

任务 5

避免交换网络环路

【任务描述】

某校园网的主干部分拓扑结构如图 5-1 所示。图中每幢楼的汇聚交换机与校园网的核心交换机之间都只有一条链路连接,如果某条单一链路出现故障,如行政办公楼和网络中心之间的链路因校园内基础工程改造施工而被挖断,该楼内的网络则瞬间无法与外界联络,影响校园的日常办公。现要求增加一些保障措施,避免此类事件发生。

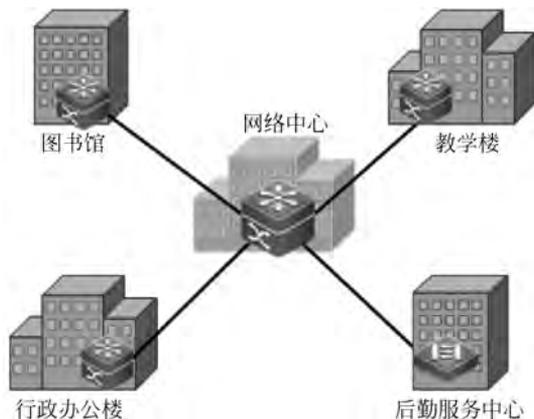


图 5-1 某校园网的主干部分拓扑结构

【任务分析】

在实际的网络环境中,为了提供可靠的网络连接,减少故障影响的一个重要方法就是“冗余”,需要网络连接提供冗余链路。所谓“冗余链路”,其实和日常走路一样,这条路不通,就走另一条路。冗余就是准备两条或两条以上的通路,当主链路不通时,马上启用备份链路,确保链路连接不中断。网络中的冗余可以起到当网络中出现单点故障时,还有其他备份的组件可以使用,使整个网络通信基本不受影响。

在交换网络中,冗余会使网络的物理拓扑形成环路,物理的环路结构很容易引起广播风暴、多帧复制和 MAC 地址表抖动等问题,这些问题会导致网络不可用。

在二层网络中配置生成树协议,可以解决环路问题,还可以在提供冗余备份的同时实现

负载均衡,从而提高网络的可靠性。

5.1 知识储备

5.1.1 “冗余链路”的危害

交换机之间具有冗余链路本来是一件很好的事情,但是它有可能引起的问题比它能够解决的问题还要多。假设两台交换机之间有两条互连链路,冗余链路使得交换机之间形成了一个环路,如图 5-2 所示,交换机并不知道如何处理环路,只是周而复始地转发帧,形成一个“死循环”。这种交换环路会导致网络中出现广播风暴、MAC 地址表抖动和多帧复制等问题,对网络性能产生极为严重的不良影响,甚至导致网络瘫痪。

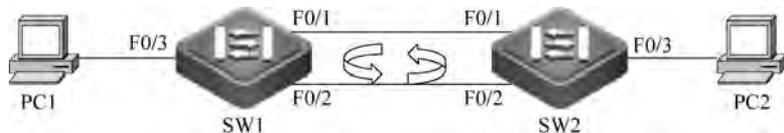


图 5-2 冗余拓扑形成交换环路

1. 广播风暴

在图 5-2 中,假设 PC1 和 PC2 要通信,由于 PC1 不知道 PC2 的 MAC 地址,因此 PC1 首先发送一个 ARP 广播帧,请求 PC2 的 MAC 地址。交换机 SW1 收到此 ARP 广播请求,根据交换机的工作原理,交换机 SW1 将向除接收端口之外的所有端口(F0/1 和 F0/2)“泛洪”这个广播帧,交换机 SW2 从自己的 F0/1 和 F0/2 端口分别收到该 ARP 广播帧,同样也会泛洪出去,该 ARP 广播帧又被转发回到 SW1,SW1 收到后仍然会继续泛洪出去,如此往返,导致该 ARP 广播帧在两台交换机之间沿着顺时针和逆时针两个方向循环转发,永无休止,环路里逆时针方向的循环转发如图 5-3 所示。

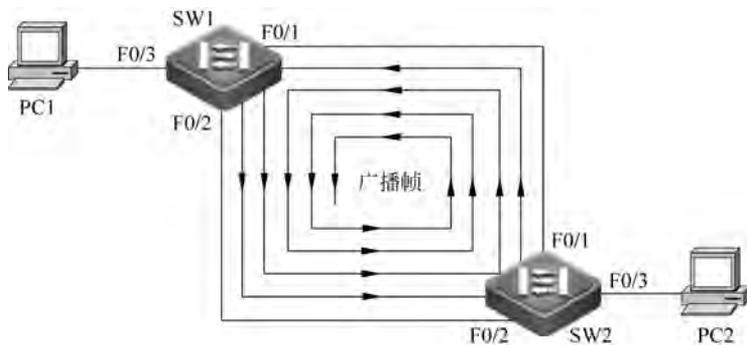


图 5-3 交换环路里广播帧的循环转发

随着其他设备间的通信增加,网络里的广播帧(如 ARP 请求、DHCP Discover 等)会越来越多,大量的广播帧在交换机之间不停循环,当流入的广播帧过多,导致所有可用带宽都被耗尽时,便形成了广播风暴。此时没有可用带宽转发正常数据流量,网络无法支持数据通



视频讲解

信。因此，一旦出现广播风暴，网络可能很快便会瘫痪。

2. MAC 地址表抖动

MAC 地址表抖动指 MAC 地址表不稳定。在图 5-2 中，假设两台交换机都是刚启动的，交换机的 MAC 地址表都为空。当交换机 SW1 收到此 PC1 发送的 ARP 广播帧（请求 PC2 的 MAC 地址）后，会从除接收端口之外的所有端口（F0/1 和 F0/2）“泛洪”出去，若交换机 SW2 从 F0/2 端口收到该 ARP 广播帧，SW2 便会根据地址学习功能在其 MAC 地址表中加入 PC1 的 MAC 地址和端口（F0/2）等信息；若交换机 SW2 从 F0/1 端口收到该帧，又会将 PC1 的 MAC 地址和端口 F0/1 关联并更新 MAC 地址表。随着该广播帧在两台交换机之间不停地循环转发，交换机 SW2 一会儿从 F0/2 收到该广播帧，一会儿又从 F0/1 收到同样的帧，造成 SW2 的 MAC 地址表中 PC1 关联的端口在 F0/2 和 F0/1 之间不停跳变，不停更新 MAC 地址表，使得 MAC 地址表无法稳定。由此可见，MAC 地址表的无法稳定，是由于交换机之间存在环路使得相同的帧在不同端口上被重复接收引起的。

这种持续的更新、刷新 MAC 地址表的过程会严重耗用内存资源，影响该交换机的交换能力，同时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。

3. 多帧复制

冗余拓扑除了带来广播风暴和 MAC 地址表不稳定外，还会引起多帧复制问题。多帧复制指目的主机收到多个相同的单播帧。

在图 5-2 中，假设 PC1 已知 PC2 的 MAC 地址，PC1 向 PC2 发送一个单播帧，交换机 SW1 收到该帧。若此时 SW1 的 MAC 地址表中没有目的主机 PC2 的 MAC 地址条目，SW1 便会将该帧从其 F0/1 和 F0/2 端口“泛洪”出去，则交换机 SW2 会从其 F0/1 和 F0/2 端口分别收到该单播帧。若 SW2 的 MAC 地址表中已有 PC2 的 MAC 地址条目，SW2 会将这两个帧都转发给 PC2。这样，PC2 就收到两个相同的单播帧，如图 5-4 所示。

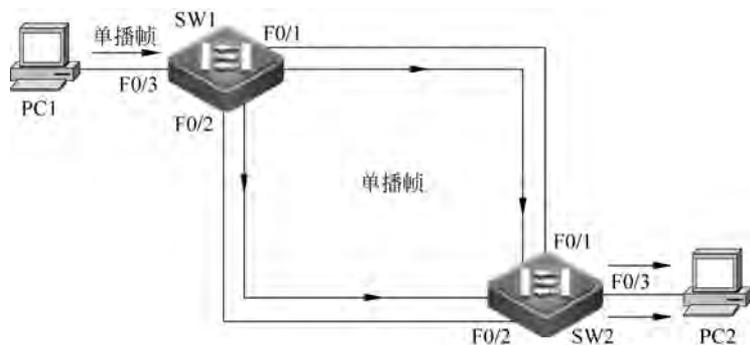


图 5-4 交换环路中多帧复制

PC1 仅发送一次单播帧，PC2 却收到两次。在工程中，重复帧复制也存在不足，比如在流量统计或计费软件的环境中，都造成不精确计算的问题。

5.1.2 生成树协议

通过冗余解决了由于单链路或单交换机故障引起的网络中断,提高了网络的可用性。但当在第二层采用冗余时,物理环路又会带来广播风暴、MAC 地址不稳定、多帧复制等问题。根据前面环路危害的分析可知,解决问题的主要途径是切断网络中的环路。

为了解决冗余链路引起的种种问题,IEEE 制定了 802.1d 协议,即生成树协议(Spanning-Tree Protocol,STP)。生成树协议通过在交换机上运行一套复杂的算法,将交换机的某些冗余端口置于“阻塞”状态(被阻塞的端口不转发数据),使得网络在通信过程中只有一条链路生效,从而使网络形成一个无环路的树状结构以消除环路,确保从源到任意目的地只有一条活动的逻辑链路。虽然逻辑上没有环路,但物理环路仍然存在,所以当活动的逻辑链路出现故障时,生成树协议将会重新计算出网络的最优链路,将处于“阻塞”状态的端口重新打开,不需要人工干预,这样,既保障了网络的正常运转,又保证了冗余能力,从而确保网络连接稳定、可靠。

虽然生成树协议并不像路由协议那样广为人知,但是它却掌管着交换机端口的转发大权,它的作用可形象比喻为“小树枝抖一抖,上层协议就得另谋生路”。

1. 生成树协议概述

生成树协议和其他协议一样,随着网络的不断发展而不断更新换代。在生成树协议的发展过程中,老的缺陷不断被克服,新的特性不断被开发出来。按照功能改进情况,可以把生成树协议的发展过程划分成 3 代。

第一代: STP(IEEE 802.1d)/RSTP(IEEE 802.1w)。

第二代: PVST/PVST+(思科网络的私有协议)。

第三代: MSTP(IEEE 802.1s)。

生成树协议的基本思想十分简单。众所周知,自然界生长的树是不会出现环路的,如果网络也能够像一棵树一样生长(扩展),就不会出现环路了。于是,生成树协议模拟自然界树的生长规律,从树根到树梢不会形成环路。所以,生成树协议通过定义根交换机、根端口、指定端口、路径开销等概念,目的就在于通过构造一棵自然树的方法,达到阻塞网络中冗余环路的目的,同时实现链路备份和路径最优化。

要实现这些功能,交换机之间必须进行一些信息的交换,在生成树协议中将这些交换的信息称为 BPDU(桥接协议数据单元)。BPDU 是运行 STP 功能的交换机之间交换的数据帧,所有支持 STP 的交换机都会接收并处理 BPDU,生成树协议的所有功能都是通过交换机之间周期性地发送的 BPDU 来实现的。理解 BPDU 中的各个字段的含义对掌握 STP 的工作原理至关重要,BPDU 包含的字段较多,这里着重介绍网桥 ID、端口 ID、路径开销和 BPDU 计时器 4 个字段。

1) 网桥 ID

在 STP 中,交换机也称网桥或桥。每个网桥都有一个唯一的网桥 ID(Bridge ID,BID),用于确定交换环路中的根网桥(根交换机)。网桥 ID 共 8 字节,由 2 字节网桥优先级和 6 字节网桥 MAC 地址组成,如图 5-5 所示。网桥优先级的取值范围是 0~65 535,默认值是 32 768。



视频讲解

网桥 ID 最小的交换机称为根网桥。首先比较网桥优先级,优先级最低的网桥将成为根网桥。如果网桥优先级相同,则比较网桥 MAC 地址,具有最低 MAC 地址的交换机将成为根网桥。

2) 端口 ID

端口 ID 参与决定到根网桥的路径。端口 ID 共 2 字节,由 1 字节的端口优先级和 1 字节的端口编号组成,如图 5-6 所示。端口优先级的取值范围是 0~255,默认值是 128 (0x80)。端口编号则是按照端口在交换机上的顺序排列的,例如,0/1 端口的 ID 是 0x8001,0/2 端口的 ID 是 0x8002。端口优先级值越小,则端口优先级越高。如果端口优先级值相同,则端口编号越小,优先级越高。

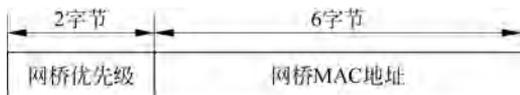


图 5-5 网桥 ID 的组成

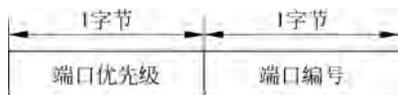


图 5-6 端口 ID 的组成

3) 路径开销

选出根网桥后,生成树算法要确定其他每台交换机到达根网桥的最短路径。生成树的根路径开销也称为根路径成本,是指到根网桥的某条路径上所有端口开销的累计和。根网桥的路径开销为零,其他交换机收到 BPDU 报文后,把报文中的根路径开销值(Root Path Cost)加上接收端口的开销值,得到该端口的路径开销,路径开销反映了端口到根网桥的“距离”。路径开销最低的路径是最短路径,会成为活动链路转发数据,而所有其他冗余路径都作为备份链路会被阻塞。

端口开销与端口的带宽有关,带宽越大,开销值越小。STP 使用的端口开销值由 IEEE 定义,IEEE 802.1d 标准修订前后端口开销如表 5-1 所示。

表 5-1 修订前后的 IEEE 802.1d 端口开销

链路带宽	成本(修订前)	成本(修订后)
10Gb/s	1	2
1000Mb/s	1	4
100Mb/s	10	19
10Mb/s	100	100

4) BPDU 计时器

BPDU 计时器决定了 STP 的性能和状态转换,生成树协议定义了 3 个计时器,这 3 个计时器的数值虽然可以修改,但一般情况下不建议修改。

(1) Hello Time(呼叫时间)。

Hello Time 是交换机之间定期发送 BPDU 的时间间隔,默认值是 2s,取值范围为 1~10s。

(2) Forward Delay(转发延迟)。

Forward Delay 是交换机从监听状态跳转到学习状态或从学习状态到转发状态的时间间隔,默认值是 15s,取值范围为 4~30s。

(3) Max Age(最大老化时间)。

Max Age 是交换机端口保存 BPDU 的最长时间。交换机收到 BPDU 会保存下来,同

时启动计时器开始倒计时。正常情况下,交换机之间每隔 2s(Hello Time)发送一次 BPDU,如果在 Max Age 内还没有收到新的 BPDU,便认为线路出现故障,邻居交换机无法到达,从而开始新的 STP 计算。Max Age 的默认值是 20s,取值范围为 6~40s。

2. STP 端口角色

在 STP 工作过程中,交换机端口会被自动配置为 4 种不同的端口角色。

(1) 根端口(Root Port)。

根交换机上没有根端口,根端口是对非根交换机而言的。根端口是交换环路中非根交换机上到达根网桥的路径开销值最小的端口。根端口从根网桥接收 BPDU 并向下发送,每个非根交换机上只能有一个根端口。根端口可以接收并转发数据。

(2) 指定端口(Designated Port)。

指定端口存在于根网桥和非根交换机上,根网桥上的所有端口都是指定端口,非根交换机上的指定端口用于转发根网桥与非根交换机之间的流量。交换机之间的每一个物理网段只能有一个指定端口。指定端口也可以接收并转发数据。

(3) 非指定端口(Non-designated Port)。

除根端口和指定端口之外的其余端口都被称为非指定端口,非指定端口处于阻塞状态。此类端口不会转发数据,也不会根据源 MAC 进行地址学习更新 MAC 地址表。

(4) 禁用端口(Disabled Port)。

禁用端口是指未开启 STP 的端口,这种端口不参与 STP 的计算过程。

3. STP 端口状态

由于交换机刚启动时并不清楚整个网络的拓扑情况,如果交换机的端口直接进入转发状态,则可能会形成暂时性的环路。因为这个原因,STP 引入了 5 种端口状态,即禁用(Disabled)、阻塞(Blocking)、监听(Listening)、学习(Learning)和转发(Forwarding)。这些状态与 STP 的运行过程以及交换机的工作原理有着重要的关系,每种端口状态的功能如表 5-2 所示。其中,“√”表示某种状态下具有某种功能,“×”表示某种状态下没有某种功能。

表 5-2 STP 端口状态的功能

状 态	功 能			
	接收 BPDU	发送 BPDU	学习 MAC 地址	转发 DATA
禁用(Disabled)	×	×	×	×
阻塞(Blocking)	√	×	×	×
监听(Listening)	√	√	×	×
学习(Learning)	√	√	√	×
转发(Forwarding)	√	√	√	√

在这 5 种端口状态中,禁用因端口未启用 STP,不参与 STP 的计算,其余 4 种为 STP 的正常状态。其中,转发和阻塞为稳定状态,根端口和指定端口处于转发状态,非指定端口处于阻塞状态;监听和学习是不稳定的中间状态,在经过一定时间后会自动跳转到其他状

态。端口状态的转换过程如下。

(1) 禁用状态。可以使用 no shutdown 命令和插入网线进行激活。

(2) 阻塞状态。链路激活，端口转换到阻塞状态，这个状态会停留大约 20s，主要用于决定该端口的角色，如果该端口是根端口或指定端口，将转换到下一个状态，即监听状态；如果该端口是非指定端口，则继续停留在阻塞状态；本来处在阻塞状态的端口，如果在最大老化时间（默认为 20s）内接收不到 BPDU，则也会自动转换到监听状态。

(3) 监听状态。除了接收 BPDU 外，还向邻居交换机发送 BPDU，以通知邻居该交换机将参与生成树的拓扑形成。在生成树拓扑形成过程中，该端口最终可能被选定为根端口或指定端口，也可能失去根端口或指定端口的地位，如果失去，则返回到阻塞状态。监听状态会停留大约一个转发延迟时间（默认为 15s）。

(4) 学习状态。一个端口在监听状态下经过一段时间后，将自动进入学习状态，开始学习 MAC 地址。学习状态会停留大约一个转发延迟时间。

(5) 转发状态。端口进入稳定状态，可以转发数据。

端口状态过渡和停留时间如图 5-7 所示。

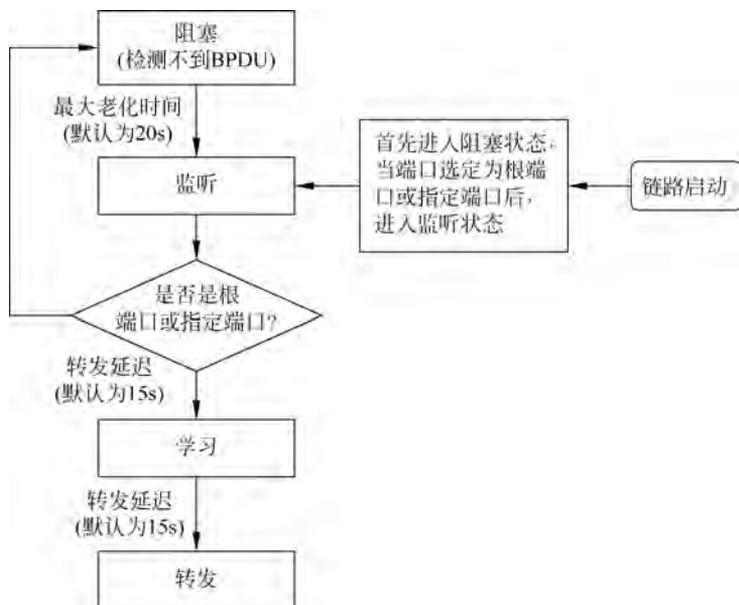


图 5-7 端口状态过渡和停留时间

从上述过程可以看出，端口在参与 STP 的计算过程中，先从阻塞状态开始，中间先后经过监听和学习状态，最后进入转发状态正常转发数据，这个过程需要耗费 30~50s 时间。这也是实际应用中，将一台计算机插入启动了 STP 功能的交换机上后，不能马上通信的原因。

4. STP 收敛

收敛是生成树的一个重要方面。STP 收敛是指网络在一段时间内确定根网桥，以根网桥为参考点计算出所有端口的角色，排除所有潜在的环路，使整个网络达到一个逻辑上没有环路的稳定状态的过程。当网络拓扑发生变化时，执行生成树算法，让网络重新收敛。生成

树的收敛经历如下 4 个过程。

(1) 选举根网桥。

在一个网络中只能存在一个根网桥,网桥 ID 最小的交换机被选举为根网桥。比较网桥 ID 时,首先比较网桥优先级,优先级最低的网桥将成为根网桥;如果网桥优先级相同,则比较 MAC 地址,具有最低 MAC 地址的交换机将成为根网桥。

在同一个广播域中的所有交换机参与选举根网桥。交换机开机时,每一台都假定自己就是根网桥,把自己的网桥 ID 写入 BPDU 中的根网桥 ID,封装 BPDU 并泛洪出去。每台交换机从邻居交换机收到 BPDU 后,都会将收到 BPDU 内的根桥 ID 与自己目前 BPDU 的根桥 ID 进行比较。如果接收 BPDU 的根桥 ID 比其目前的根桥 ID 小,就会用这个更小的值替换正在发送 BPDU 中的根桥 ID 字段,同时还修改根开销(Root Path)等参数的值,更新后再向外发送。这样,经过一段时间以后,同一个广播域中的所有交换机都会比较全部的根桥 ID,最后都有一致的根桥 ID。由此选举出具有最小桥 ID 的交换机作为根网桥(根交换机)。

STP 选举根网桥拓扑如图 5-8 所示,每台交换机的优先级值和 MAC 地址如图中所示。交换机开机时,都假定自己就是根网桥,然后开始发送 BPDU,BPDU 中的根桥 ID 等于自己的桥 ID。当 SW1 与 SW2 交换过 BPDU 后,因为 SW2 的优先级值 4096 比 SW1 的优先级值 32 768 小,因此,SW1 向 SW3 发出的 BPDU 中的 Root BID 值修改为 SW2 的 BID;SW3 接收到此 BPDU 后,读取 Root BID,发现优先级同为 4096,但其 MAC 地址比 Root BID 中的 MAC 地址大,因此 SW3 也修改 BPDU 中的 Root BID 值后再发出。以此类推,这样,经过两个周期左右的 BPDU 交换后,SW2 被选举为根网桥。根网桥上的所有端口均为指定端口,可以接收并转发数据。

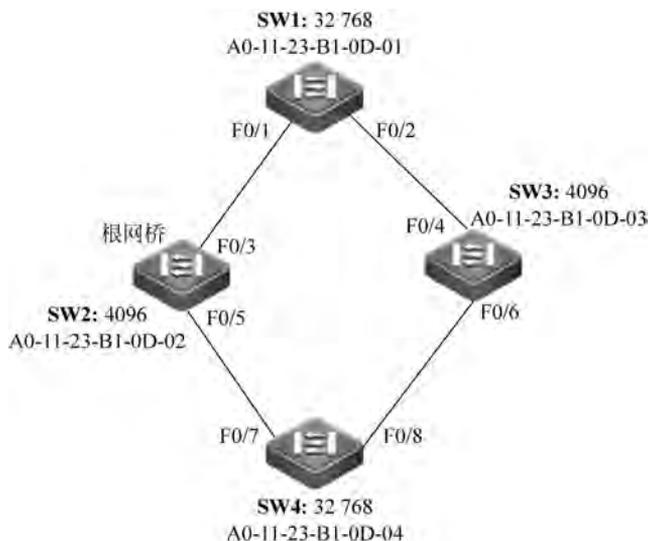


图 5-8 STP 选举根网桥拓扑

(2) 选举非根交换机的根端口。

除根网桥外的所有交换机都成为非根网桥,或称非根交换机。在交换环路中,每台非根交换机都至少有两个端口接收到来自根网桥的 BPDU,因此要在这些端口中选择出一个端

口作为转发端口,其他端口则不承担转发数据的任务,这个端口就是根端口。

每个非根交换机都需要选出一个根端口,而且仅有一个。非根交换机上根路径开销最小的端口被选举为根端口。

STP 选举根端口拓扑如图 5-9 所示,假设图中每条链路的带宽均为 100M,则交换机的端口开销值为 19,SW2 是根网桥,SW1、SW3 和 SW4 为非根交换机。根端口的选举顺序如下。

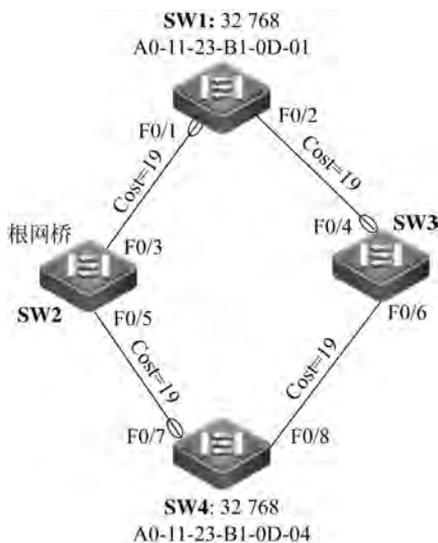


图 5-9 STP 选举根端口拓扑

① 根路径开销最小的端口选举为根端口。图 5-9 中,SW4 的 F0/7 和 F0/8 端口的根路径开销分别为 19 和 57(19+19+19),根路径开销最小的 F0/7 被选举为根端口,仿此可选举 SW1 的 F0/1 为根端口。

② 根路径开销相同的情况下,比较上游交换机的桥 ID。图 5-9 中,对于 SW3 来说,从端口 F0/4 和 F0/6 到达根交换机的根路径开销都是 38(19+19),此时端口的根路径开销相同,则比较端口对应的上游交换机(发送 BPDU 的交换机)的桥 ID。F0/4 端口对应的上游交换机是 SW1,F0/6 端口对应的上游交换机是 SW4,而 SW1 的桥 ID 小于 SW4 的桥 ID,因此,SW3 的 F0/4 被选举成根端口。

③ 在上游交换机的桥 ID 相同的情况下,比较上游交换机对应的端口 ID。在图 5-10 中,SW1 是根网桥,SW2 是非根交换机。SW2 的 F0/3 和 F0/4 到达根交换机的根路径开销相同,都是 19;上游交换机的桥 ID 也相同,都是交换机 SW1 的桥 ID。此时比较上游交换机的端口 ID。SW2 的 F0/3 对应的上游交换机端口为 SW1 的 F0/1,其端口 ID 为 128.1(此处假设端口优先级为默认值 128);SW2 的 F0/4 对应的上游交换机端口为 SW1 的 F0/2,其端口 ID 为 128.2。SW1 的 F0/1 端口 ID 更小,所以,SW2 的 F0/3 被选举为根端口。

④ 在上游交换机端口 ID 相同的情况下,比较接收交换机自身的端口 ID。图 5-11 中,SW1 是根交换机,SW2 为非根交换机,SW1 的 F0/1 端口通过集线器连接着 SW2 的 F0/2 和 F0/3。SW2 的 F0/2 和 F0/3 到达根交换机的根路径开销一样,上游交换机的桥 ID 也一样(都是 SW1 的桥 ID),上游交换机的端口 ID 也一样(都是交换机 SW1 的 F0/1 的端口

ID)。接下来比较交换机 SW2 自身的端口 ID。SW2 上 F0/2 端口的 ID 小于 F0/3 端口的 ID,所以,SW2 的 F0/2 端口被选举为根端口。

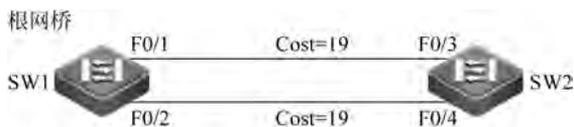


图 5-10 比较上游交换机的端口 ID

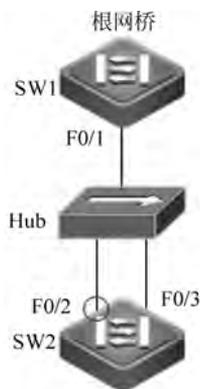


图 5-11 比较接收交换机的端口 ID

综上所述,选举根端口的依据顺序为:根开销最小、上游交换机 ID 最小、上游交换机发送端口 ID 最小、接收端口 ID 最小。

(3) 选举每个物理链路的指定端口。

在 STP 中,物理链路指的是交换机之间互连的链路。在物理链路选举指定端口,实际上是选举指定网桥(指定交换机)。每个物理链路都有一个指定交换机,该交换机负责把数据发往根交换机,指定交换机上的端口称为指定端口。选举指定端口的依据顺序为:根开销最小、交换机 ID 最小、端口 ID 最小。

① 根路径开销最小的被选定为指定网桥。在图 5-9 中,SW2 是根网桥。在 SW2-SW1 链路上,SW2 到根的路径开销为 0,SW1 到根交换机的路径开销为 19,所以 SW2 是这个链路上的指定交换机,则 SW2 上的端口 F0/3 是 SW2-SW1 链路上的指定端口。同理,SW2 是 SW2-SW4 链路上的指定交换机,SW2 上的端口 F0/5 是 SW2-SW4 链路上的指定端口。在 SW1-SW3 链路上,SW1、SW3 到根交换机的开销分别为 19、38(19+19),所以在此链路上,SW1 是指定交换机,则 SW1 上的 F0/2 端口为 SW1-SW3 链路上的指定端口。同理可选举 SW4 上端口 F0/8 为 SW4-SW3 链路上的指定端口。

② 根路径开销相同的情况下,比较交换机的桥 ID。图 5-12 中,SW2 是根交换机,在 SW1-SW3 链路上,SW1、SW3 到根交换机的开销相同,都为 19,接下来比较交换机的 ID,SW3 的桥 ID 比 SW1 的桥 ID 小,则在 SW1-SW3 链路上,SW3 是指定交换机,则 SW3 上的 F0/4 端口是 SW1-SW3 链路上的指定端口。

如果指定交换机上有多个端口连接到同一个网段,则具有最小端口 ID 的端口成为指定端口,这种情况较少出现,此处不再详述。

(4) 阻塞非指定端口。

除根端口和指定端口之外的其余所有端口被称为非指定端口。非指定端口自动被阻塞,不能转发数据。图 5-8 中的 F0/6 端口是非指定端口,被阻塞,这相当于 SW4-SW3 链路被逻辑断开,这条链路只能转发 BPDU,不能转发数据。图 5-8 所示拓扑最终的无环生成树逻辑拓扑即稳定的生成树逻辑拓扑如图 5-13 所示。

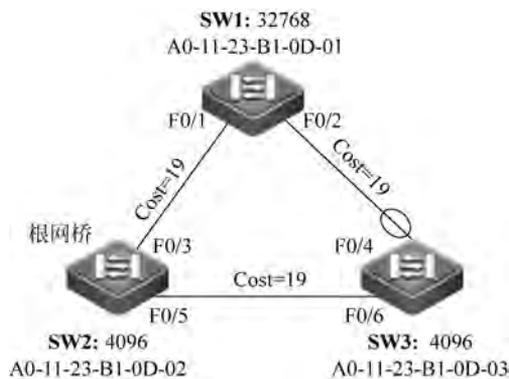


图 5-12 比较交换机的桥 ID

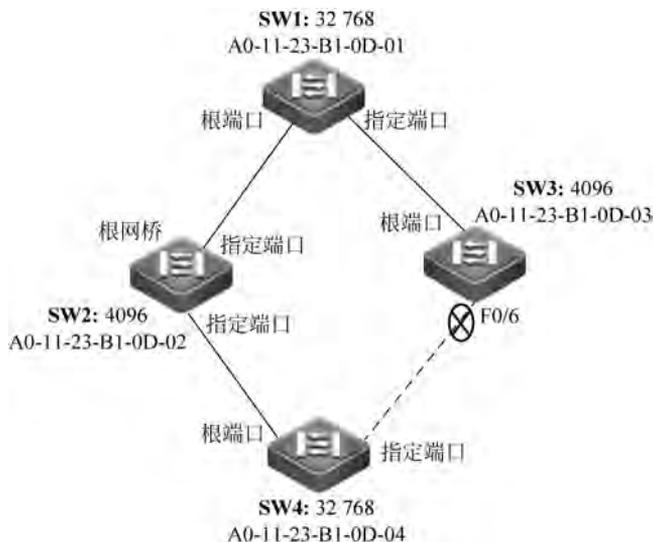


图 5-13 稳定的生成树逻辑拓扑

STP 拓扑稳定后,根交换机通过每隔 2s 的呼叫时间创建和发送 BPDU,非根交换机通过根端口接收 BPDU,并且从指定端口转发改变后的 BPDU。各交换机通过接收到的 BPDU 消息,来保持各端口状态的有效,直到拓扑发生变化。这样,通过生成树算法构造了一棵“树”,逻辑上阻断网络中存在的环路,达到冗余链路的同时还实现了链路备份和路径最优化。当主链路出现故障时,阻断的链路马上恢复工作。

5.1.3 快速生成树协议

从上面的分析可知,传统的 STP(IEEE 802.1d)虽然可以解决网络交换环路问题,但端口从阻塞状态进入转发状态需要经历监听状态和学习状态,即从阻塞状态过渡到转发状态必须经历 2 倍的转发延迟时间。所以,网络拓扑发生变化时需要经过 30~50s 的时间才能恢复连通性,这对于现在高可靠性的网络而言,已无法满足用户的需求。

针对传统的 STP 收敛慢这一弱点,一些网络设备供应商开发了针对性的 STP 增强协

议,而 IEEE 正是看到了这一广泛的需求,制定了 802.1w 标准协议。这一协议针对传统的 802.1d 收敛慢做了改进,它使得以太网的环路收敛在 1~10s 完成,所以 802.1w 又被称为快速生成树协议(Rapid Spanning Tree Protocol,RSTP)。RSTP 是从传统的 STP 发展而来的,具备 STP 的所有功能,但 RSTP 引入了新的机制,加快了网络收敛速度,提高了网络的可靠性和稳定性。

1. RSTP 的改进

相对于 STP,RSTP 能够快速收敛的原因主要是在以下 3 方面做了改进。

(1) 新增了两种端口角色。

RSTP 端口角色中的根端口和指定端口的确定方法与 STP 一致,而对阻塞端口可进一步细分为替换端口(Alternate Port)和备份端口(Backup Port)。正常情况下,这两种端口均处于阻塞状态(丢弃状态),接收 BPDU 但不转发数据。替换端口是根端口的备份,当根端口失效,替换端口立刻转换为根端口,直接进入转发状态;备份端口是指定端口的备份,若指定端口失效,备份端口立刻转换为指定端口,也直接进入转发状态。

在图 5-14 所示的网络环境中,假设 SW1 的桥 ID 最小,而 SW3 的桥 ID 最大。根据前面的分析可知,SW1 被选举为根交换机,F0/3、F0/4 被选举为根端口,图中标示为 RP,而 F0/1、F0/2、F0/6 被选举为指定端口,图中标示为 DP。在 STP 中 F0/5、F0/7 为阻塞端口,而在 RSTP 中,SW3 上的 F0/5 端口是根端口 F0/3 的替换端口,图中标示为 AP,为当前根端口 F0/3 到根网桥 SW1 提供了替代路径;SW2 上的 F0/7 端口是指定端口 F0/6 的备份端口,图中标示为 BP,为到达同网段提供备份路径,是对一个网段的冗余连接。所以当 SW3 的根端口 F0/3 断开或 SW2 的指定端口 F0/6 断开时,它们的替换端口或备份端口会无延迟进入转发状态,切换过程几乎不影响网络用户的正常通信。

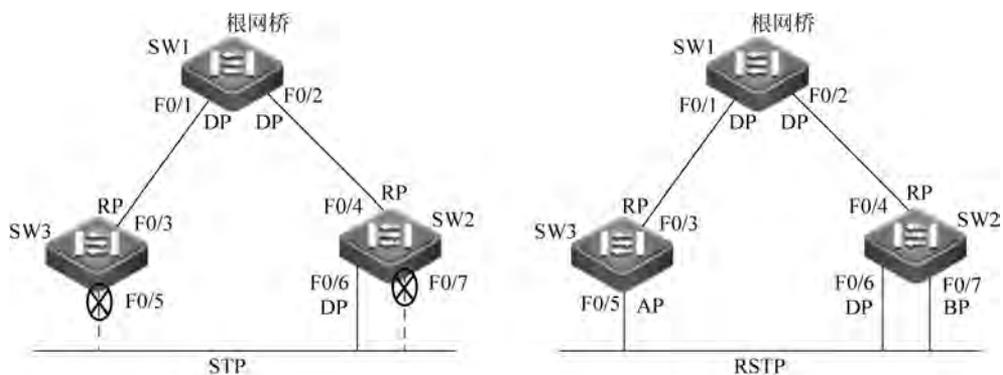


图 5-14 STP 和 RSTP 中的端口角色

(2) 引入了边缘端口的概念。

边缘端口(Edge Port)是指连接计算机、打印机等终端设备的交换机端口,这类端口通常不会产生环路,没必要也经过 30~50s 的时间才进入转发状态。边缘端口无须经过监听、学习等中间状态,可以直接进入转发状态。

(3) 区分了不同的链路类型。

RSTP 在点对点类型的链路上使用请求/同意握手机制完成快速收敛。非边缘端口能

否快速进入转发状态,取决于该端口所在的链路类型。非边缘端口的链路类型有点对点和共享两种,链路类型是自动确定的,一般全双工链路是点对点类型,半双工链路是共享类型。若是点对点链路,该端口只需要向对端交换机发送一个握手请求报文,如果对端响应了一个同意报文,则该端口可以直接进入转发状态。如果端口所在的链路是共享的,则端口状态切换同 STP,需要经过 2 倍的转发延迟才能进入转发状态。因当前的交换机端口默认情况下都工作在全双工状态,交换机之间的链路都是点对点链路,故端口都可以快速进入转发状态。

2. RSTP 的端口状态

STP 的端口状态有禁用、阻塞、监听、学习和转发 5 种,阻塞状态和监听状态在数据帧转发和 MAC 地址学习上没有区别,都是丢弃数据帧而且不学习 MAC 地址。RSTP 的端口状态只有丢弃、学习和转发 3 种,后面介绍的 MSTP 端口状态与 RSTP 相同。不同生成树协议的端口状态及功能对比如表 5-3 所示。

表 5-3 STP 和 RSTP/MSTP 端口状态及功能对比

STP 端口状态	RSTP/MSTP 端口状态	RSTP/MSTP 端口功能	
		学习 MAC 地址	转发数据
禁用	丢弃	×	×
阻塞			
监听			
学习	学习	√	×
转发	转发	√	

替换端口和备份端口处于丢弃状态,根端口和指定端口稳定状态下处于转发状态,学习状态是根端口和指定端口在进入转发状态之前的一种临时过渡状态。

5.1.4 多生成树协议

当前的交换网络往往工作在多 VLAN 环境下,不管是 STP 还是 RSTP,在进行生成树计算时,都没有考虑多个 VLAN 的情况,而是所有 VLAN 共享一棵生成树,因此在交换机的一条 Trunk 链路上,所有 VLAN 要么全部处于转发状态,要么全部处于阻塞状态,这就导致链路带宽不能充分利用,无法实现负载均衡。多 VLAN 情况下 STP/RSTP 的不足如图 5-15 所示。

在图 5-15 中,假设通过生成树计算,交换机 SW2 的 F0/4 端口被阻塞,网络中只有一棵生成树,那么所有 VLAN 的流量均无法从该端口通过,VLAN 10 的用户访问服务器 Server 的流量路径是 SW2-SW3-SW1-SW4-Server,VLAN 20 的用户访问服务器 Server 的流量路径是 SW3-SW1-SW4-Server,导致所有 VLAN 的流量均从 SW3 与 SW1 之间的链路通过,可能造成该链路发生拥堵,而同时 SW2 与 SW1 之间的链路却空闲,无任何流量通过,冗余链路的带宽完全被浪费。

虽然思科的 PVST/PVST+ 是可在每个 VLAN 中单独运行 STP 算法的协议,但由于交换机资源的因素,如果在单台交换机上实现几十个 VLAN,那么要运行几十个独立的

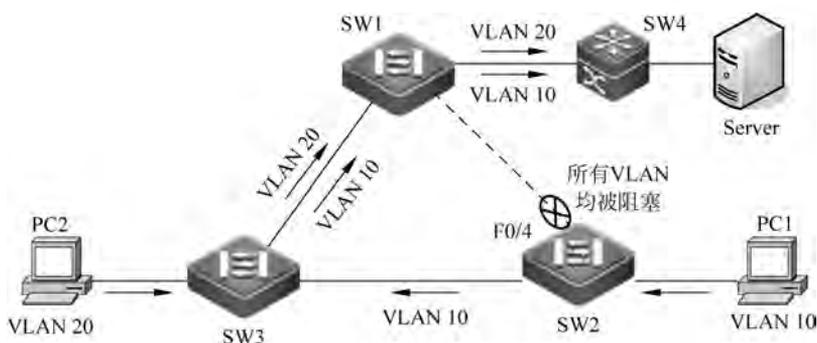


图 5-15 多 VLAN 情况下 STP 和 RSTP 的不足

STP 算法,显然会对设备的性能和资源提出严峻的挑战,并且由于其是私有的,不同厂家设备的互操作性存在问题。针对以上这些因素,IEEE 制定了 802.1s 的协议,目的是既能够实现同一台交换机内运行不同 STP 算法的协议,又可以将相同属性的 VLAN 归纳成组,在一个 VLAN 组内运行一个 STP 算法。

MSTP(Multiple Spanning Tree Protocol)是 IEEE 802.1s 中定义的一种新型多实例化生成树协议,除了具有 RSTP 的快速收敛机制外,还能实现链路的负载均衡。MSTP 将一个或多个 VLAN 映射到一个实例(Instance)中,同一个交换机上可以有多个实例,每个实例运行一棵单独的生成树,不同的实例可以有不同的生成树计算结果,这样就可以控制各 VLAN 的数据沿着不同的路径进行转发,实现基于 VLAN 的数据分流,从而充分利用链路带宽。多 VLAN 情况下 MSTP 的负载均衡如图 5-16 所示。

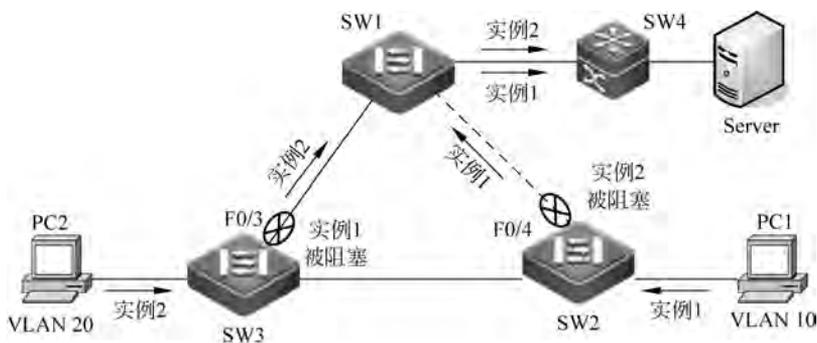


图 5-16 多 VLAN 情况下 MSTP 的负载均衡

在图 5-16 中,将 VLAN 1~VLAN 10 映射到实例 1,VLAN 11~VLAN 20 映射到实例 2,通过配置交换机在不同实例中的优先级,使得 SW2 成为实例 1 的根网桥,SW3 成为实例 2 的根网桥。假设实例 1 中生成树的计算结果是 SW3 的 F0/3 端口被阻塞,实例 2 中生成树的计算结果是 SW2 的 F0/4 端口被阻塞,这样,SW3 的 F0/3 端口阻塞实例 1 中的 VLAN 数据流通过,但允许实例 2 中的 VLAN 数据流通过,SW2 的 F0/4 端口阻塞实例 2 中的 VLAN 数据流通过,但允许实例 1 中的 VLAN 数据流通过。因此,VLAN 10 的用户访问服务器 Server 的流量路径是 SW2-SW1-SW4-Server,VLAN 20 的用户访问服务器 Server 的流量路径是 SW3-SW1-SW4-Server,从而实现了负载均衡的效果。由此可见,在

MSTP 中,同一个端口在不同实例中的端口角色及状态可以不同,如图 5-16 中的 F0/3 和 F0/4 端口,既可以阻塞某些实例的流量,同时又允许其他实例的流量通过。MSTP 可以实现不同 VLAN 的数据沿着不同的路径转发,从而实现了基于 VLAN 的负载均衡。

MSTP 可以向下兼容 RSTP 和 STP,但如果网络中存在 STP 与 RSTP/MSTP 的混用,交换机会根据“就低”原则,使用 STP 来计算生成树,从而导致无法发挥 RSTP/MSTP 的快速收敛功能。所以,网络中尽可能使用 MSTP 来消除交换环路,这样既能发挥生成树的快速收敛功能,又能实现不同 VLAN 的负载均衡。

5.1.5 配置 STP/RSTP/MSTP 常用命令

锐捷二层交换机与三层交换机配置生成树命令基本相同,常用的配置命令如表 5-4 所示。

表 5-4 STP/RSTP/MSTP 常用的配置命令

命令模式	CLI 命令	作用
全局模式	spanning-tree	打开生成树协议,默认为关闭
	no spanning-tree	关闭生成树协议
	spanning-tree mode{stp rstp mstp}	设置生成树协议类型,默认是 MSTP
	spanning-tree priority < 0-61440 >	设置 STP/RSTP 交换机的优先级。默认是 32 768
接口模式	spanning-tree port-priority < 0-240 >	设置 STP/RSTP 端口的优先级。默认是 128
	spanning-tree link-type {point-to-point shared}	设置生成树的链路类型
全局模式	spanning-tree portfast	设置端口为边缘端口
	spanning-tree portfast default	设置交换机的所有端口为边缘端口
MSTP 模式	spanning-tree mst configuration	进入 MSTP 配置模式
MSTP 模式	instance instance-id vlan vlan-range	设置 VLAN 与实例的映射关系
全局模式	spanning-tree mst instance-id priority < 0-61440 >	设置交换机在 MSTP 实例中的优先级
接口模式	spanning-tree mst instance-id port-priority < 0-240 >	设置端口在 MSTP 实例中的优先级
特权模式	show spanning-tree [summary]	显示生成树协议的全局信息
	show spanning-tree interface interface-id	显示生成树协议的端口信息
	show spanning-tree mst instance instance-id	
	show spanning-tree mst configuration	显示 MSTP 配置信息
	show spanning-tree mst [instance-id]	显示 MSTP 实例信息

5.2 任务实施

在 STP/RSTP 中,所有 VLAN 共享一棵生成树,无法在 VLAN 间实现数据流量的负载均衡,链路被阻塞后将不承载任何流量。而当前的交换网络往往都工作在多 VLAN 环境下,为了保证网络的快速收敛,又能基于 VLAN 实现流量分担,一般都配置 MSTP 以消

除交换环路并在链路上实现负载均衡。下面以图 5-17 所示的网络拓扑介绍 MSTP 的配置。

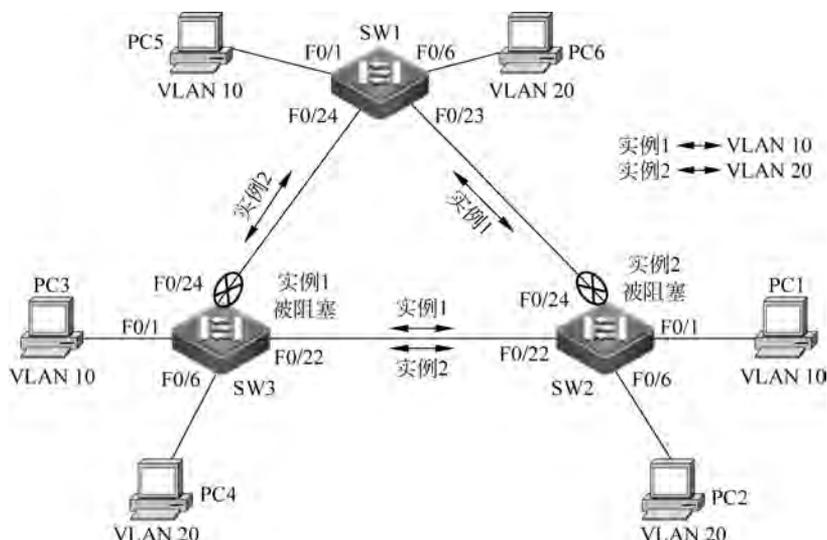


图 5-17 MSTP 的网络拓扑

图 5-17 中,每个交换机的优先级、MAC 地址及 VLAN 信息如表 5-5 所示。

表 5-5 交换机的优先级、MAC 地址及 VLAN 信息

交换机信息			VLAN 信息		
设备	优先级值	MAC 地址	VLAN ID	端口成员	Trunk 端口
SW1	32 768	00D0.9748.E3DE	VLAN 10	F0/1~5	F0/23, F0/24
			VLAN 20	F0/6~10	
SW2	32 768	00D0.58C3.872C	VLAN 10	F0/1~5	F0/22, F0/24
			VLAN 20	F0/6~10	
SW3	32 768	000A.F3C2.1A06	VLAN 10	F0/1~5	F0/22, F0/24
			VLAN 20	F0/6~10	

1. 基础配置

为了防止环路影响操作, MSTP 配置成功之前可先断开交换机之间互连的任一根线。

(1) 修改每台交换机的主机名。

依次在全局模式下用 `hostname` 命令修改每台交换机的主机名。

```
switch(config)# hostname SW1 //修改第一台交换机的主机名,其他两台类似
```

(2) 设置 PC 的 IP 地址。

此处假设图 5-17 中 6 台 PC 都处于 192.168.1.0/24 网络中,每台 PC 的 IP 地址中主机号与图中主机编号相同,如 PC1 的地址为 192.168.1.1/24。配置地址后,使用 `ping` 命令测试 6 台 PC 间的连通性,验证结果应是互通。

2. 配置 VLAN 及 Trunk

参照任务 4,按表 5-5 在每台交换机上配置 VLAN 及 Trunk。使用 ping 测试主机的互通性,VLAN 10 中的 PC1、PC3 和 PC5 互通,VLAN 20 中的 PC2、PC4 和 PC6 互通,其他不通。

3. 配置 MSTP

根据表 5-5 可知,三台交换机的优先级由高到低依次为 SW3→SW2→SW1,在不基于 VLAN 进行流量均衡时,默认情况下 SW3 被推选为根交换机,SW2 的 F0/22、SW1 的 F0/24 被推选为根端口,SW2 的 F0/24、SW3 的 F0/22 和 F0/24 被推选为指定端口,因此,SW1 的 F0/23 就为阻塞端口,VLAN 10 和 VLAN 20 的流量路径都为 SW1↔SW3↔SW2。而图 5-17 中要求每个 VLAN 的流量路径不同,VLAN 10 的流量路径为 SW1↔SW2↔SW3,VLAN 20 的流量路径为 SW1↔SW3↔SW2。根据不同 VLAN 的流量路径确定 MSTP 的配置信息,如表 5-6 所示。

表 5-6 MSTP 配置信息

实例	关联 VLAN	交换机的优先级由高到低的顺序	根交换机	阻塞端口
实例 1	VLAN 10	SW2→SW1→SW3	SW2	SW3 的 F0/24
实例 2	VLAN 20	SW3→SW1→SW2	SW3	SW2 的 F0/24

下面以 VLAN 10 为例分析为什么按表 5-6 所示的信息配置实例 1。首先分析根交换机的配置,如果 SW1 配置为根交换机,则 SW3 必为非根交换机,SW3 的 F0/24 端口为根端口,不可能为阻塞端口;如果 SW3 配置为根交换机,则 SW3 的 F0/24 端口为指定端口,也不可能成为阻塞端口,所以,只有 SW2 被配置为实例 1 的根交换机。在 SW1~SW3 的网段上,如果 SW3 的 F0/24 是阻塞端口,则 SW1 的 F0/24 应是指定端口。在 SW1~SW3 的网段推选指定端口时,要使得 SW1 的 F0/24 成为指定端口,那么交换机 SW1 的优先级必须高于 SW3。所以,对实例 1,交换机的优先级由高到低的顺序为 SW2→SW1→SW3。

(1) 配置 MSTP。

三台交换机上 MSTP 配置相同,此处只列出 SW1 上的配置命令,其他的仿此完成。

```
SW1(config)# spanning-tree //开启生成树协议
SW1(config)# spanning-tree mode mstp //配置生成树协议为 MSTP
SW1(config)# spanning-tree mst configuration //进入 MSTP 配置模式
SW1(config-mst)# instance 1 vlan 10 //将 VLAN 10 映射到实例 1
SW1(config-mst)# instance 1 vlan 20 //将 VLAN 20 映射到实例 2
```

在使用命令建立 VLAN 与实例之间的映射时,系统会提示“将 VLAN 映射到实例之前,必须创建 VLAN”,所以,要确保交换机上存在相应的 VLAN。

实例号的取值范围为 0~64,不同的实例通过实例号来区分,默认情况下所有 VLAN 均与实例 0 映射,实例 0 不能被删除。一个实例可以包含一个或多个 VLAN,但是一个 VLAN 只能映射到一个实例中。此例中只考虑 VLAN 10 和 VLAN 20 的流量均衡,其他

VLAN 映射到默认的实例 0,实例 0 的流量沿默认生成树的路径传输。

(2) 配置交换机的优先级实现负载均衡。

根据表 5-6 中交换机的优先级高低配置每个交换机在不同实例中的优先级值。交换机的优先级值默认为 32 768,值越小优先级越高,且优先级值只能设置为 0 或 4096 的倍数。

实例 1 中交换机的优先级由高到低的顺序为 SW2→SW1→SW3。

```
SW2(config)# spanning-tree mst 1 priority 0
           //配置 SW2 在实例 1 中的优先级,使其成为实例 1 的根交换机
SW1(config)# spanning-tree mst 1 priority 4096
           //配置 SW1 在实例 1 中的优先级,使其成为实例 1 的备份根交换机
```

实例 2 中交换机的优先级由高到低的顺序为 SW3→SW1→SW2。

```
SW3(config)# spanning-tree mst 2 priority 0
           //配置 SW3 在实例 2 中的优先级,使其成为实例 2 的根交换机
SW1(config)# spanning-tree mst 2 priority 4096
           //配置 SW1 在实例 2 中的优先级,使其成为实例 2 的备份根交换机
```

(3) 验证生成树信息。

① 在特权模式下使用 show spanning-tree mst configuration 命令查看 MSTP 配置信息。如:

```
SW1# show spanning-tree mst configuration
Multi spanning tree protocol : Enabled
Name :
Revision : 0
Instance      Vlans  Mapped
-----
0             : 1 - 9, 11 - 19, 21 - 4094
1             : 10
2             : 20
```

该命令显示 VLAN 与实例的映射关系。实例 0 是默认实例,没有具体指定实例的 VLAN 均属于实例 0。

② 在特权模式下使用 show spanning-tree mst 命令可以查看 MSTP 中每个实例的配置信息。下面查看每个交换机上实例 1 的相关信息。

```
SW2# show spanning-tree mst 1           //显示实例 1 的信息
###MST 1 vlans mapped : 10             //VLAN 与实例 1 的映射关系
BridgeAddr : 00D0.58C3.872C            //SW2 交换机的 MAC 地址
Priority : 0                            //SW2 交换机在实例 1 中的优先级值
TimeSinceTopologyChange : 0d:0h:17m:42s
TopologyChanges : 10
DesignatedRoot : 1001. 00D0.58C3.872C  //根网桥的 MAC 地址(从第 5 位开始)
RootCost : 0                            //根路径开销
RootPort : 0                            //根端口号
```

从 SW2 的显示信息可以看出,实例 1 中的根网桥的 MAC 地址与 SW2 的 MAC 地址相同,这说明 SW2 就是实例 1 的根网桥。所以,SW2 的根路径开销为 0,根端口号为 0,0 表示根交换机上没有根端口。

```
SW1 # show spanning-tree mst 1 //显示实例 1 的信息
# # # # # MST 1 vlans mapped : 10 //VLAN 与实例 1 的映射关系
BridgeAddr : 00D0.9748.E3DE //SW1 交换机的 MAC 地址
Priority : 4096 //SW1 交换机在实例 1 中的优先级值
TimeSinceTopologyChange : 0d:0h:18m:24s
TopologyChanges : 12
DesignatedRoot : 1001. 00D0.58C3.872C //根网桥的 MAC 地址(从第 5 位开始)
RootCost : 190000 //根路径开销
RootPort : FastEthernet 0/23 //根端口号
```

从 SW1 的显示信息可以看出,实例 1 中的根网桥的 MAC 地址与 SW1 的 MAC 地址并不相同,这说明 SW1 不是实例 1 的根网桥。从上述信息中还可以看出,F0/23 端口为根端口,这说明 VLAN 10 的流量通过 SW1 的 F0/23 端口转发出去到达根交换机 SW2,再由 SW2 发送出去,这与规划的流量路径是一致的。

```
SW3 # show spanning-tree mst 1 //显示实例 1 的信息
# # # # # MST 1 vlans mapped : 10 //VLAN 与实例 1 的映射关系
BridgeAddr : 000A.F3C2.1A06 //SW3 交换机的 MAC 地址
Priority : 32768 //SW3 交换机在实例 1 中的优先级值
TimeSinceTopologyChange : 0d:0h:19m:47s
TopologyChanges : 2
DesignatedRoot : XXXX. 00D0.58C3.872C //根网桥的 MAC 地址(从第 5 位开始)
RootCost : 200000 //根路径开销
RootPort : FastEthernet 0/22 //根端口号
```

从 SW3 的显示信息可以看出,实例 1 中的根网桥的 MAC 地址与 SW3 的 MAC 地址并不相同,这说明 SW3 不是实例 1 的根网桥。从上述信息中还可以看出,F0/22 端口为根端口,这说明 VLAN 10 的流量通过 SW3 的 F0/22 端口转发出去到达根交换机 SW2,再由 SW2 发送出去,这也与规划的流量路径是一致的。

③ 在特权模式下使用 show spanning-tree summary 命令可以查看生成树协议中每个实例的概要及端口信息。每个实例的概要中包含了前两个查看命令显示的相关信息,只是显示格式略有不同;端口信息包括了端口的角色、状态、开销、优先级、链路类型及是否是边缘端口等特性。由于显示内容比较多,此处不列出。

4. 设置边缘端口,加快网络收敛速度

在 3 台交换机上将连接主机的端口设置成边缘端口(Edge Port),使得这些端口可以无时延地进入转发状态,避免生成树计算过程中的转发延迟导致用户访问网络中断,而且边缘端口上主机的接入或插拔也不会导致生成树协议重新计算,增加了网络的稳定性。根据表 5-5 将每个业务 VLAN 的成员端口设置为边缘端口。此处只列出 SW1 上的配置命令,其他交换机的配置命令相似。

```
SW1(config)# interface range f0/1 - 10
SW1(config-if-range)# spanning-tree portfast //将端口设置为边缘端口
SW1(config-if-range)# exit
```

在使用命令 `spanning-tree portfast` 将端口设置为边缘端口时,系统会提示:“边缘端口只能连接终端主机,若连接交换机等设备会形成环路。”

所有的配置完成后,再次使用 `show spanning-tree summary` 命令查看生成树协议的全局信息,包括生成树协议中每个实例的摘要及端口信息,此处重点关注端口信息中是否是边缘端口的变化。

```
SW1# show spanning-tree summary
Spanning-tree enabled protocol mstp //生成树协议类型为 MSTP
//以下显示默认实例 0 的相关信息
MST 0 vlans map : 1 - 9, 11 - 19, 21 - 4094 //默认的实例 0 的相关信息
...
//以下显示实例 1 的相关信息
MST 1 vlans map : 10 //VLAN 10 与实例 1 映射
Region Root Priority 0 //根桥优先级
Address 00D0.58C3.872C //根桥 MAC 地址
this bridge is region root //这个桥是根桥

Bridge ID Priority 4096 //自身桥优先级
Address 00D0.9748.E3DE //自身桥 MAC 地址

Interface Role Sts Cost Prio Type OperEdge
-----
Fa0/1 Desg FWD 190000 128 P2p True
Fa0/6 Desg FWD 190000 128 P2p True
Fa0/23 Root FWD 19 128 P2p False
Fa0/24 Desg FWD 19 128 P2p False

//以下显示实例 2 的相关信息
MST 2 vlans map : 20
...
```

5.3 知识扩展

5.3.1 STP 拓扑变更机制

一个网络中所有交换机的端口都处于阻塞或者转发状态,那这个网络就达到收敛,处于稳定状态。转发状态的端口发送且接收数据和 BPDU,阻塞状态的端口仅接收 BPDU。

当网络拓扑变更时,交换机必须重新计算 STP,端口的状态也会发生改变,这样会中断用户通信,直至计算出一个重新收敛的 STP 拓扑。网络中的交换机是如何知道网络拓扑发生变化的?发生变化的交换机开始逐级向上游交换机发送 TCN (Topology Change

Notification)BPDU 即拓扑变化通知 BPDU,直至根交换机,根交换机收到后再将拓扑变化通知广播到整个网络。

在正常的 STP 操作中,一个交换机从根端口持续接收根交换机发来的 BPDU,它从不向根交换机发送 BPDU。当某台交换机发现拓扑变化时,该交换机将生成 TCN BPDU(拓扑变化通知 BPDU),并每隔呼叫时间从其根端口向上游交换机发送,直至上游交换机发回确认拓扑变化通知(TCA) BPDU 为止;上游交换机接收到 TCN BPDU,知道这是一个 TCN BPDU,给下游交换机发送 TCA BPDU,同时从其自己的根端口向上游交换机发送 TCN BPDU。生成树中从发现网络拓扑变化的交换机到根交换机的路径上每个交换机重复这个动作,直到根交换机收到这个 TCN BPDU,最后根交换机再向全网内所有交换机广播拓扑变更的通知。

在如图 5-18 所示环境中,SW5 检测到某个处于转发状态的指定端口所在链路发生了拓扑变化,它将会启动一次收敛过程,图中的编号标识了各类消息发送的顺序。

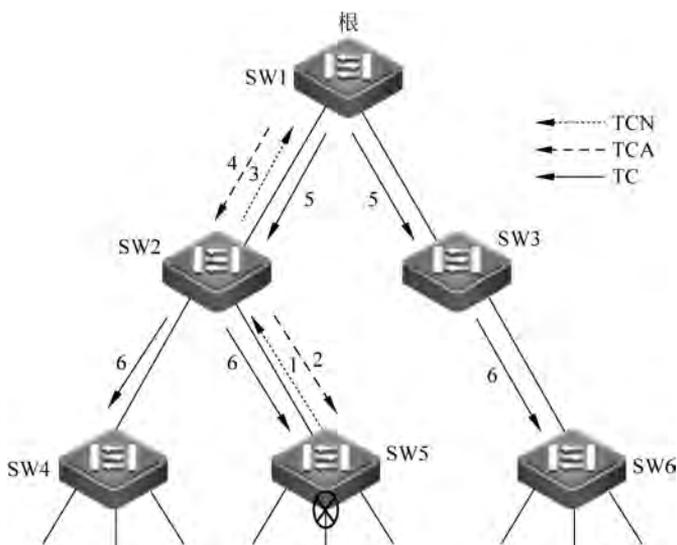


图 5-18 STP 拓扑变更

- ① SW5 检测到拓扑变化,从其根端口向它的上游交换机 SW2 发送 TCN。
- ② SW2 收到 SW5 发过来的 TCN,SW2 使用 TCA 向 SW5 确认。
- ③ SW2 产生 TCN,从其根端口发给上游交换机 SW1,也就是根交换机。
- ④ SW1 收到 SW2 发过来的 TCN,SW1 使用 TCA 向 SW2 确认。

至此完成了拓扑变化通知到根交换机的过程。一旦根交换机知道网络拓扑发生变化,它开始向外广播 TC(Topology Change,拓扑变化)比特位被设置的配置 BPDU。

- ⑤ SW1 向外广播 TC。

- ⑥ 下游交换机 SW2、SW3 收到 SW1 广播的 TC,转发该 BPDU 到各自的下游交换机。

最后,所有交换机都知道拓扑发生变化。所有的下游交换机得到拓扑改变的通知后,会把它们的 Address Table Aging(地址表老化)计时器从默认值(300s)降为 Forward Delay(默认为 15s),从而让不活动的 MAC 地址比正常情况下更快地从地址表更新。所有的交换机将重新决定根交换机、交换机的根端口以及每个物理网段的指定端口,这样生成树的拓扑

结构也就重新决定了。

5.3.2 边缘端口

由前面的 STP 工作原理可知,启用 STP 功能的交换机,一个端口从 up 到 forwarding 大约需要 50s 的时间,而普通的非网管型交换机,端口从 up 到 forwarding 瞬间就可以完成。但非网管型交换机之间一旦有环路存在,端口不仅不能瞬间变为 forwarding,而且会由于广播风暴变为 down,整个网络都会瘫痪。显然,相对于端口能否从 up 瞬间到 forwarding 而言,交换机支持 STP 更重要。

RSTP/MSTP 引入的边缘端口可以很好地解决端口从 up 到 forwarding 的瞬间转变。边缘端口是指连接计算机、打印机等终端设备的交换机端口,这类端口通常不会产生环路,没必要经过监听、学习等中间状态,可以直接进入转发状态。边缘端口的状态变化(up/down)也不会导致生成树协议重新计算,增加了网络的稳定性。但交换机无法自动识别端口是否直接与终端相连,所以需要人工配置。可在接口模式或全局模式下配置,配置方法如下。

```
ruijie(config) # interface f0/2
ruijie(config-if) # spanning-tree portfast           //将一个端口设置为边缘端口
ruijie(config) # spanning-tree portfast default     //将所有端口设置为边缘端口
```

配置过程中,系统会提示边缘端口只能连接终端主机,若连接集线器、交换机、网桥等设备会形成环路。

5.4 实践训练

实训 5 交换机生成树配置

1. 实训目标

- (1) 了解生成树协议的作用。
- (2) 熟悉生成树协议的配置。

2. 应用环境

交换机之间具有冗余链路本来是一件很好的事情,但是它有可能引起的问题比它能够解决的问题还要多。如果交换机之间有两条以上的链路,就必然形成了一个环路,交换机并不知道如何处理环路,只是周而复始地转发帧,形成一个“死循环”,这个“死循环”会造成整个网络处于阻塞状态,导致网络瘫痪。配置生成树协议可以避免环路。

3. 实训设备

- (1) 二层交换机 2 台。
- (2) PC 2 台。

- (3) Console 线 2 根。
- (4) 直通网线 4~8 根。

4. 实训拓扑

本实训生成树配置拓扑如图 5-19 所示。

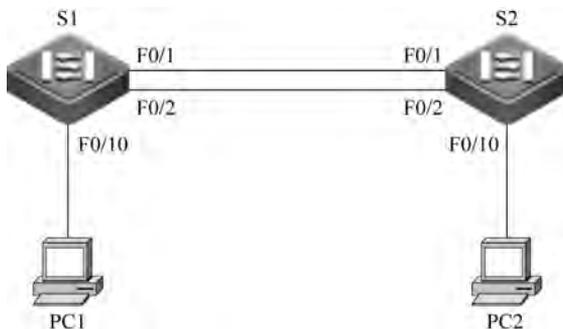


图 5-19 生成树配置拓扑

5. 实训要求

按照拓扑图连接网络设备时,先将两台交换机都配置快速生成树协议后,再将两台交换机连接起来。如果先连线再配置会造成广播风暴,影响交换机的正常工作。

配置生成树,使得 PC1 无论何时均可以 ping 通 PC2。PC 网络属性信息如表 5-7 所示。

表 5-7 PC 网络属性信息

设 备	IP 地址	子网掩码
PC1	192.168.1.2	255.255.255.0
PC2	192.168.1.3	255.255.255.0

6. 实训步骤

第 1 步:正确连接网线(注意交换机之间只连一根线),恢复出厂设置,重启交换机后,做初始配置。修改设备的系统名称(交换机 2 的配置与交换机 1 相似)。

```
Ruijie(config) # hostname S1
```

第 2 步:确认两台交换机的生成树协议关闭(两台交换机配置相同)。

```
S1(Config) # no spanning - tree //关闭 spanning - tree 协议
```

验证配置:

```
S1 # show spanning - tree
Global MSTP is disabled
```

第 3 步：通过 PC1 ping PC2-t 命令观察现象。

① ping 通。

② 两台交换机之间添加一根连线,持续观察 PC 互相 ping 的现象。

③ ping 不通(3~5min 后)。

④ 所有连接网线的端口的绿灯频繁地闪烁,表明端口收发数据量很大,已经在交换机内部形成广播风暴。

⑤ 使用 show cpu usage 命令观察两台交换机 CPU 使用率(观察到的值与以下显示的不一定相同)。

```
S1 # show cpu usage
Last 5 second CPU IDLE: 96 %
Last 30 second CPU IDLE: 96 %
Last 5 minute CPU IDLE: 97 %
From running CPU IDLE: 97 %
```

第 4 步：分别在两台交换机中启用生成树协议(两台交换机配置相同)。

```
S1(Config) # spanning - tree mode stp //打开 spanning - tree 协议
STP is starting now, please wait.....
STP is enabled successfully.
S1(Config) #
```

分别在两台交换机上验证配置,并根据输出结果划出生成树形态。

```
S1 # show spanning - tree
... //使用设备不同,此处显示不相同,省略显示内容
S2 # show spanning - tree
...
```

第 5 步：理解生成树形成过程中的制约要素。

① 改变根交换机的连接端口实现生成树拓扑形态的改变。此处可将连接根交换机的两个端口线对调一下,再次使用 show spanning-tree 命令查看生成树,并根据输出结果画出生成树形态。比较两次生成树形态的不同,分析理解生成树的形成过程。

② 改变根交换机端口的优先级实现生成树拓扑形态的改变。默认情况下,交换机端口的优先级都是 128,值越小优先级越高。注意,在端口优先级的修改中,系统支持以 16 为基数进行增减。命令如下:

```
S1(config) # interface f0/2
S1(config-if) # spanning - tree mst 0 port - priority 112 //提升端口 2 的优先级
```

再次使用 show spanning-tree 命令查看生成树,并根据输出结果画出生成树形态。比较两次生成树形态的不同,分析理解生成树的形成过程。

③ 改变非交换机的优先级实现生成树拓扑形态的改变。默认情况下,交换机的优先级都是 32 768,值越小优先级越高。假设 S2 是非根交换机,则将交换机 2 的优先级值升高或

降低交换机 1 的优先级。在设备优先级的修改中,系统支持以 4096 为基数进行增减。命令如下:

```
S2(config) # spanning-tree mst 0 priority 4096 //改变交换机优先级
```

再次使用 show spanning-tree 命令查看生成树,并根据输出结果画出生成树形态。比较两次生成树形态的不同,分析理解生成树的形成过程。

7. 课后实训

参照任务实施完成交换机多生成树配置。

5.5 习题

一、选择题

- 如果以太网交换机中某个运行 STP 的端口不接收或转发数据,接收并发送 BPDU,不进行地址学习,那么该端口应该处于()状态。
 - 阻塞
 - 监听
 - 学习
 - 转发
 - 等待
 - 禁用
- 如果以太网交换机中某个运行 STP 的端口接收并转发数据,接收、处理并发送 BPDU,进行地址学习,那么该端口应该处于()状态。
 - 阻塞
 - 监听
 - 学习
 - 转发
 - 等待
 - 禁用
- 关于 STP 说法正确的是()(选择一项或多项)。
 - 网桥 ID 值由网桥的优先级和网桥的 MAC 地址组合而成。前面是优先级,后面是 MAC 地址
 - 以太网交换机的默认优先级值是 32 768
 - 优先级值越小优先级越低
 - 优先级相同时,MAC 地址越小优先级越高
 - 网桥 ID 值大的将被选为根网桥
- 下列关于 STP、RSTP 和 MSTP 说法正确的是()(选择一项或多项)。
 - MSTP 兼容 STP 和 RSTP
 - STP 不能快速收敛,当网络拓扑结构发生变化时,原来阻塞的端口需要等待一段时间才能变为转发状态
 - RSTP 是 STP 的优化版。端口进入转发状态的延迟在某些条件下大大缩短,从而缩短了网络最终达到拓扑稳定所需要的时间
 - MSTP 可以弥补 STP 和 RSTP 的缺陷,它既能快速收敛,也能使不同 VLAN 的流量沿各自的路径转发,从而为冗余链路提供了更好的负载分担机制
- MSTP 的特点有()(选择一项或多项)。
 - MSTP 兼容 STP 和 RSTP

- B. 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,生成树间彼此独立
- C. MSTP 将环路网络修剪成为一个无环的树形网络,避免报文在环路网络中的增生和无限循环,同时还可以提供数据转发的冗余路径,在数据转发过程中实现 VLAN 数据的负载均衡
- D. 以上说法均不正确

二、简答题

1. 环路给网络带来哪些危害?
2. 生成树的作用是什么?
3. 简述 STP 的工作过程。