

# 了解快速生成树协议(802.1w)

## 目录

[简介](#)

[背景](#)

[Catalyst 交换机对 RSTP 的支持](#)

[新的端口状态和端口角色](#)

[端口状态](#)

[端口角色](#)

[新的 BPDU 格式](#)

[思科BPDU、IEEE BPDU和BPDU图的完整视图](#)

[新的BPDU 处理](#)

[每次Hello时发送BPDU](#)

[信息更快过期](#)

[接受次要 BPDU](#)

[快速转换到转发状态](#)

[边缘端口](#)

[链路类型](#)

[与802.1D的融合](#)

[与 802.1w 的收敛](#)

[建议/协议顺序](#)

[UplinkFast](#)

[新的拓扑结构更改机制](#)

[拓扑更改检测](#)

[拓扑更改传播](#)

[与802.1D兼容](#)

[结论](#)

[相关信息](#)

## 简介

本文档介绍有关RSTP添加到之前802.1D标准的增强功能的信息。

## 背景

802.1D 生成树协议 (STP) 标准的设计初衷是，每次连接都在中断一分钟左右后恢复可视为性能良好。随着LAN环境中第3层交换的出现，网桥现在与路由解决方案展开竞争，在路由解决方案中，开放最短路径优先(OSPF)和增强型内部网关路由协议(EIGRP)等协议能够在更短的时间内提供备用路径。

思科通过[Uplink Fast](#)、[Backbone Fast](#)和Port Fast等功能增强了原始802.1D规范，以加快桥接网络的收敛时间。缺点是这些机制是专有的并且需要额外配置。

快速生成树协议(RSTP;IEEE 802.1w)可看作802.1D标准的演进，而不是一场革命。802.1D 中的术

语基本上保持不变。而且大部分参数也没有变化，因此熟悉 802.1D 的用户可以轻松地快速配置新协议。在大多数情况下，RSTP 的性能比 Cisco 专有扩展的性能要好，而且不需要进行任何额外的配置。802.1w也可以恢复为802.1D，以便与旧网桥进行互操作（基于每个端口）。但这恰恰抛弃了它带来的好处。

新版 802.1D 标准 (IEEE 802.1D-2004) 收敛了 IEEE 802.1t-2001 和 IEEE 802.1w 标准。

## Catalyst 交换机对 RSTP 的支持

下表显示了一些Catalyst交换机系列中对RSTP的支持，以及支持该支持所需的最低软件。

Catalyst 平台	带RSTP的MST	RPVST+ (亦称 PVRST+)
Catalyst 2900 XL / 3500 XL	不可用。	不可用。
Catalyst 2940	12.1(20)EA2	12.1(20)EA2
Catalyst 2950/2955/3550	12.1(9)	12.1(13)EA1
Catalyst 2970/3750	12.1(14)EA1	12.1(14)EA1
Catalyst 3560	12.1(19)EA1	12.1(19)EA1
Catalyst 3750 Metro	12.1(14)AX	12.1(14)AX
Catalyst 2948G-L3/4908G-L3	不可用。	不可用。
Cisco IOS®)	12.1(12c)EW	12.1(19)EW
Catalyst 6000/6500(Cisco IOS)	12.1(11b)EX、12.1(13)E 和 12.2(14)SX	12.1(13)E
Catalyst 8500	不可用。	不可用。

## 新的端口状态和端口角色

802.1D在以下五种不同的端口状态下定义：

- 禁用
- 侦听
- 学习
- 阻塞
- 转发

有关端口状态的详细信息，请参阅本文档[端口状态](#)部分中的表。

端口的状态为混合状态（无论是阻塞还是转发流量），它在活动拓扑中的作用也是如此（根端口、指定端口等等）。例如，从运行角度看，阻塞状态的端口和监听状态的端口没有任何差异。这两种状态都丢弃帧，并且不能学习 MAC 地址。真正的区别在于生成树为端口分配的角色。可以安全地假设：监听端口为指定端口或根端口，并且正在进入转发状态。然而，在处于转发状态后，无法根据端口状态推断端口是根端口还是指定端口。这显示了该基于状态的术语的失败之处。RSTP 将端口的角色和状态分离，从而解决了此问题。

## 端口状态

RSTP 中仅保留了三种端口状态，分别对应于三种可能的运行状态。802.1D 中的禁用、阻塞和监听状态在 802.1w 中合并为唯一的丢弃状态。

STP (802.1D) 端口状态	RSTP (802.1w) 端口状态	端口是否包括在活动拓扑中？	端口是否可获知 MAC 地址
禁用	丢弃	无	无
阻塞	丢弃	无	无
侦听	丢弃	Yes	无
学习	学习	Yes	Yes
转发	转发	Yes	Yes

## 端口角色

端口角色现在是分配给给定端口的变量。根端口和指定端口的角色仍然保留，而阻塞端口的角色拆分为备份端口和替代端口角色。生成树算法(STA)根据网桥协议数据单元(BPDU)确定端口的角色。为了简化问题，关于BPDU，需要记住的一点是，总是有一种方法可以比较其中的任意两个，然后确定其中哪一个比另一个更有用。这是依据 BPDU 中存储的值，而有时依据接收它们的端口。因此，本部分中的信息阐述了确定端口角色的实用方法。

### 根端口角色

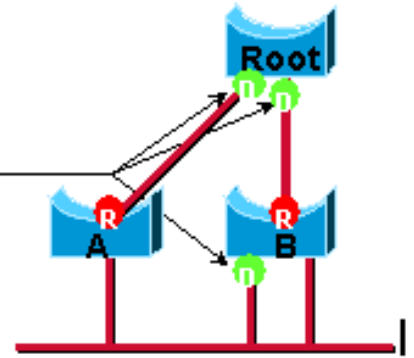
- 网桥上接收最佳 BPDU 的端口即是根端口。在路径开销方面，这是最接近根网桥的端口。STA 在整个桥接网络 ( 每个 VLAN ) 中选择一个根网桥。根网桥发送的 BPDU 比任何其他网桥发送的 BPDU 都更有用。根网桥是该网络中唯一没有根端口的网桥。所有其他网桥在至少一个端口上接收 BPDU。



### 指定端口角色

- 如果端口可以在它所连接到的网段上发送最佳 BPDU，则它是指定端口。802.1D 网桥将不同网段 ( 例如以太网段 ) 链接在一起，以创建桥接域。在给定的网段上，只能有一条路径通往根网桥。如果存在两条路径，则网络中存在一个网桥环路。连接到给定网段的所有网桥将监听每个网桥的 BPDU，并且同意将发送最佳 BPDU 的网桥作为网段的指定网桥。网桥上对应的端口是该网段的指定端口。

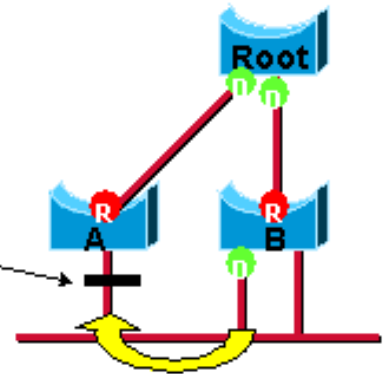
## ■ Designated Port



### 替代端口和备份端口角色

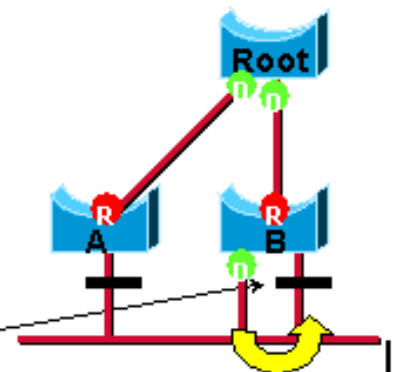
- 这两个端口角色对应于 802.1D 的阻塞状态。阻塞的端口既不是指定端口也不是根端口。阻塞的端口收到的 BPDU 比其在网段上发送的 BPDU 更有用。请记住，端口需要接收 BPDU 才能保持阻塞状态。为此，RSTP 引入了这两个角色。
- 替代端口从其他网桥接收更有用的 BPDU，并且它是阻塞的端口。如下图所示：

## — Alternate Port



- 备份端口从其所在的网桥接收更有用的 BPDU，并且它是阻塞的端口。如下图所示：

## — Backup Port



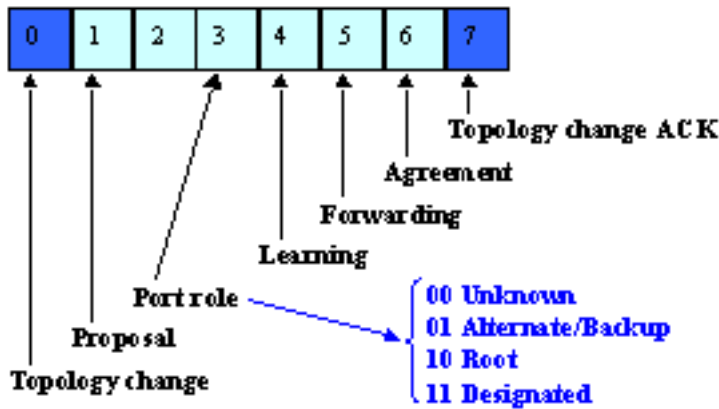
802.1D 内部已对此进行区分。这实质上是 Cisco UplinkFast 的工作原理。基本原理是替代端口提供通往根网桥的替代路径，因此可以在根端口出现故障时替换它。当然，备份端口可提供到同一网段的冗余连接，但不能保证到根网桥的替代连接。因此，上行链路组将其排除在外。

由此，RSTP 针对与 802.1D 使用相同标准的生成树计算最终拓扑。使用不同网桥和端口优先级的方式没有变化。“阻塞”这一名称用于描述 Cisco 实施中的丢弃状态。CatOS 版本 7.1 及更高版本仍会显示监听和学习状态。这提供的端口信息比 IEEE 标准要求的多。但是，新功能体现在，协议确定的端口角色与其当前状态之间存在差异。例如，它现在对于同时指定和阻塞的端口有效。虽然这种情况通常发生很短一段时间，但这意味着该端口处于面向指定转发状态的过渡状态。

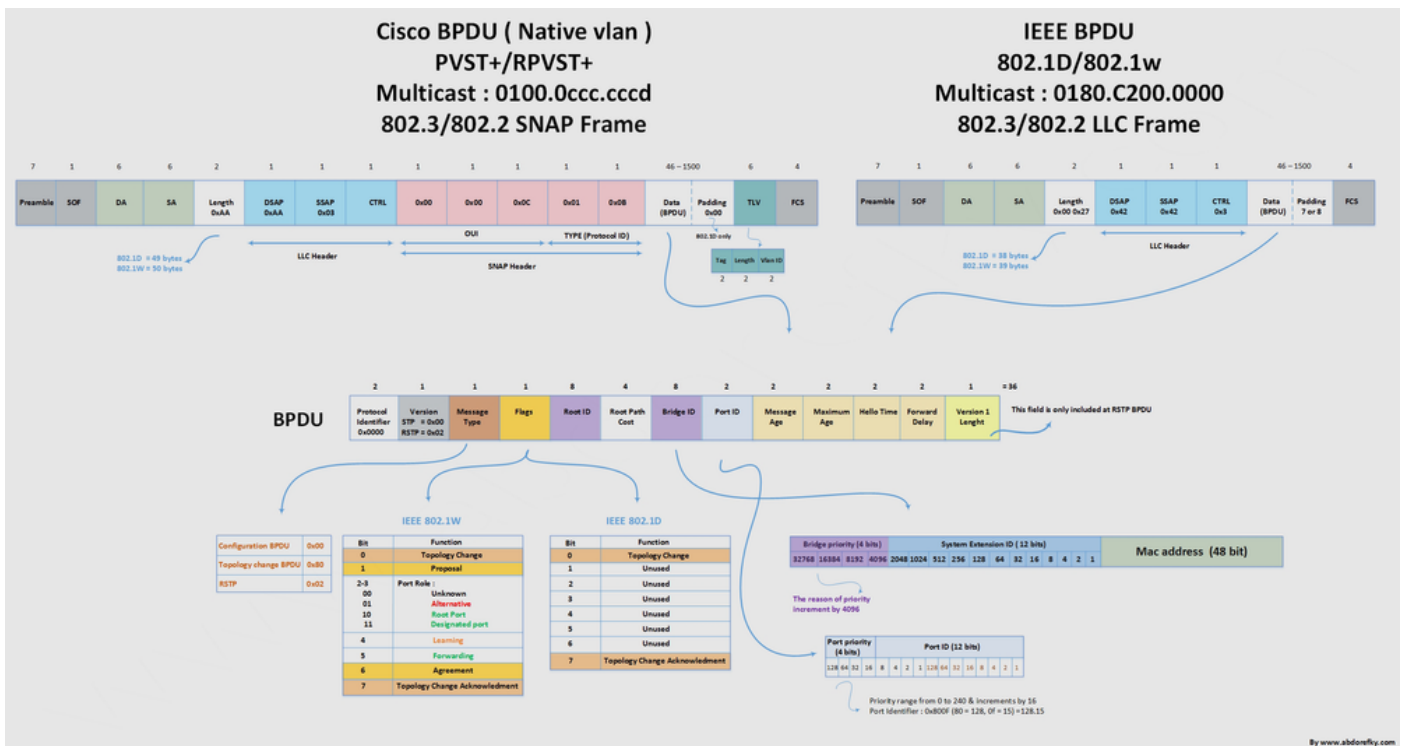
## 新的 BPDU 格式

RSTP 引入了少量对 BPDU 格式的更改。802.1D 中仅定义了两个标志，即拓扑变化 (TC) 和 TC 确认 (TCA)。但是，RSTP 会使用剩余的六个标志字节位来执行：

- 对产生 BPDU 的端口的角色和状态进行编码
- 处理建议/协议机制



### 思科BPDU、IEEE BPDU和BPDU图的完整视图



有关更高分辨率的图像，请参阅[Cisco BPDU、IEEE BPDU和BPDU图](#)。

注：位0（拓扑更改）是最低有效位。

另一项重要更改是 RSTP BPDU 现在为类型 2，版本 2。这表示传统网桥必须丢弃此新的 BPDU。此属性使802.1w网桥可以轻松检测与其连接的旧网桥。

### 新的BPDU 处理

## 每次Hello时发送BPDU

BPDU每次hello时发送一次，不再简单地中继。使用 802.1D 时，仅当非根网桥在根端口上收到 BPDU 时，才会生成 BPDU。网桥对BPDU的中继比它实际生成的更多。802.1w 并非如此。网桥每 <hello-time>秒（默认为2秒）发送一次包含其当前信息的BPDU，即使它没有从根网桥收到任何信息。

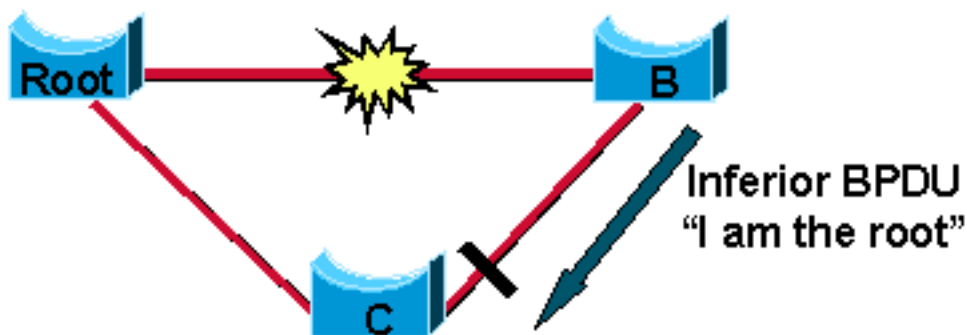
## 信息更快过期

在给定端口上，如果连续三次没有收到 Hello（或如果 max\_age 过期），则协议信息会立即过期。由于前面所述的协议修改，BPDU 现在可用作网桥之间的保持活动机制。如果网桥连续错过三个 BPDU，则它会认为与其直接相邻根网桥或指定网桥断开连接。这种信息快速过期使得可以进行快速故障检测。如果网桥无法从相邻网桥接收 BPDU，可以肯定与该相邻网桥的连接已断开。这与 802.1D相反，后者的问题可能发生在通往根桥的路径上的任何地方。

**注意：**在物理链路发生故障时，故障检测速度更快。

## 接受次要 BPDU

此概念是 BackboneFast 引擎的核心部分。IEEE 802.1w委员会在RSTP中引入了类似的机制。当网桥从其指定网桥或根网桥接收次要信息时，它会立即接收并替换以前存储的信息。



由于网桥 C 仍然识别出根网桥处于活动状态并且运行良好，因此立即发送 BPDU 到网桥 B，该 BPDU 中包含有关根网桥的信息。因此，网桥 B 不会发送其自己的 BPDU 并且接受导向网桥 C 的端口作为新的根端口。

## 快速转换到转发状态

快速转换是 802.1w 中引入的最重要功能。传统 STA 在将端口变成转发状态之前，会被动地等待网络收敛。实现更快的收敛是对保守默认参数（转发延迟和max\_age计时器）进行更改的问题，通常会使网络的稳定性受到威胁。新的快速STP能够主动确认端口可以安全转换到转发状态，而无需依赖任何计时器配置。现在，符合 RSTP 的网桥之间存在真正的反馈机制。为了在端口上实现快速收敛，协议依赖于两个新变量：边缘端口和链路类型。

## 边缘端口

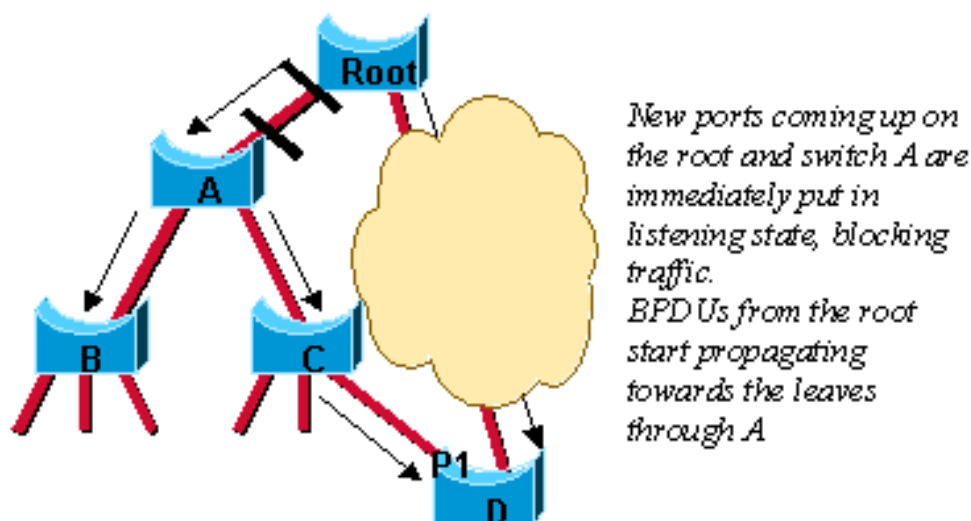
Cisco生成树用户已经知道边缘端口概念，因为它基本上对应于PortFast功能。直接连接到终端站的所有端口无法在网络中形成网桥环路。因此，边缘端口直接转换到转发状态，并且跳过监听和学习阶段。当链路切换时，边缘端口和启用了Portfast的端口都不会生成拓扑更改。接收BPDU的边缘端口会立即失去边缘端口状态并成为正常的生成树端口。此时，边缘端口状态有一个用户配置的值和一个运行值。Cisco实施仍保留PortFast关键字用于边缘端口配置。这使转换为RSTP变得更简单。

## 链路类型

只有在边缘端口和点对点链路上，RSTP才能快速转换到转发状态。链路类型会从端口的双工模式自动派生。以全双工模式运行的端口视为点对点端口，而半双工端口默认视为共享端口。此自动链路类型值可由显式配置覆盖。在当今的交换网络中，大多数链路都在全双工模式下运行，RSTP将这些链路视为点对点链路。这使它们能够快速转换到转发状态。

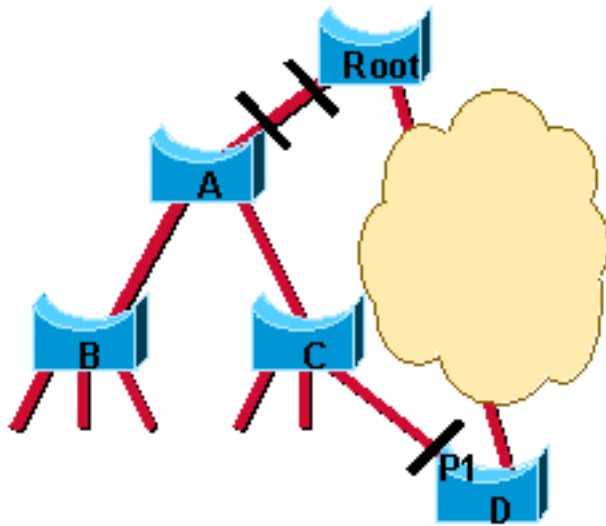
## 与802.1D的融合

下图说明了802.1D处理添加到桥接网络的新链路的方式：



在此方案中，根网桥和网桥 A 之间添加了一个链路。假设网桥A和根网桥之间已经有间接连接（图中的C - D）。STA阻塞端口并禁用网桥环路。首先，当根网桥和网桥 A 启动时，它们之间的链路上的两个端口会处于监听状态。网桥 A 现在能直接监听根网桥。它会立即在指定端口上向树的分支传播其 BPDU。一旦网桥B和C收到来自网桥A的这个新的上级信息，它们立即将该信息转发到枝叶。几秒钟后，网桥 D 便收到来自根网桥的 BPDU，并立刻阻塞端口 P1。





*Very quickly, the BPDUs from the root reach D that immediately blocks its port P1. The topology has now converged, though the network is disrupted for twice forward\_delay.*

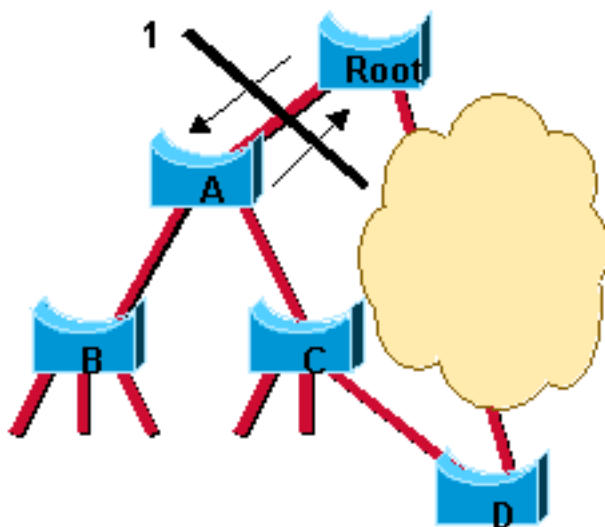
|

生成树可以非常高效地计算网络的新拓扑。现在唯一的问题是，只有经过两倍的转发中继时间之后，根网桥和网桥 A 之间的链路才能最终进入转发状态。这意味着流量将中断 30 秒（网络的整个 A、B 和 C 部分被隔离），原因是 802.1D 算法缺乏反馈机制，无法清晰地通告网络将在几秒钟后收敛。

### 与 802.1w 的收敛

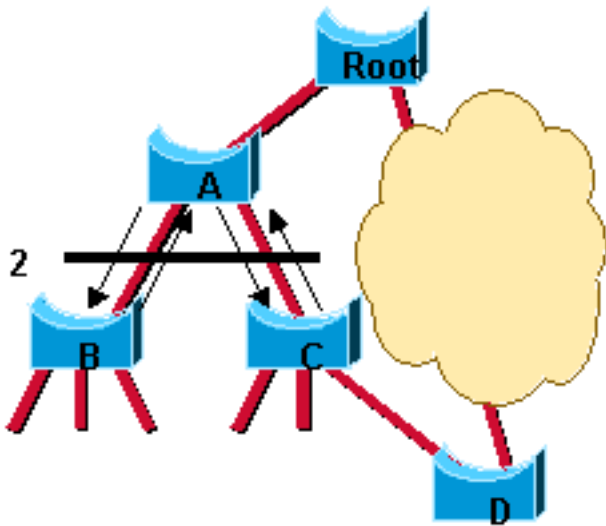
现在，您可以看到 RSTP 处理类似情况的方式。请记住，最终拓扑与 802.1D 计算的拓扑完全一样（即，与以前相同的位置上有一个阻塞的端口）。只是实现此拓扑的步骤发生了变化。

交换机 A 和根之间的链路上的两个端口一进入就被置于指定阻塞状态。至此，一切情况与纯 802.1D 环境中完全一样。但是，在此阶段，交换机 A 和根网桥之间会发生协商。交换机 A 收到根的 BPDU 后，便会阻塞非边缘指定端口。此操作称为同步。完成此操作后，网桥 A 会明确授权根网桥将其端口置于转发状态。下图说明了网络上该过程的结果。交换机 A 和根网桥之间的链路被阻塞，两个网桥可交换 BPDU。

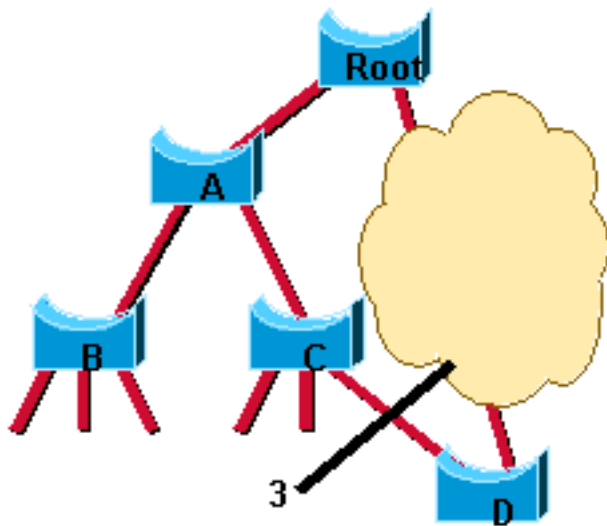


交换机 A 阻塞其非边缘指定端口后，交换机 A 和根网桥之间的链路便进入转发状态，您将遇到以下情况：





仍然不能有环路。现在，网络在交换机A之后阻塞，而不是在交换机A之前阻塞。但是，潜在网桥环路是在另一个位置被切断的。这种切断与根网桥产生的新BPDU一起通过交换机A沿树向下传递。在此阶段，交换机A上新阻塞的端口还会与都启动同步操作的交换机B和交换机C上与其相邻的端口进行协商，从而快速转换到转发状态。除了指向A的根端口之外，交换机B只有边缘指定端口。因此，它没有要阻塞的端口来授权交换机A进入转发状态。同样，交换机C只需将其指定端口阻塞为D。现在可达到下图所示的状态：



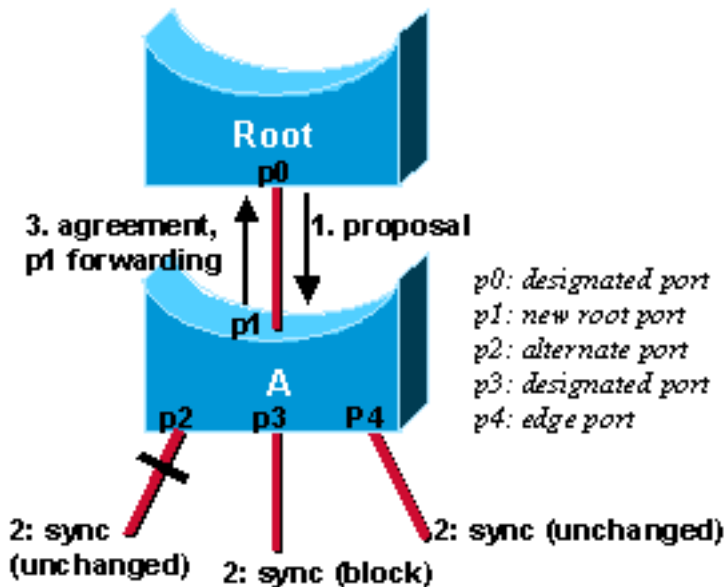
请记住，最终拓扑与802.1D示例完全相同，这表示D上的端口P1最终会进入阻塞状态。这表示在新的BPDU沿树向下传递所需的必要时间后，将实现最终网络拓扑。在这种快速收敛中没有涉及计时器。RSTP唯一引入的新机制是确认交换机可以在其新的根端口上进行发送，以便授权立即转换到转发状态，从而绕过长达两倍转发延迟时间的监听和学习阶段。管理员只需要记住以下几点就可以受益于快速收敛：

- 只有当网桥通过点对点链路（即全双工链路，除非显式端口配置）连接时，网桥之间的这种协商才可能实现。
- 在802.1D中端口上启用Portfast后，边缘端口更加重要。例如，如果网络管理员未能正确配置交换机B上的边缘端口，则交换机A与进入网络的根之间的链路会影响它们的连接。

## 建议/协议顺序

当STA选择某个端口作为指定端口时，802.1D仍会等待两次<forward delay>秒（默认情况下为2 x 15），然后才能将其转换为转发状态。在RSTP中，此情况与具有指定角色但处于阻塞状态的端口

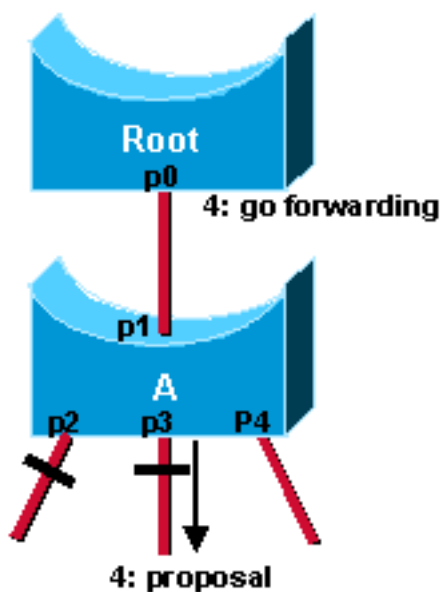
相对应。下图说明了如何逐步实现快速转换。假设在根网桥和交换机 A 之间创建了一条新链路。此链路上的两个端口都置于指定的阻塞状态，直到它们从对等的端口收到 BPDU 为止。



当指定端口处于丢弃或学习状态时（并且只有在这种情况下），它才会在其发送的 BPDU 中设置建议位。这是根网桥的端口 p0 所发生的情况，如上图的步骤 1 所示。由于交换机 A 收到高级信息，它会立即识别出 p1 是新的根端口。然后，交换机 A 开始同步，以确保其所有端口均与此新信息同步。如果端口满足以下标准之一，则端口是同步的：

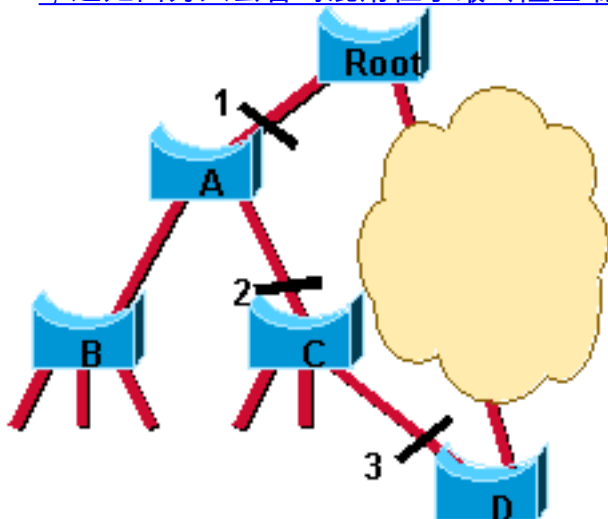
- 端口处于阻塞状态，意味着稳定拓扑中的丢弃状态。
- 端口是边缘端口。

为了说明同步机制对不同端口的影响，假设交换机 A 上存在替代端口 p2、指定转发端口 p3 和边缘端口 p4。请注意 p2 和 p4 已经满足其中一个条件。为了保持同步（参见上图的步骤 2），交换机 A 只需阻塞端口 p3，并为其分配丢弃状态。既然其所有的端口都保持同步，交换机 A 就可以取消阻塞其新选择的根端口 p1 并发送协议信息以回复根网桥。（请参阅步骤 3。）此消息是建议 BPDU 的副本，其中设置了协议位而非建议位。这可确保端口 p0 确切地知道它接收的协议与哪个建议对应。



p0 收到该协议后，它可立即转换到转发状态。这是上图的步骤 4。请注意，同步之后，端口 p3 将继续保持指定的丢弃状态。在步骤 4 中，该端口所处的情况与步骤 1 中端口 p0 所处的情况完全相同。然后，它会开始向其相邻端口提出建议，并且尝试快速转换到转发状态。

- 由于建议协议机制不依赖任何计时器，因此其速度非常快。此握手浪潮会快速传播到网络边缘，并在拓扑发生变化后快速恢复连接。
- 如果指定的丢弃端口在发送建议之后未收到协议，它将缓慢转换到转发状态，并退回到传统的 802.1D 监听学习顺序。如果远程网桥不了解 RSTP BPDU，或者如果远程网桥的端口处于阻塞状态，可能会发生这种情况。
- Cisco 在同步机制中引入了增强功能，允许网桥在进行同步时只将其以前的根端口置于丢弃状态。有关该机制的工作原理的详细信息不属于本文档的讨论范畴。但是，可以安全地假设，该机制在大多数常见再收敛案例中调用。本文档的[与 802.1w 的收敛部分中描述的方案非常有效，这是因为只会暂时混淆位于最终阻塞端口的路径上的端口。](#)



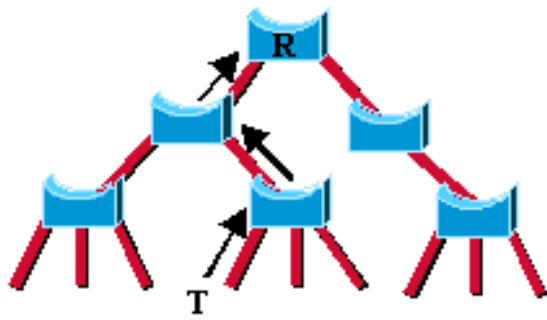
## UplinkFast

RSTP 中包含的快速转换到转发状态的另一种形式类似于 Cisco 的 UplinkFast 专有生成树扩展。基本上，当网桥丢失其根端口时，它可以将其最佳备用端口直接置于转发模式（新根端口的外观也由 RSTP 处理）。选择替代端口作为新的根端口会产生拓扑更改。802.1w 拓扑更改机制会清除上游网桥的内容可寻址存储器 (CAM) 表中的相应条目。这就消除了对 UplinkFast 的虚拟多播生成进程的需要。

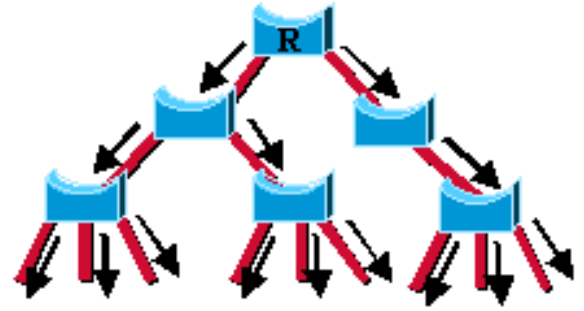
因为 RSTP 中自带 UplinkFast 机制并会自动启用，因此无需对该机制进行进一步配置。

## 新的拓扑结构更改机制

当 802.1D 网桥检测到拓扑更改时，它会使用可靠的机制首先通知根网桥。如下图所示：



A topology change is generated on point T.  
1<sup>st</sup> step: A TCN is going up to the root.



2<sup>nd</sup> step: the root advertises the TC for max-age+ forward delay.

一旦根网桥知道网络拓扑发生了变化，它就会在其发出的BPDU上设置TC标志，然后中继到网络中的所有网桥。当网桥收到设置了TC标志位的BPDU时，它会将其网桥表老化时间缩短到转发延迟秒数。这可确保相对快速地刷新过期信息。此拓扑更改机制在 RSTP 中进行了很高层次的再造。拓扑更改的检测及其通过网络的传播都在不断演变。

## 拓扑更改检测

在 RSTP 中，只有进入转发状态的非边缘端口会导致拓扑更改。这意味着连接断开将不再视为拓扑更改，这与 802.1D 正好相反（即，进入阻塞状态的端口不再生成 TC）。当 RSTP 网桥检测到拓扑更改时，会发生以下操作：

- 如有必要，它将其所有非边缘指定端口和根端口启动 TC While 计时器，计时器的值为 hello-time 值的两倍。
- 它将刷新与所有这些端口关联的 MAC 地址。

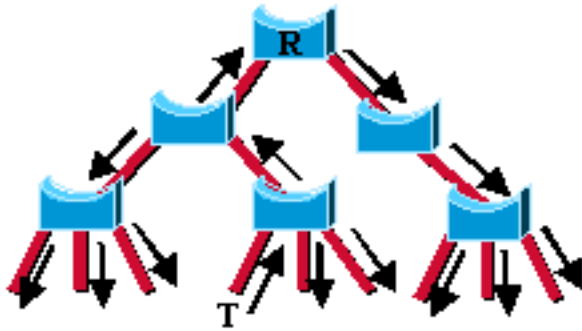
注：只要 TC While 计时器在端口上运行，从该端口发送的 BPDU 就会设置 TC 位。当计时器处于活动状态时，根端口上也会发送 BPDU。

## 拓扑更改传播

当网桥从相邻网桥收到设置了 TC 位的 BPDU 时，会发生以下操作：

- 它将清除其所有端口（接收拓扑更改的端口除外）上识别的 MAC 地址。
- 它启动 TC While 计时器，并发送在所有指定端口和根端口上设置了 TC 位的 BPDU（RSTP 不再使用特定的 TCN BPDU，除非传统网桥需要通知）。

这样，TCN 会快速传播到整个网络。TC 传播现在是一个单步过程。事实上，拓扑更改的发起方会将此信息泛洪到整个网络，而 802.1D 则只有根网桥会泛洪此信息。此机制比 802.1D 等效机制更快。无需等待通知根网桥，然后针对整个网络保持拓扑更改状态长达  $\langle \text{max age plus forward delay} \rangle$  秒钟。



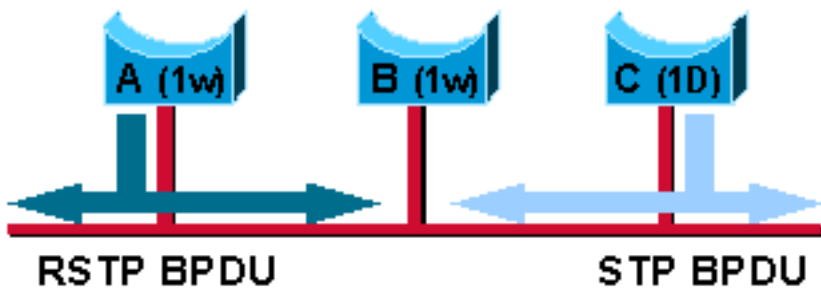
The originator of the TC directly floods this information through the network

仅在几秒钟或者几倍 hello-time 后，整个网络 (VLAN) 的 CAM 表中的大部分条目都将刷新。此方法会导致更多潜在的临时泛洪，但另一方面它清除了阻止快速恢复连接的潜在过期信息。

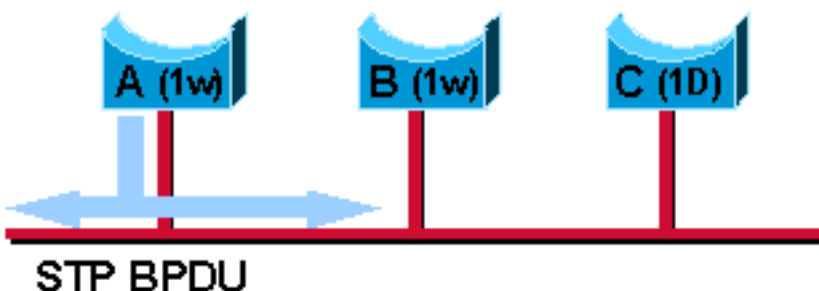
## 与802.1D兼容

RSTP 可以与传统 STP 协议交互操作。但是，务必注意，与传统网桥交互时，802.1w 固有的快速收敛优点将会丧失。

每个端口都维护一个变量，该变量定义可在对应网段上运行的协议。当端口启动时，值为三秒的迁移延迟计时器也将启动。当此计时器运行时，会锁定与端口关联的当前 STP 或 RSTP 模式。迁移延迟过期后，端口会适应与其接收的下一个 BPDU 对应的模式。如果端口由于收到的 BPDU 而更改其运行模式，则迁移延迟会重新启动。这会限制可能的模式更改频率。



例如，假设上图中的网桥A和B都运行RSTP，并且交换机A被指定用于网段。此链路上引入了传统 STP 网桥 C。由于 802.1D 网桥会忽略 RSTP BPDU 并丢弃它们，因此 C 认为网段中没有其他网桥，并开始发送其次要 802.1D 格式的 BPDU。交换机 A 将收到这些 BPDU，并在最多两倍于 hello-time 的秒数之后，只在该端口上将其模式更改为 802.1D。结果，C 现在识别出交换机 A 的 BPDU 并且接受 A 作为该网段的指定网桥。



请注意，在此特定情况下，如果删除网桥C，则网桥A会在该端口上以STP模式运行，即使它能够与其唯一的邻居B在RSTP模式下更有效地运行。这是因为A不知道网桥C已从网段中删除。对于这种特殊（少见）情况，需要用户介入，以便手动重新启动端口的协议检测。

当端口处于802.1D兼容模式时，它还可以处理拓扑更改通知(TCN)BPDU以及设置了TC或TCA位的BPDU。

## 结论

RSTP (IEEE 802.1w) 本身包括大多数针对 802.1D 生成树的 Cisco 专有增强功能，例如 BackboneFast、UplinkFast 和 PortFast。RSTP 可以在正确配置的网络中实现更快的收敛，有时只需几百毫秒。传统802.1D计时器（例如转发延迟和max\_age）仅用作备份，如果管理员正确识别并设置了点对点链路和边缘端口，则不需要这些计时器。此外，如果没有与旧网桥的交互，则不需要计时器。

## 相关信息

- [在运行 CatOS 的 Catalyst 系列交换机上配置 MST \(802.1s\)/RSTP \(802.1w\)](#)
- [了解和配置 Cisco Uplink Fast 功能](#)
- [工具和资源](#)
- [技术支持和文档 - Cisco Systems](#)

## 关于此翻译

思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。