# 生成树协议



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 发明人 Radia Perlman

- 1983年，发明生成树协议STP (Spanning Tree Protocol)
- 分布式协同机制：避免复杂配置并能适应拓扑变化
- 打破物理环，维护逻辑无环的树
- 从数学到CS：David Clark学生
- 美国工程院院士、互联网名人堂入选者，Sigcomm终身成就奖
- 被人尊称为互联网之母



I think that I shall never see  
A graph more lovely than a tree.  
A tree whose crucial property  
Is **loop-free** connectivity.  
A tree which must be sure to span.  
So packets can reach every LAN.  
First the **Root** must be selected  
By **ID** it is elected.  
**Least cost paths** from Root are traced  
In the tree these paths are placed.  
A mesh is made by folks like me  
Then bridges find **a spanning tree**.

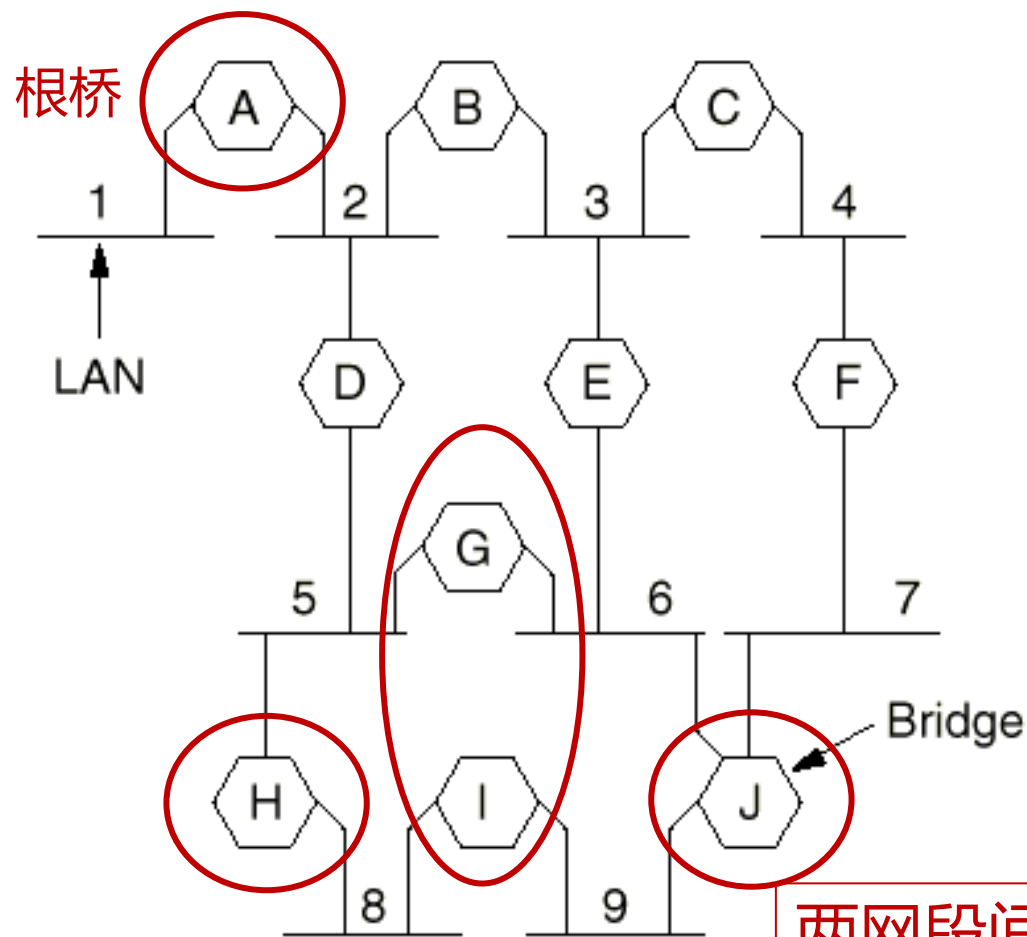


# 生成树协议



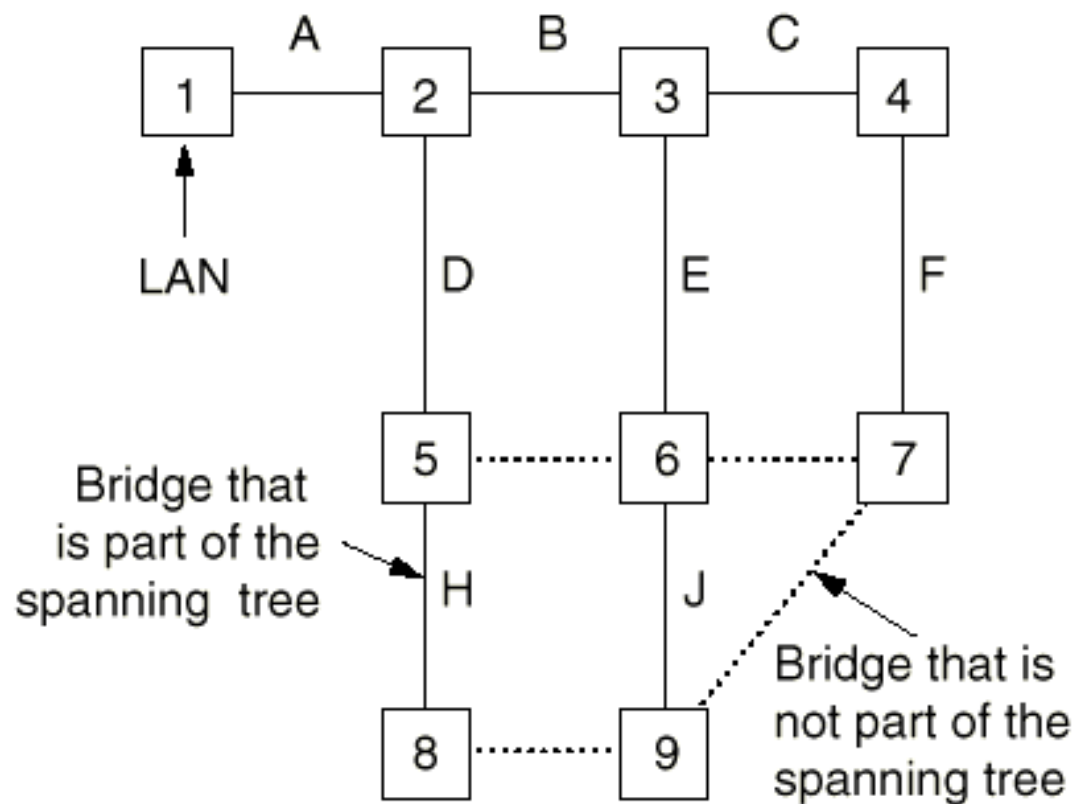
清华大学  
Tsinghua University

计算机网络教案社区



(a)

两网段间的路径是最短吗？  
当然不是，如何优化？



(b)



# 源路由网桥

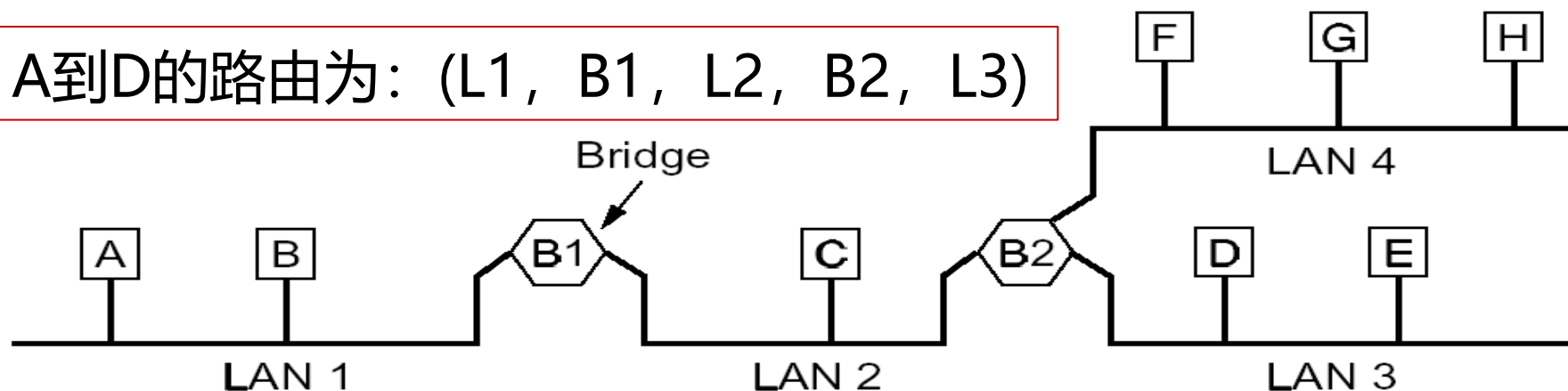


清华大学  
Tsinghua University



计算机网络教案社区

A到D的路由为: (L1, B1, L2, B2, L3)



## ➤ 源路由 (Source Routing) 网桥的需求分析

如何设计实现?

- **目的地定位**: 发送者需要知道目的主机是否在自己的LAN内, 以及在哪里
- 如果不在, 在发出的帧头内构造一个准确的**路由序列**, 包含要经过的网桥、LAN的编号
- 网桥对接收到的帧, 决定是否采用源路由

设置: 帧的**源地址最高位置**为1

判定: 源地址的最高位为1的

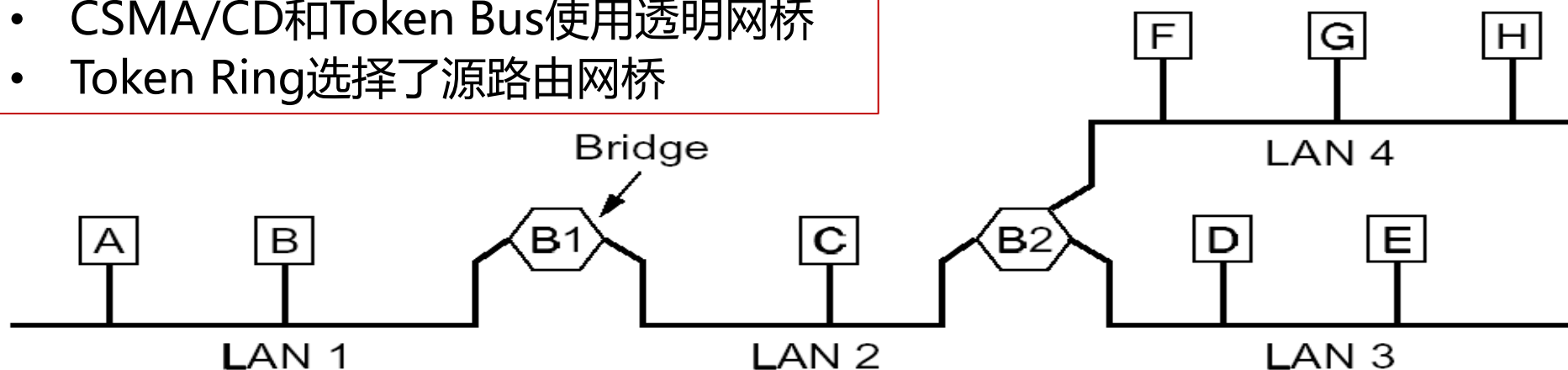


# 源路由网桥（续）



- 通过**广播“发现帧”**获得最佳路由
  - 源站若需要定位目的地，则发送“发现帧”，每个网桥收到后广播
  - 该帧经过网桥时被**加上桥标识（接口标识）**，目的站收到后发应答帧
  - 源站收到后就知道了到目的站的最佳路由
- 优点：对拓扑和带宽进行最优使用
- 缺点：网桥的插入对于网络是不透明的（要进行端-网协同）

- CSMA/CD和Token Bus使用透明网桥
- Token Ring选择了源路由网桥





# 网桥的比较



清华大学  
Tsinghua University



计算机网络教案社区

Issue	Transparent bridge	Source routing bridge
Orientation	Connectionless	Connection-oriented
Transparency	Fully transparent	Not transparent
Configuration	Automatic	Manual
Routing	Suboptimal	Optimal
Locating	Backward learning	Discovery frames
Failures	Handled by the bridges	Handled by the hosts
Complexity	In the bridges	In the hosts





# 主要内容与学习目标



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 主要内容

- 经典以太网
- 链路层交换原理
- 生成树协议
- 链路层交换机
- 虚拟局域网

## ➤ 学习目标

- 了解经典以太网的帧结构
- 掌握链路层交换原理
- 掌握生成树协议与MAC表维护
- 掌握交换机的三种工作模式
- 了解虚拟局域网的使用场景和实现方法



# 链路层交换机



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 交换机

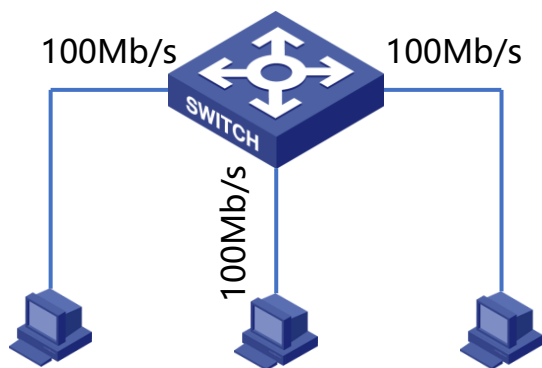
- 即插即用网络设备，执行数据链路层交换算法

## ➤ 交换方式：从带宽的角度

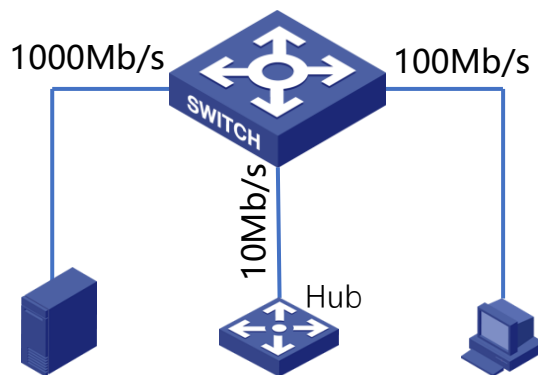
- 对称交换：出和入的**带宽相同**
- 非对称交换：出和入的**带宽不同**



交换机：多端口透明网桥



对称交换



非对称交换





# 链路层交换机



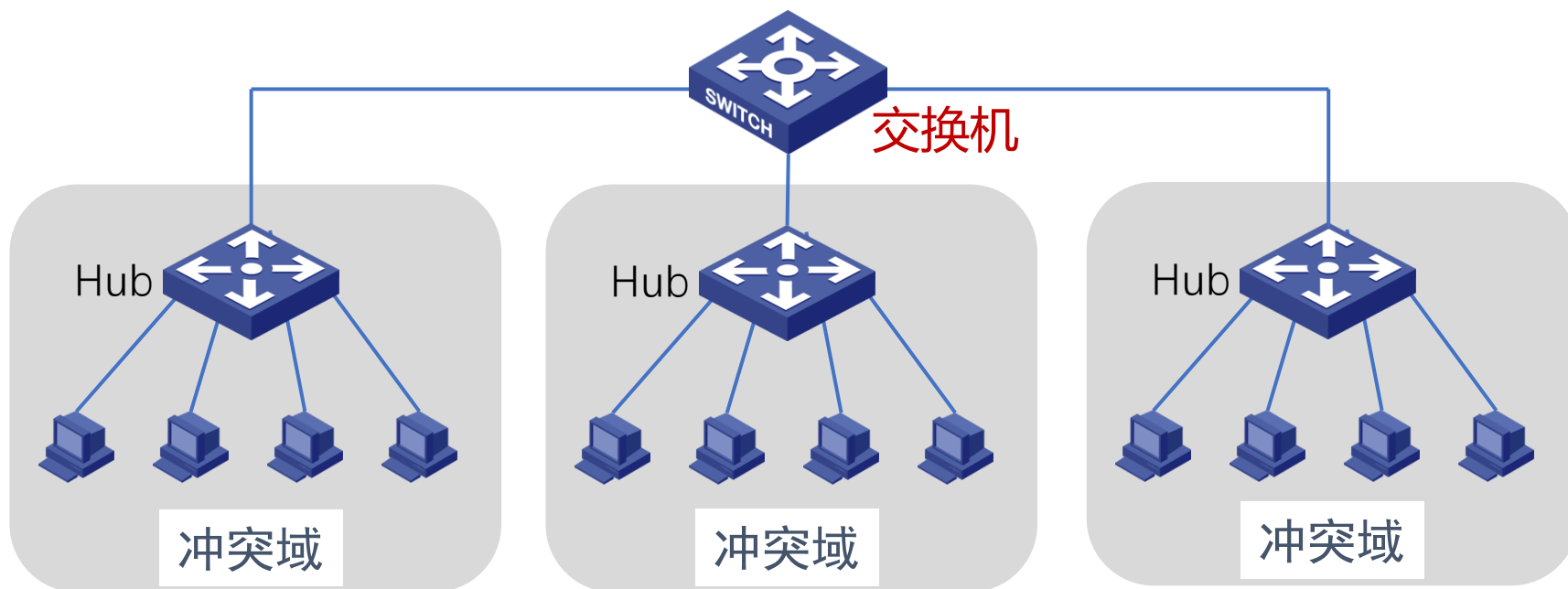
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 传统LAN分段

- 交换机端口通常与集线器连接
- 使用交换机把LAN分段为更小的冲突域





# 链路层交换机



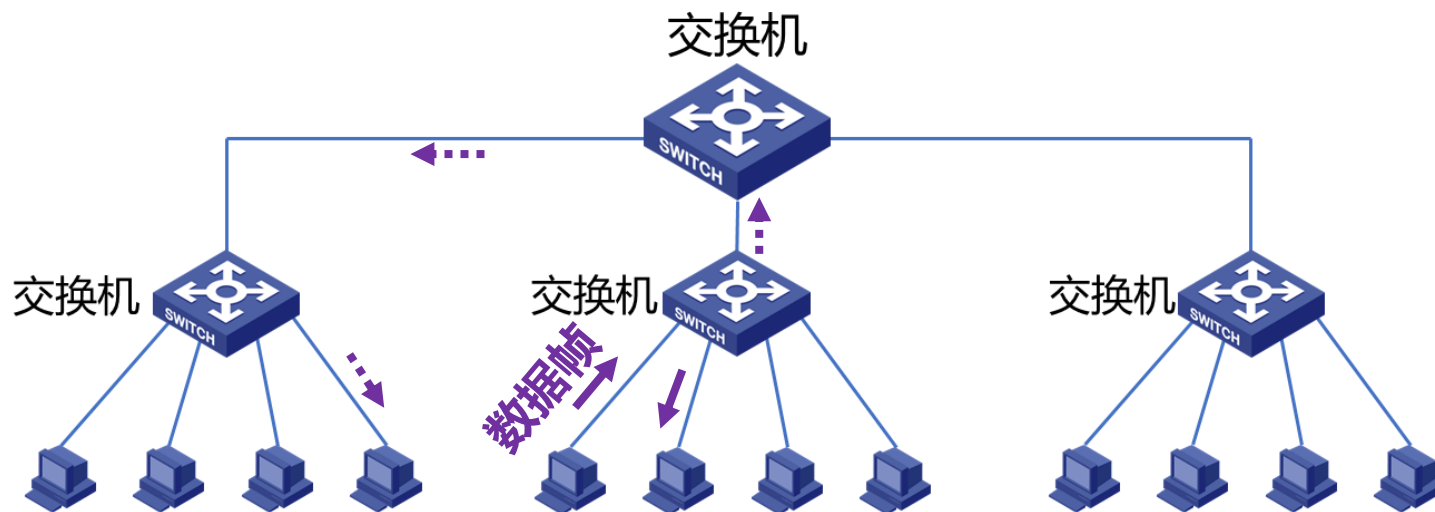
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 现代LAN分段

- 直连PC，微分段，创建无冲突域



优点

- 不发送不必发送的帧
- 发送时阻断冲突域
- 提升信道利用率

缺点

网络时延  
交换机比集线器  
增加了转发时延



# 链路层交换机



清华大学  
Tsinghua University



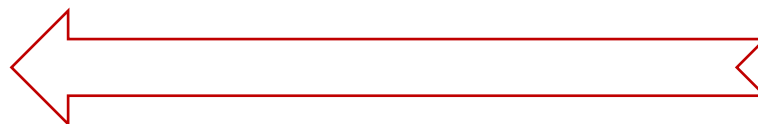
计算机网络教案社区

## ➤ 交换模式1：存储转发

- 特点：转发前必须接收**整个帧**、执行CRC校验
- 优点：不转发出错帧、支持非对称交换
- 缺点：转发延迟大

7字节	1字节	6字节	6字节	2字节	46~1500字节	4字节
Preamble	SFD	Destination	Source	Length/ Type	Data and Pad	FCS

能否减少转发延迟？



存储转发模式  
高延迟



# 链路层交换机



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 交换模式2：直通交换

- 特点：一旦接收到帧的目的地址，就开始转发
- 优点：延迟非常小，可以边入边出
- 缺点：可能转发错误帧，不支持非对称交换

7字节	1字节	6字节	6字节	2字节	46~1500字节	4字节
Preamble	SFD	Destination	Source	Length/ Type	Data and Pad	FCS

直通模式  
低延迟、无错误检查

如何优化？  
减少转发延迟 & 不转发错误帧



# 链路层交换机



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 交换模式3：无碎片交换

- **特点：**接收到**帧的前64字节**，即开始转发
- **优点：**过滤了冲突碎片，延迟和转发错帧介于存储转发和直通交换之间
- **缺点：**仍可能转发错误帧，不支持非对称交换

7字节	1字节	6字节	6字节	2字节	46~1500字节	4字节
Preamble	SFD	Destination	Source	Length/ Type	Data and Pad	FCS

帧的前64字节

无碎片模式  
较低延迟  
过滤冲突导致的碎片帧



# 链路层交换机



清华大学  
Tsinghua University



计算机网络教案社区

- 交换模式：从转发时机的角度
  - 1) 存储转发模式 (Store and Forward)
  - 2) 直通模式 (Cut-through)
  - 3) 无碎片模式 (Fragment-free)





# 链路层交换机



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 练习

在交换机的三种交换模式中，正常情况下( )模式的转发延迟最大。

- A、直通
- B、存储转发
- C、无碎片
- D、都一样





# 主要内容与学习目标



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 主要内容

- 经典以太网
- 链路层交换原理
- 生成树协议
- 链路层交换机
- 虚拟局域网

## ➤ 学习目标

- 了解经典以太网的帧结构
- 掌握链路层交换原理
- 掌握生成树协议与MAC表维护
- 掌握交换机的三种工作模式
- 了解虚拟局域网的使用场景和实现方法





# 虚拟局域网



清华大学  
Tsinghua University



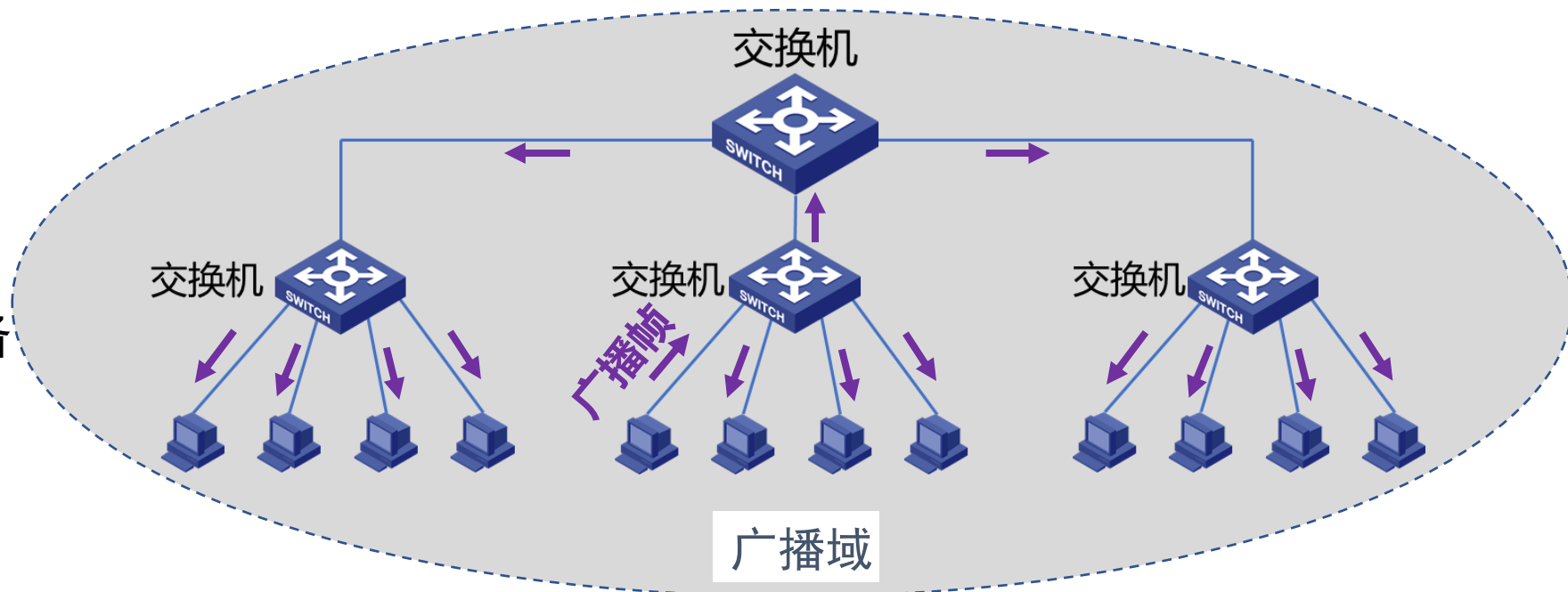
计算机网络教案社区

## ➤ 广播域 (Broadcasting Domain)

- 广播域是广播帧能够到达的范围
- 缺省情况下，交换机所有端口同属于一个广播域，无法隔离广播域
- 广播帧在广播域中传播，占用资源，降低性能，且具有安全隐患

不是冲突域！

图中某个站点  
发送了一个广播帧  
能够收到该广播帧的设备  
同处于一个广播域





# 虚拟局域网



清华大学  
Tsinghua University



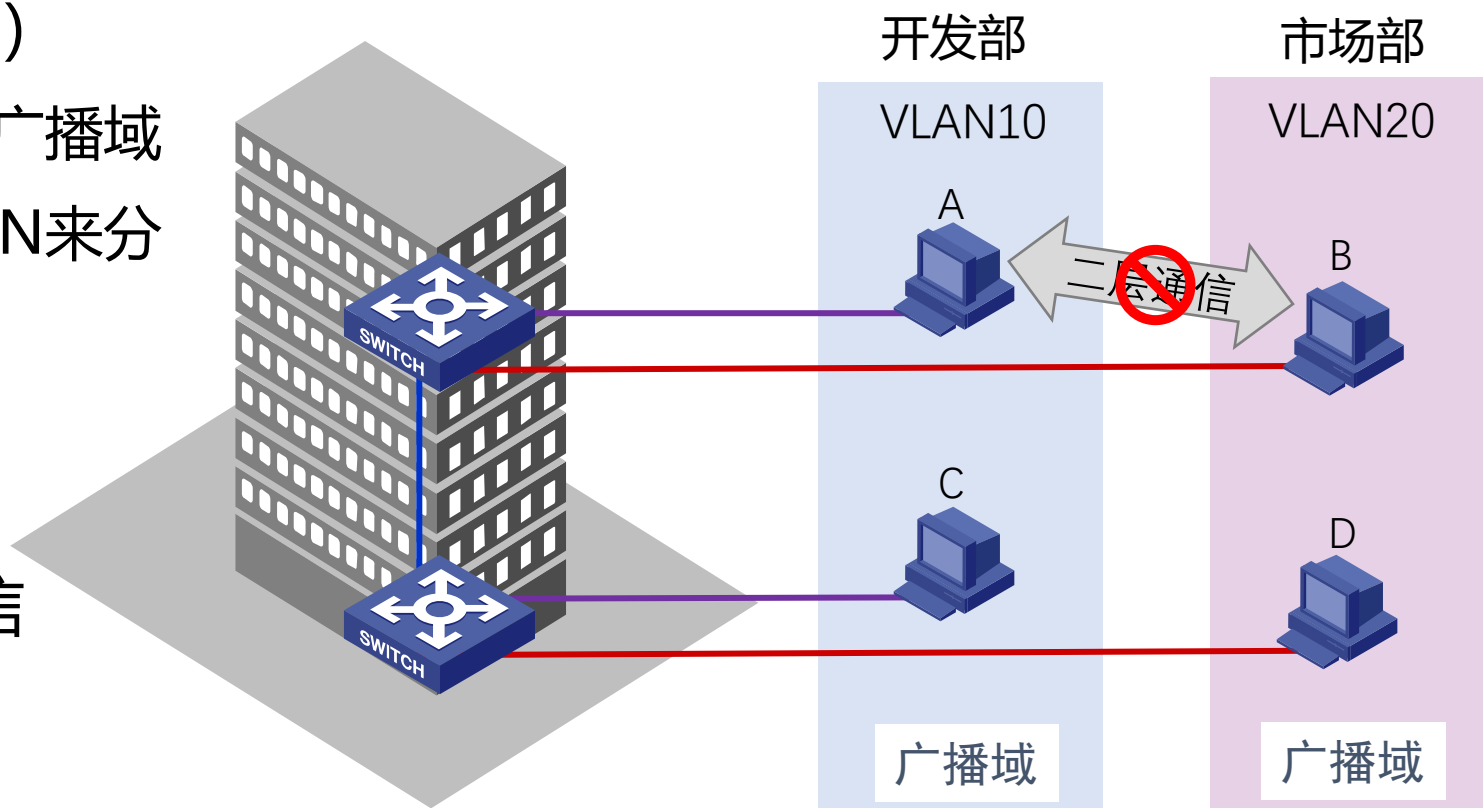
计算机网络教案社区

- 需求分析：需要在物理网络上根据用途、工作组、应用等来进行**逻辑划分**的局域网，与用户的物理位置没有关系

- VLAN (Virtual LAN)

- VLAN是一个独立的广播域
- 交换机通过划分VLAN来分隔广播域

不同VLAN的成员  
不能直接进行二层通信





# 虚拟局域网



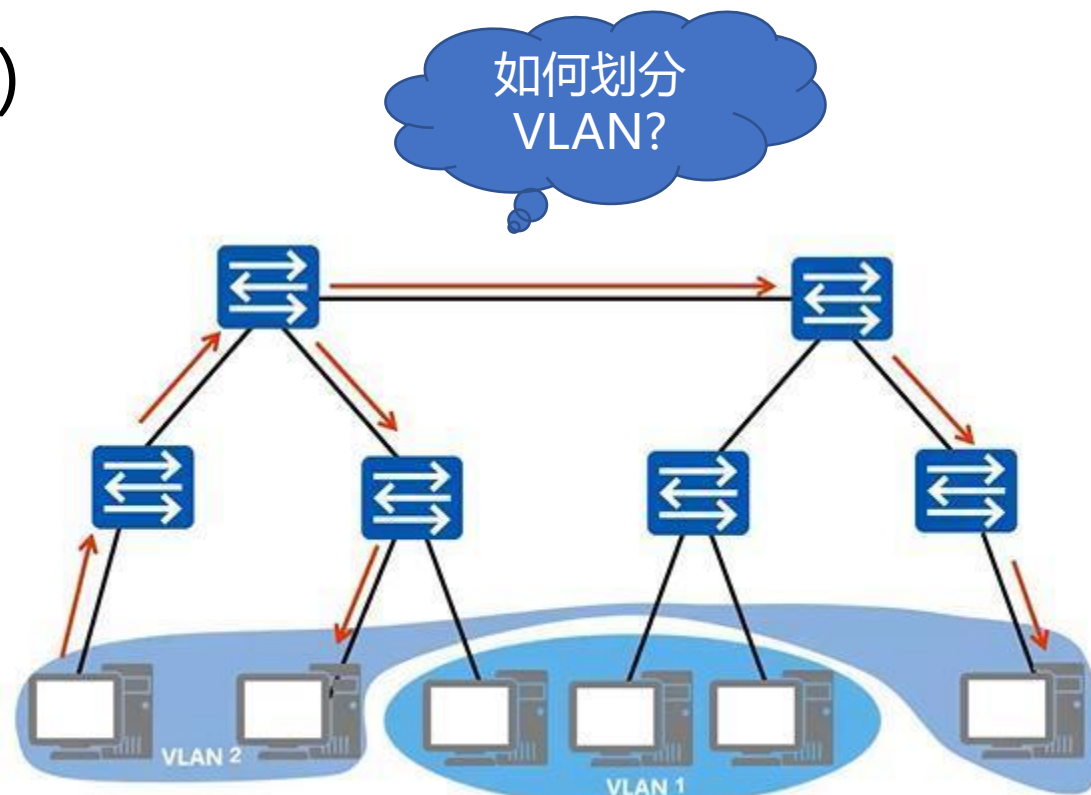
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ VLAN类型

- 基于端口的VLAN (**最常见**)
- 基于MAC地址的VLAN
- 基于协议的VLAN
- 基于子网的VLAN





# 虚拟局域网



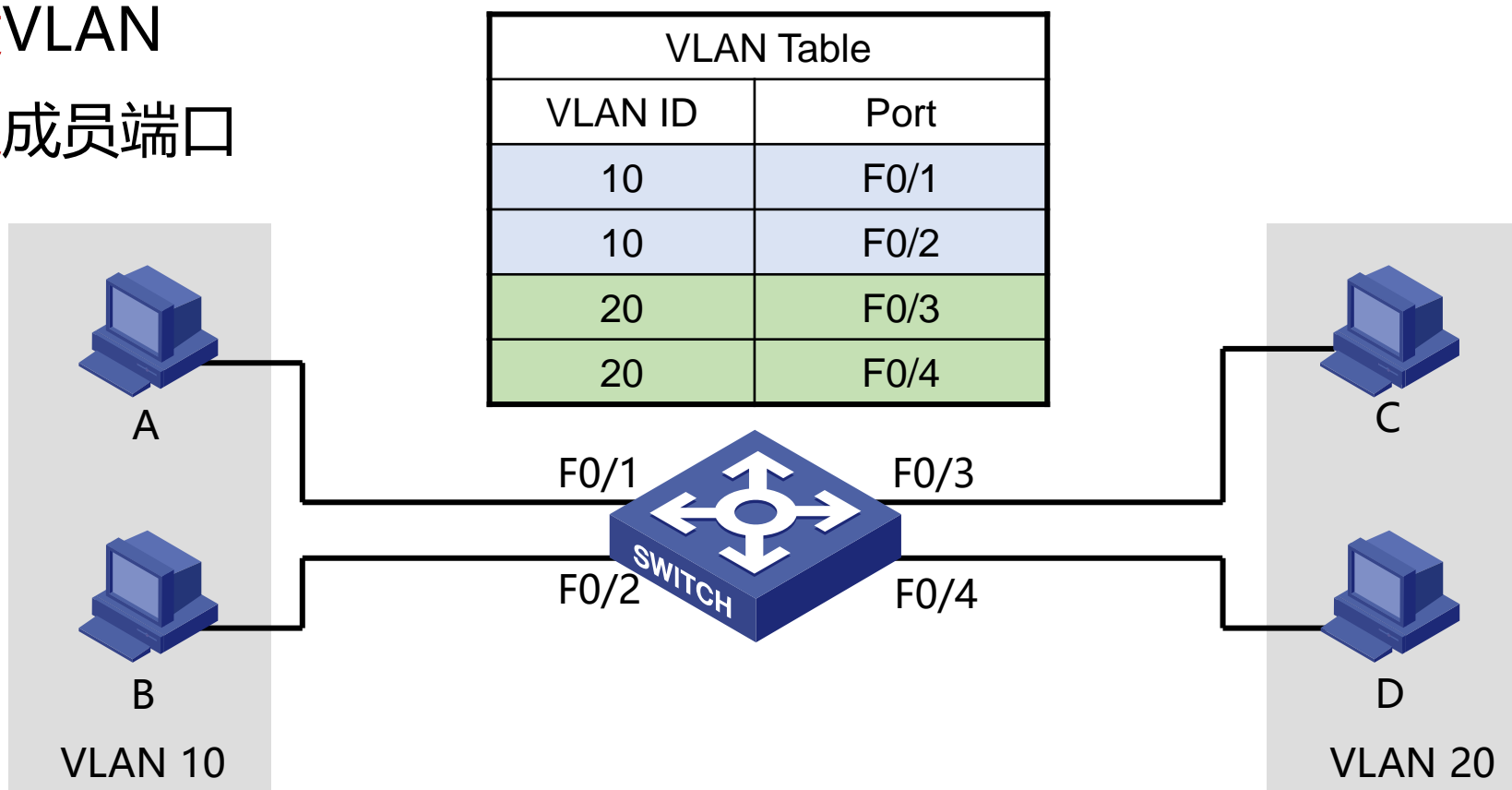
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 基于端口的VLAN

- 创建VLAN
- 指定成员端口





# 虚拟局域网

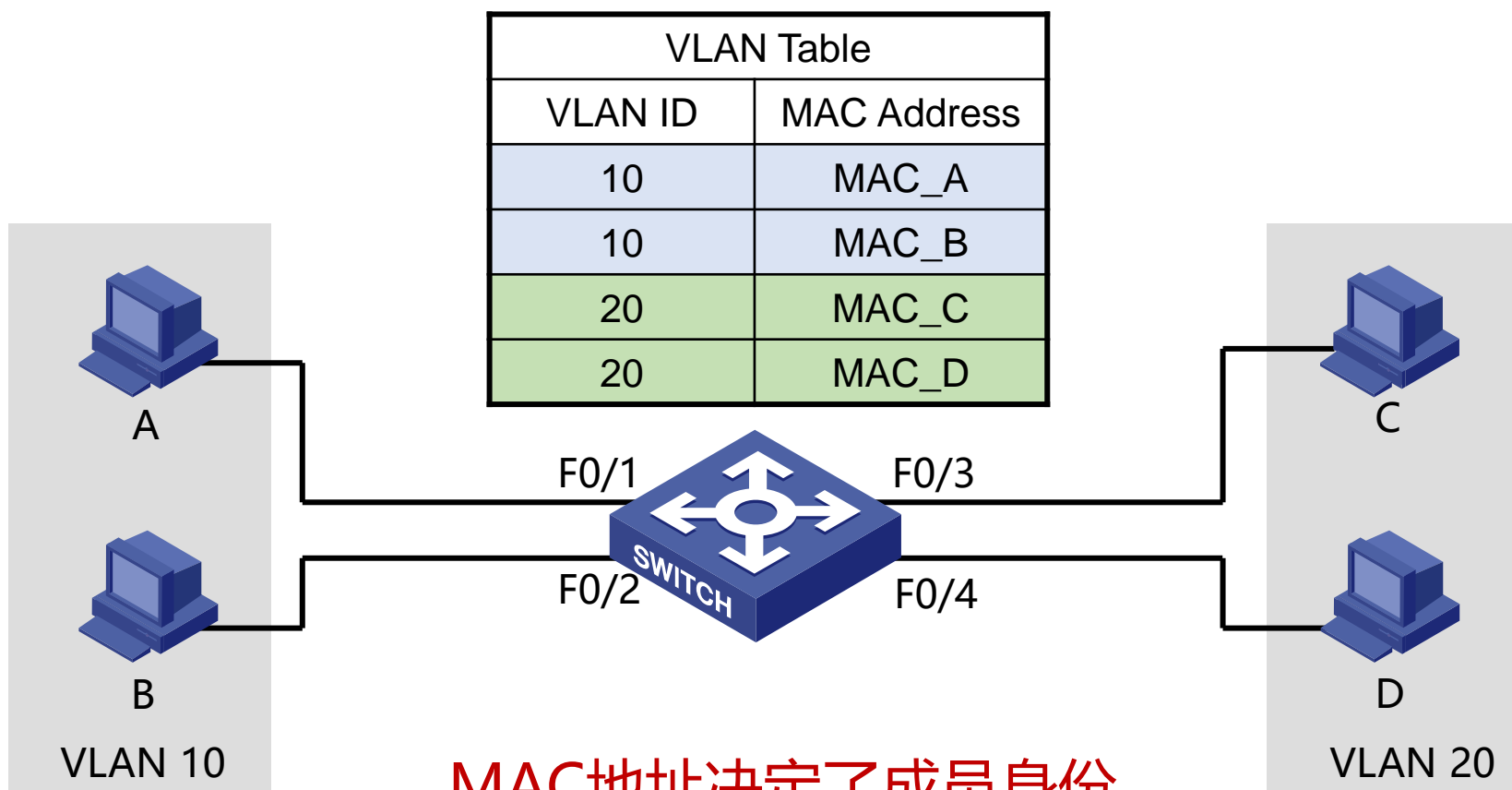


清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 基于MAC地址的VLAN





# 虚拟局域网

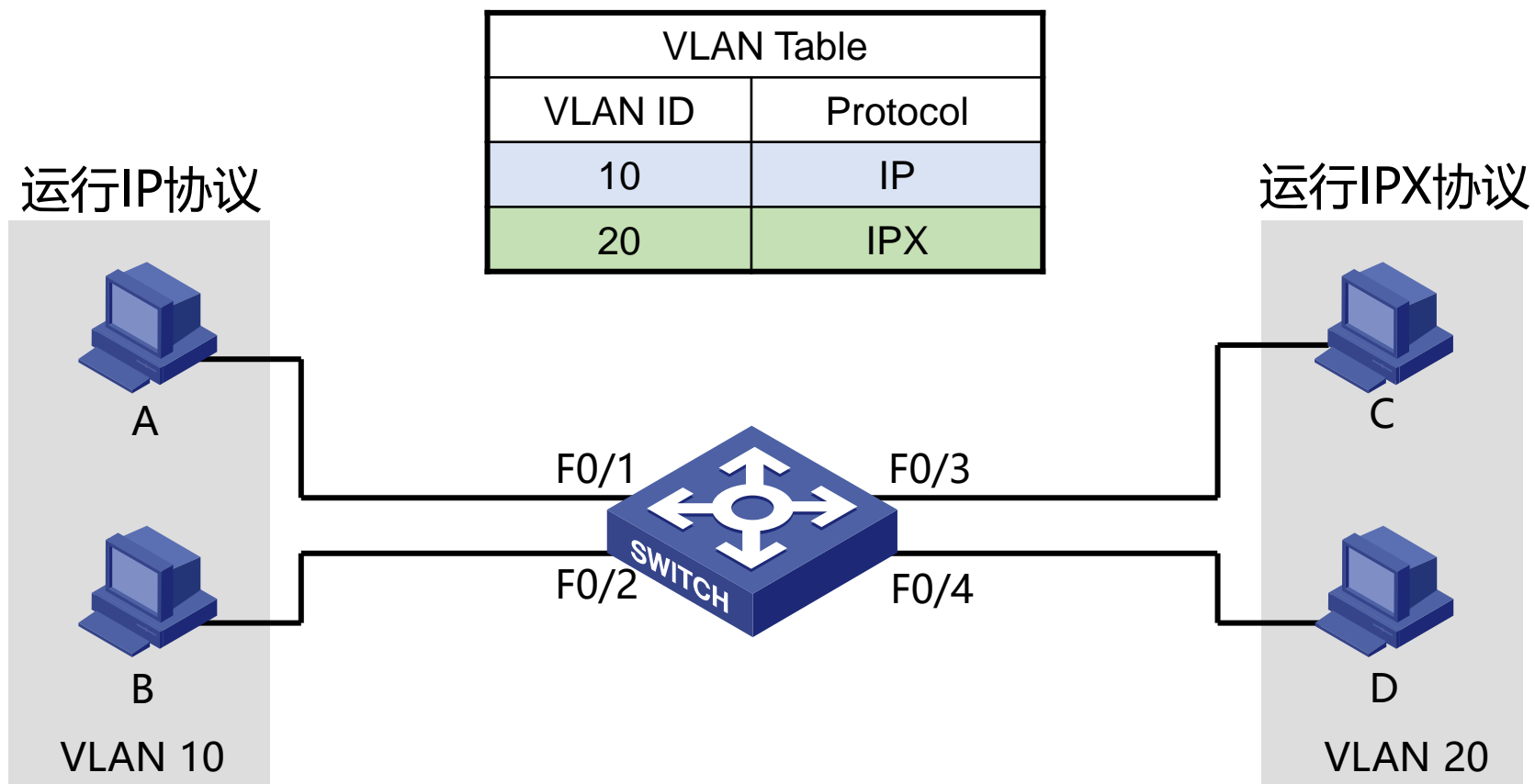


清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 基于协议的VLAN





# 虚拟局域网

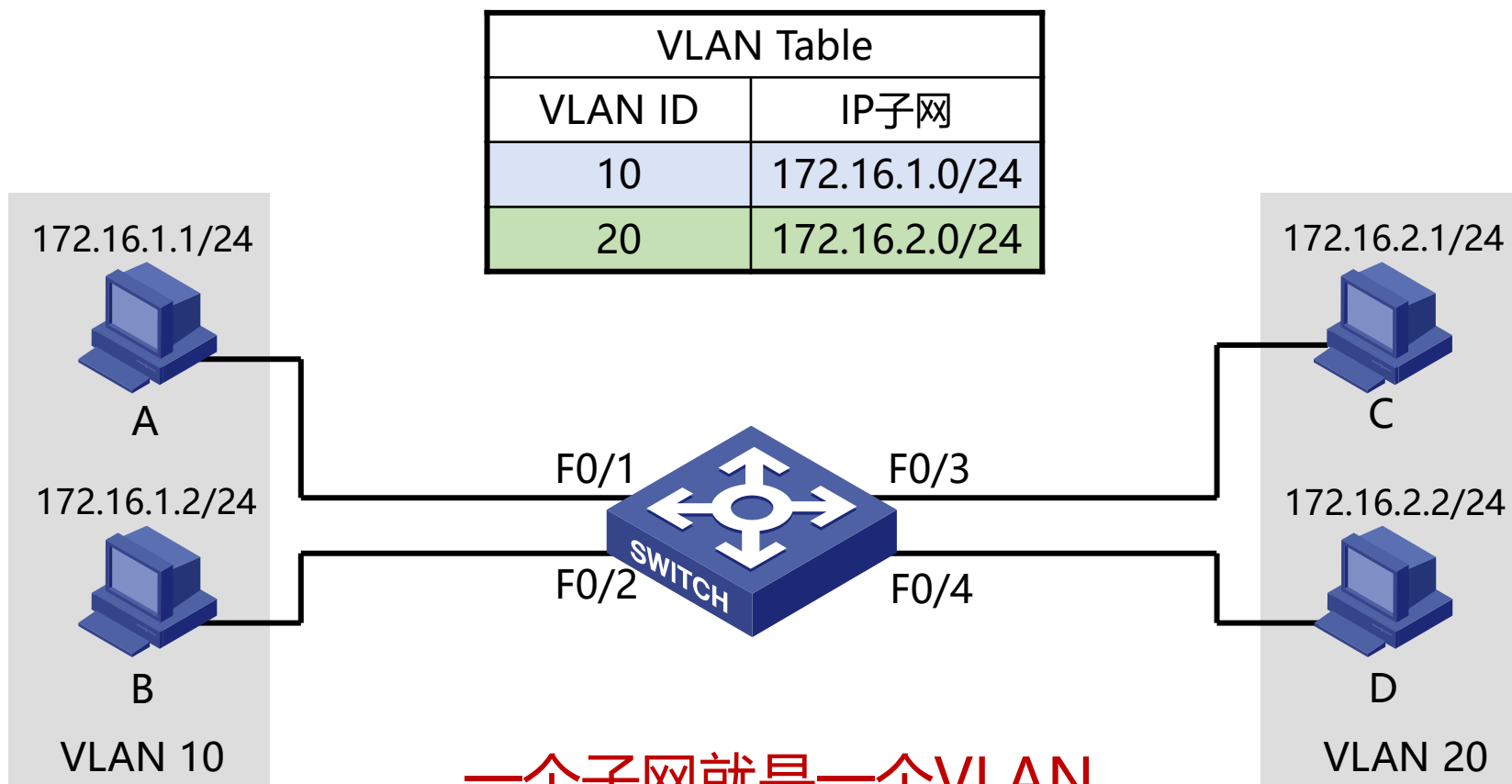


清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 基于子网的VLAN



一个子网就是一个VLAN



# 虚拟局域网



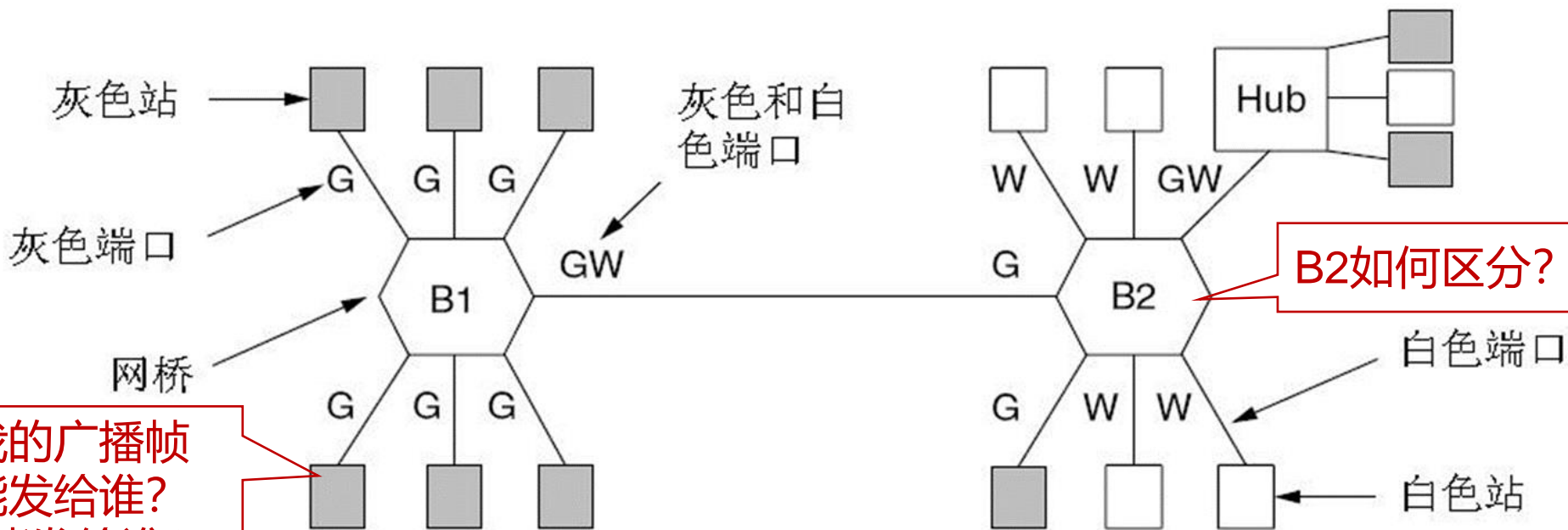
清华大学  
Tsinghua University



计算机网络教案社区

- 如何能让交换机区分不同VLAN的数据帧？
  - 在数据帧中需要携带VLAN标记
  - VLAN标记由交换机添加/剥离，对终端站点透明

交换机需要区分VLAN，  
但希望对主机透明







# 虚拟局域网



清华大学  
Tsinghua University

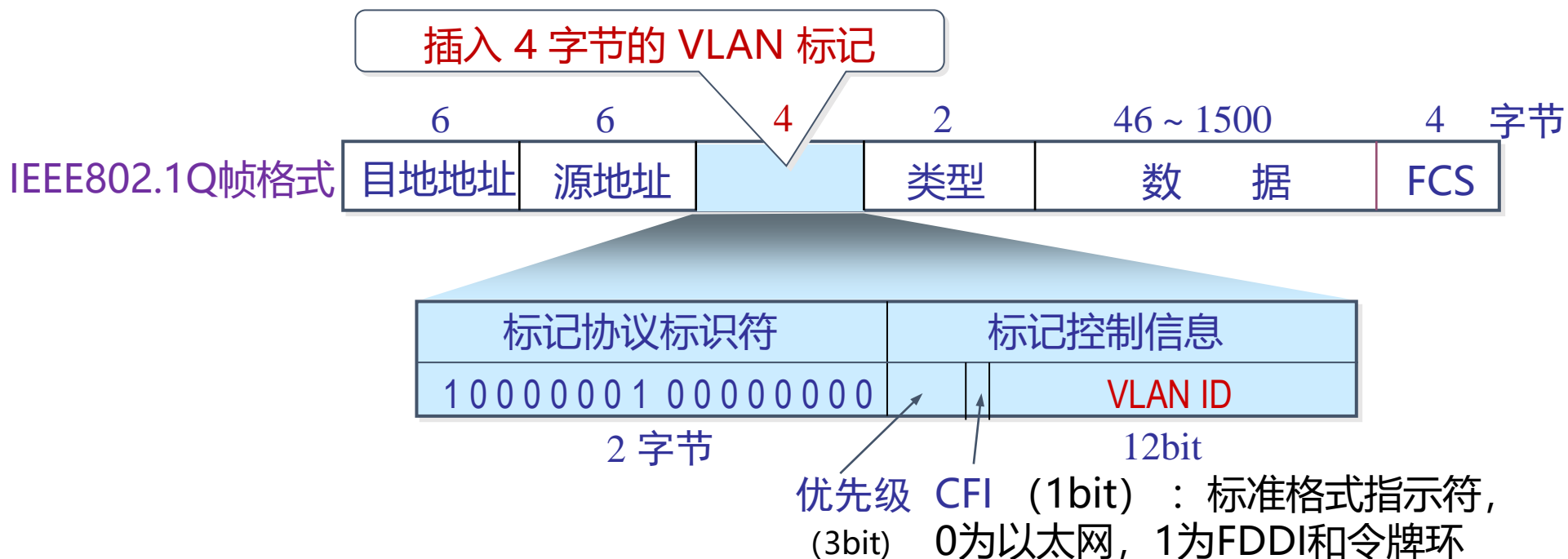


计算机网络教案社区

## ➤ 帧标记标准：IEEE802.1Q

- 带VLAN标记的帧称为**标记帧** (Tagged Frame)
- 不携带VLAN标记的普通以太网帧称为**无标记帧** (Untagged Frame)

交换机需要区分VLAN，  
但希望对主机透明





# 虚拟局域网



清华大学  
Tsinghua University

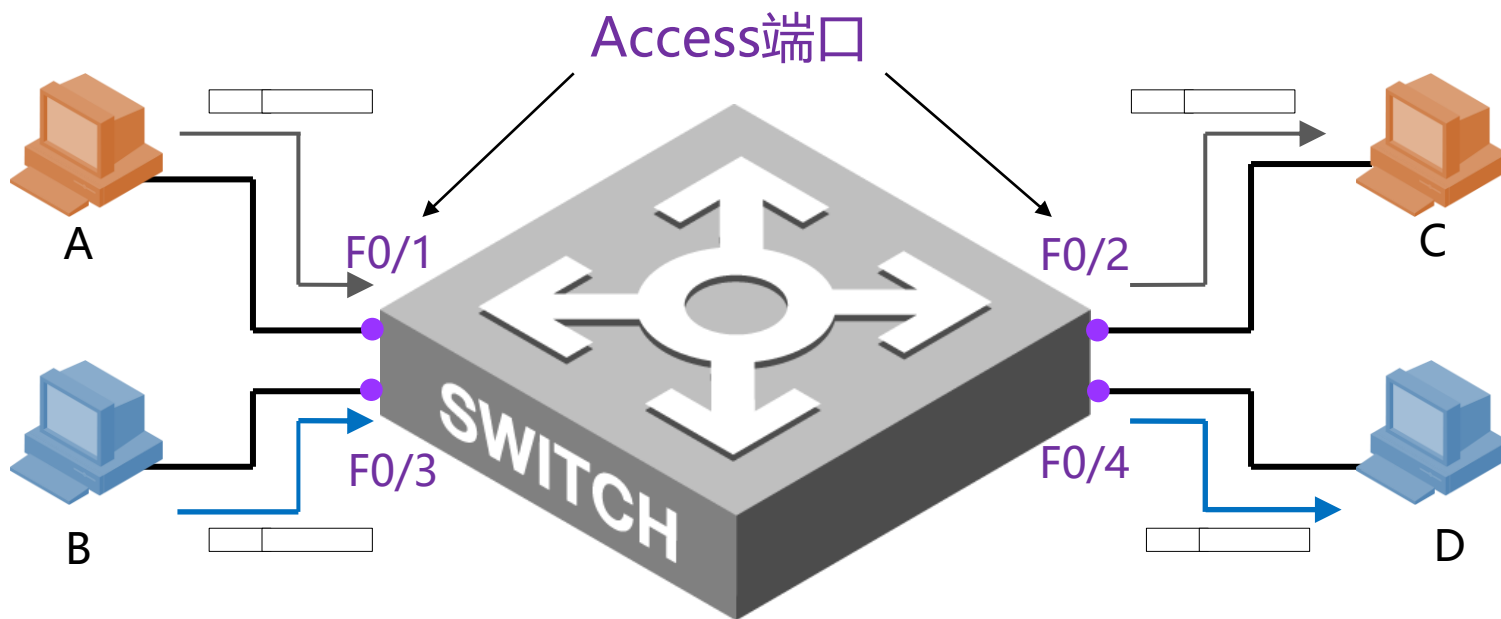


计算机网络教案社区

## ➤ 接入链路（Access）类型端口

- 交换机连接主机的端口
- 在单交换机环境中，所有端口都是Access类型端口
- 对采用基于端口的VLAN划分, Access端口只能加入一个VLAN
- 一旦Access端口加入了特定的VLAN，连接在该端口的设备被视为属于该VLAN

VLAN Table	
VLAN ID	Port
10	F0/1
10	F0/2
20	F0/3
20	F0/4





# 虚拟局域网



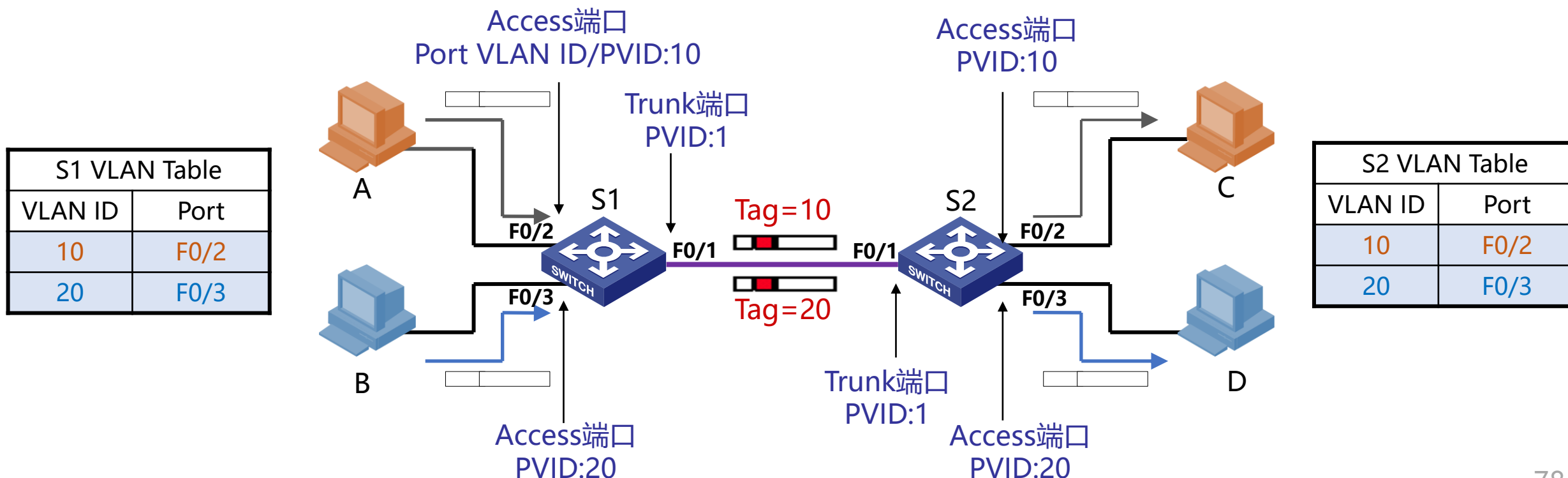
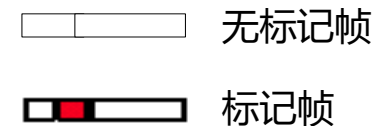
清华大学  
Tsinghua University

计算机网络教案社区

## ➤ 干道 (Trunk) 类型端口

- Trunk端口一般用于交换机之间连接
- 干道链路允许多个VLAN的流量通过
- Port VLAN ID/PVID: trunk端口默认为1

VLAN标记的  
添加/剥除





# 虚拟局域网

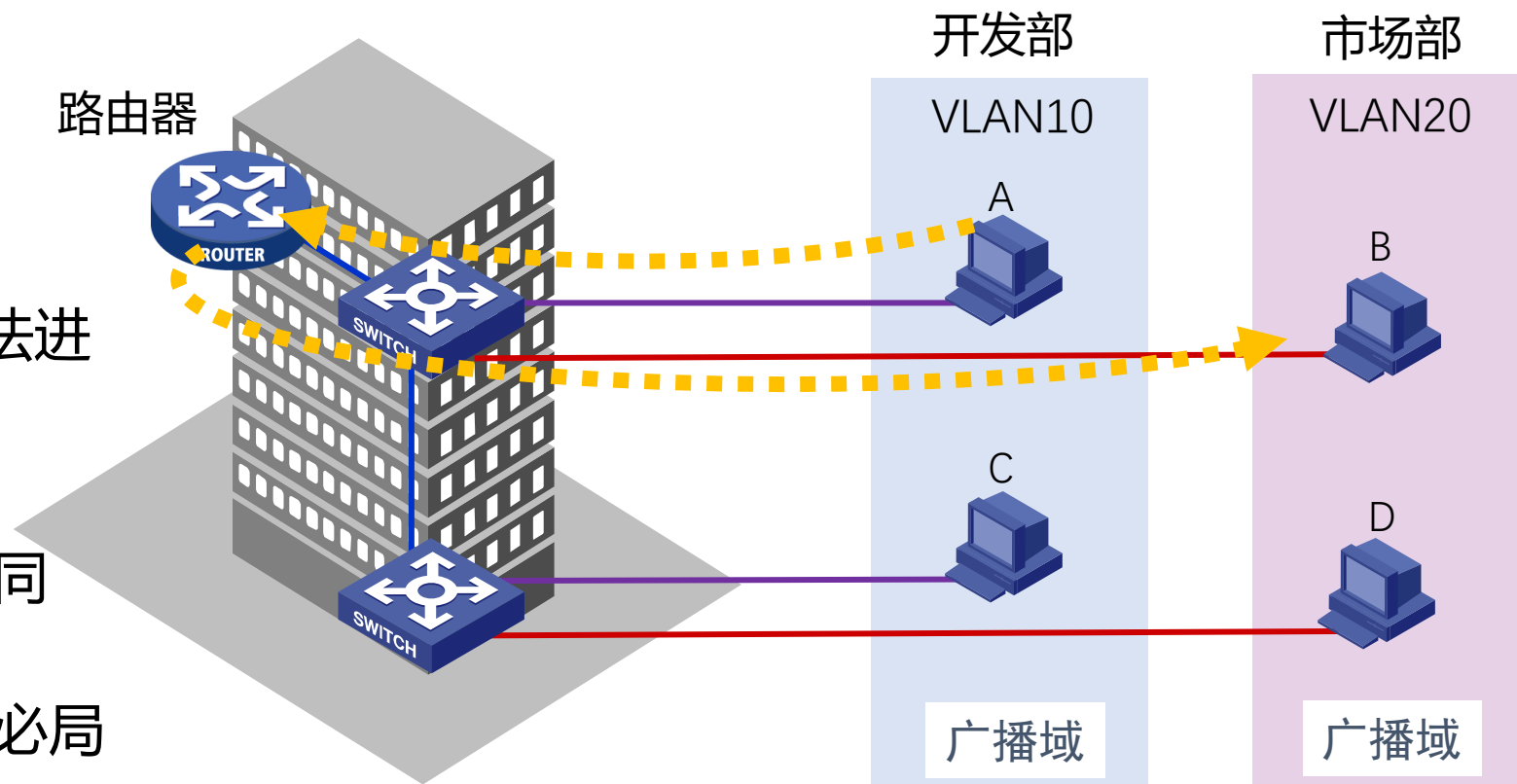


清华大学  
Tsinghua University



计算机网络教案社区

- VLAN控制广播域范围
  - 广播流量被限制在一个VLAN内
- 增强网络的安全性
  - VLAN间相互隔离，无法进行二层通信
- 提高网络的可管理性
  - 将不同的业务规划到不同VLAN便于管理
  - 灵活构建虚拟工作组不必局限于同一物理范围
- VLAN间用三层设备通信



全球分布的开发部，能否构成局域网？



# 总结：以太网



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 规模可扩展的以太网

- 冲突域太大效率低，如何划分多个冲突域？
- 集线器->交换机：希望交换机无配置并自适应拓扑（少做无用转发）？
- 交换机：逆向学习+转发、过滤和洪泛
- 冗余拓扑带来新问题：广播风暴、重复帧、MAC表震荡
- 生成树协议：根桥、根端口、指定端口，解决了问题但不够优化
- 源路由网桥：能够寻找最优路径

## ➤ 链路层交换机

- 三种转发模式：存储转发、直通模式、无碎片模式

时延v.s.错误

## ➤ 规模太大怎么办：虚拟局域网

- 扩大局域网的同时，隔离不同的局域网
- 哪些端口属于哪个VLAN，VLAN标记的添加/剥除



# 作业

---



清华大学  
Tsinghua University



计算机网络教案社区

- 《Computer Networks-5<sup>th</sup> Edition》 章节末习题
  - CHAPTER 4: 27 (error-correcting code) , 37 (数据链路层交换) , 38 (交换) , 39 (交换) , 41 (VLAN)
- 截止时间: 下周三晚11:59, 提交网络学堂



# 致谢社区本章贡献者



清华大学  
Tsinghua University



计算机网络教案社区



袁华

华南理工大学

- 1、信道分配问题
- 2、多路访问协议



曹轶臻

中国传媒大学

- 2、以太网



赵婧如

西安邮电大学

- 3、数据链路层交换



徐敬东

南开大学

- 4、无线网络

《计算机网络：自顶向下方法》(原书第7版), 库罗斯 罗斯, 机械工业出版社, 2018年06月

《计算机网络 (第5版) 》, Tanenbaum & Wetherall, 清华大学出版社, 2012年3月

《计算机网络 (第7版) 》, 谢希仁, 电子工业出版社, 2017年01月

《计算机网络 (第4版) 》, 吴功宜, 清华大学出版社, 2017年04月

《计算机网络技术与应用》, 张建忠, 徐敬东, 清华大学出版社, 2019年9月

《路由交换技术详解与实践:第1卷:下册 》,新华三大学, 清华大学出版社, 2017年08月



# 特别鸣谢



清华大学  
Tsinghua University



计算机网络教案社区

- 本课程课件中的部分素材来自于：
  - (1) 库罗斯.罗斯、Tanenbaum & Wetherall、谢希仁、吴功宜、徐敬东等出版的《计算机网络》教材
  - (2) 思科网络技术学院教程
  - (3) H3C网络学院系列教程
  - (4) 网络上搜到的其他资料
- 在此，对清华大学出版社、思科网络技术学院、H3C网络学院、电子工业出版社、机械出版社以及其它提供本课程引用资料的个人表示衷心的感谢！
- 对于本课程引用的素材，仅用于教学，如有任何问题，请与我们联系！