第2章 交换机的结构与基本功能

本章提要

交换机用做网络集中设备，其端口连接网络中的主机。在转发数据帧时，端口带宽能够独享。

交换机按其工作在 OSI 参考模型的对应层次，包括第二层、第三层和第四层交换机，可管理的交换机内置了操作系统软件。

第二层交换机采用帧交换转发数据，帧交换的方式有 3 种，即存储转发、伺机通过和自由分段。

交换机通过学习进入自己端口的数据帧源 MAC 地址，记下地址与端口的对应关系，生成 MAC 地址与交换机端口的对应关系表——MAC 地址表；通过 MAC 地址表，交换机可以实现数据帧的单播转发。

使用备份连接是提高网络可靠性的常用方法,但所形成的环路可能会导致广播风暴和引起多帧副本问题。STP 协议的应用可消除环路问题，使冗余备份得以实现。

【学习目标】

（1）了解以太网交换机的基本结构与功能。
（2）学会查看 MAC 地址表。
（3）理解链路冗余的作用和环路的危害，学会生成树协议 STP 的配置。

2.1 交换机的作用与组成

本章所指的交换机若无特别说明，均指以太网交换机。在以太网络中，交换机起信息中转站的作用，它把从某个端口接收到的数据从其他端口转发出去。不同厂家、不同型号的交换机，其外观和内部组成有一定的差异。

2.1.1 交换机的外观

交换机前面板上的多个 RJ-45 接口是以太网口，用来连接计算机或其他交换机。后面板或前面板上的串口是交换机的配置口，用串口线缆（Console 线）将其与计算机的

串口连接起来，可实现对交换机的配置操作。

前面板上有若干个指示灯，其亮、灭或闪烁可以反映交换机的工作状态是否正常，此外还有电源插口、电源开关等。

可上机架（柜）式交换机的标准长度为 48.26cm（19in）。如图 2-1 所示的是 Cisco Catalyst 4500 和 Cisco Catalyst 2900 系列交换机的外观。

Cisco Catalyst 4500 系列

Cisco Catalyst 2900 系列

图 2-1　交换机的外观

2.1.2　交换机的内部组成

交换机的内部组成包括以下部分。

CPU：交换机使用特殊用途集成电路芯片 ASIC，以实现高速的数据传输。

RAM/DRAM：主存储器，存储当前运行的配置文件。

NVRAM：存储备份配置文件等。

Flash ROM：存储系统软件、映像文件等，是可擦可编程的 ROM。

ROM：存储开机诊断程序、引导程序和系统软件。

接口电路：交换机各接口的内部电路。

2.2　交换机的分类

可按多种方式对交换机进行分类，若参照开放系统互联（OSI）参考模型，则交换机属于第 2 层～第 4 层的设备。

2.2.1　OSI 参考模型与数据通信设备

开放系统互联参考模型分为 7 层，其层次及其相应设备如表 2.1 所示。

根据 OSI 参考模型，每一层都使用相应的协议实现特定的功能，完成数据交换。每一层数据逻辑上在源主机与目标主机对应层之间进行传输，屏蔽下层的细节。而数据实际的传输过程是：在发送端，应用层数据经过下面各层，依次被各层进行封装，最后通过物理层下的传输介质完成物理层比特流的传输，到达接收端的物理层；在接收端，各层依次拆封并向上层提交数据，最后送达应用层。

表 2.1　OSI 参考模型的层次及其相应设备

层　数	名　称	协议数据单元名称	相应设备及其作用
第七层	应用层	Data	计算机；处理相应数据
第六层	表示层	Data	计算机；处理相应数据
第五层	会话层	Data	计算机；处理相应数据
第四层	传输层	Segment	四层交换机、计算机；处理数据字段
第三层	网络层	Packet	路由器、三层交换机；处理数据包
第二层	数据链路层	Frame	交换机、网桥、网卡等；处理数据帧
第一层	物理层	Bit	各种接口线缆、网卡等；确定与传输媒体接口有关的特性，处理 Bit 流

交换机可以工作在第 2 层～第 4 层，对应的技术称为第二层、第三层和第四层交换技术，第二层和第三层交换机是目前使用最多的交换机。

本书主要介绍第二层交换技术和第二层交换机的应用，同时也介绍涉及工程应用的第三层交换机的配置和应用。

2.2.2　交换机的简单分类

这里对以太网交换机按配置可否改变或者按在 OSI 参考模型中的对应层次来进行简单的分类。

1．模块式与固定配置式

按交换机的配置可否改变，可把交换机分为模块式和固定配置式交换机。

模块式交换机的模块可以拔插，模块通常是 100Mbit/s 或 1000Mbit/s 光纤接口模块、1000Mbit/s RJ-45 接口模块，或者是堆叠模块。交换机上设有相应的插槽，使用时，将模块插入插槽之中。模块式交换机的配置灵活，模块可按需要购买。一般来说，模块式交换机的档次较高，模块插槽结构可最大限度地保护模块。

固定配置式交换机的接口固定，硬件不可升级。

2．第二层、第三层与第四层交换机

第二层交换机工作在 OSI 参考模型的第二层，它的每个端口拥有自己的冲突域。如果第二层交换机具有虚拟局域网（Virtual Local Network，VLAN）功能，则每一个 VLAN 称为一个广播域。第二层交换机采用 3 种方式转发数据帧，即直通（Cut Through）、存储－转发（Store and Forward）和自由分段（Fragment Free）。

第三层交换机根据目的 IP 地址转发数据包，与路由器一样，它也必须创建和动态维护路由表。但是，第三层交换机能做到“一次路由，多次交换”，即第三层交换机能够把报文转发到不同的子网，并在后续的通信中采用比路由更快的交换方式转发数据包。

第四层交换机可以解释第四层的传输控制协议（TCP）和用户数据报协议（UDP）信息，允许设备为不同的应用（使用端口号区分）分配各自的优先级。这样，第四层交换机可以“智能化”地处理网络中的数据，最大限度地避免阻塞，提高带宽利用率。

2.3 交换机在网络中的连接及作用

2.3.1 交换机的端口

交换机的端口又称接口，两种都是合规术语。是 8 个引脚的 RJ-45 电接口或 SC、ST、FC 等光纤接口。其电接口种类通常有 10Base-T、10Base-F、100Base-TX、100Base-T4、1000Base-T、1000Base-CX 等，光纤接口种类通常有 100Base-FX 和 1000Base-FX 等。

其中，Base 指的是采用基带传输技术，10、100 和 1000 分别代表传输速率为 10Mbit/s、100Mbit/s 和 1000Mbit/s，通常把对应的技术分别称为以太网、快速以太网和千兆位以太网。

交换机的各种接口如表 2.2 所示。

表 2.2 交换机的各种接口

标准类型		传输速率（Mbit/s）	接口标准	传输介质	传输距离（m）	备注
10Base-T		10	RJ-45	UTP（非屏蔽双绞线）	100	
10Base-F		10	光纤接口	62.5/125MMF（多模光纤）	2000	
100Base-TX		100	RJ-45	UTP	100	
100Base-T4		100	RJ-45	UTP（4 对芯线）		
100Base-FX		100	光纤接口	62.5/125MMF	412	半双工
				62.5/125MMF	2000	全双工
				9/125SMF（单模光纤）	10000	
1000Base-CX		1000	RJ-45	STP（屏蔽双绞线）	25	
1000Base-T		1000	RJ-45	UTP（4 对芯线）	100	
1000Base-FX	-SX（780nm 短波）	1000	光纤接口	62.5/125MMF	260	使用 1550nm 波长的单模，最大传输距离为 120km
				50/125MMF	525	
	-LX（1300nm 长波）			62.5/125MMF	550	
				50/125MMF	550	
				9/125SMF	3000～10000	

2.3.2 共享式与交换式网络

采用双绞线或光纤作为传输介质的网络，使用集线器或交换机作为网络的中心。计算机之间的通信，通过集线器或交换机进行数据的转发。

1. 集线器与共享式局域网

集线器通常称为 Hub，按其使用的技术可分为被动式与主动式两种。前者只提供简单的集中网线转发数据的工作，后者可对数据做一定的处理。

按集线器按端口的传输速率（或称带宽）来分，有 10Mbit/s 和 100Mbit/s 两种。通

常所说的集线器是指共享式集线器，其带宽是所有端口共享的。例如，一台 16 端口的 100Mbit/s 集线器，当全部端口都使用时，每一端口的带宽就只有 100Mbit/s 的 1/16。由集线器作为中心设备的局域网和总线型拓扑的局域网称为共享式局域网。

集线器的全部端口属于同一个冲突域，集线器在端口之间转发数据帧时采用向所有端口广播的方式进行，因此其全部端口又属于同一个广播域。单一的冲突域和广播域使网络在通信繁忙时容易产生阻塞和广播风暴。可以使用多台集线器连接或堆叠起来以增加总的端口数，但不能用此方法来延伸网络的距离。

随着交换机价位的降低，共享式集线器正逐渐淡出局域网领域。

2．交换机与交换式局域网

交换机可以看成功能增强型的集线器，有时也称为交换式集线器。它采用了许多新技术，如其端口之间的通信可全双工进行，能实现数据的线速转发等。其最显著的特点之一是端口带宽的独享。

例如，一台 100Mbit/s 交换机在使用时，每一个端口的数据传输速率都是 100Mbit/s，数据传输速率不会随着使用端口数的增加而减少，即所说的端口带宽独享。

应当注意的是，在网卡和交换机端口或交换机和交换机端口相连时，只有当相连接的两者的带宽为同一值时，才能实现以该速率传输数据。例如，只有网卡和交换机都是 1000Mbit/s 时，才能实现 1000Mbit/s 的传输速率；否则只能按两者中较低的速率传输，这一特性称为带宽的自动协商或者带宽的自适应。

光纤能支持 1000Mbit/s 以上的传输速率，但使用光纤的网络未必都是千兆位以太网，最初的光纤以太网就是 10Mbit/s 的以太网。

通常把由交换机作为中心设备的局域网称为交换式局域网。

交换机的端口按其带宽可分为 10Mbit/s、100Mbit/s、10/100Mbit/s 自适应和 1000Mbit/s（1Gbit/s）、10Gbit/s。现在市面上的交换机端口带宽一般都有自适应功能，有的交换机上只有以上端口之一，更多的则是兼有两种或多种端口，而具有 10Gbit/s 端口的交换机则在大型公司网络或 ISP 网络使用。

交换机的每一个端口都独立转发数据（各属不同的冲突域），只要背板带宽足够宽，就不会因端口的使用数增加而降低端口的传输速率。但交换机的所有端口仍属于同一个广播域，当网络中的广播信息增多时，也会导致网络传输速率的降低。

如果采用虚拟局域网（VLAN）技术，则每一个 VLAN 具有各自的广播域，这样交换机就有了多个广播域。广播数据帧被局限在各自的域内，可有效地防止广播风暴的发生。

与集线器一样，也可使用多台交换机连接或堆叠起来增加总的端口数。然而，交换机的连接却可以用来延伸网络的距离，如图 2-2 所示的连接可使网络距离扩展至 400m。

最廉价的交换机可能不支持网络管理功能，只用于简单的网络环境。支持网络管理功能的交换机称为可管理或可配置的交换机。

若用在小型、简单的网络中，可管理的交换机也不需配置（实际是使用了默认配置）即可工作；而当网络规模较大或者较为复杂时，就需要对其进行配置和管理。

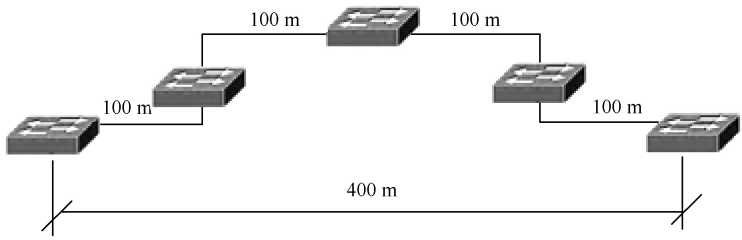

图 2-2　连接交换机以扩展网络距离

2.4　ARP 协议与 MAC 地址表

2.4.1　ARP 协议

连接在交换机端口上的主机通过 ARP 地址解析协议查询目标网卡 IP 地址所对应的物理地址（又称数据链路层地址或 MAC 地址，即 Media Access Control 地址），以便进行相互间数据帧的传输。

MAC 地址是固化在网卡内部用于唯一确定网卡身份的标识，是网卡在生产时被永久写入芯片的固定值。全球的网卡生产厂商按照其 MAC 地址范围制造网卡，因此不会有相同 MAC 地址的网卡。这样，MAC 地址就可作为唯一标识设备的地址。

网络层中的数据包在主机查得目标的链路层地址后被封装成数据帧，在链路层交给交换机，交换机则通过帧头的 MAC 目标地址找到目标主机。

由于交换机在数据传输过程中不用检查第三层（网络层）的包头信息，而是直接由第二层帧结构中的 MAC 地址来决定数据的转发目标，因此，数据的交换过程几乎没有软件的参与，从而大大提高了交换进程的速率。

2.4.2　MAC 地址表

在交换式网络中，各主机的 MAC 地址是存储在交换机的 MAC 地址表（也称 MAC 地址数据库）中的。因 MAC 地址表记录的是各主机 MAC 地址与对应的交换机端口号，所以有时也称为 MAC 地址-端口表。简单地说，MAC 地址表记录了哪台或哪些主机连接在哪台交换机的哪个端口上。交换机在工作过程中，会向 MAC 地址表不断写入新学到的 MAC 地址，一旦交换机掉电，或连接交换机的各主机在一定时间（地址表的老化时间）内没有相互访问，则其 MAC 地址表会被自动清除。

1．MAC 地址表的建立

在如图 2-3 所示的网络中，若 PC1 首次访问 PC3，则交换机的 MAC 地址表建立过程如下。

（1）发送 ARP 请求帧。PC1 向 PC3 发送查询其 MAC 地址的消息帧，该消息帧包含 PC1（源）MAC 地址、IP 地址和 PC3 的 IP 地址等信息。该帧经交换机 Switch 的 F0/0

端口先发送到交换机上。该帧是要查询 IP 地址为某个值的目标主机的 MAC 地址，是由地址解析协议 ARP 构造而成的，故也称为 ARP 请求帧。

（2）交换机生成第一条地址表记录。交换机 Switch 在端口 F0/1 收到该查询消息帧后，从该帧的 MAC 地址信息中得知 PC1 的 MAC 地址，就将该端口和该 MAC 地址记录在自己的 MAC 地址表中，这样就有了第一条记录。

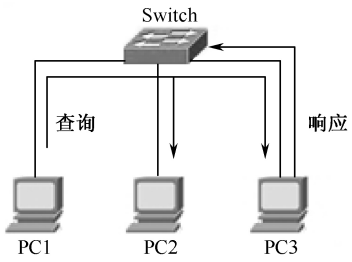

图 2-3　MAC 地址表的建立

（3）交换机广播 ARP 请求帧。由于交换机 Switch 此时还没有 PC3 的 MAC 地址表记录，不确定 PC3 连接在哪个端口上，于是就向除 F0/0 端口外的其他所有端口转发（广播）该 ARP 请求帧。

（4）ARP 应答帧。PC3 接收到该 ARP 请求帧后，发现自己是被查询对象（询问的是自己的 IP 地址所对应的 MAC 地址），就会应答该帧。从该帧中，PC3 得知了 PC1 的 MAC 地址和 IP 地址并写入自己的 ARP 表中；同时，构造并发送以 PC1 的 MAC 地址为目标，以自己的 MAC 地址为源，还包含 PC1 和 PC3 的 IP 地址等信息的 ARP 应答帧。该帧经交换机 Switch 的 F0/7 端口先发送到交换机上。

（5）交换机生成第二条 MAC 地址记录。交换机 Switch 从 F0/7 端口接收到 ARP 应答帧后，从该帧的源 MAC 地址信息中得知 PC3 的 MAC 地址，就将该端口和该 MAC 地址记录在自己的 MAC 地址表中，这样就有了第二条记录。

（6）交换机转发 ARP 应答帧。由于交换机此时已经有了 PC1 的 MAC 地址表记录，故此次转发不再需要广播，而是查看 MAC 地址表，得知该帧的目标 MAC 地址主机位于 F0/1 端口，就会把该应答帧从该端口转发出去，送达 PC1。PC1 收到该 ARP 应答帧，从该帧得知 PC3 的 MAC 地址和 IP 地址并写入自己的 ARP 表中。

经过以上过程，交换机中生成了 MAC 地址表，而 PC1 和 PC3 中生成了记录目标 IP 地址和 MAC 地址的 ARP 表。

在交换机 Switch 广播查询消息帧时，PC2 也会收到。PC2 从该消息帧查询的目标 IP 地址得知自己不是被查询对象（与自己的 IP 地址不符合），就会丢弃该帧，不予以响应。

2．计算机之间的通信

此后，PC1 和 PC3 根据 ARP 表由 IP 地址查找 MAC 地址进行链路层的通信，交换机按照接收到数据帧的 MAC 地址查找地址表中对应的端口，把数据转发出去。

3．MAC 地址表更新

如果主机在一定时间（称为老化时间）内未进行通信，交换机将会清除相应端口对应的 MAC 地址记录，再次通信时需重新通过步骤（1）～（6）生成 MAC 地址记录，这一过程称为 MAC 地址表的更新。

如果是主机之间第一次通信，或者超过 MAC 地址表更新时间后继续通信，交换机

都会广播 ARP 表进行查询。所以，以太网中有的广播是不可避免的，也是必需的。

任务 2–1　MAC 地址表的查看与人工指定地址表项

【主要设备】

Cisco 2960 交换机 2 台，计算机 3 台。

【网络拓扑】

用网线连接计算机和交换机，如图 2-4 所示，此任务使用 Switch1 的 F0/1 端口连接 Switch0 的 F0/4 端口。

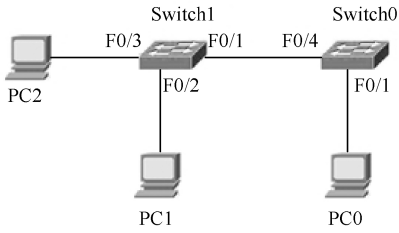

图 2-4　连接计算机和交换机

【操作步骤】

1. 查看 MAC 地址表

（1）在计算机间无访问时查看 MAC 地址表。给交换机上电，用超级终端或 Telnet 方式登录交换机，使用 show mac-address-table 命令查看，其中 Switch1 的 MAC 地址表显示如下：

```
Switch#show  mac-address-table
      Mac Address Table
-------------------------------------------
Vlan  Mac Address       Type        Ports
----  ----------        --------    -----
1     00e0.b057.d004    DYNAMIC     Fa0/1
Switch#
```

该表中有一项记录，表示的意义是交换机 Switch1 的 F0/1 端口上所连接的设备（Switch0 的 F0/4 端口）的 MAC 地址是 00e0.b057.d004。F0/1 端口属于 VLAN1（默认）。该记录的 Type 为 DYNAMIC，表示该表项是交换机动态学习到的，关于动态的含义会在以后的内容中介绍。

读者可以自己设想，此时 Switch0 的 MAC 地址表应该是怎样的。

相互连接的交换机之所以启动完成后会有 MAC 地址表记录，是因为交换机要彼此发送桥协议数据单元（BPDU，Bridge Protocol Data Unit）进行联系，所以会查询出对方端口的 MAC 地址。

（2）计算机间访问后 MAC 地址表的查看。让 3 台计算机相互访问，如在 PC0 上分别 ping PC1 和 PC2，然后查看 Switch1 和 Switch0 的 MAC 地址表。

Switch1 的 MAC 地址表如下：

```
Switch#show  mac-address-table
      Mac Address Table
-------------------------------------------
Vlan  Mac Address       Type      Ports
----  ----------        --------  -----
```

```
    1  0002.4ad5.5bb5   DYNAMIC   Fa0/1
    1  0005.5ea9.0c91   DYNAMIC   Fa0/3
    1  0010.1145.1004   DYNAMIC   Fa0/2
    1  00e0.b057.d004   DYNAMIC   Fa0/1
Switch#
```

Switch0 的 MAC 地址表如下：

```
Switch#show  mac-address-table
      Mac Address Table
------------------------------------------
Vlan  Mac Address       Type      Ports
----  ----------        --------  -----
   1  0002.4ad5.5bb5    DYNAMIC   Fa0/1
   1  0005.5ea9.0c91    DYNAMIC   Fa0/4
   1  0010.1145.1004    DYNAMIC   Fa0/4
   1  00e0.a3b9.d601    DYNAMIC   Fa0/4
Switch#
```

读者可以先判断表中的 MAC 地址是属于谁的，然后使用命令 ipconfig/all 查看计算机的 MAC 地址，验证自己的判断。

（3）MAC 地址表的老化时间。MAC 地址表建立后，如果计算机之间在一定的时间内不通信，则相应的表项就会被自动删除，当新的访问开始时再建立，这称为 MAC 地址表的自动更新，也是表项中动态（Dynamic）的含义之一。

表项从建立到自动删除的时间间隔称为老化时间。默认动态的 MAC 地址表项老化时间是 300s，其值可以修改。查看和修改老化时间的命令如下：

```
Switch#show mac-address-table aging-time   //查看老化时间
Switch(config)# mac-address-table aging-time 150
                                           //修改老化时间为 150s
```

2．人工配置 MAC 地址表项

有时基于安全或排他性的考虑，在交换机的某些端口上，我们可能不希望交换机自动学习 MAC 地址，而是由网络管理员人工配置，这称为端口-MAC 地址绑定。这样的地址表项不会自动刷新，故与动态（Dynamic）表项对应，也称为静态（Static）MAC 地址表项。例如，为交换机 Switch1 的 f0/3 端口绑定 MAC 地址 0005.5ea9.0c99 的命令如下：

```
Switch(config-if)#switchport port-security mac-address 0005.5ea9.0c99
```

注意，交换机只允许在 Access 或 Trunk 上绑定且默认只允许绑定一个 MAC 地址。要进行绑定操作，需要设置接口模式为 Access 或 Trunk（Cisco3560 交换机默认为动态协商）且必须激活端口安全。

在大型网络中，端口数量成千上万，如果让管理员手动将每一台计算机的 MAC 地址绑定到端口，将是一项十分烦琐、低效的工作；而且对于网络的需求复杂，这样做也缺乏足够的灵活性。可以使用端口的 sticky（粘连）特性来解决这个问题，sticky 让交

换机在某端口动态学习 MAC 地址并固定下来，命令如下：

```
Switch(config-if)#switchport port-security mac-address sticky
```

保存配置并重启交换机，则不需要再对已经固定下来的地址交换机端口重新学习，这些地址称为 sticky MAC 地址。它类似于管理员手动配置的静态地址，却不用管理员逐条去配置，大大提高了工作效率。sticky 同样只能在 Access 或 Trunk 接口配置且要激活端口安全。

特别要注意的是，在交换机激活端口安全后，动态学习到的 MAC 地址类型也显示为静态，但是与人工配置静态 MAC 或 sticky MAC 地址不同的是，该地址会在老化时间到来时或重启交换机后被清除。

2.4.3 局域网的帧交换方式

以太网交换机在传送数据时，数据被封装成帧，采用帧交换（Frame Switching）技术。该技术包括 3 种主要的交换方式，即存储转发（Store and Forward）、伺机通过（Cut Through）和自由分段（Fragment Free）。

1. 帧交换方式

（1）存储转发方式是最基本的交换技术之一。在进行转发数据帧前，该数据帧将被完全接收并存储在缓冲器中，数据帧从头到尾全部接收完毕才进行转发。其间，交换机需要解读数据帧的目的地址与源地址，以根据 MAC 地址表进行正确的转发。

在存储转发过程中还要进行高级别的冗余错误检测（CRC）工作，如果所接收到的数据帧存在错误、太短（小于 64B）或太长（大于 1518B），最终都会被抛弃。

采用这种转发方式的交换机在接收数据帧时延迟较大，且数据帧越大，延迟时间越长，但是对错误的检测能力强。

（2）伺机通过（也称 Fast Forward 或 Real Time 模式）技术是交换机在接收整个数据帧之前读取数据帧的目的地址到缓冲器，随后在 MAC 地址列表里查询目的地址所对应的端口，转发该帧。简而言之，它读取到帧的目标地址以后就立即进行转发。

采用这种转发方式，在完全接收整个数据帧之前就会转发。这种方法减少了传输的延迟，但由于不对帧进行错误检测，传送到目标主机帧的误码率（码元错误发生率）可能较高。

还有一些交换机可以把存储转发与伺机通过这两种技术合并在一起使用。它们首先在交换机里设置一个错误检测的门限值，当误码率低于该值时使用伺机通过的交换方法以减少数据的传输延迟；当误码率高于该门限值时，交换机将自动改为存储转发交换方式，从而保证数据的正确性，在链路恢复正常后，误码率下降到低于该门限值后，系统将再次回到伺机通过方式工作。

（3）自由分段（也称 Modified cut-through 模式）技术是在伺机通过交换方式的基础上调整而成的。自由分段在转发数据帧之前，检测可能有错误发生的冲突分段（长度为 64 个字节）。这是因为通常数据帧的错误发生在刚开始的 64 个字节内的概率最大。简而言之，它读取到帧数据字段前的 64 个字节，然后进行转发。自由分段交换方式的

错误检测级别要高于伺机通过交换方式。

2．Cisco 交换机交换方式的设置

Catalyst 1900 系列交换机用 switching-mode 命令设置工作是存储转发方式还是自由分段方式，如下所示：

```
C1912(config)#switching-mode  ?
fragment-free          Fragment Free mode
store-and-forward      Store-and-Forward mode
```

2.5 VLAN 技术

2.5.1 第二层交换式网络的缺点与 VLAN 技术

1．第二层交换式网络的缺点

由于整个网络属于同一个广播域，因此任何一个广播帧或多播帧（Multicast Frame）都将被交换机广播到整个局域网中的每一台主机上。在网络通信中，广播帧是普遍存在的，这些广播帧将占用大量的网络带宽，导致网络速度和通信效率的下降，并额外增加了主机为处理广播信息所产生的负荷。而蠕虫病毒和其他一些类似的网络攻击相当泛滥，如果不进行有效的广播域隔离，一旦病毒发起泛洪广播攻击，将会很快占用完网络的带宽，导致网络阻塞和瘫痪。概括地说，第二层交换式网络存在如下缺点。

- 全网属于一个广播域，每一次广播的数据帧无论是否需要，都会到达网络中的所有设备，这就必然会造成带宽资源的极大浪费。
- 全网属于一个广播域，极易引起广播碰撞和广播风暴等问题。
- 网络的安全性不够高，在这种网络结构中，所有用户都可以监听到服务器及其他设备端口发出的广播数据帧，因此是极不安全的。

2．VLAN 技术

为了解决存在的问题，早期采用使用路由器从第三层来分隔广播域的技术，即通过把网络中的主机设置不同的子网，广播包传送范围被限制在各自的子网里，不同子网间的访问则通过路由器的端口进行转发。在第三层实现广播域分隔技术的成本高，路由器每个端口的价格远高于相同速率的交换机，因此很快被基于交换机的 VLAN（虚拟局域网，Virtual Local Area Network）技术取代。

VLAN 是将局域网从逻辑上按需要划分为若干个网段，在第二层实现分隔广播域，分隔开用户组的一种交换技术。这些网段物理上是连接在一起的，但逻辑上已经分离，即原来一个局域网被划分成多个局域网，故名为虚拟局域网。

VLAN 允许一组不限物理位置的用户群共享一个独立的广播域，可在一个物理网络中划分多个 VLAN，即可使不同的用户群属于不同的广播域。这样，通过划分用户群、控制广播范围等方式，VLAN 技术能够从根本上解决网络效率与安全性等问题。

VLAN 对广播域的划分是通过交换机软件完成的。它通过对用户分类来规划用户群，如按项目组、部门或管理权限等来进行 VLAN 划分。划分 VLAN 时能够超越地域的界限，做到真正意义上的逻辑分组。在划分 VLAN 的交换机上，每个端口都能被赋予一个 VLAN 号，相同 VLAN 号的用户同属于一个独立的广播域。广播被限制在各自的 VLAN 之内，因此，VLAN 能够控制广播的影响范围，减少由于共享介质所造成的安全隐患。

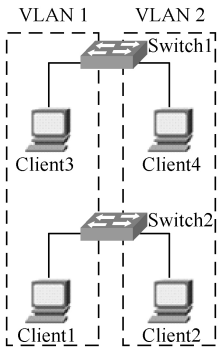

图 2-5　划分 VLAN 后的网络

划分 VLAN 后的网络如图 2-5 所示，Client1 和 Client3 属于 VLAN1，Client2 和 Client4 属于 VLAN2。在同一个 VLAN 内，计算机之间可以正常访问；而不同 VLAN 间的单播帧和广播帧都不能直接到达对方的区域。不同 VLAN 的计算机之间的访问需通过三层设备，如路由器或三层交换机来实现。从图中可看到，虽然计算机所处的物理位置不同，但却可以划归在同一个 VLAN 中。

2.5.2　划分 VLAN 的好处

（1）广播控制（Broadcast Control）。通过将一个网络划分成多个 VLAN，可以实现广播范围的控制，能够有效减少广播风暴、广播冲突和网络带宽资源的浪费等问题。

（2）灵活性（Flexibility）。VLAN 技术能够在逻辑上将不同地理位置的计算机划分在同一个广播域内，而无 VLAN 技术时，在更改一台主机的所属组时，必须将此主机直接接到该组所在的交换机上。这样，VLAN 可以非常灵活地添加或删除域内的主机而不受主机物理位置的限制，这为网络管理带来了极大的方便，如对网络流量的均衡性（Scalability）控制就会很容易实现。

（3）安全性（Security）。不同 VLAN 之间是不能够直接相互访问的，因此，按职责权限把用户（主机）划归在不同的 VLAN 里，就可使各自的内部信息得到保护，从而增强了安全性。

任务 2–2　在单台交换机上划分 VLAN

【主要设备】

Cisco 2960 交换机 1 台，计算机 3 台。

【网络拓扑】

网络拓扑如图 2-6 所示，PC1、PC2 和 PC3 分别连接在交换机的端口 F0/1、F0/2 和 F0/3 上,把 F0/1 划分为 VLAN 10，F0/2 和 F0/3 划分为 VLAN 11。

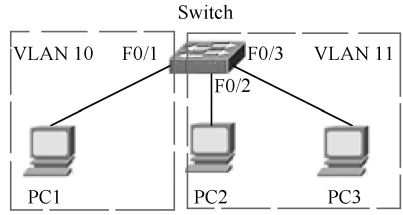

图 2-6　网络拓扑

【任务要求】

一个简单的 VLAN 配置，只需设置 VLAN 号并

把端口加入相应的 VLAN 即可。验证方法：可查看 VLAN 信息，还可在 PC1、PC2 和 PC3 上互 ping，属于同一个 VLAN 的 PC 能够 ping 通，否则不能 ping 通。

【操作步骤】

1. 增加相应的 VLAN 并验证

在交换机 Switch 上划分 VLAN 10 和 VLAN 11，假如 VLAN 10 属于技术部，VLAN 11 属于财务部，可以给相应的 VLAN 命名，以便于记忆和管理。

```
Switch1(config)#vlan 10   //在全局配置模式下增加一个 VLAN，编号为 10，
                          //并进入 VLAN 配置模式
Switch1 (config-vlan)#name TECH //在 VLAN 配置模式下，将 VLAN 命名为 TECH
Switch1 (config-vlan)#exit      //退出 VLAN 配置模式
Switch1 (config)#vlan 11        //在全局配置模式下增加一个 VLAN，编号为 11，
                                //并进入 VLAN 配置模式
Switch1 (config-vlan)#name FINANCE
                                //在 VLAN 配置模式下，将 VLAN 命名为 FINANCE
Switch1 (config-vlan)#exit      //退出 VLAN 配置模式
Switch1 (config)#end
Switch1#
Switch1#show vlan brief         //显示交换机当前 VLAN 配置的简要信息
VLAN Name                Status    Ports
---- ----------------    -------   -------------------------
1    default             active    Fa0/1, Fa0/2, Fa0/3, Fa0/4
                                   Fa0/5, Fa0/6, Fa0/7, Fa0/8
                                   Fa0/9, Fa0/10, Fa0/11, Fa0/12
                                   Fa0/13, Fa0/14, Fa0/15, Fa0/16
                                   Fa0/17, Fa0/18, Fa0/19, Fa0/20
                                   Fa0/21, Fa0/22, Fa0/23, Fa0/24
10   TECH                active
11   FINANCE             active
1002 fddi-default        active
1003 token-ring-default  active
1004 fddinet-default     active
1005 trnet-default       active
```

可以看到交换机 Switch1 上已经增加了两个 VLAN，分别是 VLAN 10 名称为 TECH，VLAN11 名称为 FINANCE。最后一列 Ports 显示交换机上的端口分别属于哪个 VLAN，默认情况下，以太网交换机的所有端口都属于 VLAN1。

2. 将交换机的相应端口加入 VLAN 并查看配置结果

第 1 台计算机连接到交换机的 F0/1，属于技术部 TECH；第 2、3 台计算机分别连接到交换机的 F0/2 和 F0/3，属于财务部 FINANCE，配置命令如下：

```
Switch1(config)#interface F0/1          //进入端口 F0/1
```

```
Switch1(config-if)#switchport mode access
                //配置此端口的模式为Access模式(默认为Dynamic动态协商模式)
Switch1(config-if)#switchport access vlan 10 //将此端口加入VLAN 10
Switch1(config-if)#interface range F0/2 - 3  //进入端口F0/2和F0/3
Switch1(config-if)#switchport mode access
                //配置这两个端口的模式为Access模式
Switch1(config-if)#switchport access vlan 11//将这两个端口加入VLAN 11
Switch1(config-if)#end
Switch1#
Switch1#show vlan brief     //显示交换机当前VLAN配置的简要信息

VLAN Name                 Status      Ports
---- -------------------- -------     --------------------------
1    default              active      Fa0/4, Fa0/5, Fa0/6
                                      Fa0/7, Fa0/8, Fa0/9, Fa0/10
                                      Fa0/11, Fa0/12, Fa0/13, Fa0/14
                                      Fa0/15, Fa0/16, Fa0/17, Fa0/18
                                      Fa0/19, Fa0/20, Fa0/21, Fa0/22
                                      Fa0/23, Fa0/24
10   TECH                 active      Fa0/1
11   FINANCE              active      Fa0/2, Fa0/3
1002 fddi-default         active
1003 token-ring-default   active
1004 fddinet-default      active
1005 trnet-default        active
```


完成以上配置后，不同部门的计算机就会被隔离，可以尝试将这 3 台计算机配置为同一个 IP 子网，然后用 ping 命令测试它们是否能互相通信，结果一定是 PC2 和 PC3 之间可以 ping 通，而 PC1 与 PC2、PC3 之间不能 ping 通。

3．另外一种配置交换机 VLAN 的方法

可以通过 VLAN 数据库配置模式完成，配置实例如下。

（1）进入 VLAN 数据库配置模式。

```
Switch1#vlan database
% Warning: It is recommended to configure VLAN from config mode,
  as VLAN database mode is being deprecated. Please consult user
  documentation for configuring VTP/VLAN in config mode.
```

以上提示信息推荐在全局配置模式而不是在该模式下配置 VLAN。

（2）新建 VLAN 并命名。

```
Switch1(vlan)#vlan 10 nameTECH
                          //新建一个VLAN10名称为TECH，下面两行为屏幕提示
VLAN 10 added:
```

```
    Name:TECH
Switch1(vlan)#vlan 11 name FINANCE
                    //新建一个 VLAN11 名称为 FINANCE，下面两行为屏幕提示
VLAN 11 added:
    Name: FINANCE
Switch1(vlan)#exit  //退出并自动应用配置
APPLY completed.
Exiting....
```

这种在 VLAN 数据库配置模式下配置的方式现在用得比较少，大部分 VLAN 的配置都是在全局配置模式下完成的。

在全局配置模式下想要删除已创建的 VLAN，可参照以下配置实例：

```
Switch1(config-if)#no switchport access vlan 10 //将端口从 VLAN10 退出
Switch1(config)#no vlan 10          //删除 ID 号为 10 的 VLAN
```

2.6 链路冗余与生成树协议

2.6.1 冗余备份与环路

在许多交换机组成的大/中型网络环境中，通常都使用一些备份连接，以提高网络的稳定性。备份连接也称备份链路、冗余链路等，如图 2-7 所示，交换机 Switch1 的端口 Port7 与交换机 Switch3 的端口 Port6 之间的链路就是一个备份连接。在主链路（图中 Port1 与 Port3 之间的链路）发生故障时，备份链路将自动启用，从而提高网络的整体可靠性。

但是，备份连接会使网络存在环路，图 2-7 中的交换机和连接链路就构成了一个环路。环路问题是备份连接面临的所有负面影响中最为严重的问题，它在网络中直接导致广播风暴、出现多个帧副本、MAC 地址表混乱。

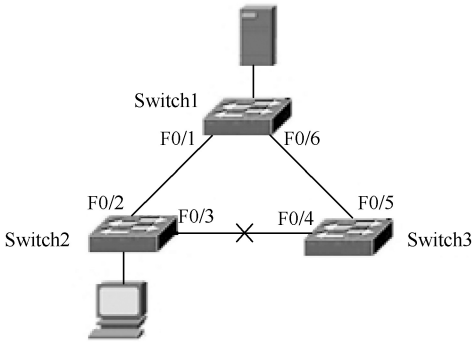

图 2-7　备份连接

1. 广播风暴

在一些较大型的网络中，当大量广播流（如 MAC 地址查询信息等）同时在网络中

传播时，便会发生数据包的冲突，随后，网络试图缓解这些冲突并重传更多的数据包，广播流量会充斥网络，这一现象称为广播风暴。在网络存在环路的情况下，如无特别的措施，广播风暴必然发生。其结果导致全网的可用带宽阻塞，并最终使得网络失去连接而瘫痪。

网络中，一台设备能够将数据包转发给网络中所有其他站点的技术称为广播。由于广播能够穿越由于交换机连接的多个局域网段，因此，几乎所有局域网的网络协议都优先使用广播方式来进行管理与操作。广播使用广播帧来发送、传递信息，广播帧没有明确的目的地址，它所发送的对象是网络中的所有主机，也就是说网络中的所有主机都将接收到该数据帧。它一般用来发送网络中的公共信息，如服务通告、地址查询等信息。

广播是引起广播风暴的主要原因。但是，在正常的网络环境中，网络广播是无所不在的，MAC 地址查询、路由协议通信、ICMP 控制报文及大量的服务通告等信息都属于网络中正常的广播。因此，需要在保证网络正常使用广播的情况下，有效地减少广播风暴的发生。

广播风暴的形成：

在如图 2-7 所示的网络中，本来是要提供冗余备份，增加一条 Switch3 到 Switch2 的通路，但若不采取其他措施，这样做的结果会导致不能正常工作，因为这是一个存在循环的连接，如果 Switch1 收到一个广播帧，下面的过程（a）～（f）会被反复执行。

（a）Switch1 通过 F0/1 转发广播帧。

（b）Switch2 通过 F0/2 收到广播帧。

（c）Switch2 通过 F0/3 转发广播帧。

（d）Switch3 通过 F0/4 收到广播帧。

（e）Switch3 通过 F0/5 转发广播帧。

（f）Switch1 通过 F0/6 再次收到原来的广播帧。

上述过程周而复始，同样的广播帧被不断复制，最后形成广播风暴，耗尽网络资源。

在一个较大规模的网络中，由于拓扑结构的复杂性，会造成许多大大小小的环路产生，由于以太网的第二层协议没有控制环路数据帧的机制，各环路产生的广播风暴将不断扩散到全网，进而造成网络瘫痪。

与广播概念相类似的还有组播（Multicast，或称多播），组播是一点对多点的通信，是一种比较有效的节约网络带宽的方法。例如，在视频点播等多媒体应用中，当把多媒体信号从一个节点传输到多个节点时，采用广播方式会浪费带宽，重复采用点对点传播也会浪费带宽，而组播能够把帧发送到组地址，而不是单个主机，也不是整个网络。由于它的发送范围明显小于广播，因而减少了对网络带宽的占用。

网络运行时，应当了解网络里运行的有协议及这些协议的主要特点，这样才能更有利于对广播流量的控制。通常，交换机对网络中的广播帧或组播帧不会进行任何数据过滤，因为这些帧的地址信息不会出现在 MAC 层的源地址字段中。交换机总是直接将这些信息广播到所有端口，如果网络中存在环路，这些广播信息将在网络中不停地转发，直至导致交换机出现超负荷运转（如 CPU 过度使用、内存耗尽等），最终耗尽所有资源，阻塞全网通信。

Cisco 第二层交换机支持广播风暴控制功能，它定义交换机端口的广播门限值，当端口接收的广播帧数量超过了该值时，该端口便会立刻处于挂起状态，不再接收广播数据帧，从而避免出现循环广播状态。该功能默认为禁用，需要通过手动配置打开。而在第二层实现控制广播风暴的有效方法则是使用生成树技术和 VLAN 技术，前者能够从逻辑上消除环路，后者则可限制广播的范围。

2．多个帧副本

网络中如果存在环路，目标主机可能会收到某个帧的多个副本，而在正常情况下，除收到的帧有错误而要求重传外，则不会有同一帧的多个副本到达目标主机。多个帧会导致上层协议在处理这些数据帧时无从选择。

3．MAC 地址表混乱

当交换机有环路连接时，将会出现通过不同端口接收到同一个广播帧的多个副本的情况。这样，在 MAC 地址表里，同一个 MAC 地址将出现在同一个交换机的不同的端口上，使得 MAC 地址表混乱，导致不能正常转发数据帧。同时，这一过程也会同时导致 MAC 地址表的多次刷新。这种持续的更新、刷新过程会耗用资源，影响该交换机的交换能力，降低整个网络的运行效率。严重时，将耗尽整个网络资源，并最终造成网络瘫痪。

2.6.2 STP 简介

要实现冗余备份，提高网络的可靠性，必须解决环路拓扑结构为网络带来的致命的负面影响。

1．生成树协议的功能

生成树协议（Spanning Tree Protocol，STP）的主要功能就是为了解决由于备份连接所产生的环路问题。本节将介绍基本的 STP 机制，实际工程应用的配置将在第 3 章中介绍。

STP 的主要思想就是当网络中存在备份链路时，只允许主链路激活，在主链路因故障而被断开后，备用链路才会激活。

STP 的基本做法就是生成“一棵树”，树的根是一个称为根桥的交换机。以根为参考，所有运行 STP 的交换机都执行生成树算法（Span Tree Algorithm，STA），使得交换机的所有链路在逻辑上形成树状结构，这棵树就是生成树。树上的链路处于工作状态，其他的链路都将被暂时阻塞。

根据设置不同，不同的交换机会被选为根桥，但任意时刻只能有一个根桥。由根桥开始，逐级形成一棵树，根桥定时发送配置数据包，非根桥接收配置数据包并转发，如果某台交换机能够从两个以上的端口接收到配置数据包，则说明从该交换机到根桥的路径不止一条，这样便构成了循环回路。此时，该交换机就会选出一个端口并把其他的端口阻塞，消除循环。而当某个端口超过一定时间不能接收到配置数据包时，交换机认为

该端口的配置超时，网络拓扑可能已经改变。此时，就重新计算网络拓扑，重新生成“一棵树”。

2．STP 相关的概念

为了理解 STP，必须熟悉以下几个概念，即网桥 ID（交换机早期的名称下叫做网桥）、路径开销、Port ID、BPDU。

（1）网桥 ID（包括网桥优先级和 MAC 地址）：生成树算法的第一个参数。STP 用网桥 ID 来标识网络中的交换机，其值最小者称为根网桥。网桥 ID 的数据结构共有 8 个字节，高位的 2 个字节是交换机的 STP 优先级，取值范围为 0～65535，默认值为 32768；其余 6 个字节的部分为交换机的 MAC 地址。

（2）路径开销：生成树算法的第二个参数。路径开销是从非根交换机到根交换机的方向按链路叠加的。通路径开销是确定非根交换机到达根交换机最短路径选择的首要参数。早期的路径开销等于参考带宽 1000Mbit/s 除以当前链路的带宽。如 100Mbit/s 的链路路径开销是 10，10Mbit/s 的链路路径开销是 100。这种计算方法后来遇到了问题，因为如果一个网络链路是 10Gbit/s，则这个链路的路径开销将不再是整数。所以对这种算法做了一定的修订，按照 802.1D 计算，100Mbit/s 的链路路径开销是 19，1000Mbit/s 的链路路径开销是 4，即带宽越大，路径开销越小。

（3）Port ID（包括端口优先级与端口号）：生成树算法的第三个参数，也是决定到达根交换机路径选择的参数。端口优先级与端口号长度都是 1 个字节，端口优先级取值为 0～255，默认值为 128；端口号取值为 1～255（交换机接口不编 0 号）。

（4）BPDU：BPDU（Bridge Protocol Data Unit，桥协议数据单元），运行 STP 的交换机之间通过交换 BPDU 消息，完成无环路的树状结构生成。BPDU 可以帮助运行 STP 的交换机选出整个生成树的根，探测到冗余链路并阻塞端口。根选出来之前，所有参与选择的交换机都发送配置 BPDU，配置 BPDU 包含网桥 ID、路径开销和端口 ID 等信息，都宣称自己是根；根选出来之后，只有根交换机才能发送配置 BPDU，默认是每隔 2s 就会发送最新的配置 BPDU。

3．STP 端口状态

运行 STP 的交换机，每个端口都会处于某种 STP 端口状态下。STP 端口状态有以下几种类型，即禁用状态、阻塞状态、侦听状态、学习状态和转发状态，如表 2.3 所示。

表 2.3　STP 端口状态

状　态	属　性
禁用（Disabled）	不能接收 BPDU，不能学习 MAC 地址，不能转发数据帧
阻塞（Blocking）	可接收 BPDU，不能转发数据帧，不能学习 MAC 地址
监听（Listening）	可监听和接收 BPDU，不能学习 MAC 地址，不能转发数据帧
学习（Learning）	可接收和发送 BPDU，可学习 MAC 地址，不能转发数据帧
转发（Forwarding）	可接收和发送 BPDU，可学习 MAC 地址，可转发数据帧