

○、虚拟网问题

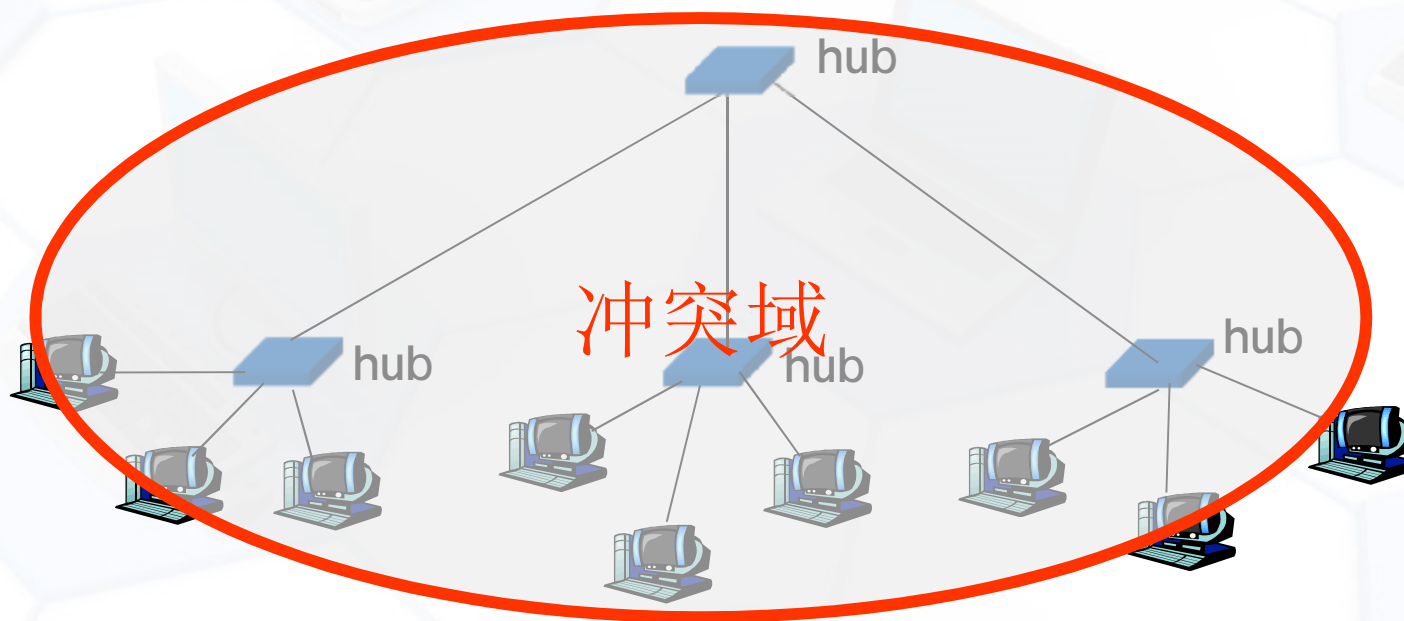
1. 使用了交换机后冲突域还存在吗？
2. 交换机如何知道哪些数据应该转发给新接入的主机？
3. 为什么帧目的地通过该帧的源端口可以到达反而被丢弃？
4. 不同的虚拟网之间能通讯吗？
5. 生成树协议如何知道连接中存在回路？



一、交换机基础

1. 集线器的主要问题

- § 无法分割LAN的流量以提高性能
- § 无法建立混合速率的LAN（如10M/100M）
- § 扩展LAN范围的同时也扩展了冲突域



使用集线器无法避免冲突域扩大

2. 交换机的主要功能

- § 分割LAN的流量来提高性能
- § 能够建立混合速率的LAN（如10M/100M）
- § 扩展LAN的范围
- § 阻止特定类型的帧发送到LAN的其他部分
- § 通过备份交换机和交换机间备份链路消除故障
- § 聚合交换机间一组链路
- § 设置帧优先权
- § 支持VLAN和SNMP

§ 链路层设备

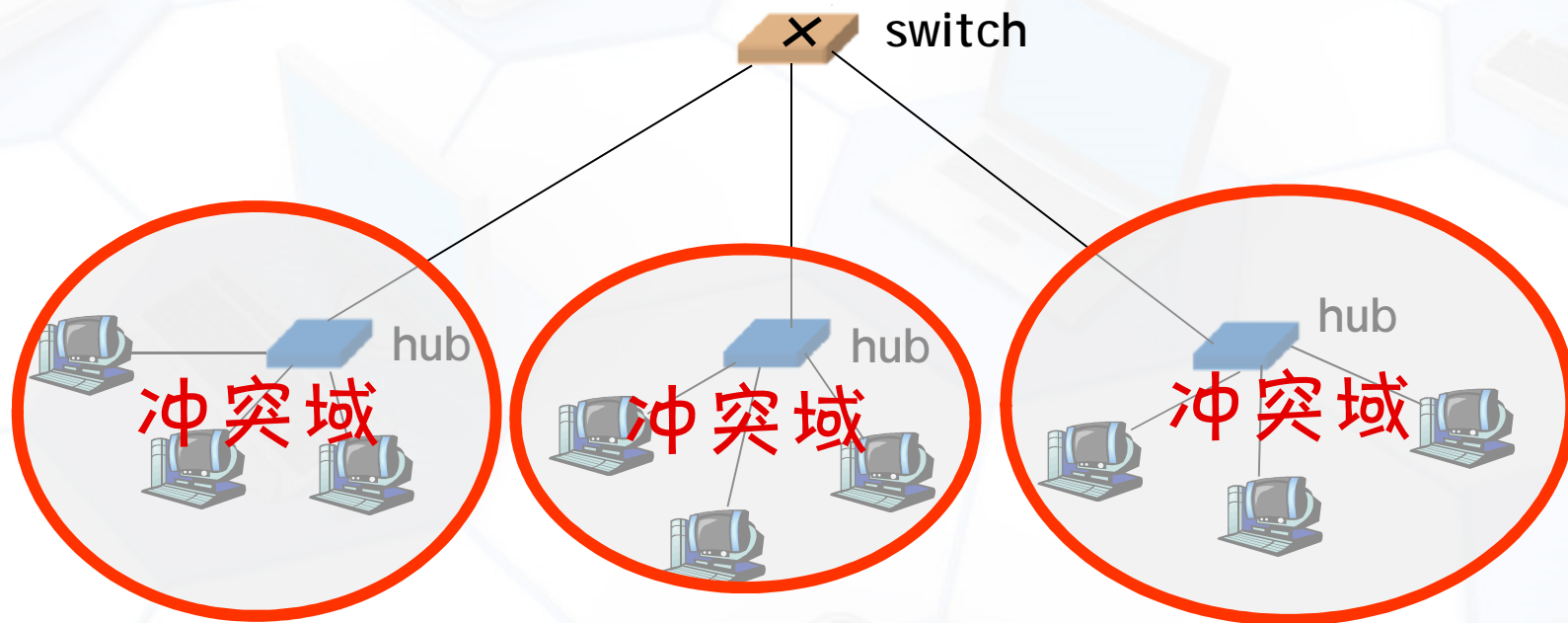
- Ø 存储并转发数据帧

- Ø 检查帧头部并按目的MAC地址选择性的转发

- Ø 在转发网段上采用CSMA/CD

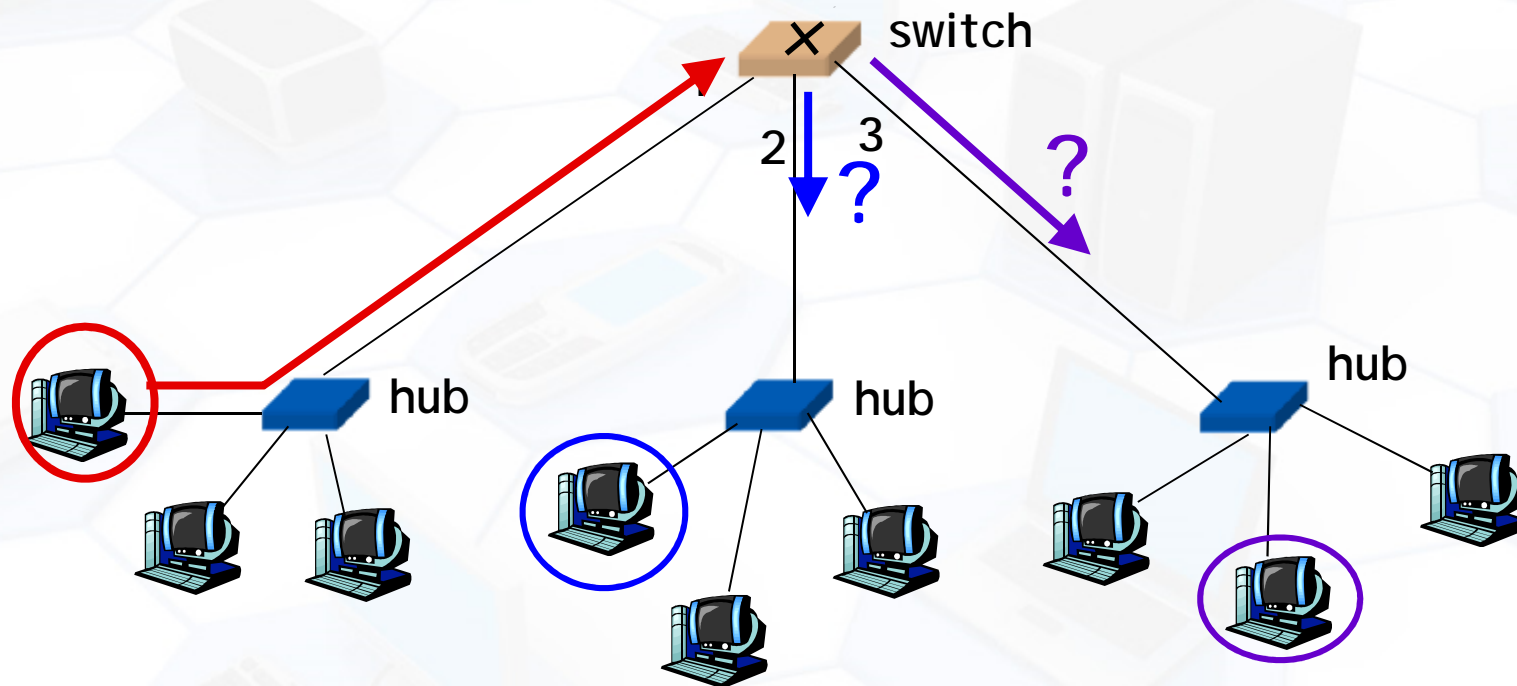
§ 透明接入，站点不必知道接入的是否为交换机

§ 即插即用，自我学习，不需要配置



通过交换机分割冲突域

3. 自我学习



交换机如何决定帧发往哪个网段？

3. 自我学习（续）

§ 交换机有一个转发表

§ 转发表项组成内容(MAC地址, 端口, 时间戳)

§ 老化的表项将被丢弃

§ 交换机自己学习通过哪个端口能够到达哪些主机

Ø 帧到达时, 交换机记录发送方的位置 (来源网段)

Ø 转发表中记录发送方和端口的对应关系

§ 交换机的转发表

MAC地址	端口	状态角色
00-60-08-1E-AE-42	2	动态
00-60-08-BD-7D-1A	3	动态
25-A9-74-0E-00-23	5	静态
00-90-27-AE-B9-1D	1	老化

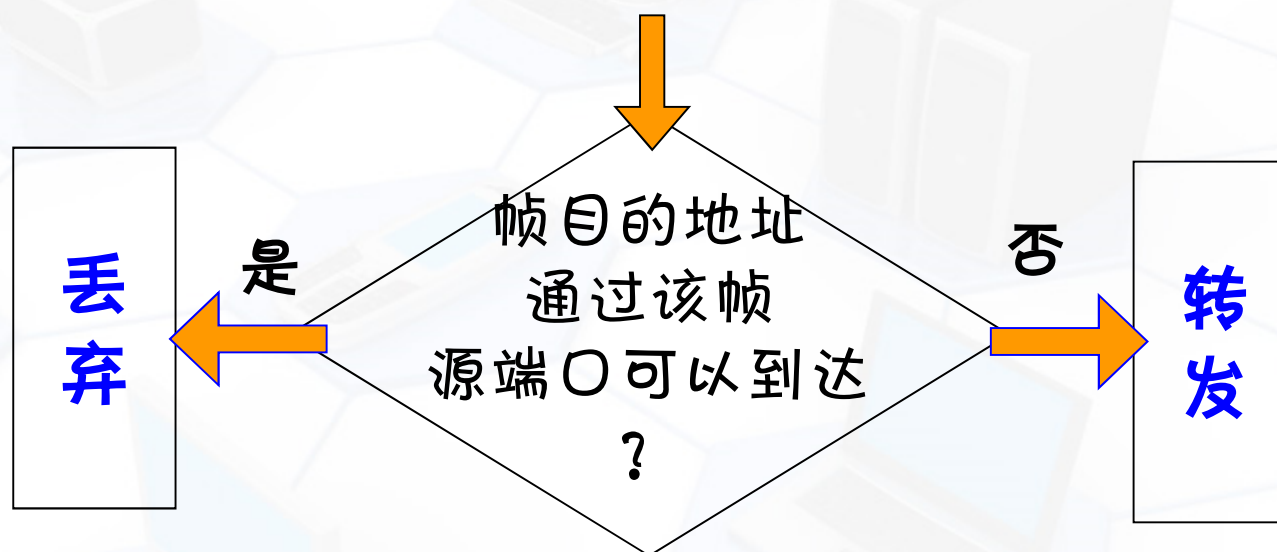
- ü 动态表项说明该地址项刚更新（或刚记录）
- ü 静态表项属于人工添加
- ü 老化标明该表项已长期未更新，即将被淘汰

§ 静态表项的目的

- ü 加快到专有端口通信量的转发
- ü 需要安全限制
- ü 阻塞无关帧以提高性能

4. 交换机工作原理

§ 交换机主要功能为转发数据帧

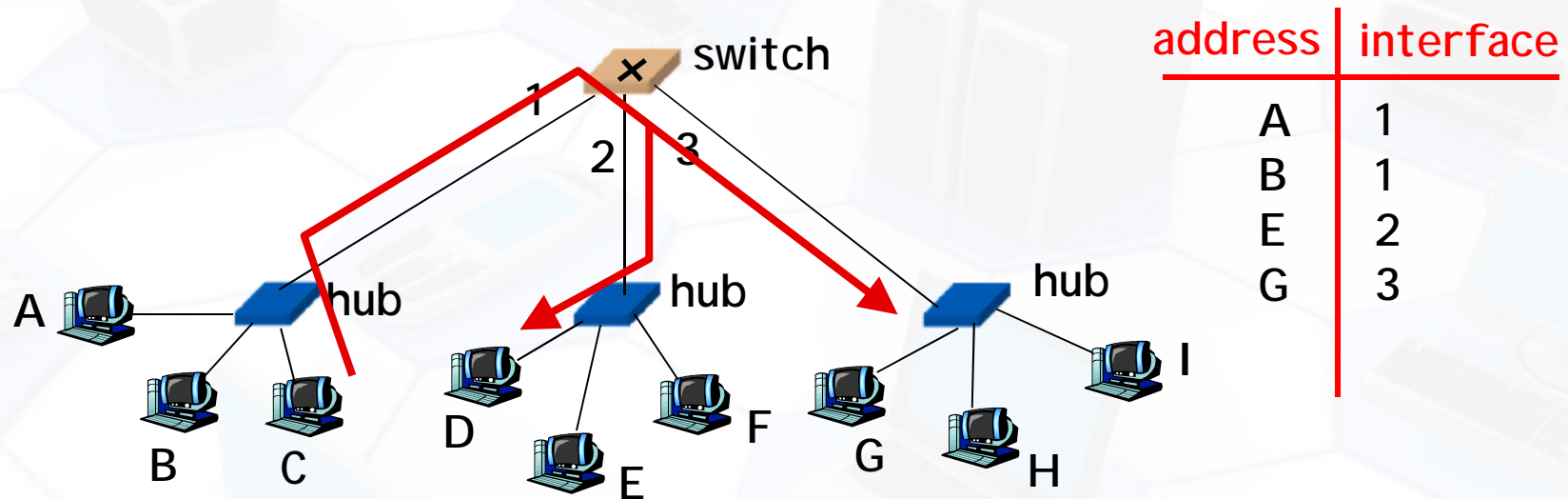


§ 交换机学习过程

- ü 观察每个端口上的所有流量
- ü 记录每一帧的源MAC地址和端口关系
- ü 将学习内容添加到过滤表

交换过程举例--A

§ 如果C发送给D



§ 交换机收到来自C的帧

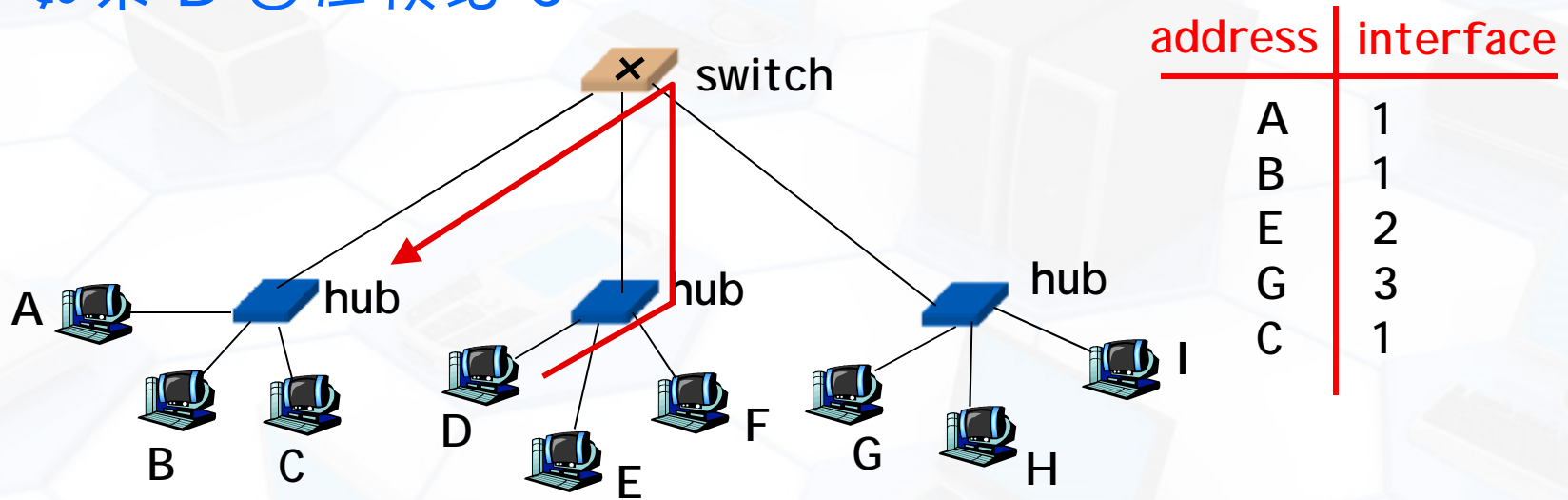
Ø 在转发表中记录C帧来自端口1

Ø D的地址在表中不存在，交换机将帧发往端口2和3（不明目的的帧被广播）

§ 站点 D 收到帧

交换过程举例--B

§ 如果 D 回应帧给 C



§ 交换机收到来自 D 的帧

Ø 在转发表中记录 D 位于端口 2

Ø 由于 C 已在表中, 交换机只将帧转发到对应的端口 1

§ 站点 C 收到帧

IP Address	MAC Address	VLAN ID	Port Name	Aging	Type
172.168.202.97	00d0-f8f9-8345	300	GigabitEthernet5/0/3	11	Dynamic
172.168.202.93	00d0-f8f9-81f4	300	GigabitEthernet5/0/3	13	Dynamic
172.168.202.110	00d0-f8f9-81a5	300	GigabitEthernet5/0/3	13	Dynamic
172.168.202.96	00d0-f8f9-8326	300	GigabitEthernet5/0/3	14	Dynamic
172.168.202.111	00d0-f8f9-8186	300	GigabitEthernet5/0/3	14	Dynamic
172.168.202.90	00d0-f8f9-8605	300	GigabitEthernet5/0/3	14	Dynamic
172.168.202.92	00d0-f8f9-835b	300	GigabitEthernet5/0/3	14	Dynamic
172.168.202.11	00e0-fc00-0001	300	GigabitEthernet5/0/3	17	Dynamic
172.168.202.10	00e0-fc07-a4e1	300	GigabitEthernet5/0/2	18	Dynamic
172.168.202.95	00d0-f8f9-835d	300	GigabitEthernet5/0/3	19	Dynamic
172.168.202.91	00d0-f8f9-8323	300	GigabitEthernet5/0/3	20	Dynamic
172.168.202.161	00e0-fc7d-a22c	300	GigabitEthernet5/0/5	20	Dynamic
172.168.202.160	00e0-fc7d-a1d5	300	GigabitEthernet5/0/4	20	Dynamic
172.168.202.94	00d0-f8f9-82f8	300	GigabitEthernet5/0/3	20	Dynamic
172.168.202.162	00e0-fc7d-a2e3	300	GigabitEthernet5/0/6	20	Dynamic
172.168.202.163	00e0-fc7d-a1c3	300	GigabitEthernet5/0/7	20	Dynamic
210.35.88.124	0090-fb09-703a	204	Ethernet3/0/18	13	Dynamic
210.35.88.126	0090-2774-d87e	204	Ethernet3/0/18	20	Dynamic
210.35.88.250	00e0-fc29-c717	203	Ethernet3/0/17	15	Dynamic
172.168.200.22	0012-79cf-74d7	200	Ethernet3/0/36	20	Dynamic
172.168.200.254	0013-21b1-5840	200	Ethernet3/0/36	14	Dynamic
10.168.101.97	0017-a4a9-cf90	101	GigabitEthernet1/0/3	3	Dynamic

---- More ----

交换机地址表举例

二、交换式以太网

- § 交换机（switch）源自于多端口网桥（bridge），采用存储-转发方式在各端口之间进行数据帧的交换
- § 交换机检测每个到达数据的帧头
- § 根据数据帧的目的地址查找输出端口
- § 如果地址查找表中没有该表项，交换机就向所有端口（除接收端口）转发
- § 地址查找表是通过帧的源地址与到达端口的对应关系建立的

1. 交换机的用法

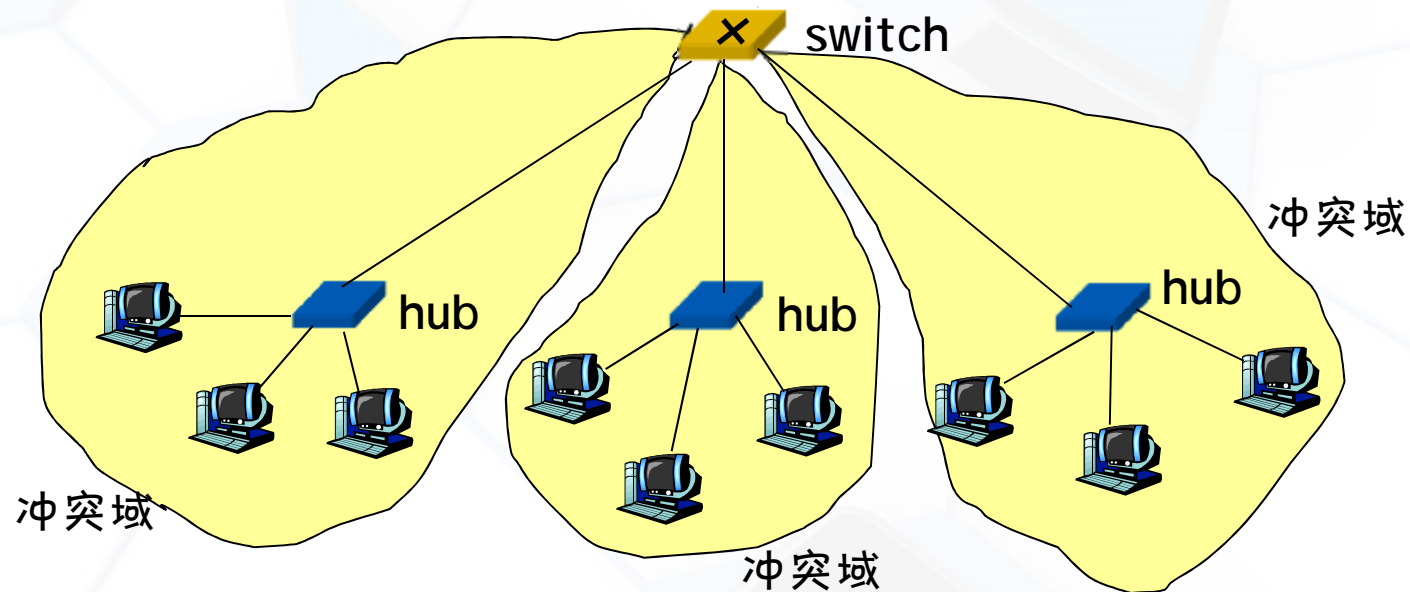
1) 分割流量

§ 交换机分割流量，并将LAN网段化

§ 交换机过滤数据帧

Ø 相同网段的帧不会被送到其他网段

Ø 各网段成为不同的冲突域

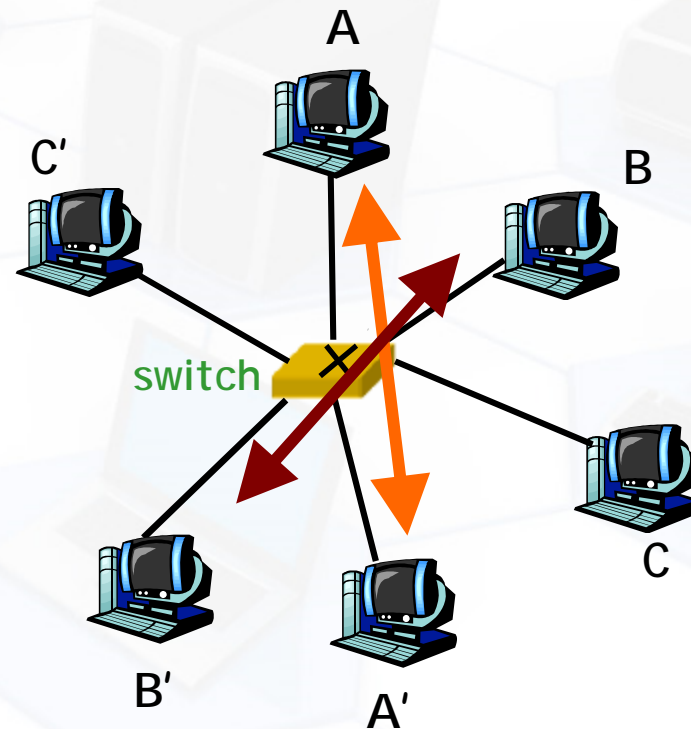


2) 专用访问通道

§ 主机能够直接连接到交换机端口

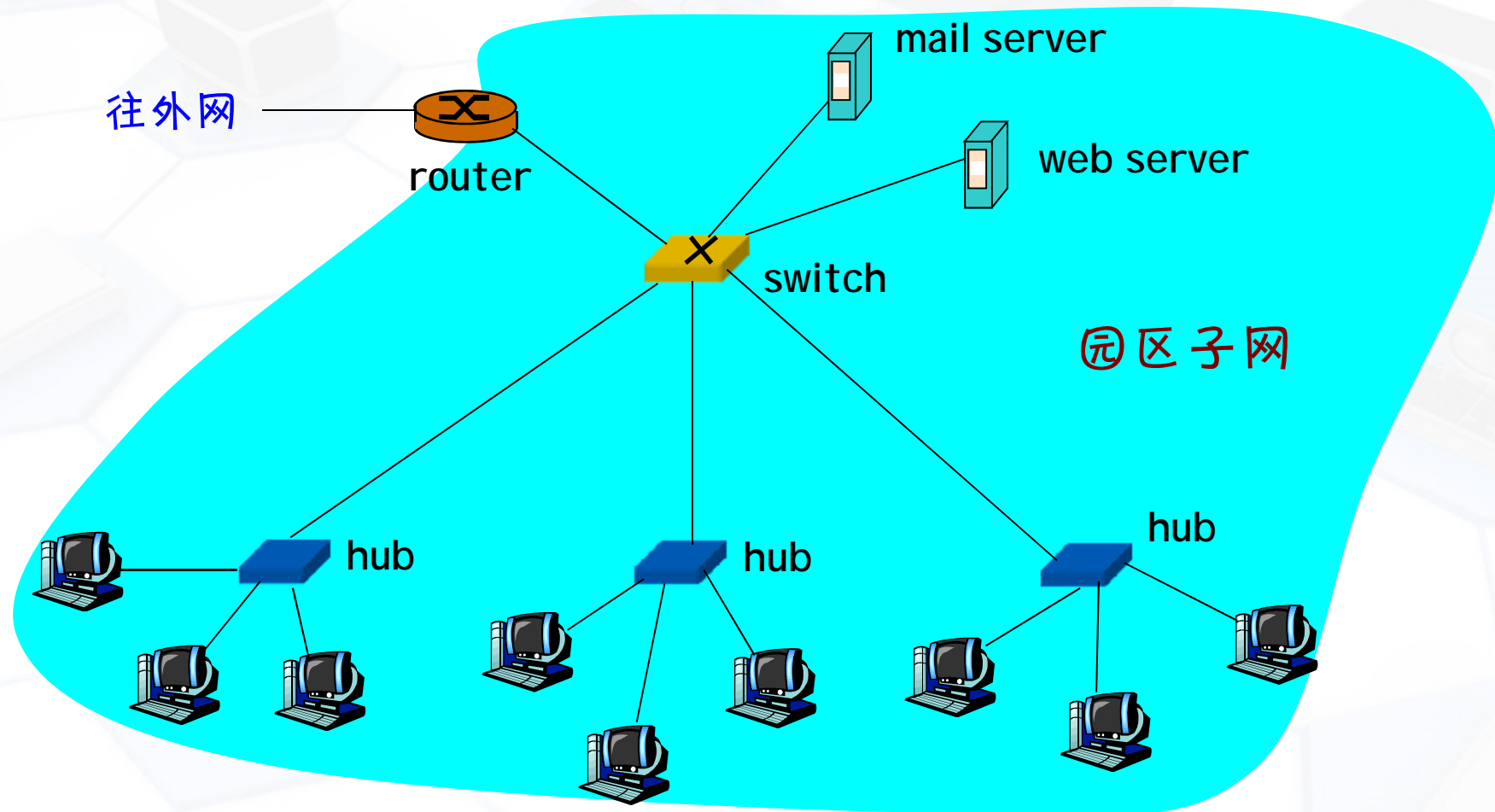
§ 端口数量充足

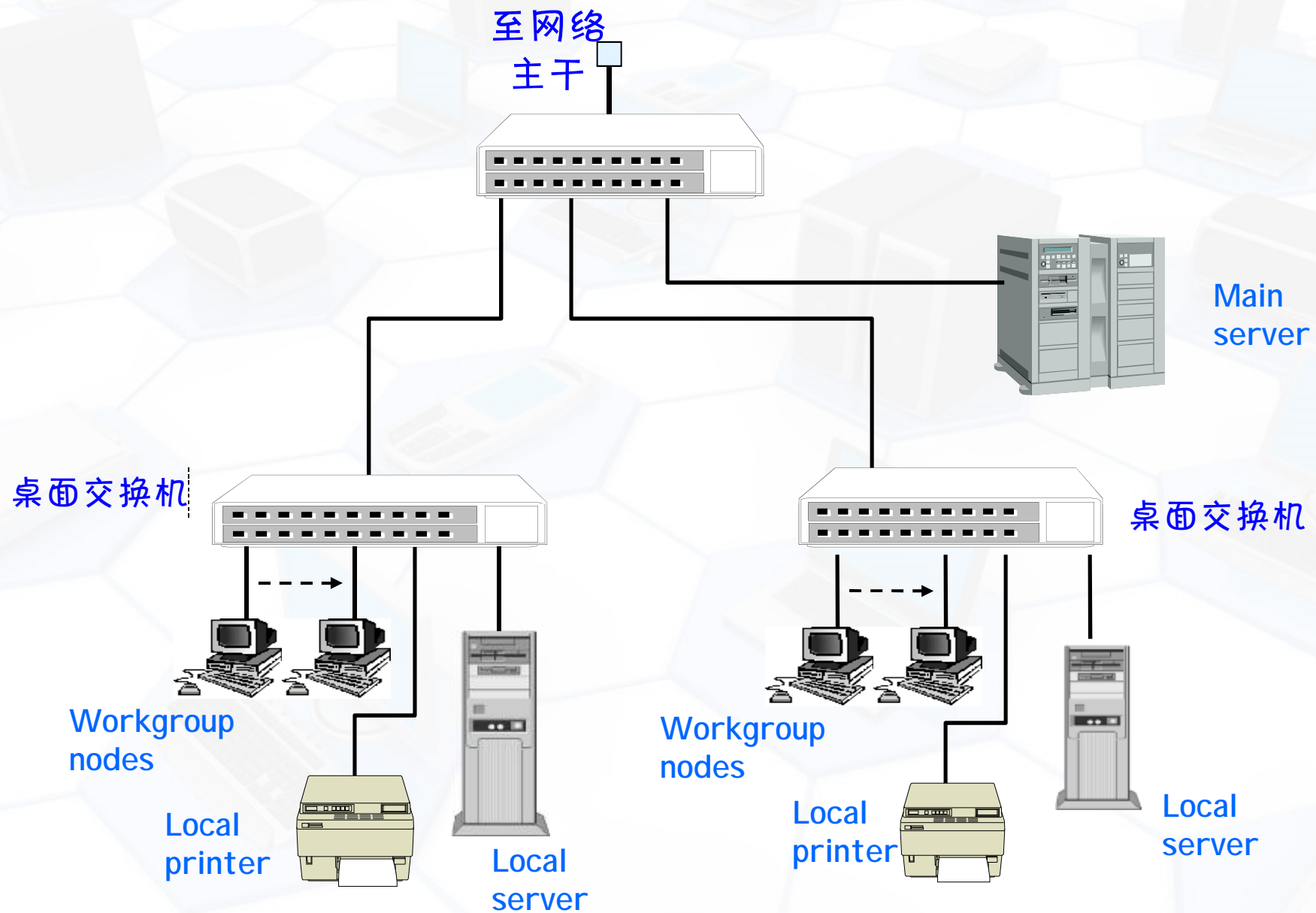
§ 不会有冲突、支持全双工



A-to-A' 与 B-to-B' 能够同时进行，无冲突

3) 园区网的基础





交换机的连接方式举例

2. 交换机的三种交换方式

§ 直通方式 (Cut & Through)

ü只参照MAC帧头的目的地址，接收到数据帧的头6个字节后就可查找内部MAC地址表，找到输出端口后立即输出。因此从帧的接收到转发它的内部时延很小

ü缺点是不检查帧的正确与否，并且不能匹配不同速率的传输端口，也不能连接异种类型的链路

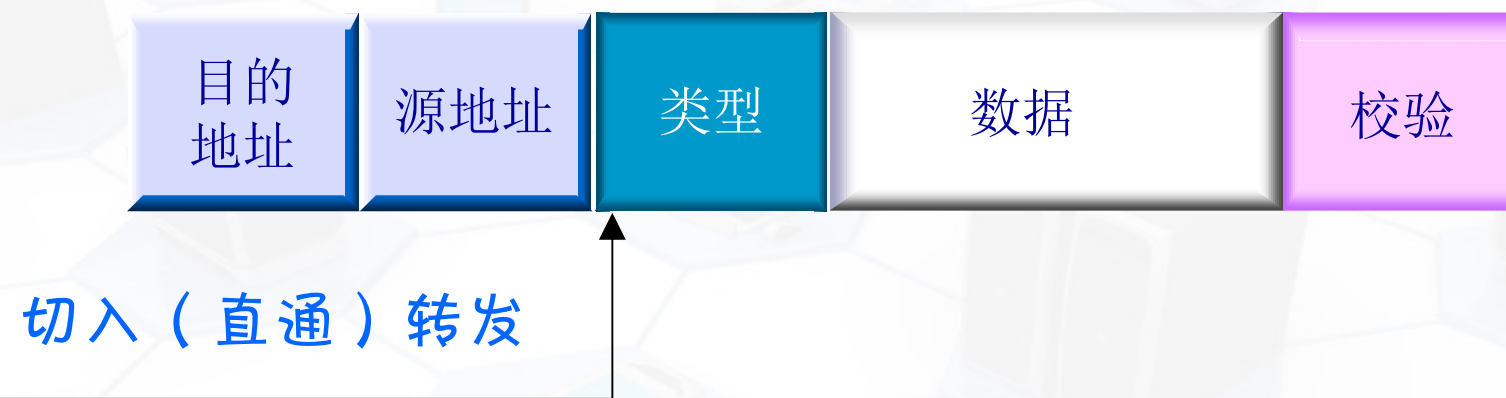
§ 存储转发方式 (Store & Forward)

ü将数据帧从头到尾全部接收后先存放到内存，检查帧的FCS后再进行转发，抛弃错误的数据帧。能与不同传输速率的链路和不同类型的链路进行互连

ü缺点是延时相对要大一些

§ 碎片隔离方式 (Fragment Free)

ü碎片隔离方式只检查帧长小于64字节的短数据帧的FCS，因为根据统计资料，短数据帧的比例较高，所以使用这种方法兼顾了前两种方法的长处，但和直通方式一样，这种方法也不能连接传输速率不同或类型不同的异种链路



交换机的转发方式图解

三、虚拟网(VLAN)

- § 提供物理局域网流量控制的方法
- § 由需要交换消息的系统组成
- § VLAN中的流量通常不会跨越界限
- § VLAN具有自己独立并唯一的标识
- § 不同的VLAN可以共用同一个物理网络
- § 需要支持VLAN的硬件设备

注意

不是所有交换设备都支持VLAN

注意

不同交换设备的VLAN工作方式不同

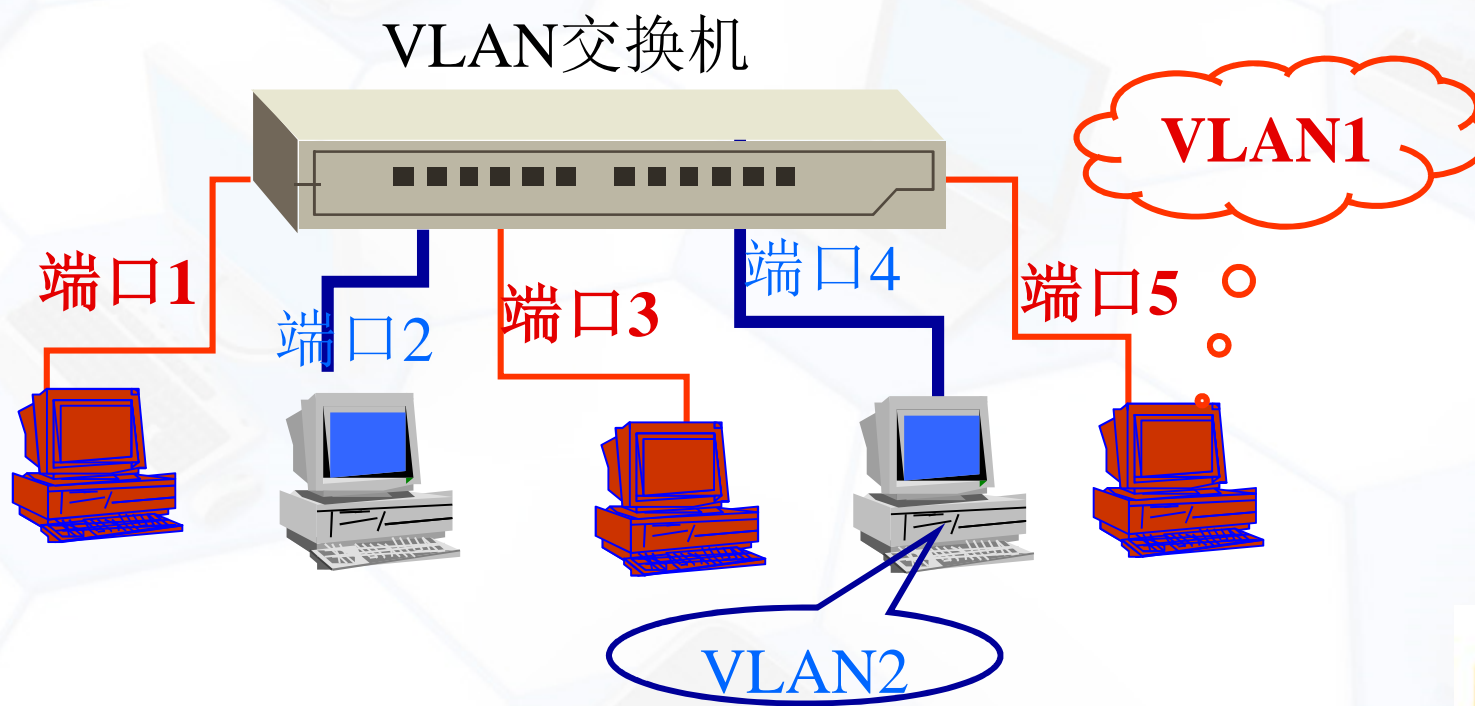
1. 虚拟网基本概念

VLAN成员类型

- § 工作组或特定关系的所有用户
- § 服务器及访问服务器的群体
- § 使用特定协议的系统群
- § 属于组播组的用户群
- § 需要进行安全隔离的一组系统

2. VLAN设备和端口

- § 实现VLAN功能的交换设备称为VLAN知晓交换设备
- § 实现VLAN通过相应交换设备的端口来配置
- § 终端不知道自己属于哪个VLAN



§ VLAN 举例

VLAN2
端口1、2、7

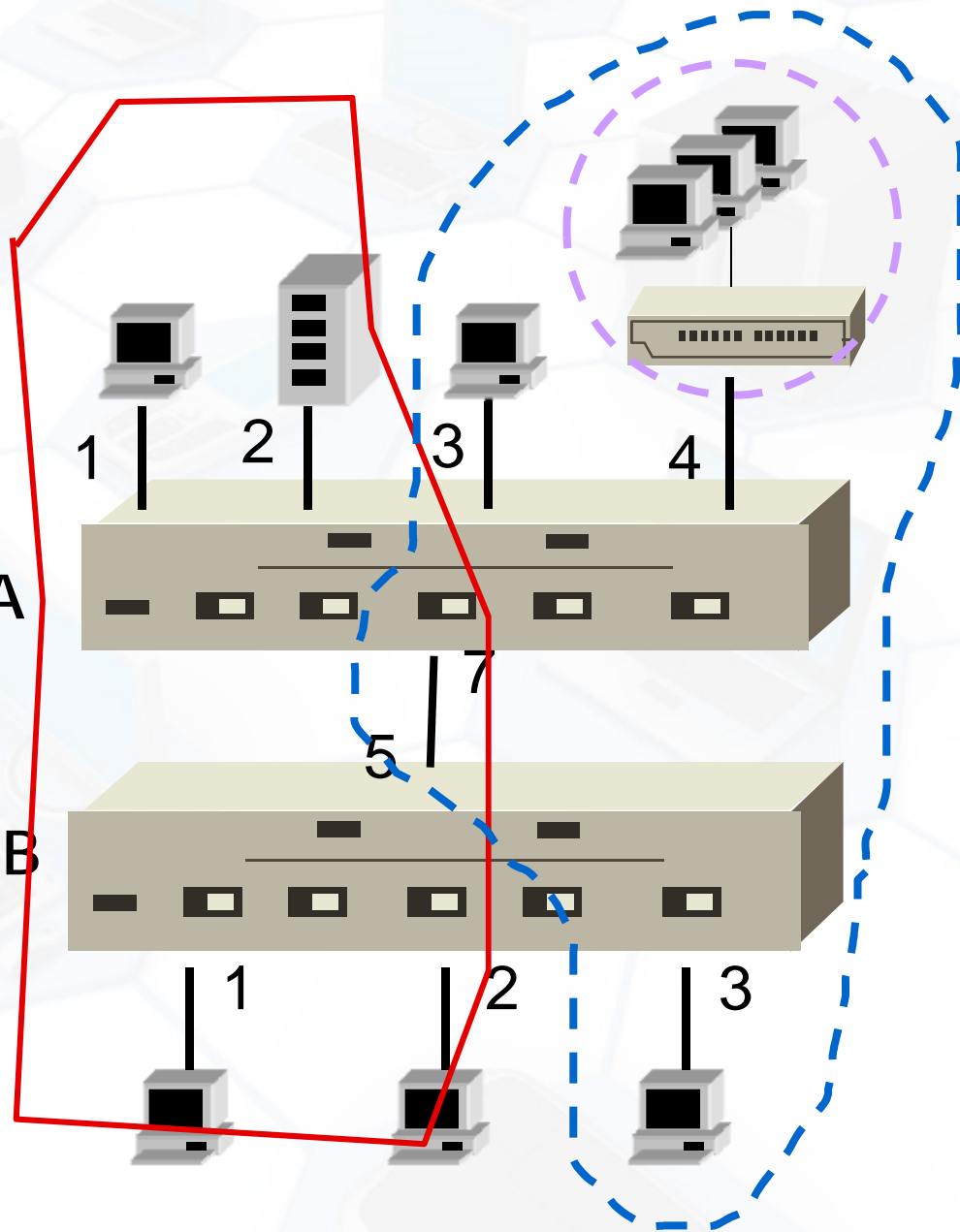
交换设备A

交换设备B

VLAN2
端口1、2、5

VLAN3
端口3、4、7

VLAN3
端口3、5



§ VLAN的隔离作用



§ VLAN重新定义

公共成员

VLAN2
端口1、2、7

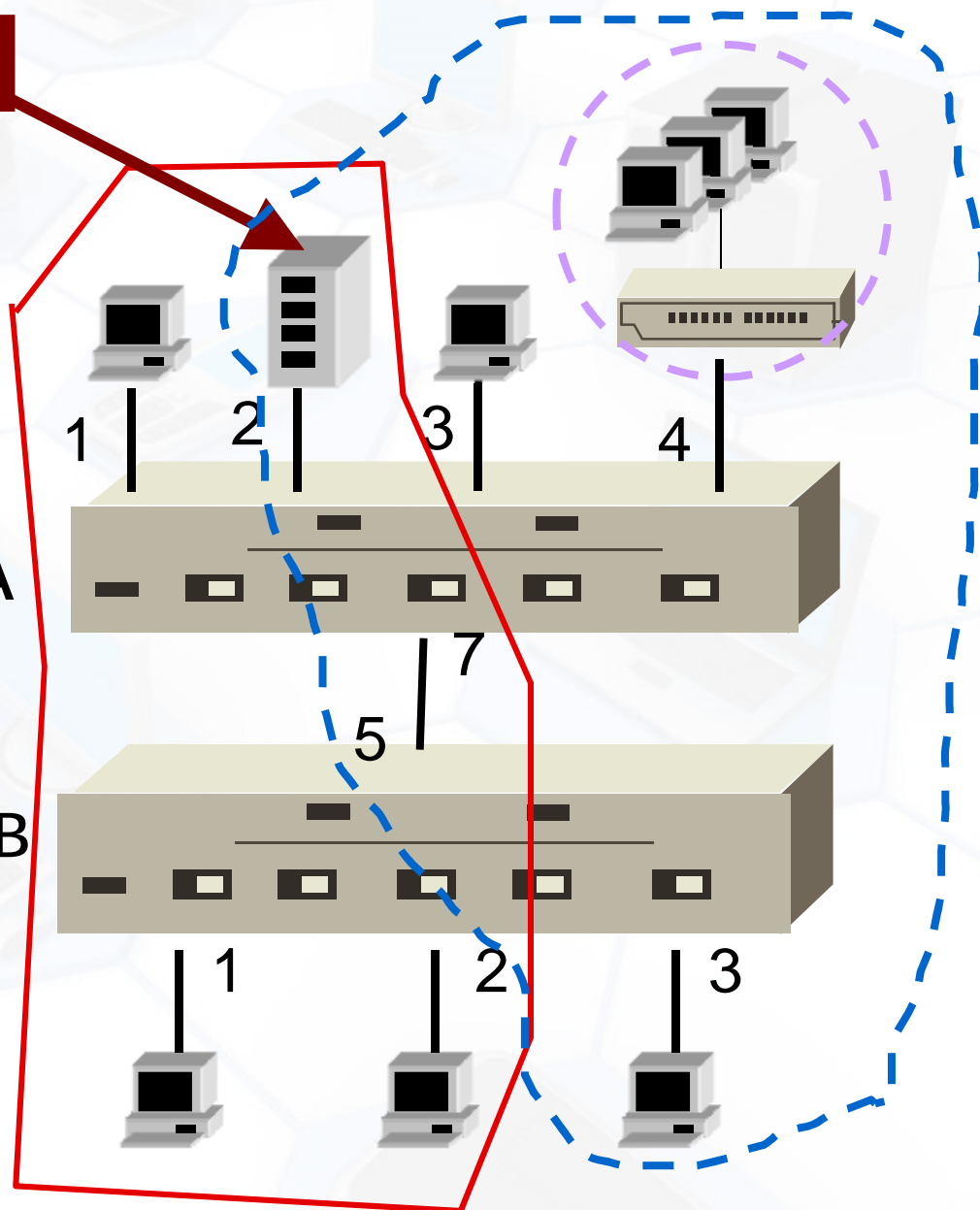
交换设备A

交换设备B

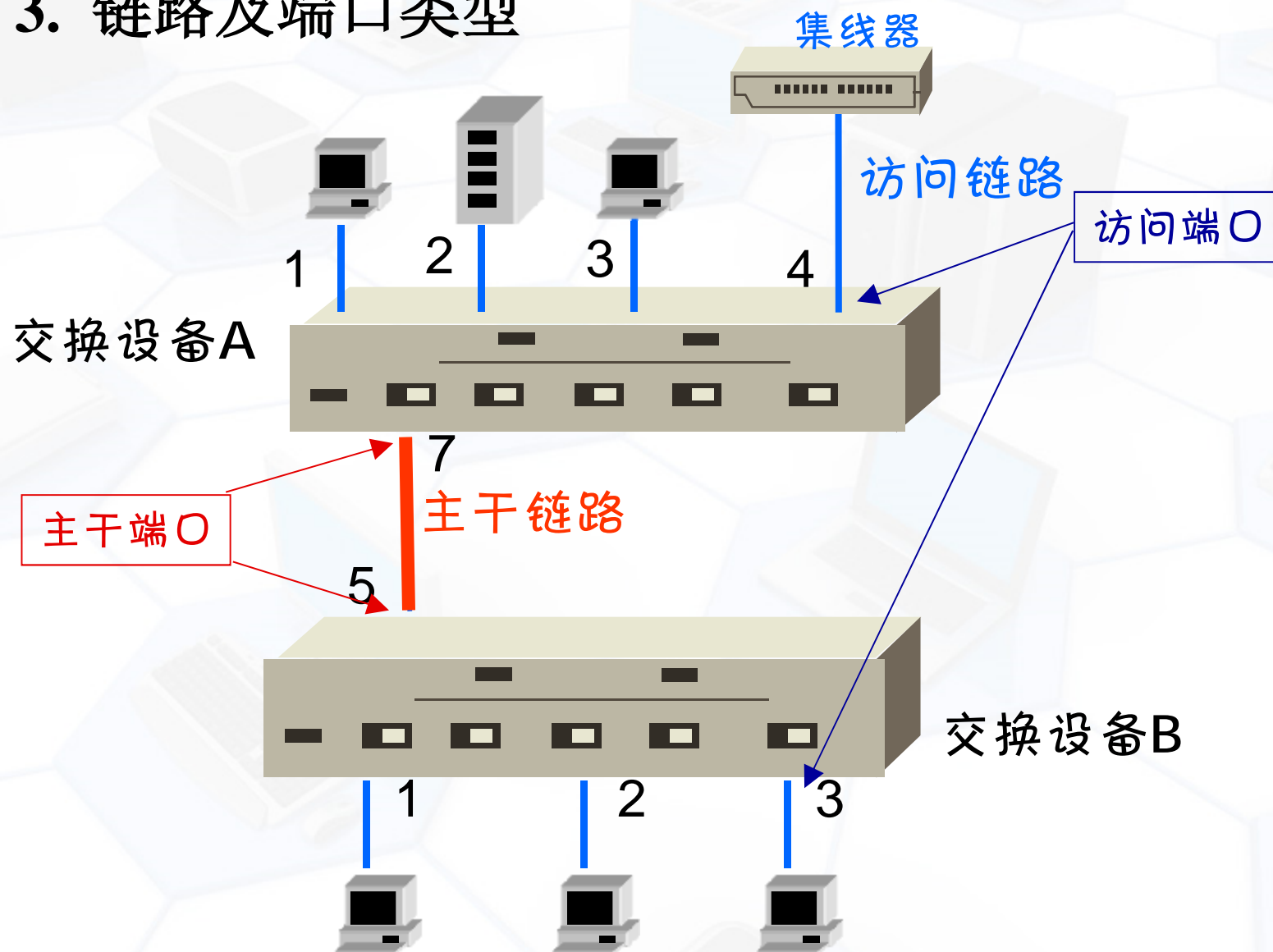
VLAN2
端口1、2、5

VLAN3
端口2、3、4、7

VLAN3
端口3、5

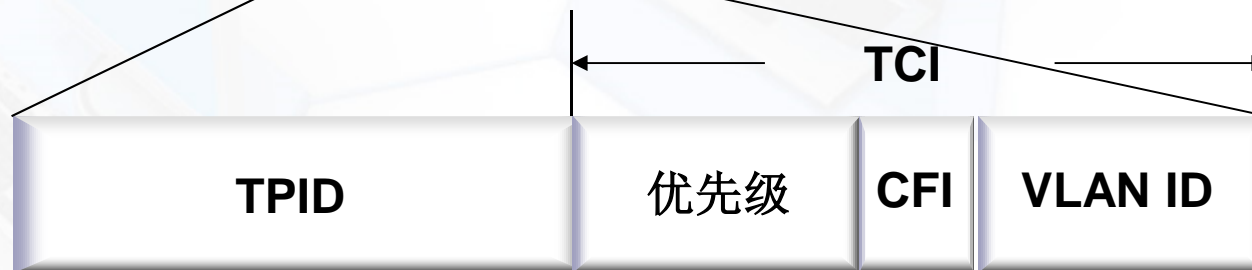


3. 链路及端口类型



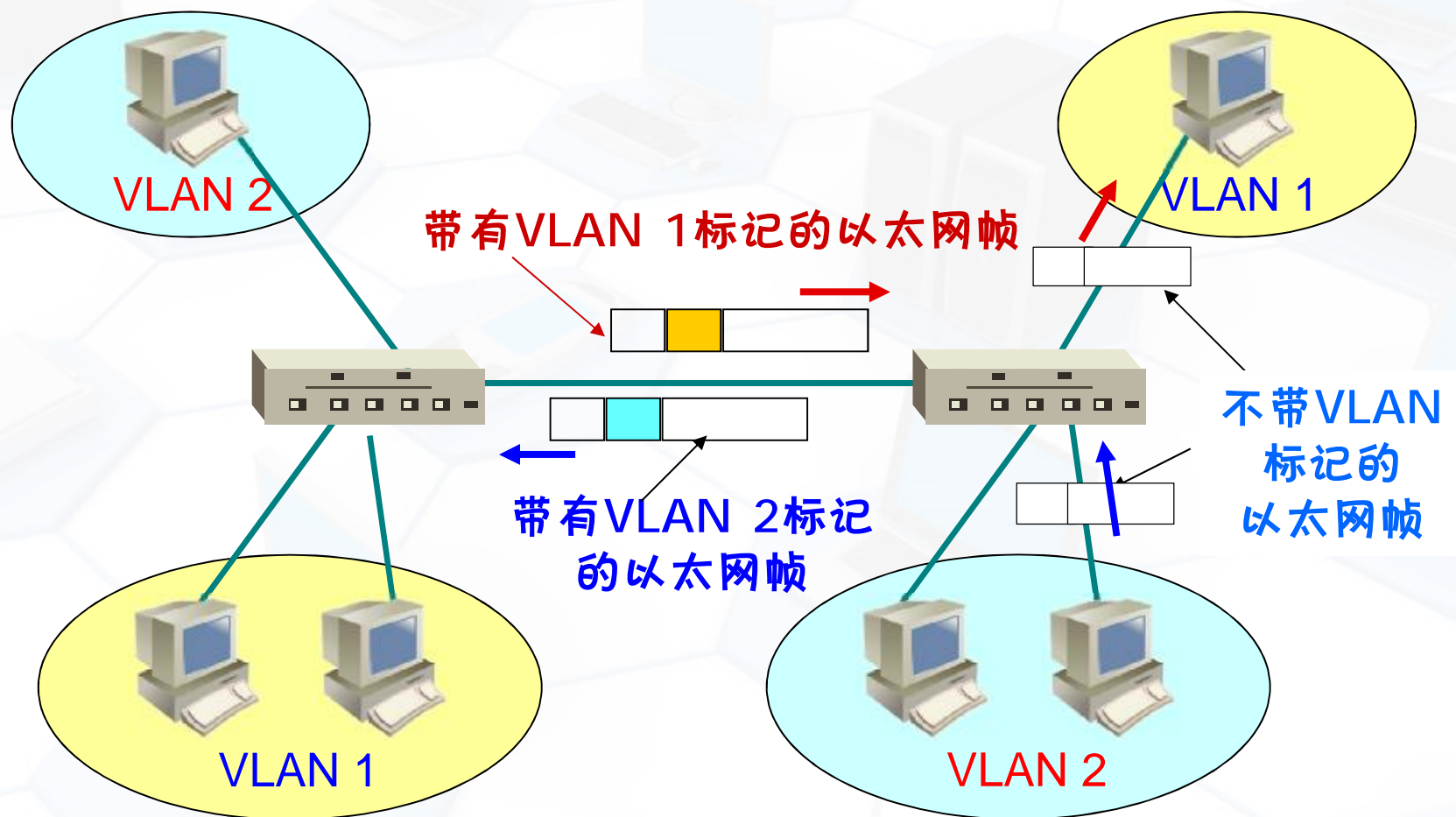
4. VLAN标记

标准以太网帧



带有IEEE802.1Q标记的以太网帧

VLAN中的帧格式



帧在VLAN间通信时的变化

5. 实现虚拟网

p 创建VLAN的必要信息

- § 分配VLAN标识符
- § 定义VLAN成员
- § 列出VLAN成员到达的端口
- § 明确每个VLAN端口的帧是否具有标志

p 创建VLAN成员的方法

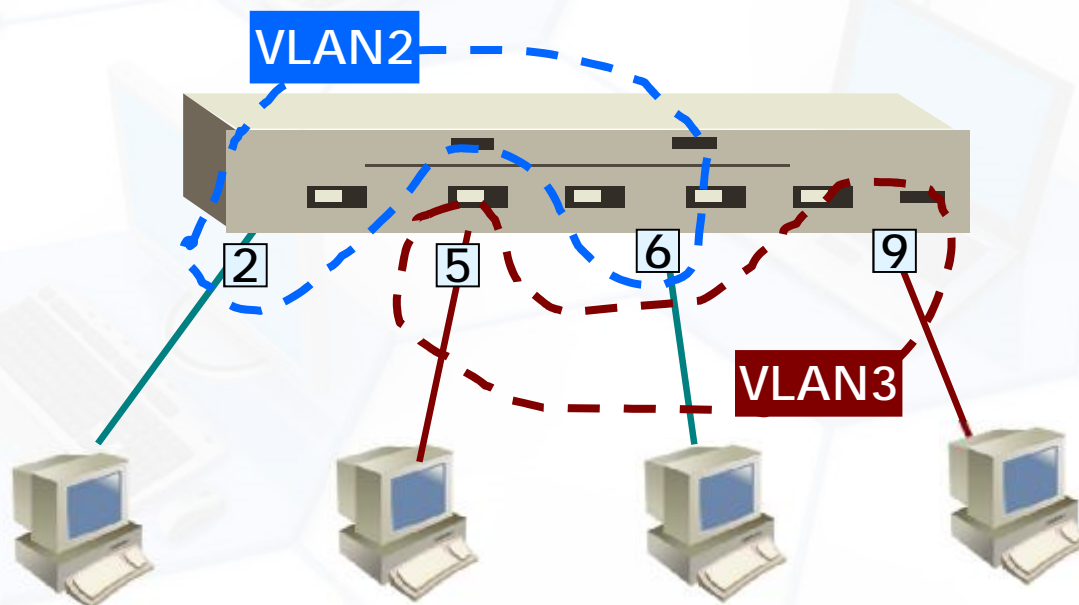
- § 默认——所有端口属同一VLAN
- § 端口——指定VLAN中的成员端口
- § 端口和协议——指明VLAN的端口和协议类型
- § IP地址——列明VLAN的端口和IP地址前缀
- § 端口和MAC地址——明确VLAN的端口和相应MAC地址
- § 端口和GVRP——由VLAN标识符创建VLAN

1) 默认方法

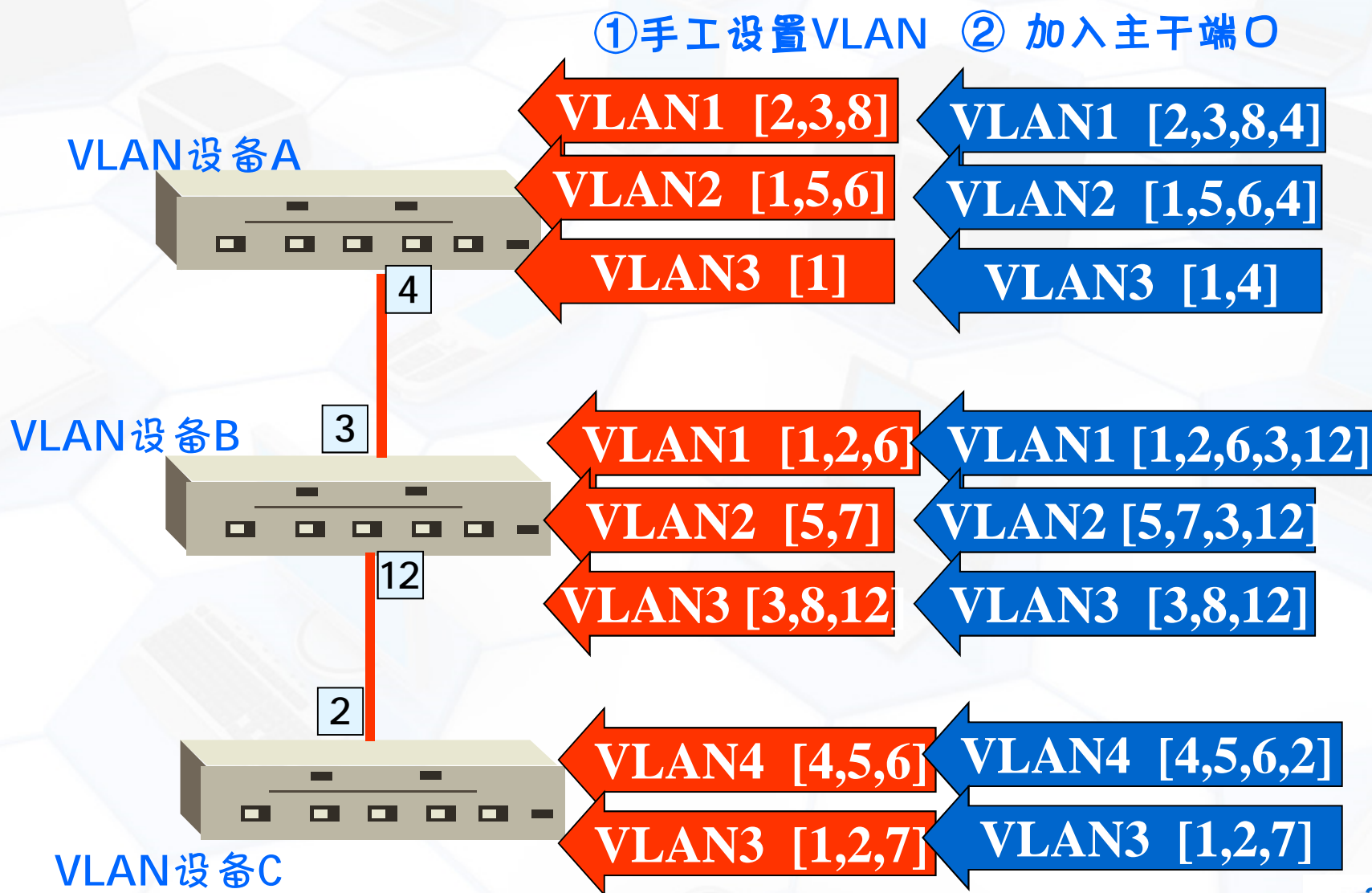
- § 交换设备初始化时进行
- § 所有端口均属同一个VLAN
- § 默认VLAN的标识符为“1”
- § 默认VLAN可以更改

2) 创建基于端口的静态VLAN

- § 最广泛被采用的方法
- § 配置过程中需要指定VLAN标识符
- § 指明属于每个VLAN标识符的端口
- § 连接到端口的终端不必知道自己属于哪个VLAN



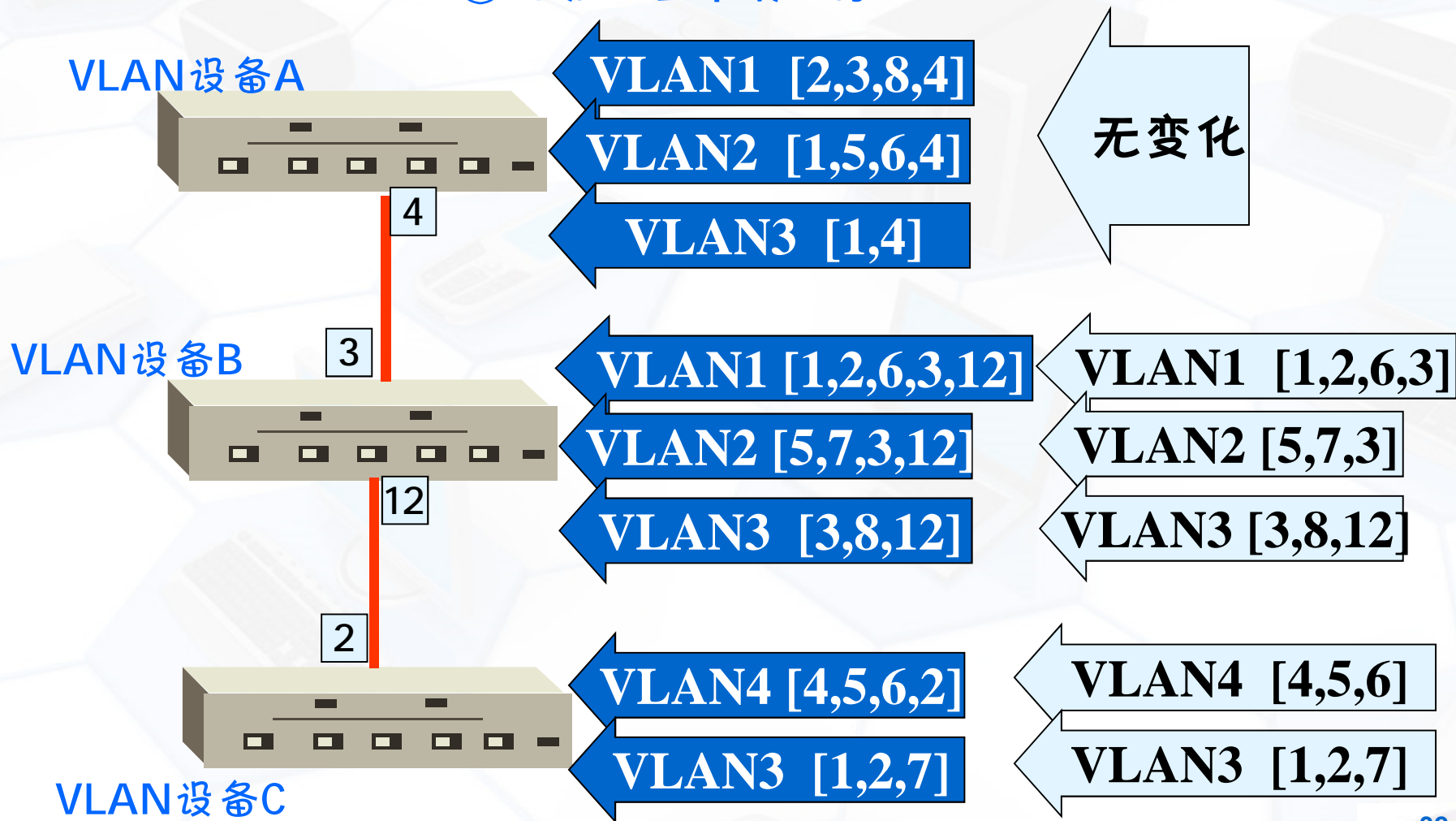
p 静态VLAN的配置举例



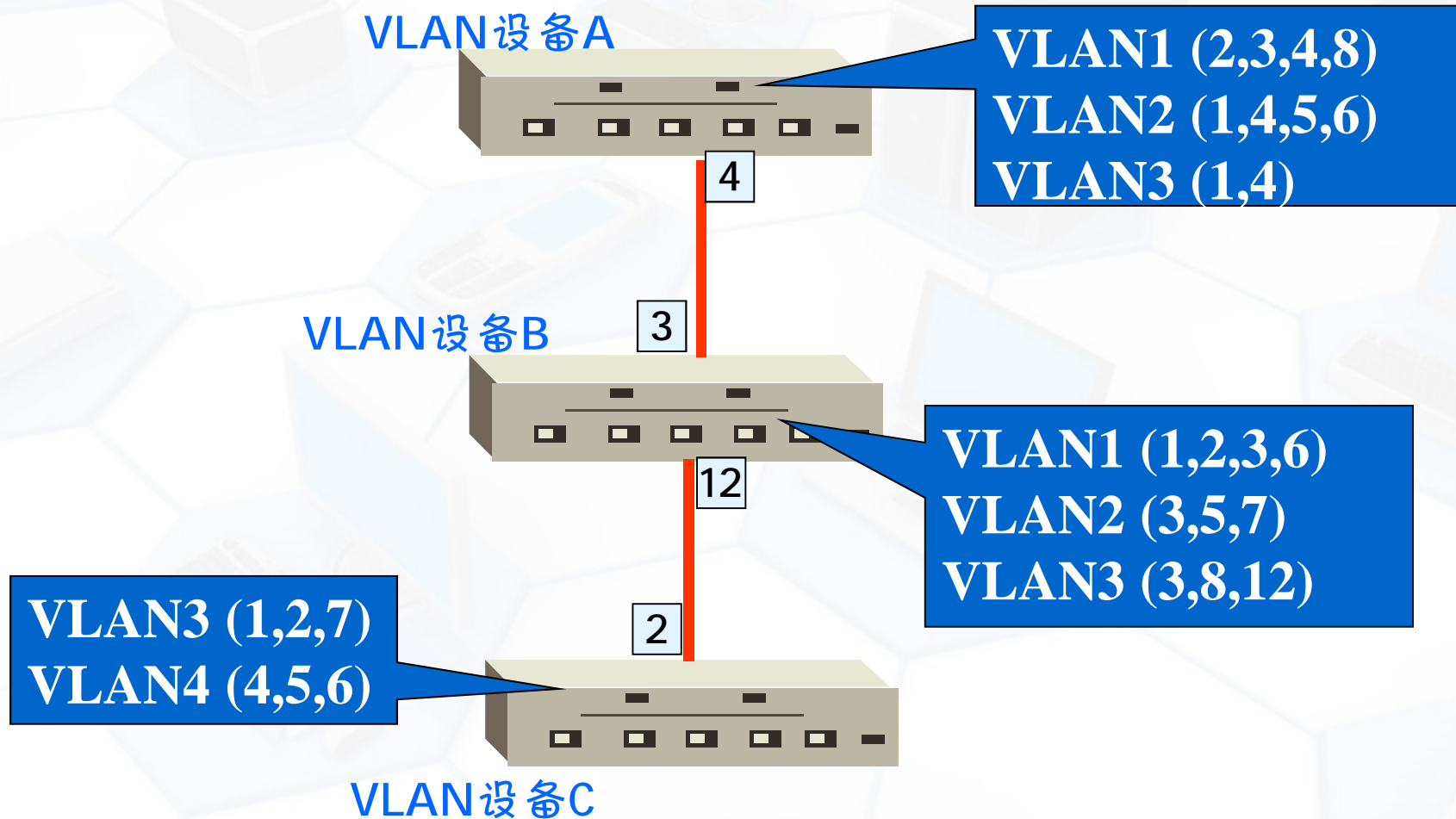
p 静态VLAN的配置举例（续）

③ 已加入主干端口的VLAN

④ 删去不必要的端口

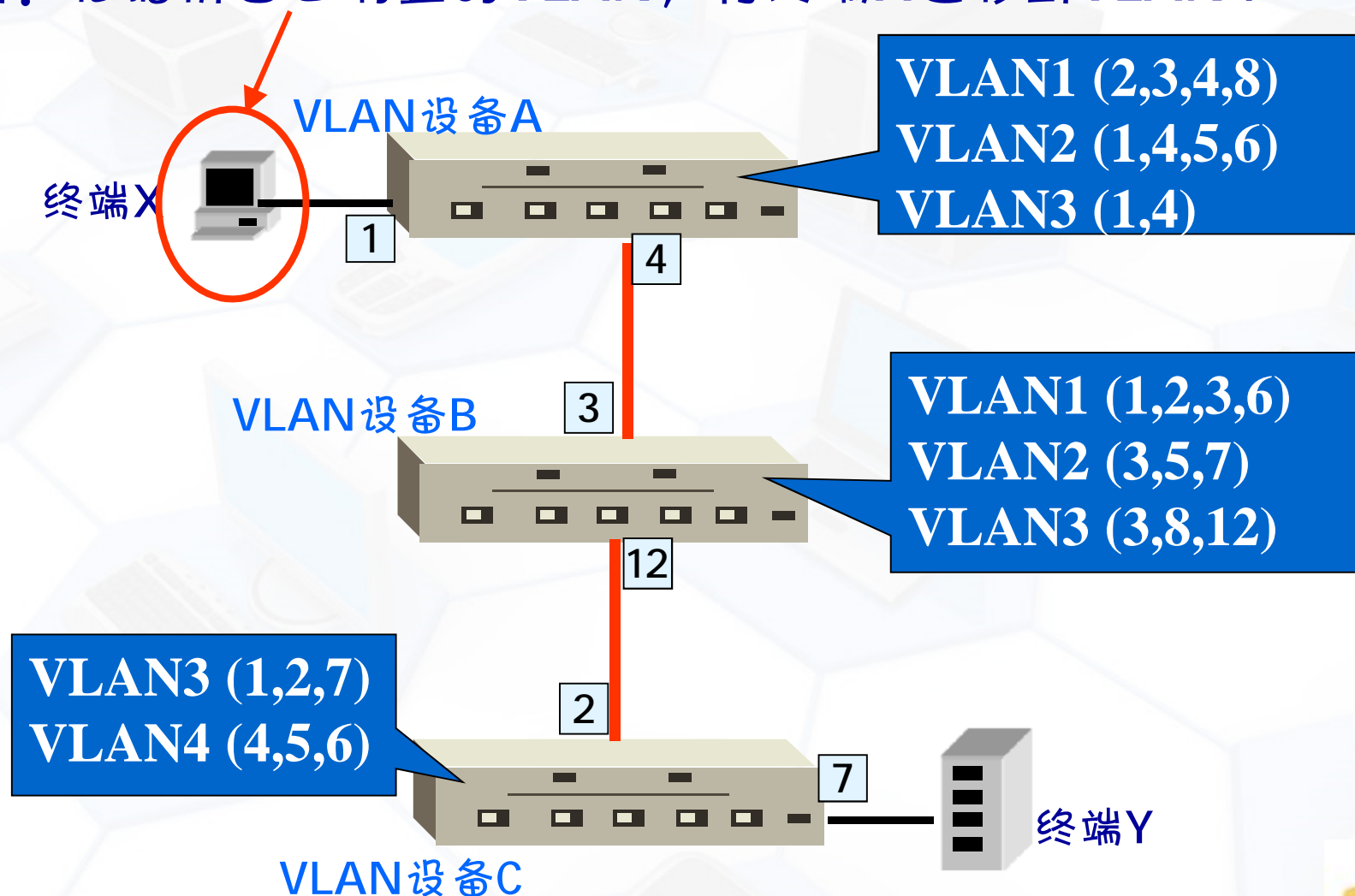


p 经过修改后更有效的VLAN

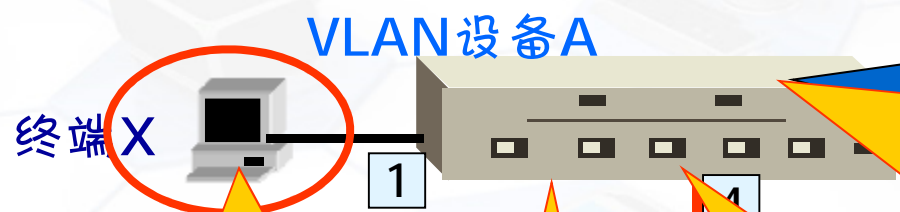


p VLAN的成员变化举例

例：根据前面已调整的VLAN，将终端X迁移到VLAN4



p VLAN的成员变化举例 - 交换机A变化



1. 终端X的物理位置不发生变化

2. 相关的VLAN定义发生变化

3. 具体步骤：
创建VLAN4（原来没有），加入端口1
删除VLAN3（已无成员）
VLAN1,2不发生变化

4. 变化后：
VLAN1 (2,3,4,8)
VLAN2 (1,4,5,6)
VLAN4 (1,4)

交换设备A

p VLAN的成员变化举例 - 交换机B变化

1. 交换设备B需要做相应变化

2. VLAN设置修改步骤：
创建VLAN4（原来没有）
添加端口3和端口12
将VLAN3的端口3删除（交换设备A中已没有VLAN3）

交换设备B

3. 变化后：
VLAN1 (1,2,3,6)
VLAN2 (3,5,7)
VLAN3 (8,12)
VLAN4 (3,12)

p VLAN的成员变化举例 - 交换机C变化



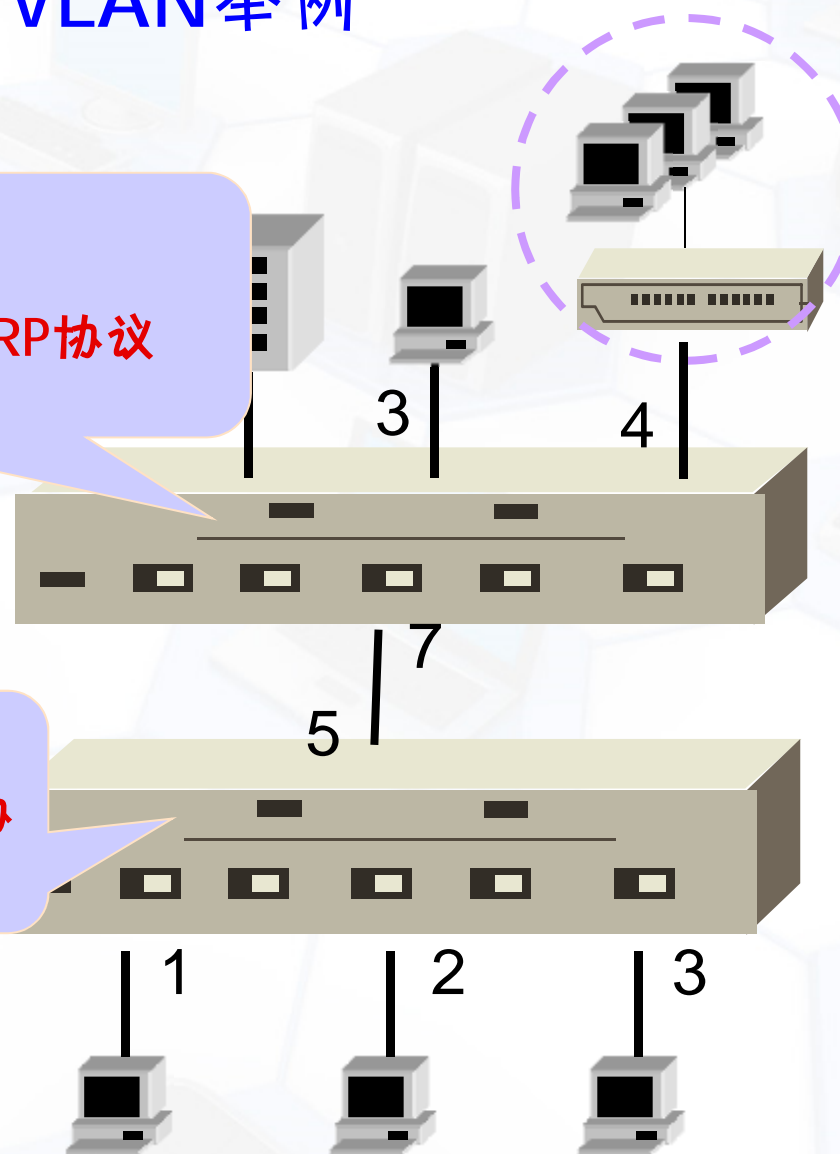
3) 创建基于端口和协议的VLAN

- § 帧到达的端口属于某个VLAN
- § 帧使用的协议属于该VLAN指明的
- § 同样的端口可以定义不同的VLAN
- § 具有协议约定的VLAN中对于属于不同VLAN的端口可以不使用VLAN标识符

基于端口和协议的VLAN举例

VLAN2 (0, 2, 3, 6)+IPX协议
VLAN3 (1, 2, 3, 4, 5, 7)+IP/ARP协议

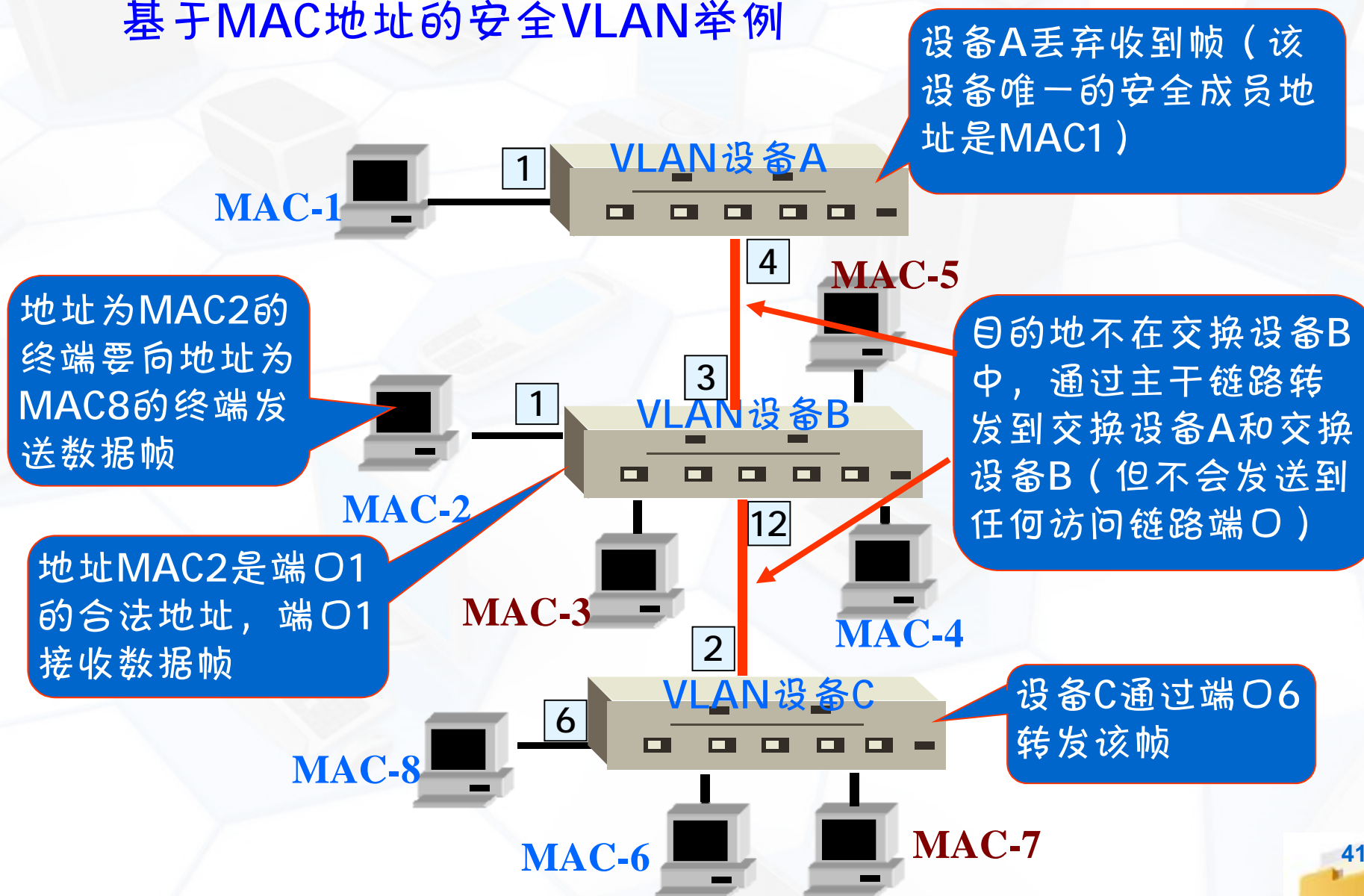
VLAN4 (2, 4, 6)
VLAN3 (0, 1, 3, 7)+IP/ARP协议



4) 创建基于MAC地址的安全VLAN

- § 明确VLAN中每个端口和其对应的MAC地址
- § 源MAC地址与到达端口无关的输入帧将被丢弃
- § 单播帧转发时需要提供目的MAC地址和端口号
- § 不明目的的单播帧绝不发送
- § 非目的地终端不可能收到任何数据

基于MAC地址的安全VLAN举例



6. VLAN帧的处理

VLAN交换机处理每个帧的步骤

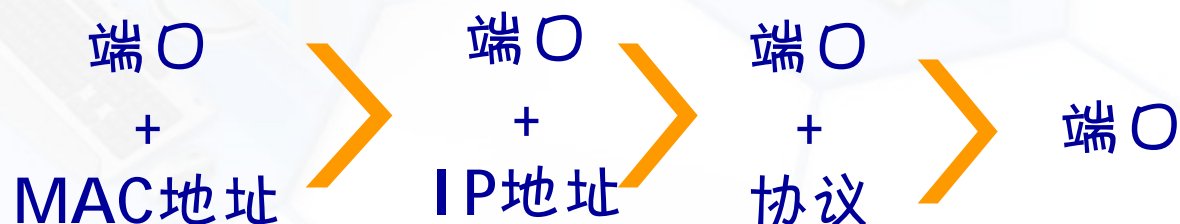
- § 依据入口规则分析每个输入帧，并发送到对应VLAN
- § 检查过滤表以确定输出端口
- § 根据出口规则针对每个输出端口验证其是否为该帧所在之VLAN的成员
- § 检查每个帧是否需要加标记

1) VLAN的入口规则

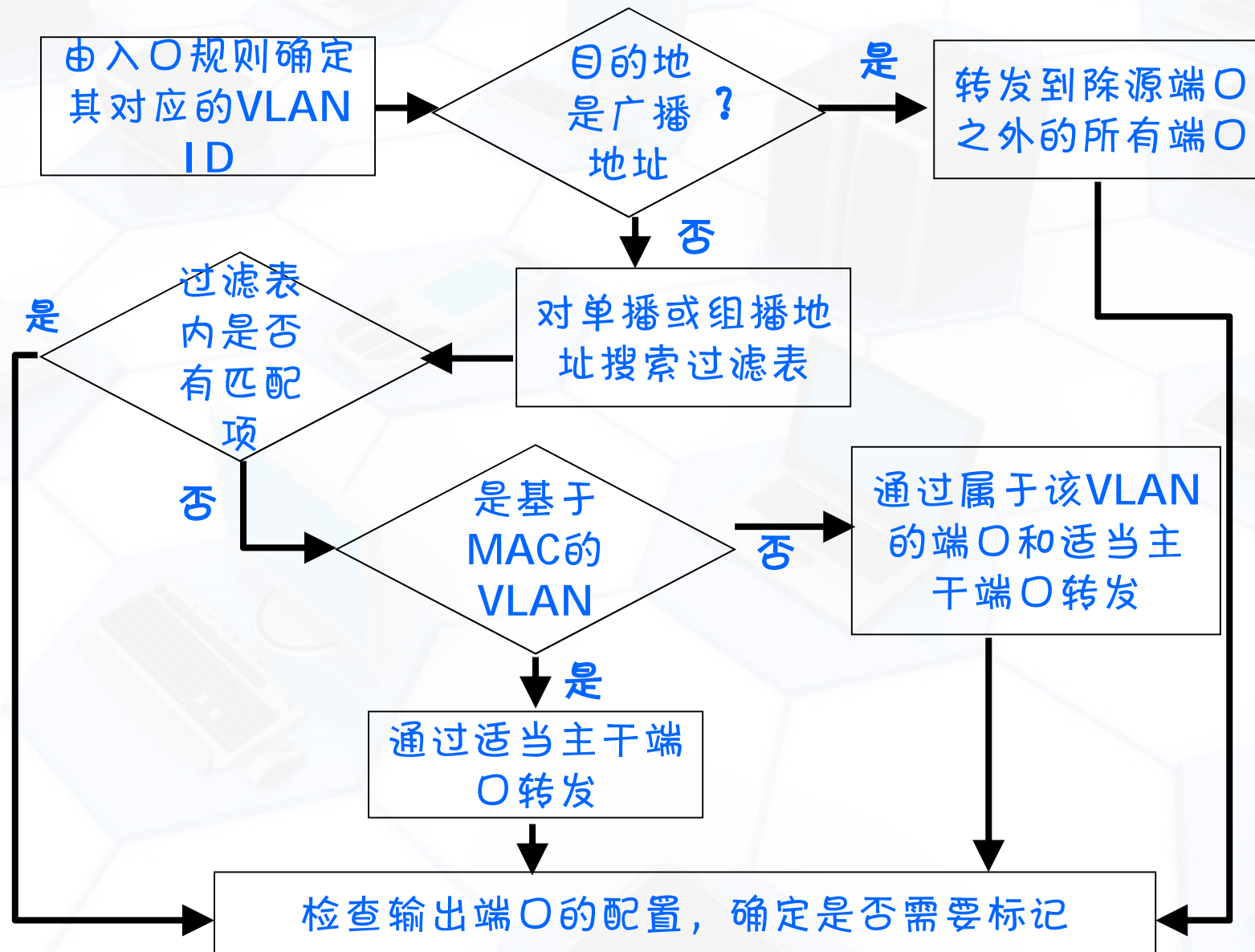
对于没有VLAN标记的帧

- § 如到达端口属于多个VLAN，检查其满足哪个VLAN成员的要求；不满足则丢弃
- § 只满足某一个VLAN成员的要求，转发到该VLAN
- § 满足多个VLAN成员的要求则转发到匹配程度最高者

匹配等级如下：

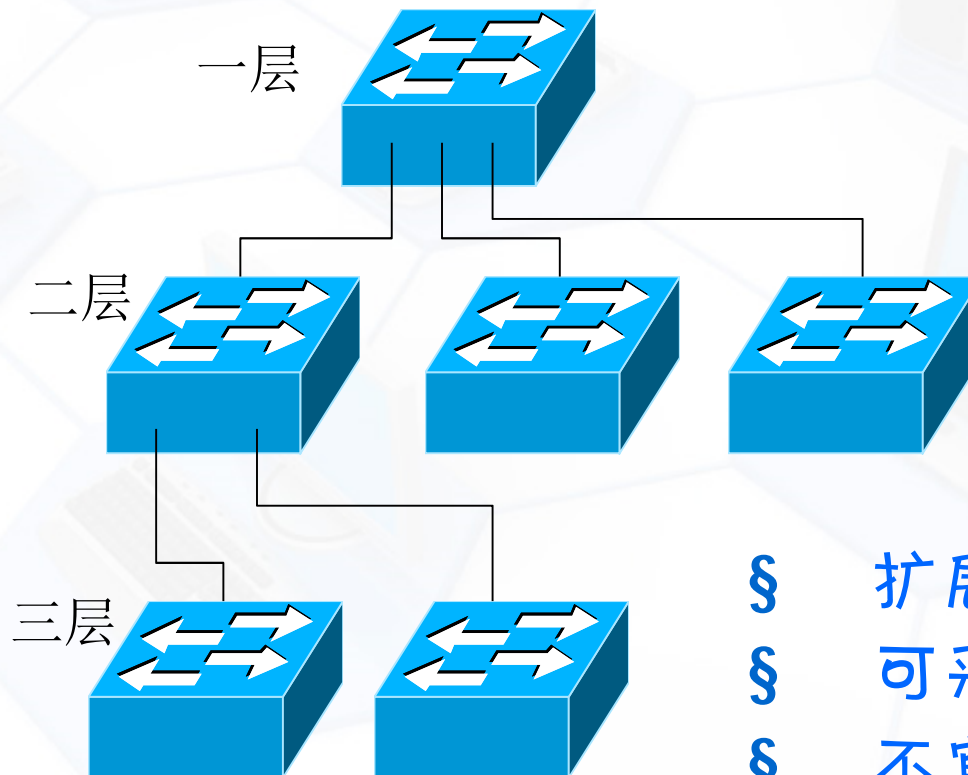


2) VLAN中帧的转发流程



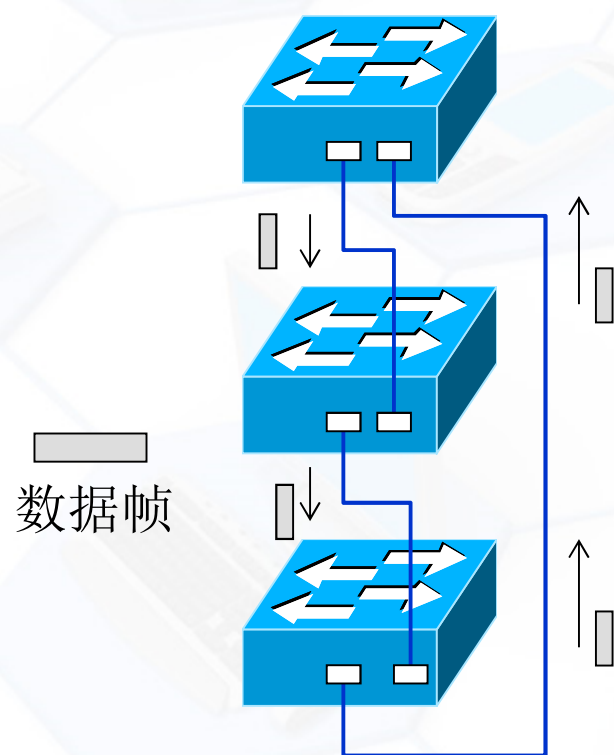
7. 交换机的扩展

1) 级连扩展

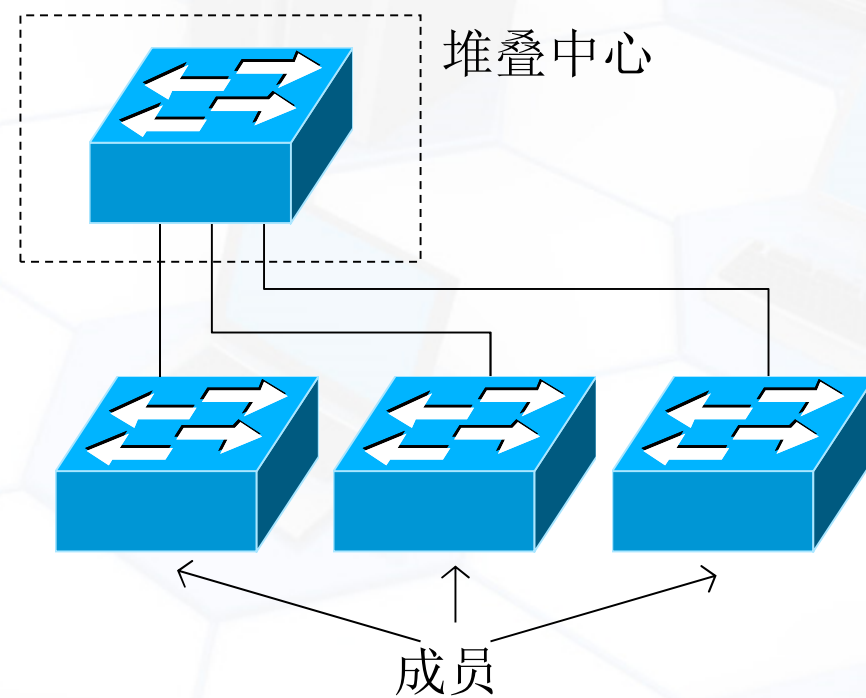


- § 扩展方式简便
- § 可采用一定的冗余技术
- § 不宜超过四层

2) 堆叠扩展



菊花链式



星型堆叠

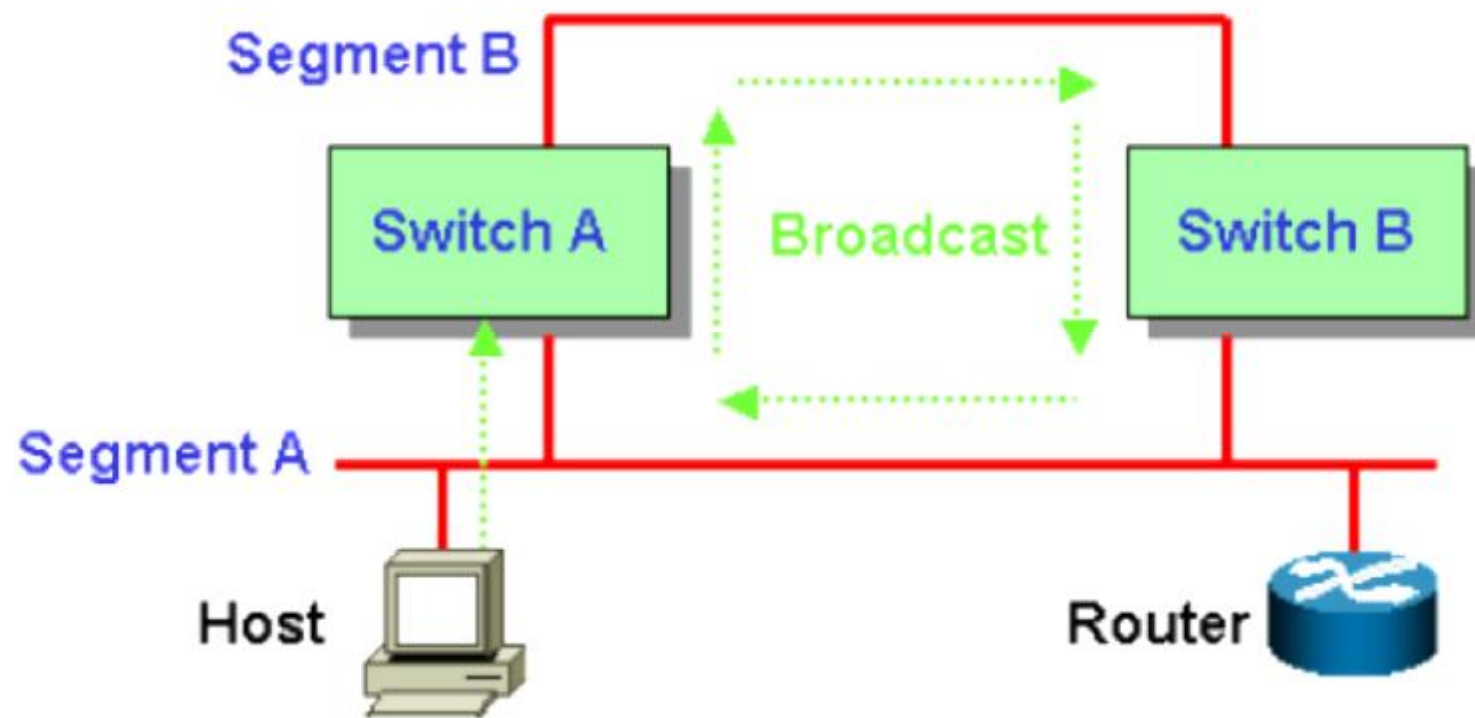
四、生成树协议

1. 生成树协议的作用

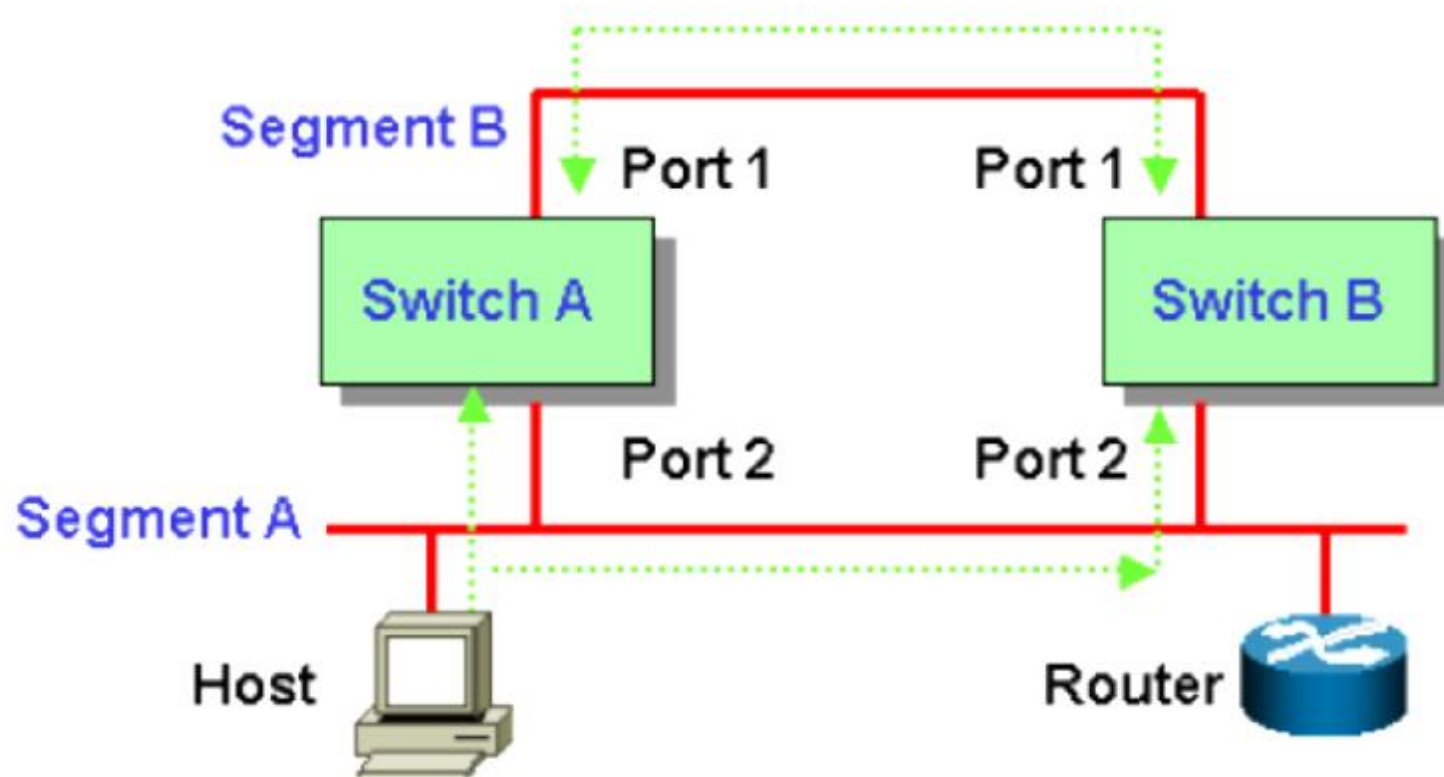
- § 自动控制局域网的拓扑结构
- § 启用局域网的最佳树型结构
- § 发现故障并自行恢复，保持最佳拓扑结构

注意：生成树协议（802.1D）根据交换机配置参数自动实现，树型优劣取决于配置参数

注：根据生成树协议中的命名方式，后续篇幅中仍使用“网桥”来称呼交换机，但两者指同一设备

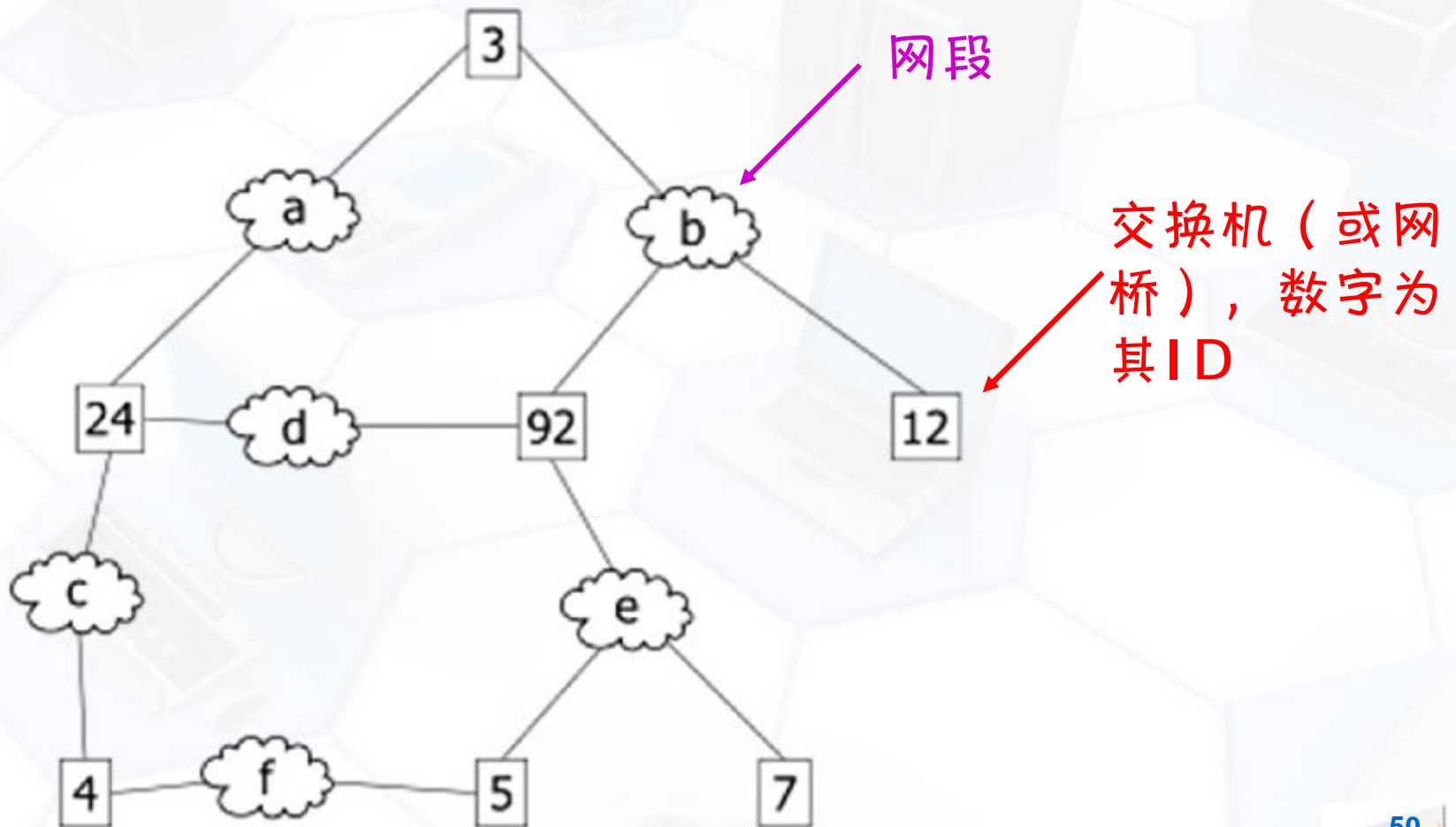


④路连接产生广播风暴



④路连接影响交换机的学习功能

2. 树形结构的组成



3. 生成树协议要素

- § 交换机间交换BPDU（网桥协议数据单元）
- § BPDU传递建立树型必须之信息
- § 使用组播地址01-80-C2-00-00-00
- § 将通路成本分配到每个端口
- § 计算各交换机成本：接收端口成本之和
- § 每个交换机需要确定自己的根端口、根通路成本、指定端口

4. 生成树协议的工作过程

第1步 推选根网桥

- r 网桥的MAC地址：可使用编号最小端口的地址
- r 每个网桥分配一个优先权
- r 网桥标识符 = 网桥地址 + 优先权
- r 具有最小网桥标识符的为根网桥
- r 网管人员将最小优先权值分配给最合适做根的网桥，其次是备份根网桥
- r 网桥优先权的默认值是32768
- r 根网桥通常在树的中心位置及主干上

举例：

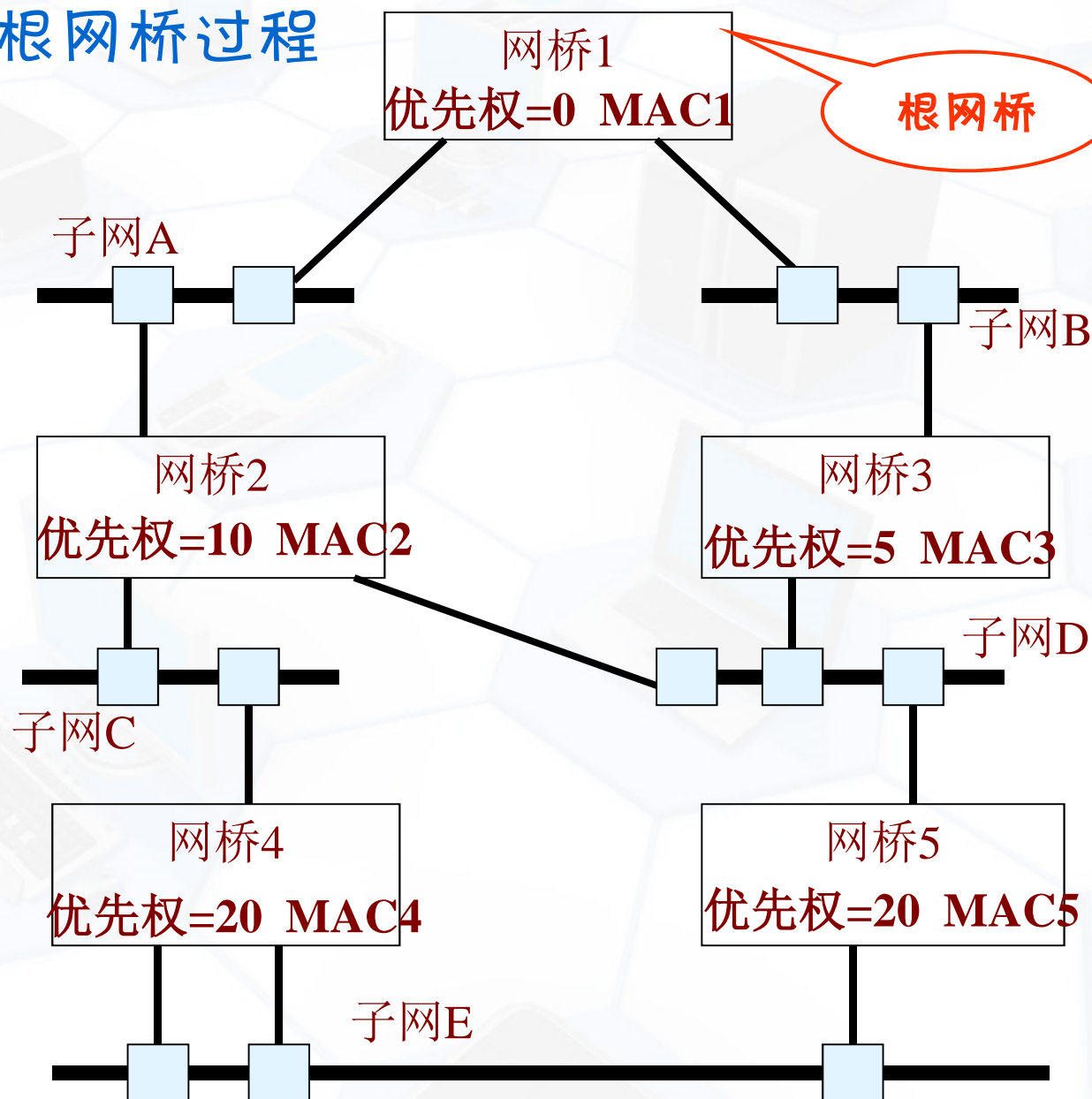
比较网桥标识时，先对比优先权，若优先权一致，再比较MAC地址。如果交换机A (MAC=0000.0000.1111)和交换机B (MAC=0000.0000.2222)的优先权都是10, 则交换机A因地址较小而被选为根交换机；如果希望交换机B做为根，则必须给它设一个更高的优先权（小于10）

r 多个端口的网桥选择最小优先权的端口，阻塞其他端口

r 认为自己为根的网桥向所有子网告知

r 拥有更好条件的网桥会提出反馈

p 推选根网桥过程



第2步 选择最佳通路

r 为每个子网分配成本

r 计算经过的每个子网的通路成本

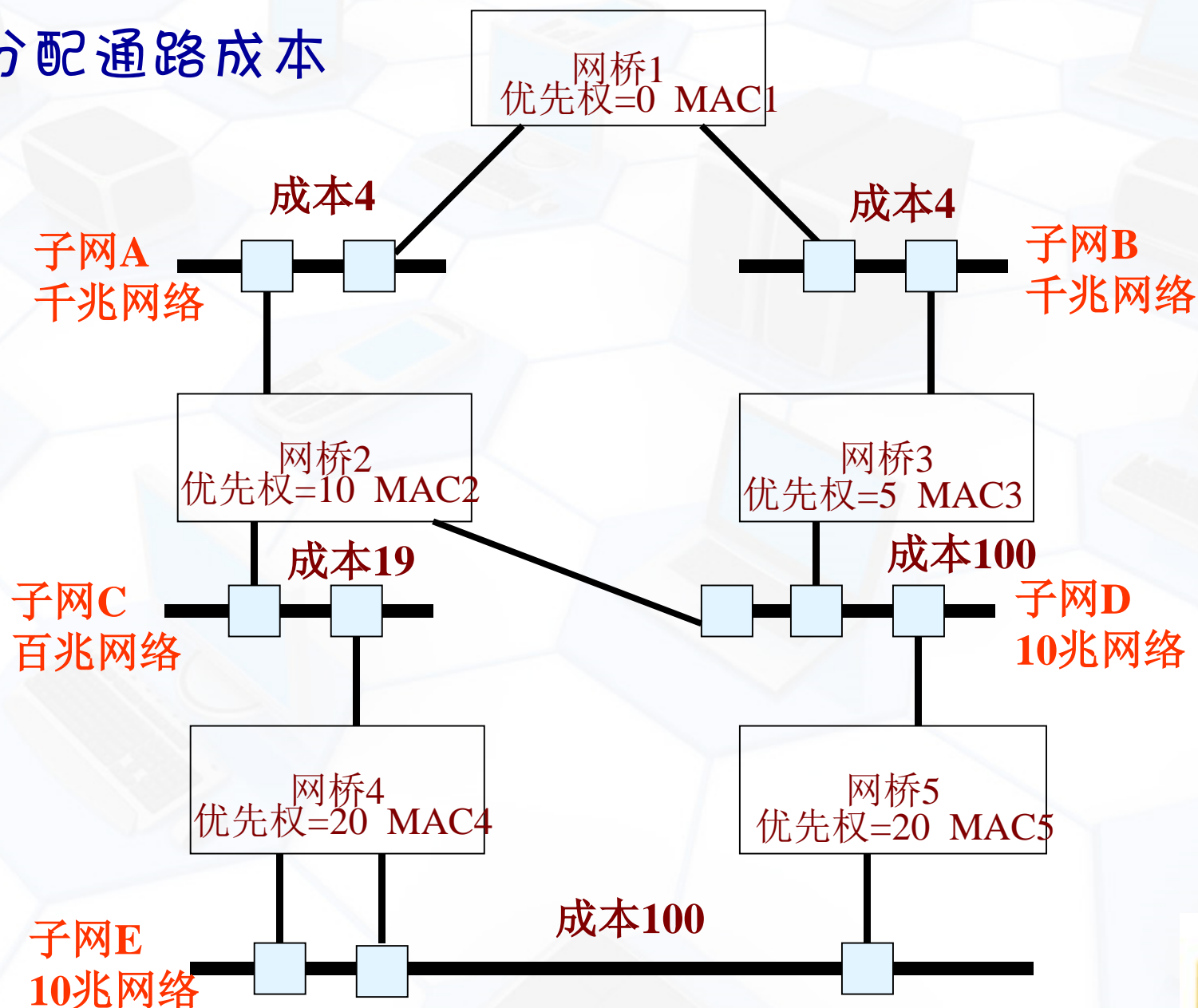
r 通路成本 = 途经所有子网成本之和

r 最佳通路 = 成本最小通路

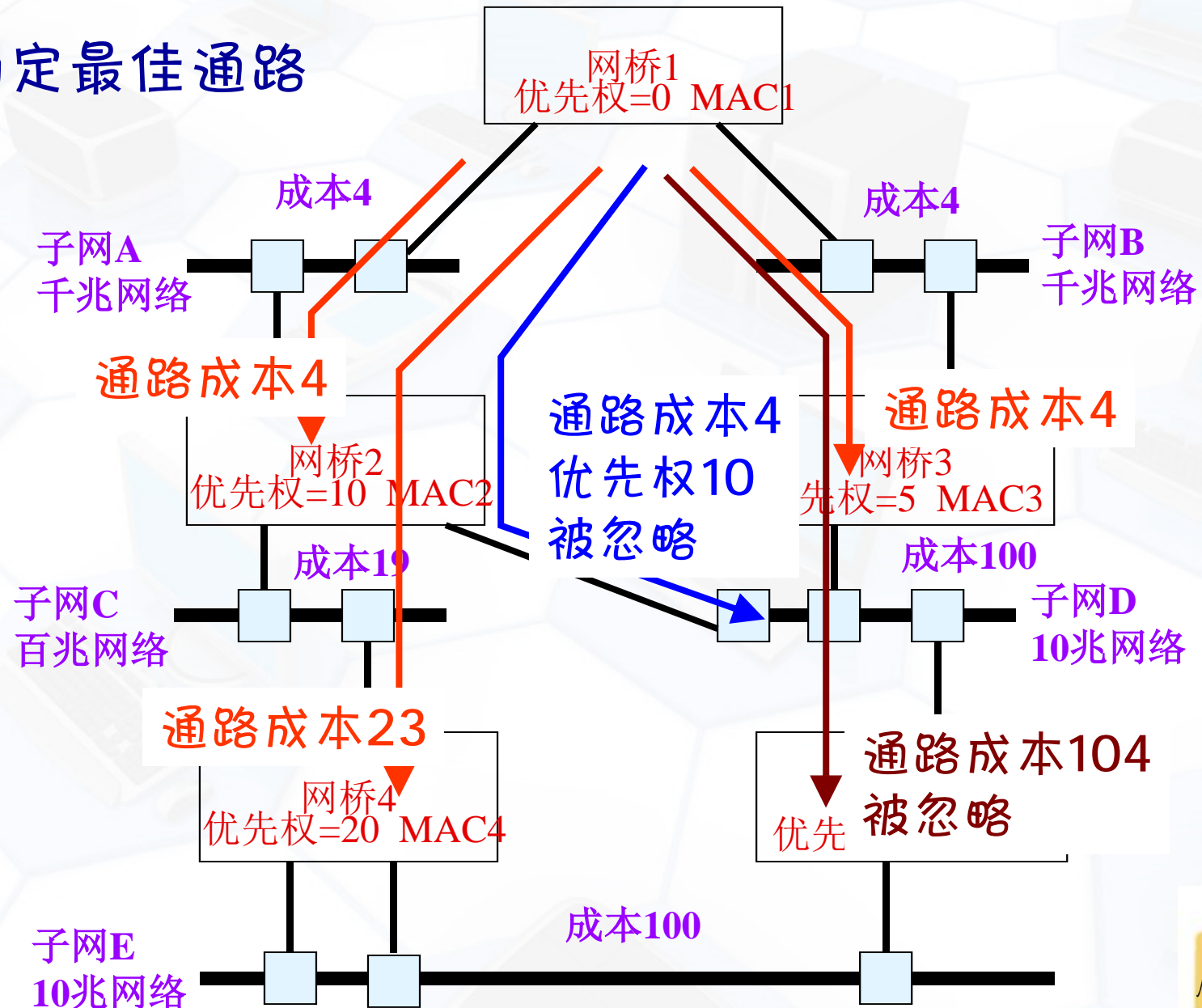
IEEE推荐的子网成本值

链路速率	推荐值	推荐值的范围
4Mb/s	250	100-1000
10Mb/s	100	50-600
16Mb/s	62	40-400
100Mb/s	19	10-60
1Gb/s	4	3-10
10Gb/s	2	1-5

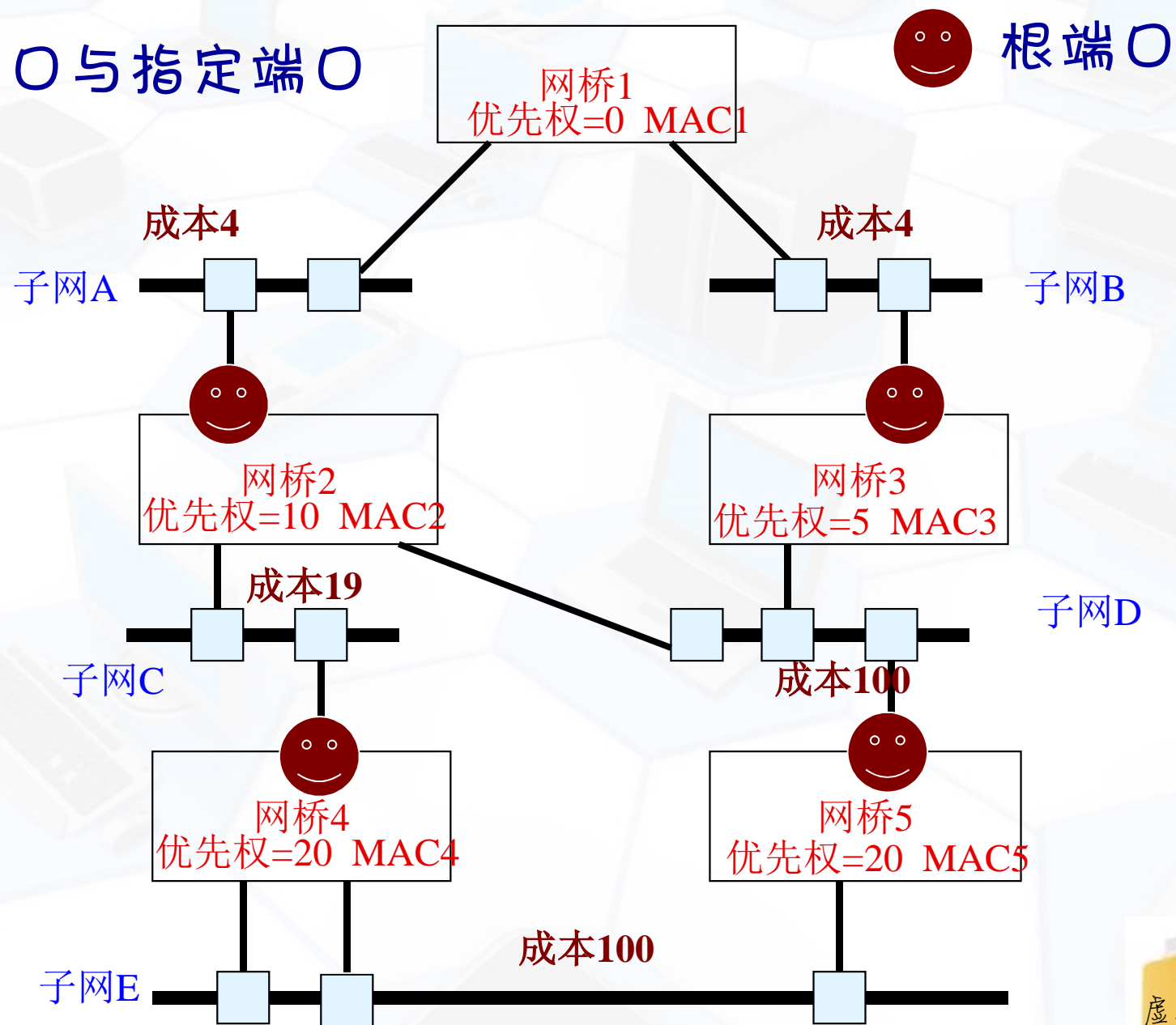
① 分配通路成本



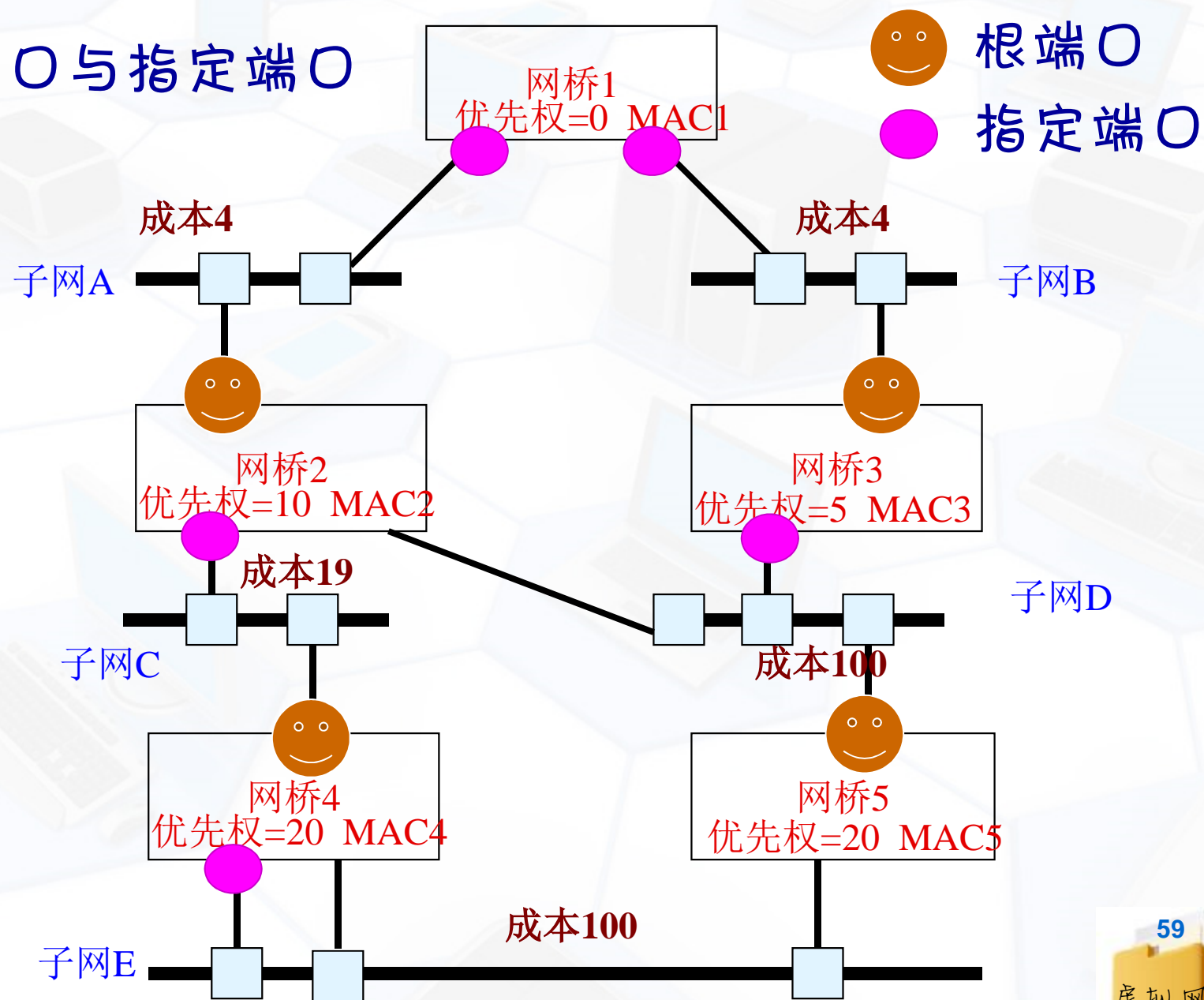
② 确定最佳通路



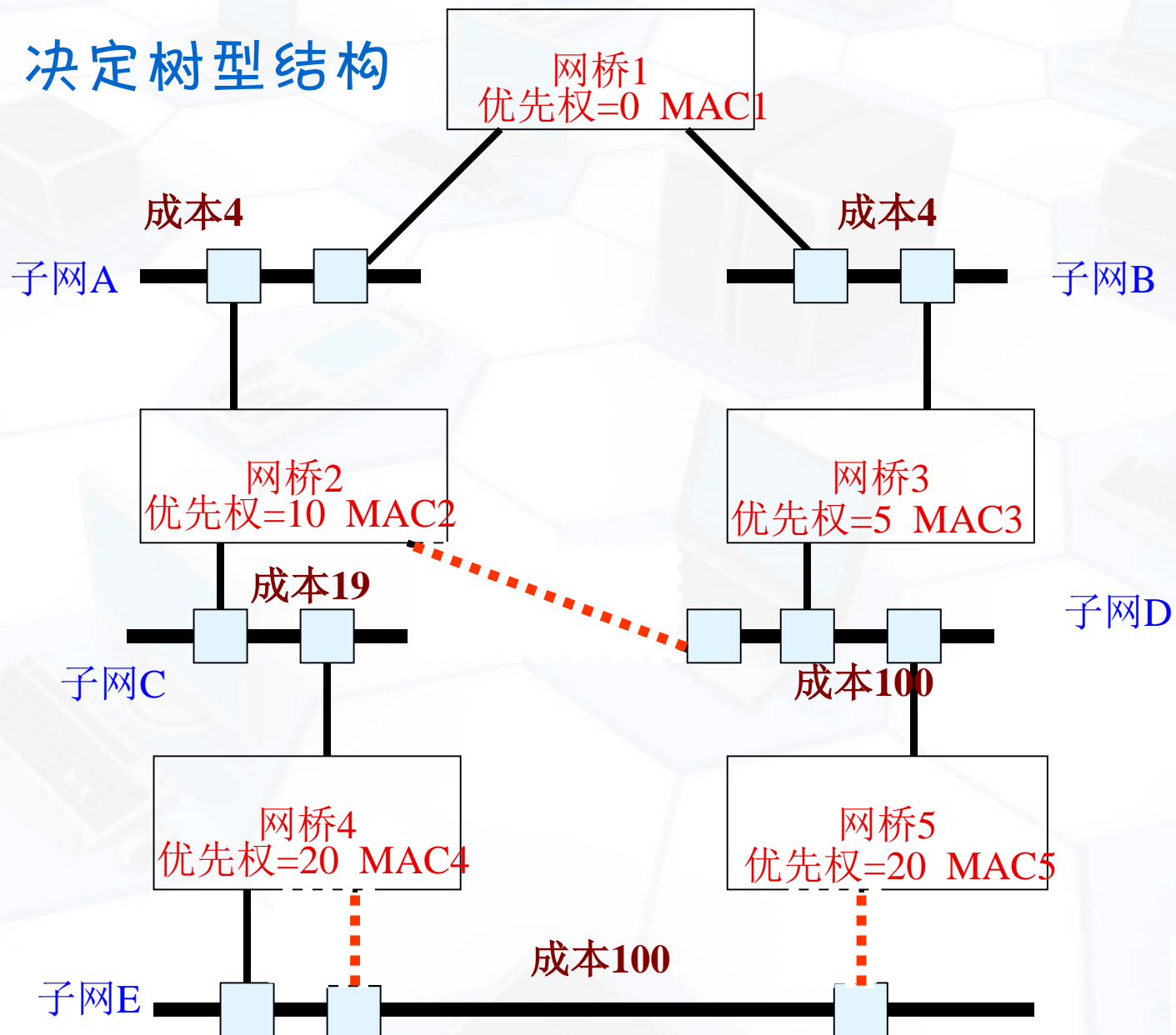
③ 根端口与指定端口



③ 根端口与指定端口



第3步 决定树型结构

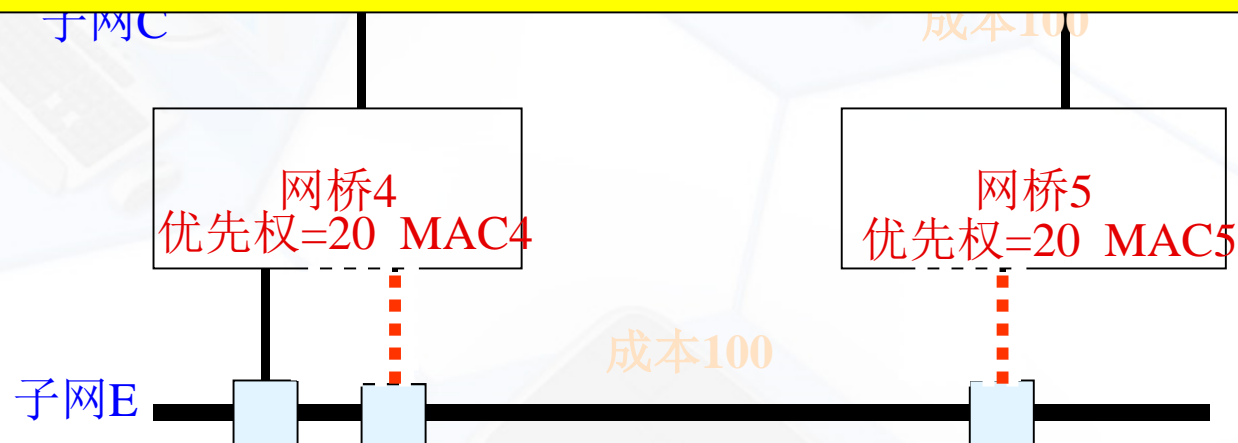


第3步 决定树型结构

网桥1
优先权=0 MAC1

原则

所有的非根端口或指定端口均被阻塞

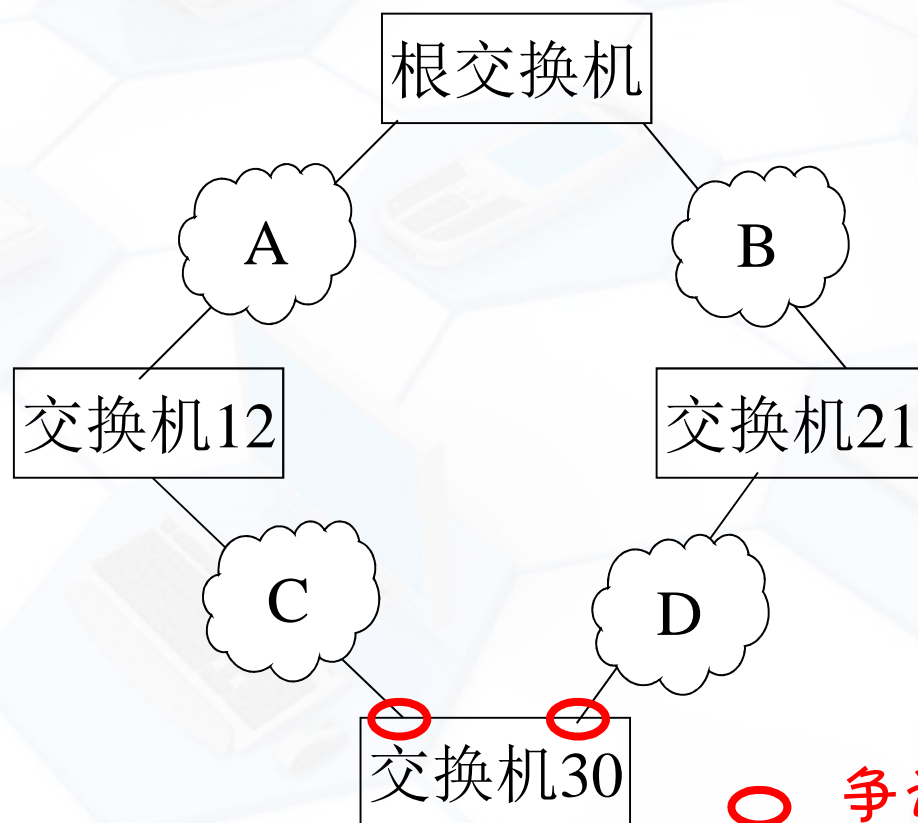


5. 交换机端口状态

状态	功能
禁用	端口不能运行
阻塞	端口只能发送和接收BPDU
侦听	设置定时器且等待。端口继续发送和接收BPDU
学习	学习MAC地址信息。端口继续发送和接收BPDU
转发	定时器满期后，端口已准备接受和转发帧。端口继续学习且能够发送和接收BPDU

6. 协议中的争议解决

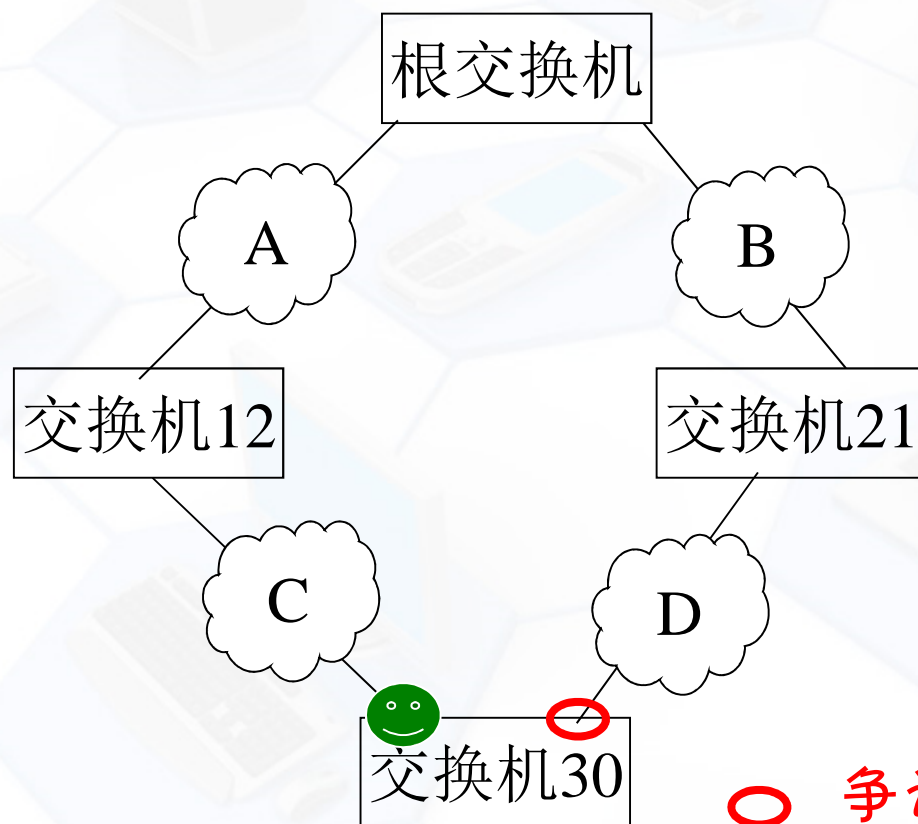
1) 根端口争议的情况



在通路成本相同的情况下，与具有较低ID标识的交换机连接的端口被选为根端口

○ 争议的根端口

1) 根端口争议的情况(结果)

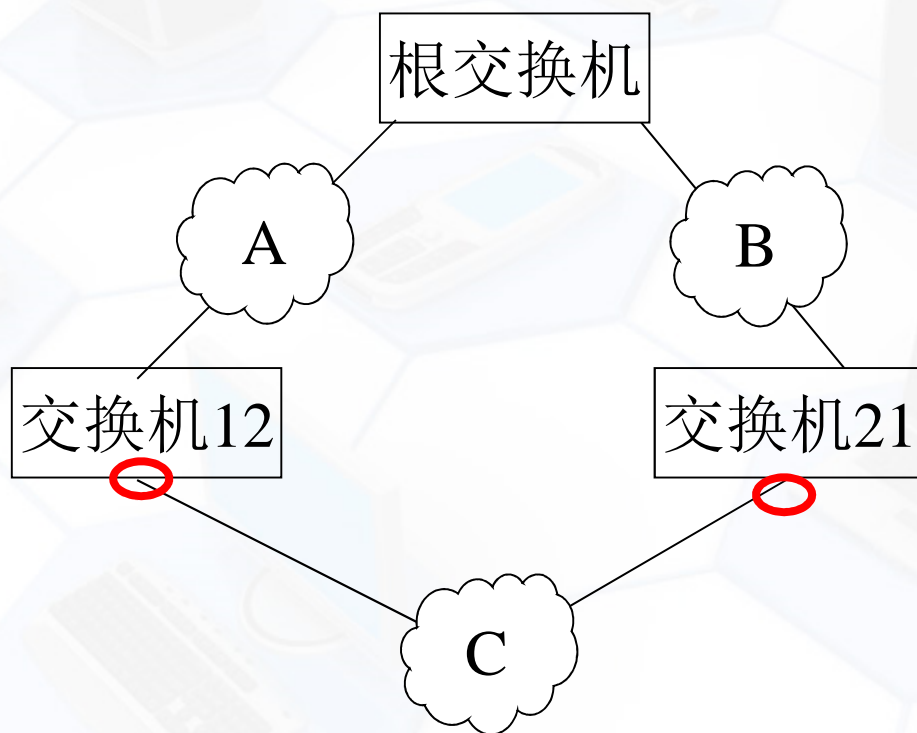


在通路成本相同的情况下，与具有较低ID标识的交换机连接的端口被选为根端口

○ 争议的根端口

😊 最终的根端口

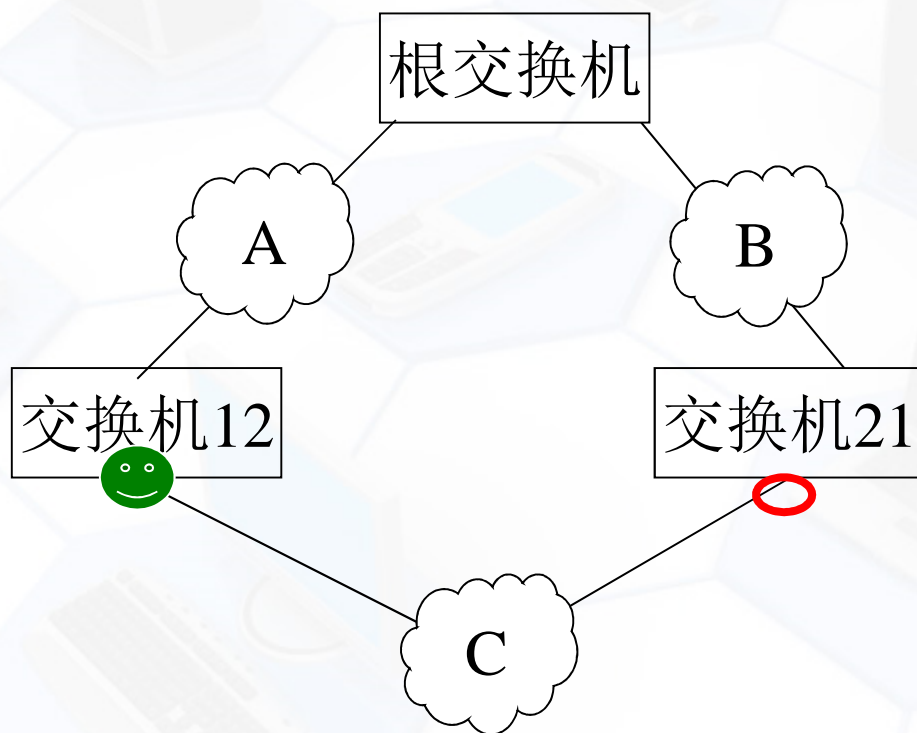
2) 指定端口争议的情况



在通路成本相同的情况下,
具有较低ID
标识的交换机
连接的网段,
其端口被选为
指定端口

○ 争议的指定端口

2) 指定端口争议的情况(结果)

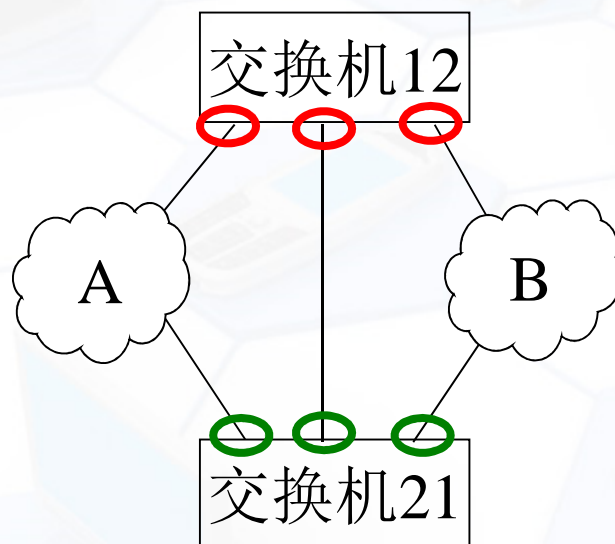


在通路成本相同的情况下，具有较低ID标识的交换机连接的网段，其端口被选为指定端口

○ 争议的指定端口

😊 最终的指定端口

3) 多端口争议的情况



具有较低优先权值的
交换机端口被选为指
定端口或根端口

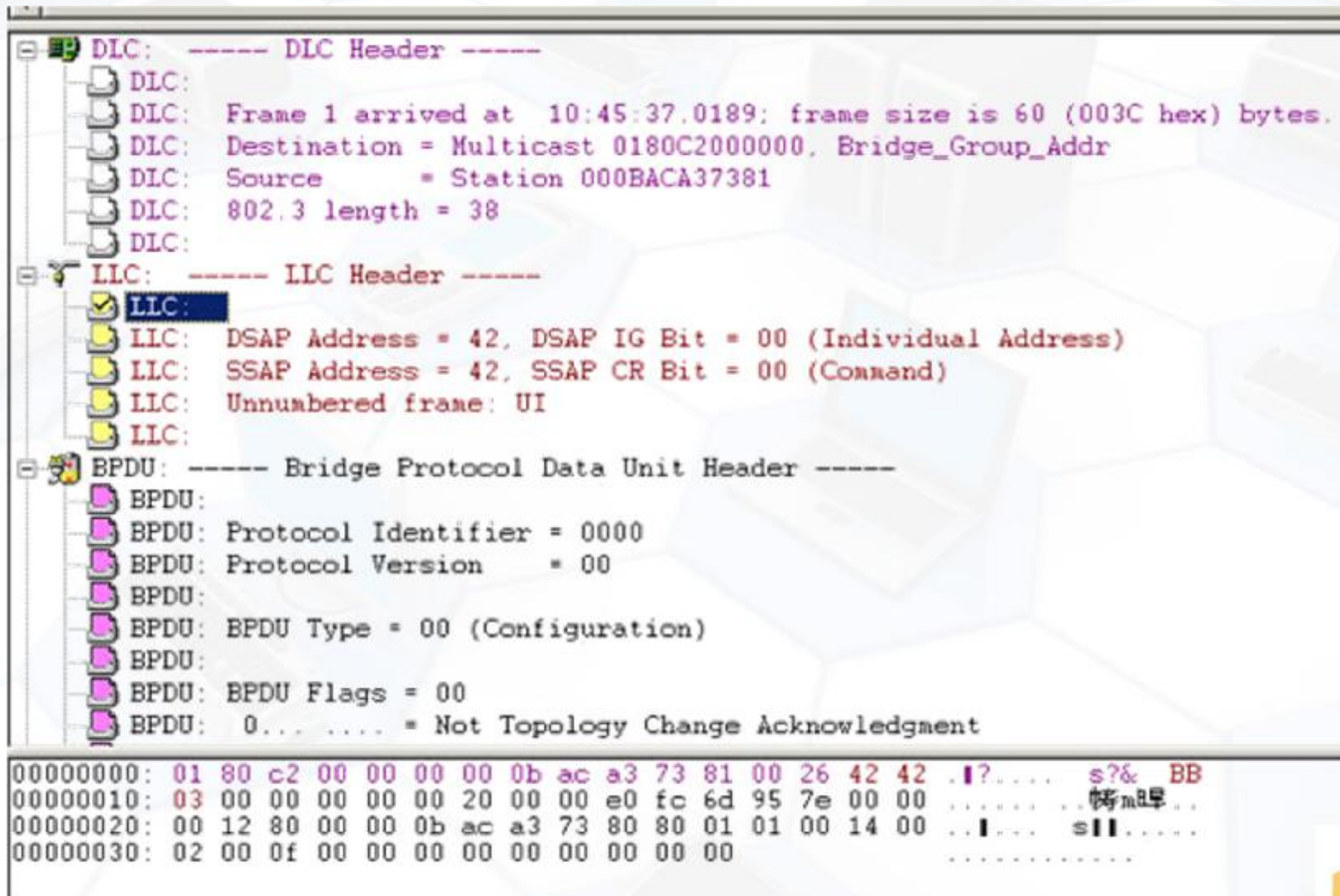
○ 争议的指定端口

○ 争议的根端口

7. BPDU的类型

- r 配置BPDU (CBPDU) - 用于生成树的计算
 - r 拓扑结构改变通告BPDU(TCN) - 发布网络拓扑结构的变化信息
 - r 拓扑结构变化通告确认(TCA)
-
- ü 拓扑结构的变化经由相关的交换机（非根交换机）将TCN信息发送到网络中
 - ü 根交换机收到此TCN后，会设置自己发出的BPDU中的拓扑结构变化标志
 - ü 其他交换机根据收到BPDU中标志的通告重新修正自己的数据转发表

8. BPDU的数据分析



DLC: ----- DLC Header -----

- DLC: Frame 1 arrived at 10:45:37.0189; frame size is 60 (003C hex) bytes.
- DLC: Destination = Multicast 0180C2000000, Bridge_Group_Addr
- DLC: Source = Station 00BACA37381
- DLC: 802.3 length = 38

LLC: ----- LLC Header -----

- LLC: DSAP Address = 42, DSAP IG Bit = 00 (Individual Address)
- LLC: SSAP Address = 42, SSAP CR Bit = 00 (Command)
- LLC: Unnumbered frame: UI

BPDU: ----- Bridge Protocol Data Unit Header -----

- BPDU: Protocol Identifier = 0000
- BPDU: Protocol Version = 00
- BPDU: BPDU Type = 00 (Configuration)
- BPDU: BPDU Flags = 00
- BPDU: 0... .. = Not Topology Change Acknowledgment

00000000: 01 80 c2 00 00 00 00 0b ac a3 73 81 00 26 42 42 .!?.... s?& BB

00000010: 03 00 00 00 00 00 20 00 00 e0 fc 6d 95 7e 00 00 4m8P...

00000020: 00 12 80 00 00 0b ac a3 73 80 80 01 01 00 14 00 ..!... sll.....

00000030: 02 00 0f 00 00 00 00 00 00 00 00 00

一个捕获的配置BPDU实例

1) 配置BPDU

Protocol Identifier	Octet 1
Protocol Version Identifier	2
BPDU Type	3
Flags	4
Root Identifier	5
	6
	7
	8
	9
	10
	11
	12
Root Path Cost	13
	14
	15
	16
	17
Bridge Identifier	18
	19
	20
	21
	22
	23
	24
	25
Port Identifier	26
	27
Message Age	28
	29
Max Age	30
	31
Hello Time	32
	33
Forward Delay	34
	35

```

protocol id: 0000 IEEE 802.1d
version id: 00
bpdu type: 00 config bpdu, 80 tcn bpdu
bit field: 1 byte
  1 : topology change flag
  2 : unused 0
  3 : unused 0
  4 : unused 0
  5 : unused 0
  6 : unused 0
  7 : unused 0
  8 : topology change ack
root priority 2 bytes
root id: 6 bytes
root path cost: 4 bytes
bridge priority: 2 bytes
bridge id: 6 bytes
port id: 2 bytes
message age: 2 bytes in 1/256 secs
max age: 2 bytes in 1/256 secs
hello time: 2 bytes in 1/256 secs
forward delay: 2 bytes in 1/256 secs
  
```

802.1D对BPDU的定义

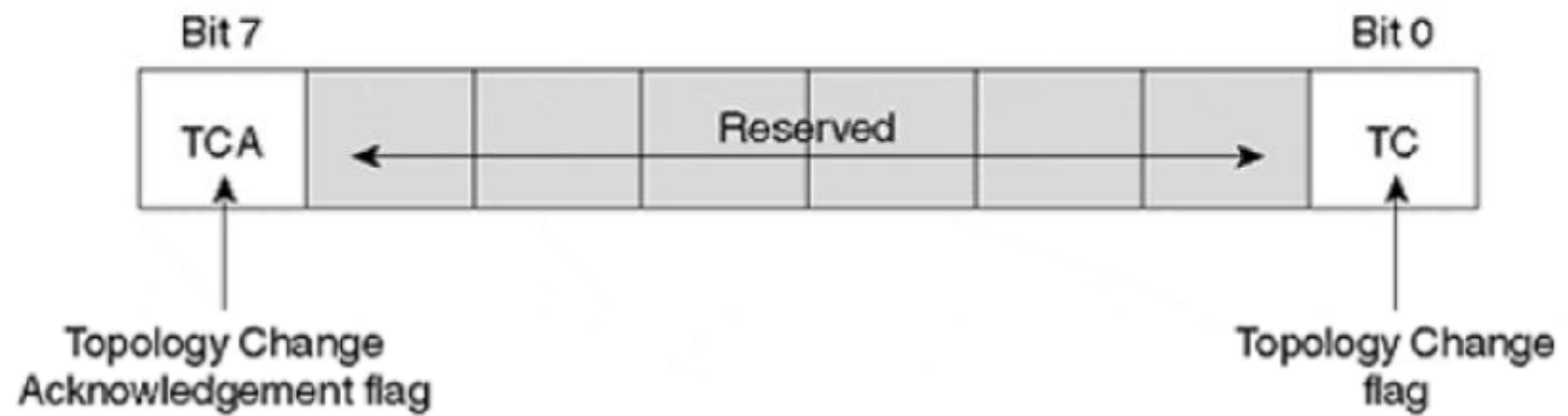
2) 配置BPDU——数据的分析

```
IEEE 802.3
  Address: 00-10-14-23-A4-08 --->01-80-C2-00-00-00
  Length: 38
  Logical Link Control
    SSAP Address: 0x42, CR bit = 0 (Command)
    DSAP Address: 0x42, IG bit = 0 (Individual address)
    Unnumbered frame: UI
  Bridge Protocol Data Unit
    Protocol ID: 0
    Version: 0
    Type: 0 (Configuration)
    Flags: 0x00
    Root ID: 800000101423A408
      Priority: 0x8000
      MAC Address: 00-10-14-23-A4-08
    Root Path Cost: 0
    Bridge ID: 800000101423A408
      Priority: 0x8000
      MAC Address: 00-10-14-23-A4-08
    Port ID: 0x8002
    Message Age: 0
    Max Age: 20
    Hello Time: 2
    Forward Delay: 15
    Frame Padding : (8 bytes)
  Calculate CRC: 0x26544023
```

3) 配置BPDU——常用字段的取值

Table 6-6. Configuration BPDU Fields		
Field	Octets	Use
Protocol ID	2	Always 0
Version	1	Always 0
Type	1	Type of current BPDU 0 = Configuration BPDU
Flags	1	LSB = Topology Change (TC) flag MSB = Topology Change Acknowledgment (TCA) flag
Root BID	8	Bridge ID of current Root Bridge
Root Path Cost	4	Cumulative cost to Root Bridge
Sender BID	8	Bridge ID of current bridge
Port ID	2	Unique ID for port that sent this BPDU
Message Age	2	Time since Root Bridge-created BPDU used to derive current BPDU
Max Age	2	Period to save BPDU information
Hello Time	2	Period between BPDUs
Forward Delay	2	Time spent in Listening and Learning states

配置BPDU的标志字段定义



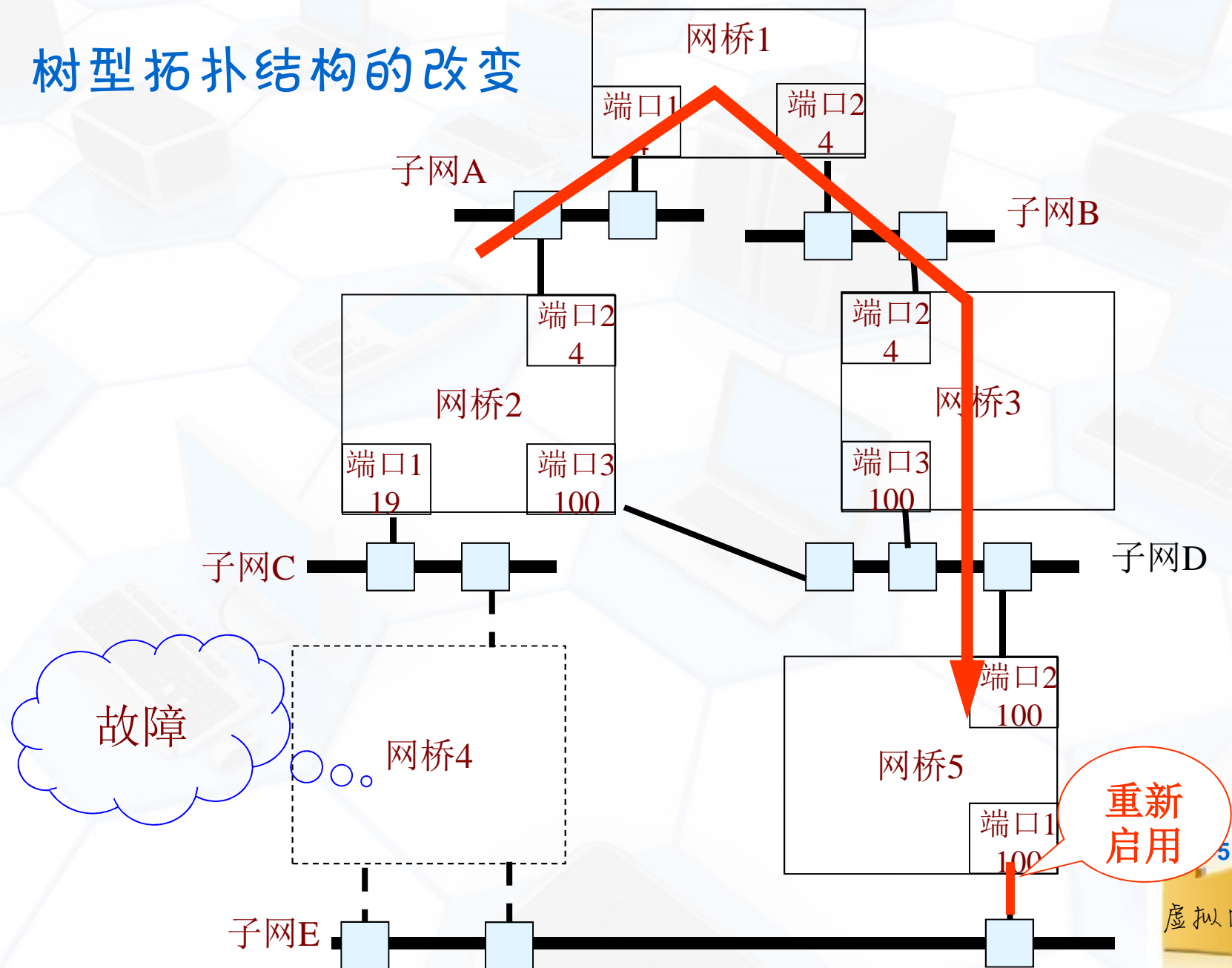
4) 拓扑结构变化BPD

	Octet
Protocol Identifier	1
	2
Protocol Version Identifier	3
BPD Type	4

0x00: Configuration BPD
0x80: TCN BPD

IEEE 802.2
Address: 00-10-14-31-08-08 --->01-80-C2-00-00-00
Length: 7
Logical Link Control
SSAP Address: 0x42, CR bit = 0 (Command)
DSAP Address: 0x42, IG bit = 0 (Individual address)
Unnumbered frame: UI
Bridge Protocol Data Unit
Protocol ID: 0
Version: 0
Type: 0x80 (Topology Change Notification)
Frame Padding : (39 bytes)
Calculate CRC: 0x65ad743d

9. 树型拓扑结构的改变



10. 快速生成树协议

Rapid Spanning Tree Protocol (RSTP)

- ü 也称为802.1w
- ü 二层交换中不能使用路由协议，只能靠STP消除回路，提供冗余
- ü STP恢复故障需要约50 ~ 60s
- ü 在要求更快速的网络融合环境中STP有缺陷
- ü 在三层交换网络中可以利用OSPF/EIGRP等路由协议用更短的时间进行路径选择
- ü RSTP在STP基础上作了改进，但保留了大部分STP的要旨

1) RSTP重新定义端口的状态

ü 802.1D的四种端口状态：阻塞、侦听、转发、学习，端口的角色和状态是混合在一起的

ü 阻塞和侦听状态的端口都不接收帧，也不学习MAC地址；而转发状态的端口又区别不出是根端口还是指定端口

ü RSTP分割了端口的角色和状态，仅有三种端口状态：放弃、学习和转发

ü 端口原有的阻塞、侦听、禁用状态均被定义为RSTP的放弃状态

RSTP与STP的端口状态对比

STP (802.1D) Port State	RSTP (802.1w) Port State	Is Port Included in Active Topology?	Is Port Learning MAC Addresses?
Disabled	Discarding	No	No
Blocking	Discarding	No	No
Listening	Discarding	Yes	No
Learning	Learning	Yes	Yes
Forwarding	Forwarding	Yes	Yes

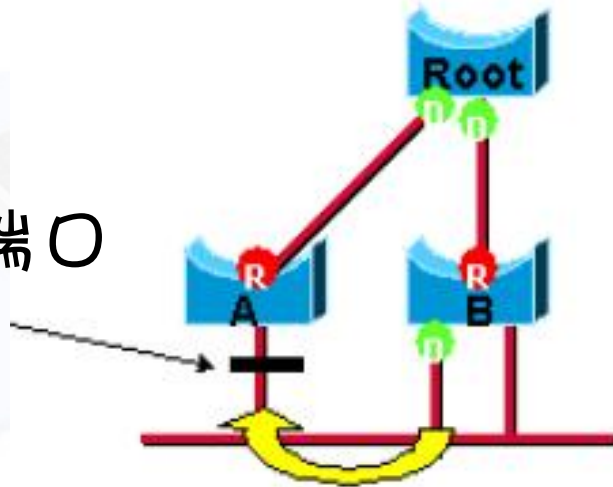
2) RSTP的端口角色

- ü RSTP中保留了根端口、指定端口两种角色
 - ü RSTP将阻塞状态细分为备份和替代两种情况
 - ü 端口角色的变化依据所传递的BPDU
-
- ü 根端口的定义与性质与STP中一样
 - ü 指定端口的定义与性质同STP中一致

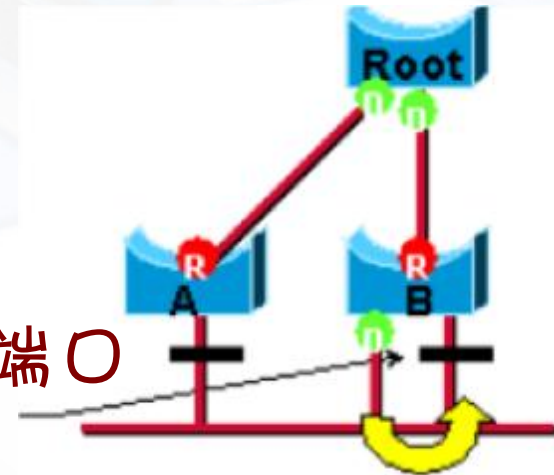
3) RSTP的替代与备份端口

- ü 相当于802.1D中的阻塞端口
- ü 阻塞端口需一直接收BPDU以保持阻塞状态
- ü 替代端口因由于收到其它网桥更优的BPDU而被阻塞（是到根的备选路径）
- ü 备份端口因相同交换机上由于收到自己发出的更优的BPDU而被阻塞（到同网段的备选路径）

替代端口

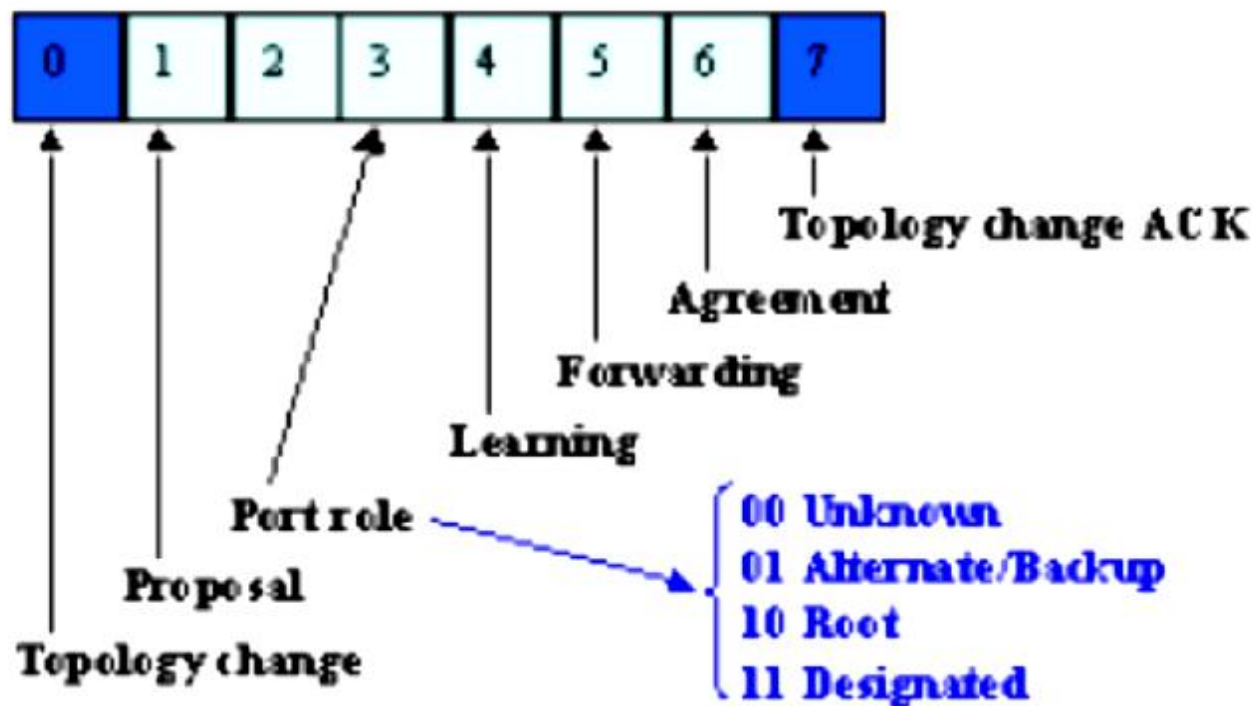


备份端口



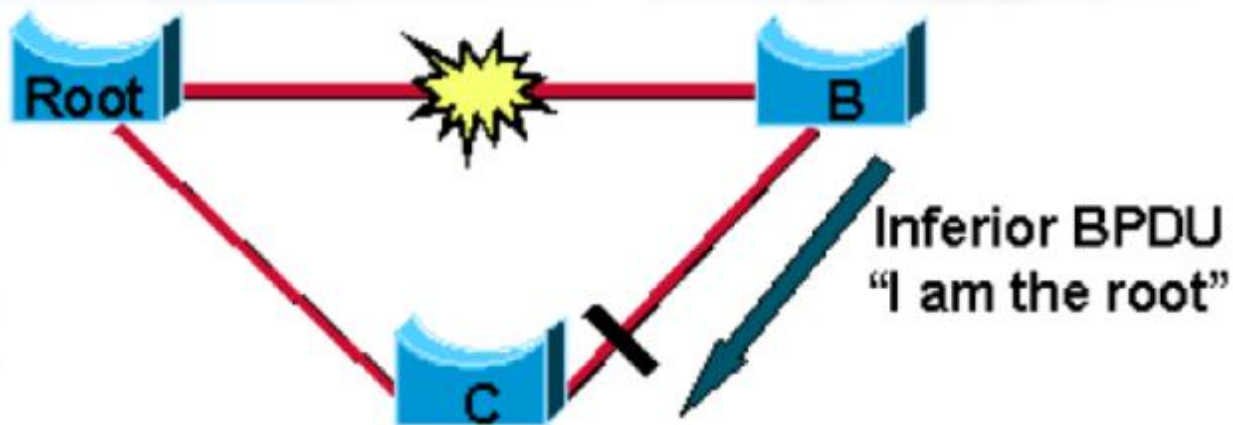
4) BPDU的新格式

- ü STP的BPDU格式中仅有两个标志：Topology Change (TC)/TC Acknowledgment (TCA)
- ü RSTP还使用余下的六位保留标志
- ü RSTP使用的BPDU为类型2、版本2



5) RSTP对BPDU的处理

- ü STP中的非根网桥仅在其根端口收到BPDU时才转发；RSTP则要求每个交换机定期发送反映当前情况的BPDU（默认为2秒，称为Hello time）
- ü 如果连续三次未收到定期的BPDU，则视为相邻节点失效（以便快速发现故障链路）
- ü 接受来自指定网桥或根网桥的次优BPDU



6) 快速转换到转发状态

- ü STP中需要等待网络收敛后才改变端口状态
- ü RSTP采用边缘端口和链路类型的概念实现快速转换

边缘端口

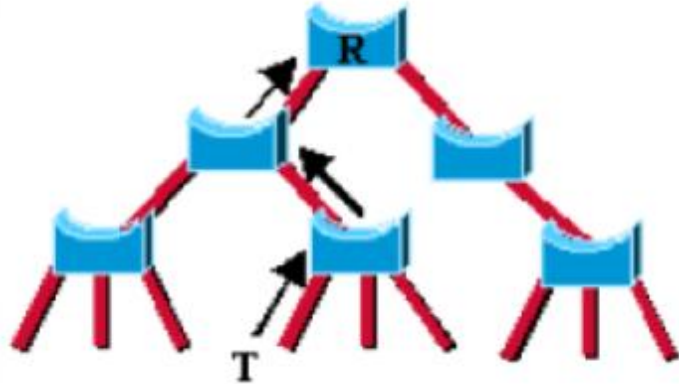
- ü 与终端直接连接的端口不会形成回路，可以跳过学习和侦听阶段直接转换到转发状态
- ü 边缘端口接收到BPDU后直接进入转发状态

链路类型

- ü 能够实现状态快速转换的是点对点链路
- ü 工作在全双工模式的端口被默认为点对点链路

7) 对比STP与RSTP的实例

① 802.1D的树型结构重新构成

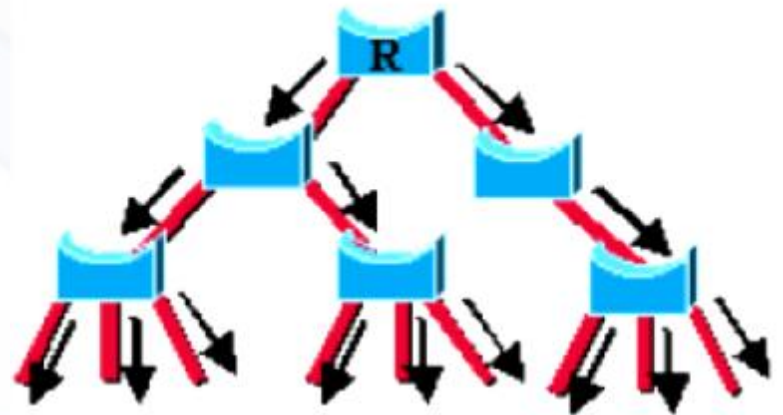


ü T处产生了拓扑改变

ü 一个TCN向上发送给根桥

ü 一旦根得知网络拓扑有改变，它会在它所发出的BPDU中设置TC标志，传播给网络中的所有网桥

ü 当一个网桥收到一个设置了TC标志位的BPDU，它会将其转发表的老化时间减少到转发延时的时间



② 802.1w的树型结构重新构成

ü RSTP中仅非边缘端口转变为转发状态会引起拓扑改变

如果一个RSTP交换机检测到拓扑变化时

ü 如果必要，它会启动一个TC等待定时器，定时器的长度等于它所有非边缘端口和根端口的hello-time的两倍

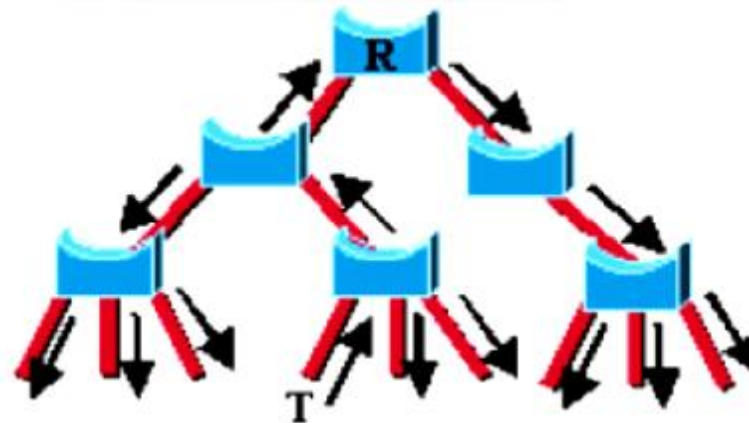
ü 它会刷新所有和这些端口相关的MAC地址

ü 只要端口运行TC等待定时器，从这个端口发送出去的BPDU都会设置TC标志位。当TC等待定时器激活时，根端口也会发送BPDU

② 802.1w的树型结构重新构成（续）

如果网桥从邻接的网桥收到设置了TC位的BPDU

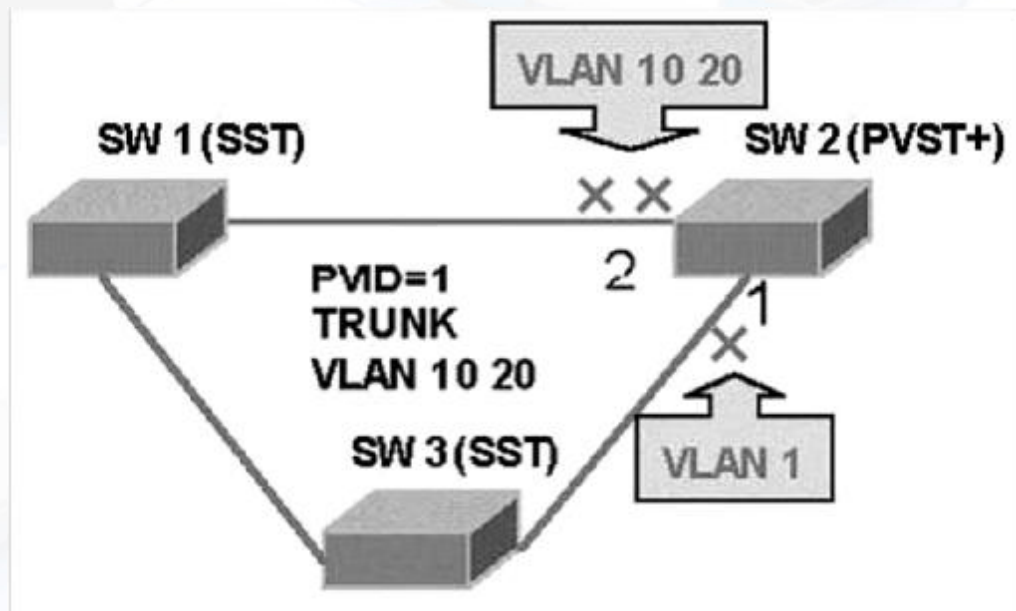
- ü 它会清除所有端口上学到的MAC地址（除到达端口）
- ü 它会启动一个TC等待定时器，并在其所有指定和根端口发送带有TC标志的BPDU（RSTP不再用特定的TCN BPDU）
- ü TCN将快速的洪泛到整个网络，TC传播仅为一个一步的过程（由拓扑改变的发起者来洪泛该信息，而在802.1D中仅根网桥可以洪泛拓扑改变信息）



11. 多生成树协议

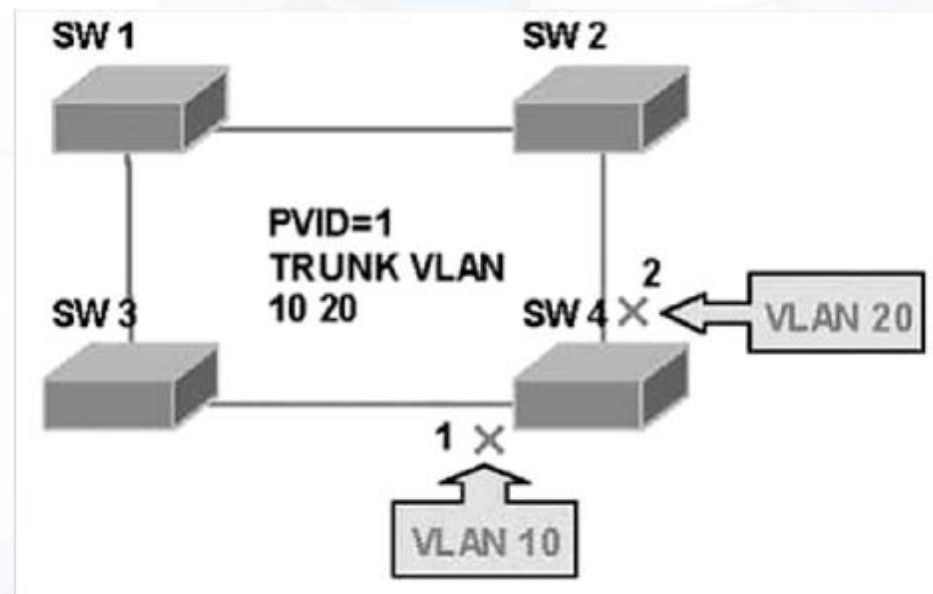
1) Per-VLAN Spanning Tree (PVST)

- ü Cisco 技术，其交换机默认使用PVST+
- ü PVST与STP不兼容，但PVST+能与STP互通
- ü 为每个 VLAN 构建一个不同的逻辑拓扑结构
- ü 通过允许每个 VLAN 拥有不同的转发链接，实现了负载均衡



PVST+与STP的兼容

PVST+的负载均衡



PVST/PVST+的缺点

- ü 由于每个VLAN都需要生成一棵树，PVST BPDU的通信量将正比于主干上的VLAN个数
- ü 在VLAN个数比较多的时候，维护多棵生成树的计算量和资源占用量将急剧增长。尤其是主干集中了很多VLAN的接口状态变化的时候，所有生成树的状态都要重新计算，CPU将是瓶颈
- ü 由于协议的私有性，不同厂家的设备不能直接互通，只能通过一些变通的方式实现

2) Multiple Spanning Tree Protocol (MSTP)

a) 单一生成树协议的缺陷

- ü RSTP和STP在局域网内所有网桥共享一棵生成树

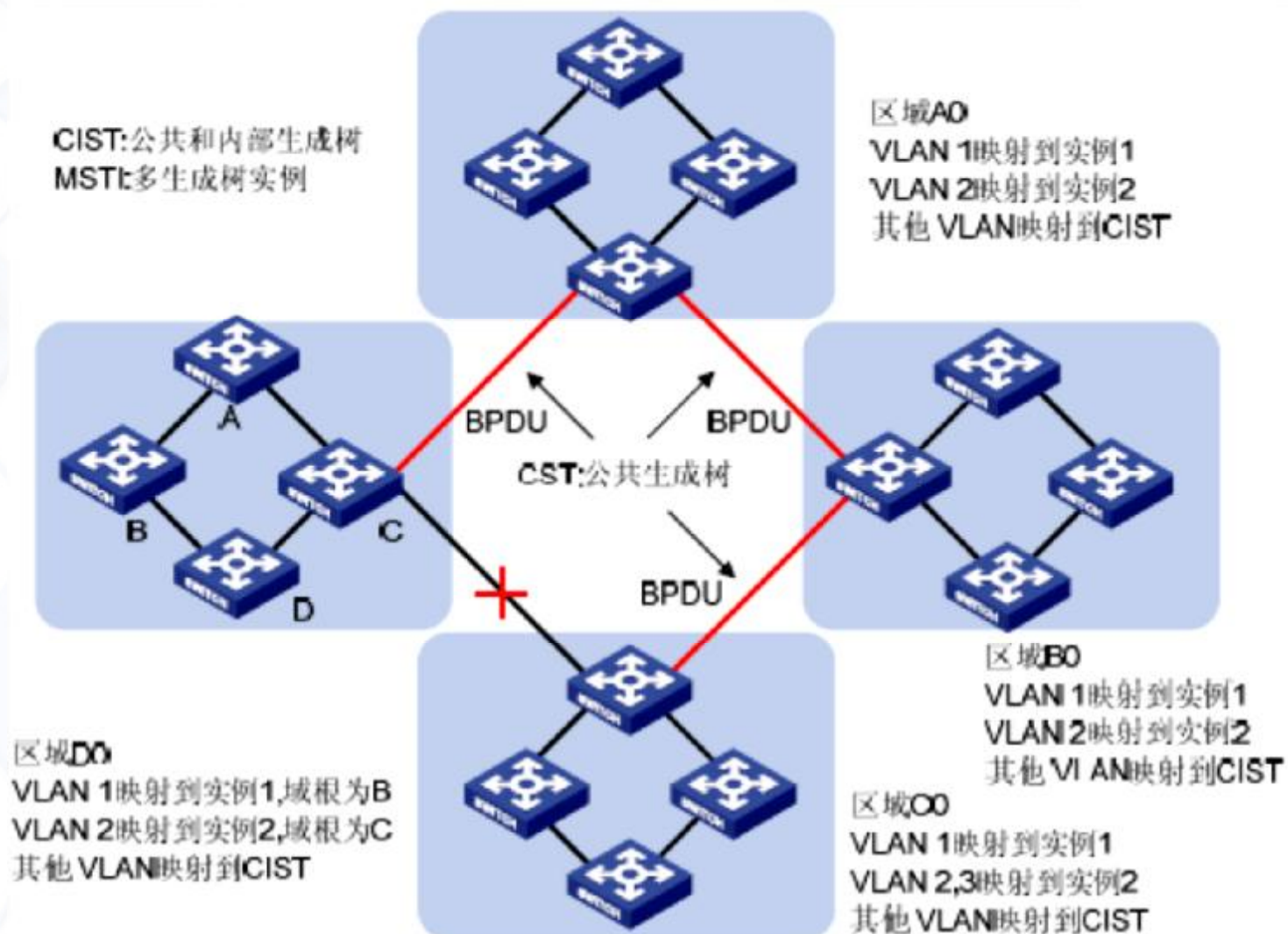
- ü 不能按vlan阻塞冗余链路

- ü 所有vlan的报文都沿着一棵生成树进行转发

b) MSTP的特点

- ü MSTP 设置VLAN 映射表（即VLAN 和生成树的对应关系表），把VLAN 和生成树联系起来
- ü 通过增加“实例”（将多个VLAN 整合到一个集合中），以节省通信开销和资源占用率
- ü MSTP 把一个交换网络划分成多个域，每个域内形成多棵生成树，生成树之间彼此独立
- ü MSTP 将环路网络修剪成为一个无环的树型网络，避免报文在环路网络中的增生和无限循环，同时还提供了数据转发的多个冗余路径，在数据转发过程中实现VLAN 数据的负载分担
- ü 兼容STP 和RSTP

c) MSTP的基本概念图解



ü mst域（multiple spanning tree regions，多生成树域）——由交换网络中的多台交换机以及它们之间的网段构成。这些交换机都启动mstp、有相同的域名、相同的vlan到生成树映射配置和相同的mstp修订级别配置，并且物理上有链路连通

ü 一个交换网络可以存在多个mst域。用户可以通过mstp配置命令把多台交换机划分在同一个mst域内

ü msti（multiple spanning tree instance，多生成树实例）是指mst域内的生成树

ü 一个mst域内可以通过mstp生成多棵生成树，各棵生成树之间彼此独立

ü vlan映射表是mst域的一个属性，用来描述vlan和msti的映射关系

ü ist（internal spanning tree，内部生成树）是mst域内的一棵生成树

ü cst（common spanning tree，公共生成树）是连接交换网络内所有mst域的单生成树。如果把每个mst域看作是一个“交换机”，cst就是这些“交换机”通过stp协议、rstp协议计算生成的一棵生成树

ü ist和cst共同构成整个交换机网络的生成树cist（common and internal spanning tree，公共和内部生成树）

ücist是连接一个交换网络内所有交换机的单生成树，由ist和cst共同构成

ü ist是cist在mst域内的片段，是一个特殊的多生成树实例

ü 域根是指mst域内ist和msti的树根。mst域内各棵生成树的拓扑不同，域根也可能不同

ü 总根（common root bridge）是指cist的树根

d) MSTP端口角色

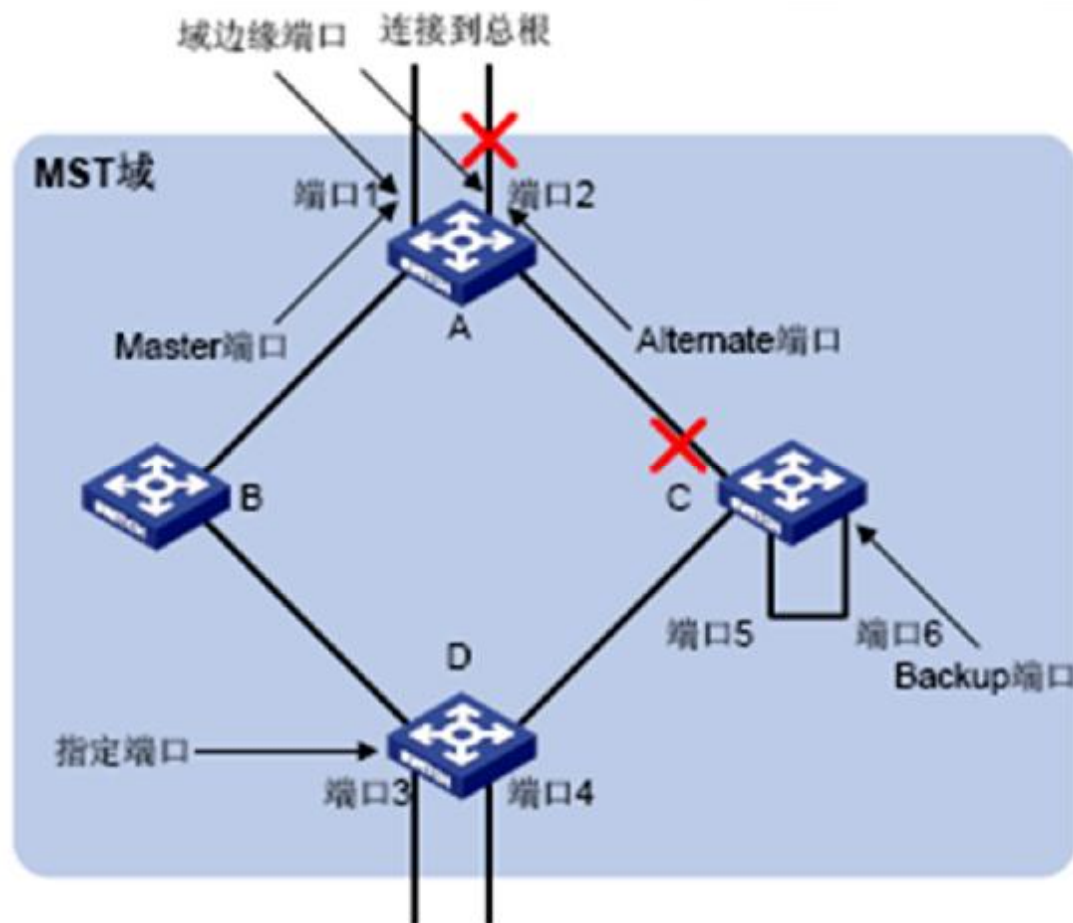
ü MSTP的中端口角色也有根端口、指定端口、替代端口、备份端口以及master端口和域边缘端口

ü master端口是连接mst域到总根的端口，位于整个域到总根的最短路径上

ü 域边缘端口是连接不同mst域、mst域和运行stp的区域、mst域和运行rstp的区域的端口，位于mst域的边缘

ü MSTP的替代端口是master端口的备份端口，若master端口被阻塞后，它将成为新的master端口

e) MSTP的端口状态图解



u 设备A、B、C、D构成一个MST域

u 设备A的端口1、端口2向总根方向连接

u 设备C的端口5、端口6构成了环路

u 设备D的端口3、端口4向下连接其他的MST域

f) MSTP基本原理

ü 将整个二层网络划分为多个mst域，各个域之间通过计算生成cst

ü 域内通过计算生成多棵生成树，每棵生成树都是一个msti

ü 同RSTP一样使用配置消息进行生成树的计算，但配置消息中携带的是各交换机上MSTP的配置信息

cist生成树的计算

- ü 通过比较配置消息，在整个网络中选择一个优先级最高的交换机作为cist的树根
- ü 在每个mst域内通过计算生成ist；同时mstp将每个mst域看作单个交换机，经计算在mst域间生成cst，再由cst和ist构成整个交换网络的cist

msti的计算

- ü 在mst域内，根据vlan和生成树实例的映射关系，针对不同的vlan生成不同的生成树实例
- ü 每棵生成树独立计算，计算过程与STP/RSTP计算生成树的过程相似