

STP（生成树协议）_百度百科

（生成树协议）

编辑^[1] 锁定^[2]

本词条由“科普中国”百科科学词条编写与应用工作项目^[3] 审核。

STP（Spanning Tree Protocol）是生成树协议^[4]的英文缩写。该协议可应用于在网络中建立树形拓扑，消除网络中的**环路**，并且可以通过一定的方法实现路径冗余^[5]，但不是一定可以实现路径冗余。生成树协议适合所有厂商的网络设备^[6]，在配置上和体现功能强度上有所差别，但是在原理和应用效果是一致的。

中文名
生成树协议^[7]
外文名
全 称
Spanning Tree Protocol
适合对象
适合所有厂商的网络设备^[8]

编辑^[9]

STP的基本原理是，通过在交换机^[10]之间传递一种特殊的协议报文^[11]，网桥协议数据单元^[12]（Bridge Protocol Data Unit，简称BPDU），来确定网络的拓扑结构^[13]。BPDU有两种，配置BPDU（Configuration BPDU）和TCN^[14] BPDU。前者是用于计算无环的生成树^[15]的，后者则是用于在二层网络拓扑发生变化时产生用来缩短MAC表项的刷新时间的（由默认的300s缩短为15s）。

Spanning Tree Protocol(STP)在IEEE802.1D文档中定义。该协议的原理是按照树的结构来构造网络拓扑，消除网络中的环路，避免由于环路的存在而造成广播风暴^[16]问题。

Spanning Tree Protocol(STP)的基本思想就是按照"树"的结构构造网络的拓扑结构^[17]，树的根是一个称为根桥的桥设备，根桥的确立是由交换机^[18]或网桥的BID（Bridge ID）确定的，BID最小的设备成为二层网络中的根桥。BID又是由网桥优先级和MAC地址构成，不同厂商的设备的网桥优先级的字节个数可能不同。由根桥开始，逐级形成一棵树，根桥定时发送配置BPDU，非根桥接收配置BPDU，刷新最佳BPDU并转发。这里的最佳BPDU指的是当前根桥所发送的BPDU。如果接收到了下级BPDU（新接入的设备会发送BPDU，但该设备的BID比当前根桥大），接收到该下级BPDU的设备将会向新接入的设备发送自己存储的最佳BPDU，以告知其当前网络中根桥；如果接收到的BPDU更优，将会重新计算生成树^[19]拓扑。当非根桥在离上一次接收到最佳BPDU最长寿命（Max Age，默认20s）后还没有接收到最佳BPDU的时候，该端口将进入监听状态，该设备将产生TCN BPDU，并从根端口转发出去，从指定端口接收到TCN BPDU的上级设备将发送确认，然后再向上级设备发送TCN BPDU，此过程持续到根桥为止，然后根桥在其后发送的配置BPDU中将携带标记表明拓扑已发生变化，网络中的所有设备接收到后将CAM表项的刷新时间从300s缩短为15s。整个收敛的时间为50s左右。

编辑^[20]

生成树协议^[21]最主要的应用是为了避免局域网中的单点故障、网络回环，解决成环以太网^[22]网络的“广播风暴^[23]”问题，从某种意义上说是一种网络保护技术，可以消除由于失误或者意外带来的循环连接。STP也提供了为网络提供备份^[24]连接的可能，可与SDH保护配合构成以太环网的双重保护。新型以太单板支持符合IEEE 802.1d标准的生成树协议STP及IEEE 802.1w规定的快速生成树^[25]协议RSTP，收敛速度可达到1s。

但是，由于协议机制本身的局限，STP保护速度慢（即使是1s的收敛速度也无法满足电信级的要求），如果在城域网^[26]内部运用STP技术，用户网络的动荡会引起运营商网络的动荡。目前在MSTP 组成环网中，由于SDH保护倒换时间比STP协议收敛时间^[27]快的多，系统采用依然是SDH MS-SPRING或SNCP^[28]，一般倒换时间在50ms以内。但测试时部分以太网^[29]业务的倒换时间为0或小于几个毫秒，原因是内部具有较大缓存。SDH保护倒换动作对MAC层是不可见的。这两个层次的保护可以协调工作，设置一定的“拖延时间”（hold-off），一般不会出现多次倒换问题。

VLAN影响

编辑^[30]

L3, L4交换已经非常成熟。Internet中也越来越广泛地应用了交换技术，全交换网络已经非常普遍。在这些网络中，VLAN的使用是必不可少的。

VLAN是一个根据作用、计划组、应用等进行逻辑划分的交换式网络。与用户的物理位置没有关系。举个例子来说，几个终端可能被组成一个部分，可能包括工程师或财务人员。当终端的实际物理位置比较相近，可以组成一个局域网（LAN）。如果他们在不同的建筑物中，就可以通过VLAN将他们聚合在一起。同一个VLAN中的端口可以接受VLAN中的广播包^[31]。但别的VLAN中的端口却接受不到。

- 1、网络容错能力不强；
- 2、报文在环路网络中容易增生和无限循环；
- 3、不利在VLAN中实现流量负载均衡^[32]

生成树协议^[33]运行生成树算法(STA).生成树^[34]算法很复杂，但是其过程可以归纳为以下3个步骤：

- （1）选择根网桥
- （2）选择根端口
- （3）选择指定端口

关于选择根网桥：选择根网桥的依据是网桥ID，网桥ID由网桥优先级和网桥MAC地址组成。网桥的默认优先级是32768。使用show mac-address-table时，显示在最前面的MAC地址就是计算时所使用的MAC地址。网桥ID值小的为根网桥，当优先级相同时，MAC地址小的为根网桥。

关于选择根端口：每个非根交换机^[35]选择一个根端口。选择顺序为：到根网桥最低的根路径成本→发送BPDU的网桥ID较小→端口ID较小的。端口ID由端口优先级与端口编号组成。默认的端口优先级为128。

关于选择指定端口：每个网段上选择一个指定端口。选择顺序为：根路径成本较低→发送BPDU的交换机的网桥ID值较小→本端口的ID值较小。另外，根网桥的接口皆为指定端口，因为根网桥上端口的根路径成本为0。

编辑^[36]

- 1、拓扑收敛慢，当网络拓扑发生改变的时候，生成树协议需要50-52秒的时间才能完成拓扑收敛。

2、不能提供负载均衡^[37]的功能。当网络中出现环路的时候，生成树协议^[38]简单的将环路进行Block，这样该链路就不能进行数据包^[39]的转发，浪费网络资源。

编辑^[40]

生成树协议运行生成树算法（STA）。生成树算法很复杂，但是其过程可以归纳为以下三个部分。

- （1）选择根网桥
- （2）选择根端口
- （3）选择指定端口（也有书籍称为转发端口）

选择根网桥的依据是交换机^[41]的网桥优先级，网桥优先级是用来衡量网桥在生成树^[42]算法中优先级的十进制数，取值范围是0~65535.默认值是32768，网桥ID=网桥优先级+网桥MAC地址组成的，共有8个字节。由于交换机的网桥优先级都是默认，所以在根网桥的选举中比较的一般是网卡MAC地址的大小，选取MAC地址小的为根网桥。

根端口依据

编辑^[43]

- （1）到根网桥路径开销^[44]最低。

根路径开销是两个网桥间的路径上所有链路的开销之和，也就是某个桥网到达根网桥的中间所有链路的路径开销之和。

附IEEE标准路径开销表：

开销（最新修订）	开销（以前）
10Gbps	
1Gbps	
100Mbps	
10Mbps	

- （2）最低的发送方网桥ID。
- （3）最低的端口ID。

由于端口的ID是由优先级和端口号^[45]组成，保证了根端口的唯一性。

指定端口依据

编辑^[46]

- （1）根路径开销^[47]较低
- （2）所在的交换机网桥ID值最小
- （3）端口号最小

注：根桥上的所有端口都是指定端口

编辑^[48] 五种状态

Blocking（阻塞状态）：此时，二层端口为非指定端口，也不会参与数据帧的转发。该端口通过接收BPDU来判断根交换机的位置和根ID，以及在STP拓扑收敛结束之后，各交换机端口应该处于什么状态，在默认情况下，端口会在这种状态下停留20秒钟时间。

Listening（侦听状态）：生成树此时已经根据交换机所接收到的BPDU而判断出了这个端口应该参与数据帧的转发。于是交换机端口将不再满足于接收BPDU，而同时也开始发送自己的BPDU，并以此通告邻接的交换机该端口会在活动拓扑中参与转发数据帧的工作。在默认情况下，该端口会在这种状态下停留15秒钟的时间。

Learning(学习状态)：这个二层端口准备参与数据帧的转发，并开始填写MAC表。在默认情况下，端口会在这种状态下停留15秒钟时间。

Forwarding（转发状态）：这个二层端口已经成为了活动拓扑的一个组成部分，它会转发数据帧，并同时收发BPDU。

Disabled（禁用状态）：这个二层端口不会参与生成树，也不会转发数据帧。

STP潜在故障

编辑^[49]

双工不匹配：在点到点链路上，双工不匹配是一种常见的配置错误。当链路的一段采用手工的方式配置为了全双工模式，而另一侧却使用自动协商的默认配置的时候，那么就可能发生双工不匹配的情况。

单向链路失效：单向链路是产生桥接环路的一个非常常见的原因。如果光纤链路存在没有检测出来的故障或收发器故障，通常导致单向链路。在启用STP来提供网络冗余的情况下，对于两个链路伙伴之间所连接的物理链路，如果由于某种原因导致这条链路工作在单向通信的状态下，那么就可能导致桥接环路或路由选择黑洞，进而对维护网络稳定非常有害。

帧破坏：帧破坏是导致STP故障的另外一种原因。如果接口正在经受高速的物理错误，其结果有可能就会导致BPDU丢失，而这会使处于阻塞状态的接口过渡到转发状态。虽然如此，但因为STP默认参数是非常保守的，所以很少会发生上述情况。

资源错误：即使在通过专门的ASIC硬件执行大部分交换功能的高端交换机中，STP仍然由CPU来执行。这就意味着：如果处于某种原因而过度使用了网桥的CPU，那么就可能导致CPU没有足够的资源来发出BPDU。通常情况下，STP不是一种密集调用处理器的应用，而且STP的优先级高于其他进程。因此，出现资源问题的可能性并不大。

PortFast配置错误：如果管理员在端口上启用了PortFast特性，那么当链路启动的时候，端口就会绕过STP的监听和学习状态，并且直接过渡到转发状态。但是如果在错误的端口上配置了PortFast特性，那么这种快速过渡就有可能导致桥接环路。

中国通信学会^[50]，

Links

1. javascript;;
2. <http://baike.baidu.com/view/10812319.htm>
3. <http://www.cast.org.cn/n35081/>