

# 了解快速生成树协议 (802.1w)

## Contents

### [Introduction](#)

[在Catalyst交换机的对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](#)

[全视图Cisco BPDU, IEEE BPDU和BPDU图表](#)

[新BPDU处理](#)

[发送BPDU每个呼叫时间](#)

[信息更快过期](#)

[接受次要 BPDU](#)

[对转发状态的迅速转换](#)

[边缘端口](#)

[链路类型](#)

[与802.1D的收敛](#)

[与 802.1w 的收敛](#)

[建议/协定的序列](#)

[UplinkFast](#)

[新的拓扑更改机制](#)

[拓扑更改检测](#)

[拓扑更改传播](#)

[与802.1d的兼容性](#)

[结论](#)

[Related Information](#)

## Introduction

本文档提供有关 RSTP 给旧版 802.1D 标准带来的增强功能的信息。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，桥接解决方案现在能够与路由解决方案抗衡，桥接解决方案中的开放最短路径优先 (OSPF) 和增强型内部网关路由协议 (EIGRP) 等协议能够在更短时间内提供替代路径。

Cisco 快速地提高与功能的原始 802.1D 规格例如 [快速上行链路](#)、[快速骨干网](#) 和端口加速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) 比一次革命能被看到作为 802.1D 标准的演变更多。802.1D 术语依然是主要同样。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 也能恢复到 802.1D，以基于每个端口与传统网桥进行交互操作。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。

# 在Catalyst交换机的对RSTP的支持

下表显示了 Catalyst 交换机中对 RSTP 的支持，以及该支持功能所需的最低软件版本。

Catalyst 平台	采用 RSTP 的 MST	RPVST+ ( 亦称 PVRST+ )
Catalyst 2900XL/3500XL	不可用。	不可用。
Catalyst 2940	12.1(20)EA2	12.1(20)EA2
Catalyst 2950/2955/3550	12.1(9)EA1	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。
Catalyst 4000/2948G/2980G (CatOS)	7.1	7.5
Catalyst 4000/4500 (IOS)	12.1(12c)EW	12.1(19)EW
Catalyst 5000/5500	不可用。	不可用。
Catalyst 6000/6500	7.1	7.5
Catalyst 6000/6500 (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在这五个不同的端口状态被定义：

- 失效
- 监听
- 了解
- 阻拦
- 转发

有关详细信息，请参阅本文档的[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。实际区别在于生成树分配给端口的角色。可能安全假设，一个监听端口被指派或根并且在其途中对转发状态。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
阻拦	丢弃	无	无
监听	丢弃	是	无
了解	了解	是	是

## 端口角色

现在，该角色是分配到给定端口的一个变量。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。为了简化起见，关于 BPDU 需要记住的是，始终有一种方法可以比较其任意两者并确定其中更加有用的一项。这在 BPDU 存储的值和偶尔地根据他们被接受的端口。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。就路径成本而言，根端口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。根网桥是在没有一个根端口的网络的唯一的网桥。所有其他网桥在至少一个端口上接收 BPDU。



### 选定的端口角色

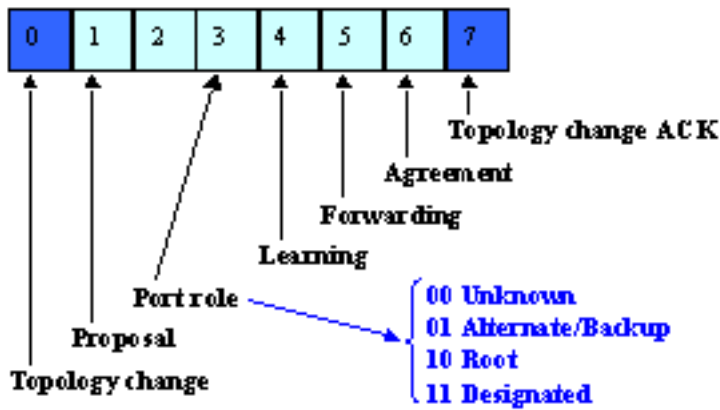
- 端口被选定，如果能发送在被连接的分段的最佳的 BPDU。802.1D 网桥将不同网段（例如以太网段）链接在一起，以创建桥接域。在给定的网段上，只能有一条路径通往根网桥。如果有两条路径，则网络中会有桥接环路。连接到给定网段的所有网桥将监听每个网桥的 BPDU，并且同意将发送最佳 BPDU 的网桥作为网段的指定网桥。该网桥上对应的端口是该网段的指定端口。



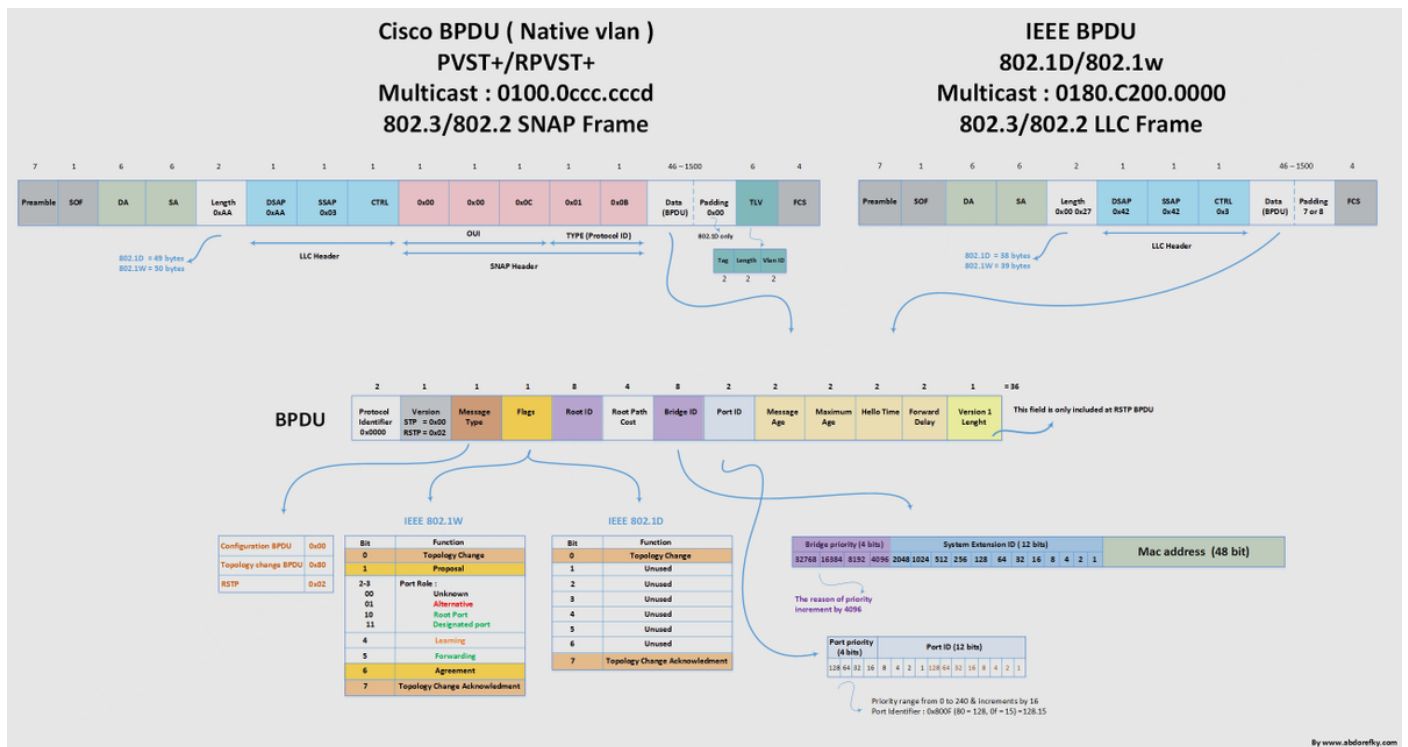
### 备选和备份端口角色

- 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，端口必须接收 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：





## 全视图Cisco BPDU , IEEE BPDU和BPDU图表



关于更加高分辨率的镜像，请参阅[Cisco BPDU , IEEE BPDU和BPDU图表](#)。

**Note:**位0 (拓扑更改)是最少有效位。

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。此属性使容易对802.1w网桥发现传统网桥被连接到它。

## 新BPDU处理

### 发送BPDU每个呼叫时间

每 Hello-Time 发送 BPDU，而不再只是进行中继。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。实际上，网桥会中继 BPDU，而不仅是实际上生成他们。802.1w 并

非如此。现在，即使网桥未从根网桥收到任何 BPDU，也会每 <hello-time> 秒（默认情况下是 2 秒）将 BPDU 与其最新信息一起发送。

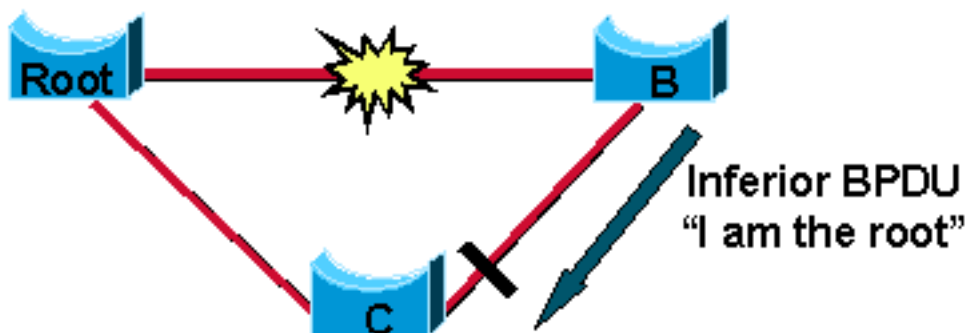
## 信息更快过期

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。由于以前被提及的协议修改，BPDU 当前使用作为在网桥之间的一个保活机制。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，在 802.1D 中问题可能发生在通往根网桥的路径上的任何位置。

**Note:** 如果是物理链路故障，会更快地检测出来。

## 接受次要 BPDU

此概念是 BackboneFast 引擎的核心部分。IEEE 802.1w 委员会决定在 RSTP 中融入相似的机制。当网桥从其选定或根网桥时获得下等信息，立即接受它并且替换那个以前存储。



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。

## 对转发状态的迅速转换

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。以前，实现更快收敛通过调整保守默认参数（转发延迟和 max\_age 计时器）完成，并且经常会危及网络稳定性。新的快速 STP 能够主动确认端口是否可安全转换到转发状态，而不需要依靠任何计时器配置。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，协议依靠两个新的变量：边缘端口和链路类型。

## 边缘端口

Cisco 生成树用户已经对边缘端口概念非常熟悉，因为它基本上与 Portfast 功能对应。所有直接连接到终端站的端口都不能在网络中创建桥接环路。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

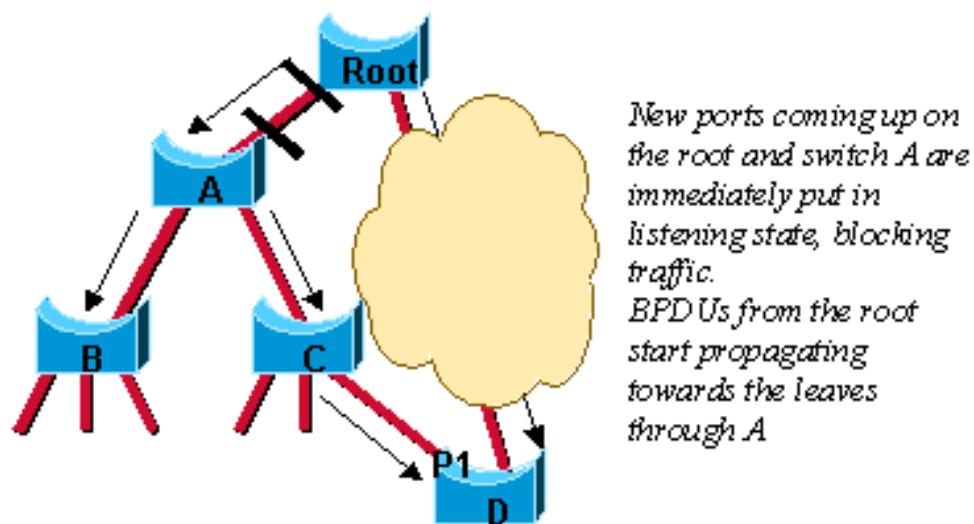
BPDUs 的边缘端口会立即失去边缘端口状态并成为正常的生成树端口。此时，边缘端口状态有一个用户配置的值和一个运行值。Cisco 实施仍保留 *PortFast* 关键字用于边缘端口配置。这使转换为 RSTP 变得更简单。

## 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。可以使用明确配置覆盖此自动链路类型设置。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。

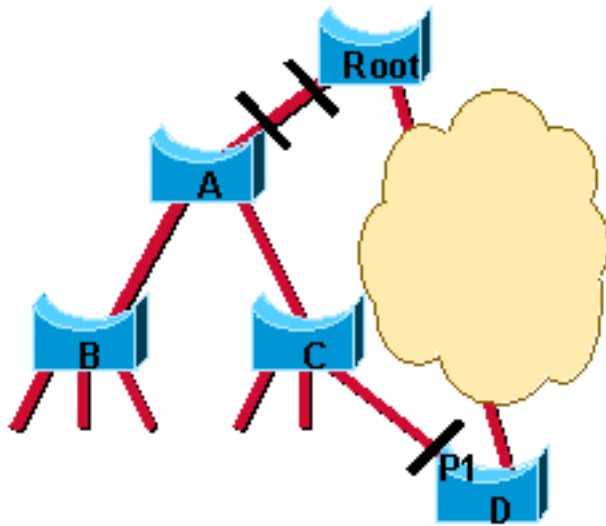
## 与802.1D的收敛

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一条间接连接（通过图中的 C-D）。STP 会阻塞端口并且禁用桥接环路。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 从网桥 A 收到此新的高级信息，它们就会立即向分支中继该信息。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.*

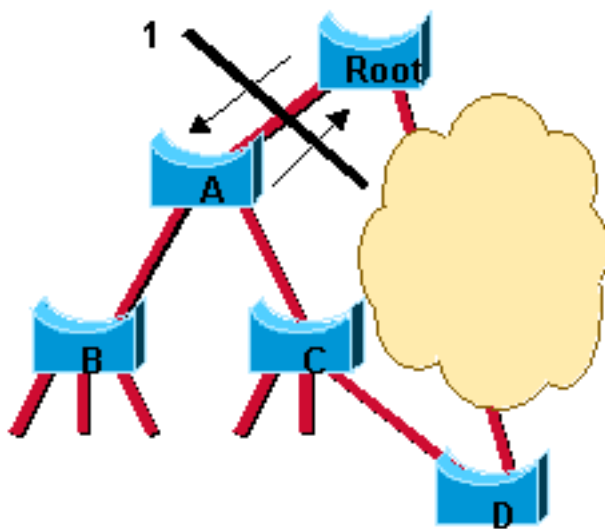
|

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

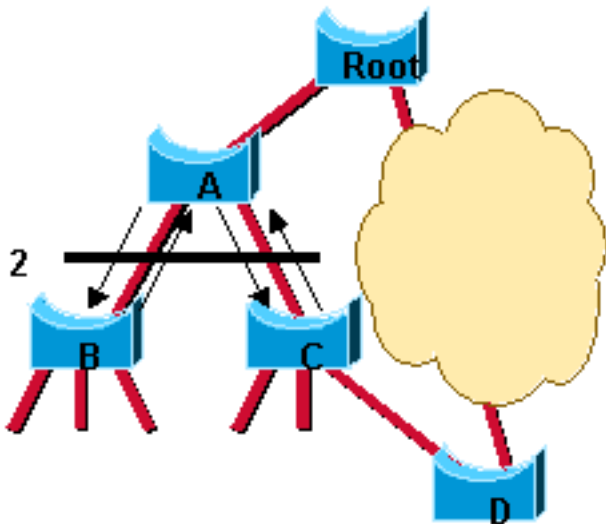
现在，您可以看到 RSTP 处理类似情况的方式。请记住，最终拓扑与 802.1D 计算的拓扑完全一样（即，与以前相同的位置上有一个阻塞的端口）。只是实现此拓扑的步骤发生了变化。

当他们出现，在链路的端口A和根之间在选定的阻塞放置。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。一旦网桥 A 收到根网桥的 BPDU，它就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拦，并且两个网桥交换BPDU。

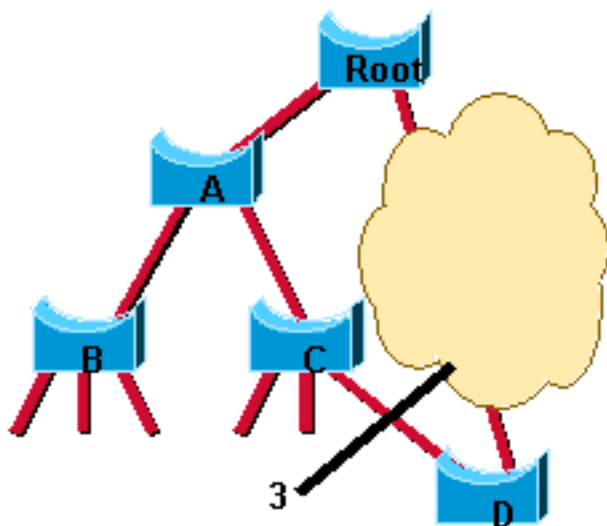


交换机 A 阻塞其非边缘指定端口后，交换机 A 和根网桥之间的链路便进入转发状态，您将遇到以下情况：





仍然不能有环路。现在，网络不是阻塞交换机 A 以上的部分，而是阻塞交换机 A 以下的部分。但是，会在其他位置切断可能的桥接环路。此剪切与根在此阶段产生的新的BPDU一起移动树通过交换机A，在交换机A的新堵塞的端口也协商一快速转换对转发状态与他们的在交换机B的邻接端口，并且交换机C该两个发起同步操作。除通向交换机 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。同样地，交换机 C 只需阻塞其通向 D 的指定端口。现在可进入下图中显示的状态：



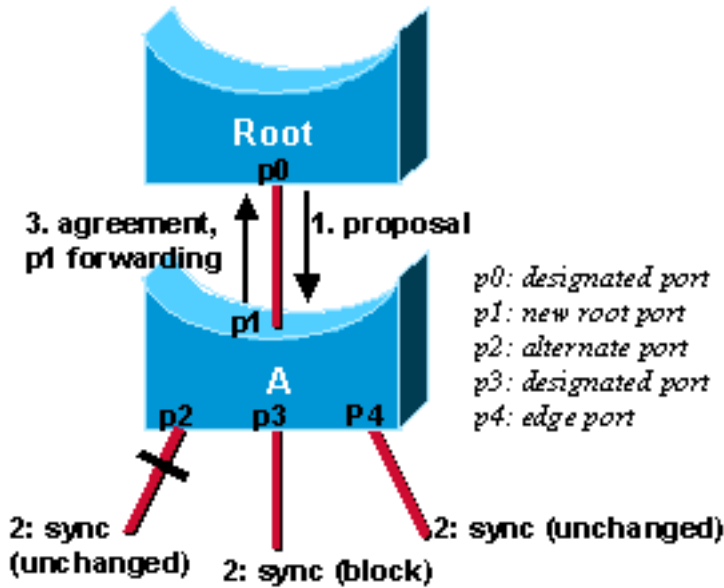
请记住，最终拓扑与 802.1D 示例完全相同，这表示 D 上的端口 P1 最终会进入阻塞状态。这表示在新的 BPDU 沿树向下传递所需的必要时间后，将实现最终网络拓扑。在这种快速收敛中没有涉及计时器。RSTP 唯一引入的新机制是确认交换机可以在其新的根端口上进行发送，以便授权立即转换到转发状态，从而绕过长达两倍转发延迟时间的监听和学习阶段。管理员只需要记住以下几点就可以受益于快速收敛：

- 在网桥之间的此协商只是可能的，当网桥由点到点链路时(即全双工链路连接，除非明确端口配置)。
- 在 802.1D 中端口上启用 Portfast 后，边缘端口更加重要。例如，如果网络管理员未正确配置 B 上的边缘端口，则其连接会受到启动的 A 和根之间的链路影响。

## 建议/协定的序列

当一个端口由 STA 选择作为指定端口时，在将其转换到转发状态之前，802.1D 还要等待 <forward delay> 秒的两倍时间（默认为  $2 \times 15$ ）。在RSTP，此情况对应于端口以一个选定的角色，但是阻塞

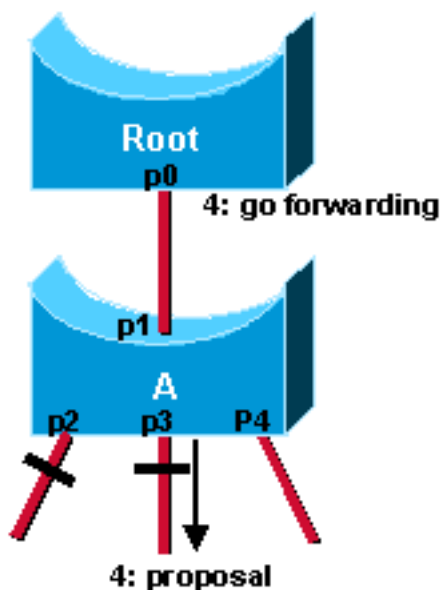
状态。下图说明了如何逐步实现快速转换。假设在根网桥和交换机 A 之间创建了一条新链路。此链路上的两个端口都置于指定的阻塞状态，直到它们从对等的端口收到 BPDU 为止。



当一个被指派的端口在一个丢弃的或学习状态(和只在这种情况下)，设置派出在BPDU的建议位。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。

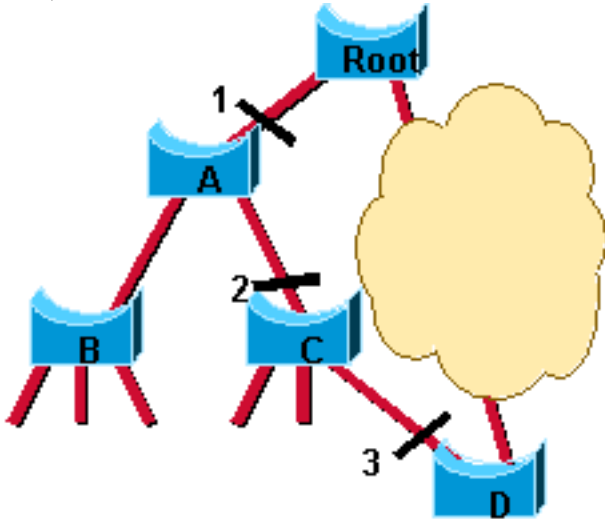
为了说明同步机制对不同种类端口的影响，假设交换机 A 上存在替代端口 p2，指定的转发端口 p3 和边缘端口 p4。请注意，p2 和 p4 已经满足其中一个标准。为了保持同步（请参阅上图的步骤 2），交换机 A 仅需要阻塞端口 p3 并向其分配丢弃状态。既然其所有的端口都保持同步，交换机 A 就可以取消阻塞其新选择的根端口 p1 并发送协议信息以回复根网桥。（请参阅步骤 3。）此消息是建议BPDU的复制，协议位设置而不是建议位。这可确保端口 p0 确切地知道它接收的协议与哪个建议对应。



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。

- 因为不依靠任何计时器，建议协议机制非常快速。握手此通知迅速传播往网络的边缘和在一个变化以后迅速恢复连接在拓扑上。
- 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。



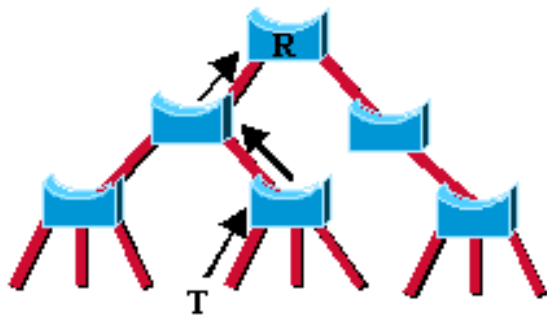
## UplinkFast

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，它能够将其最佳替代端口直接置于转发模式（RSTP 也会处理新的根端口的出现与否）。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。

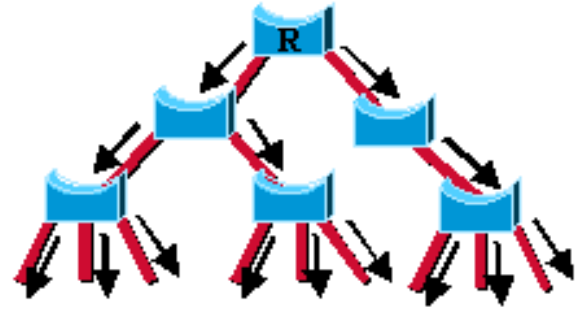
因为 RSTP 中自带 UplinkFast 机制并会自动启用，因此无需对该机制进行进一步配置。

## 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.

一旦根网桥知道一个变化在网络的拓扑上，设置派出，然后被传递到在网络的所有网桥在BPDU的TC标志位。当网桥收到设置了TC标志位的BPDU时，它会将其桥接表老化时间降低到转发延迟秒数。这可确保相对快速地刷新过期信息。有关此过程的详细信息，请参考[了解生成树协议拓扑更改](#)。此拓扑更改机制在RSTP中进行了很深层的再造。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。

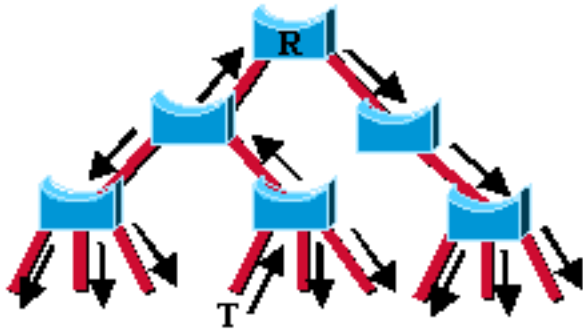
**Note:**只要TC While计时器在端口上运行，从该端口发送的BPDU就都会设置TC位。当计时器处于活动状态时，根端口上也会发送BPDU。

## 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 timer并且发送BPDU在所有其被指派的端口和根端口设置的TC (RSTP不再使用特定TCN BPDU，除非传统网桥需要被通知)。

这样，TCN会快速传播到整个网络。现在，TC传播是只需一个步骤的进程。实际上，拓扑更改的发起方会将此信息传播到整个网络中（与802.1D相反，在其中只有根网桥才能如此）。此机制比802.1D等效机制更快。没有需要等待根网桥被通知然后保持整个网络的拓扑更改状态<max使用周期的加上向前delay>秒钟。



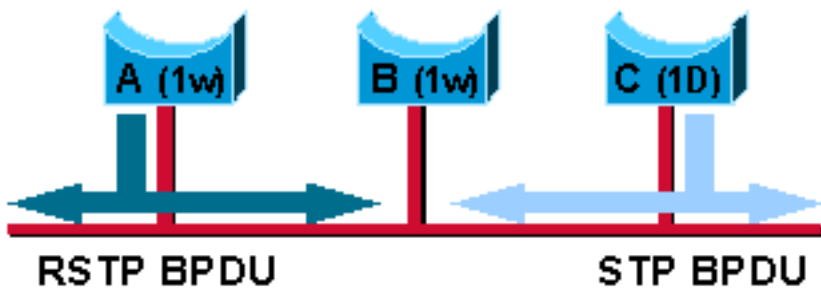
The originator of the TC directly floods this information through the network

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。此方法潜在导致更多临时泛滥，另一方面，但是清除防止迅速连接恢复性的潜在的过时的信息。

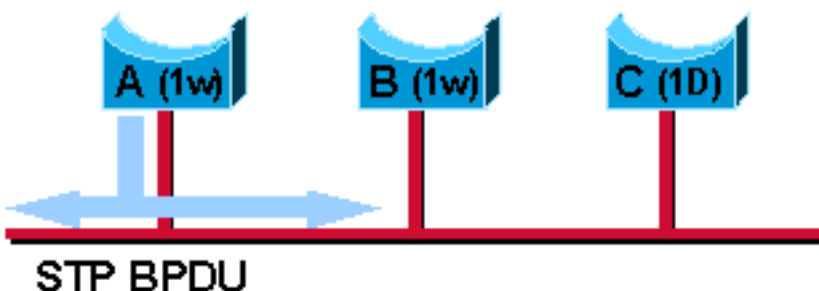
## 与802.1d的兼容性

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，并且指定交换机 A 用于该网段。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。



请注意在这个特殊情况下，如果移除网桥 C，则网桥 A 会在该端口上以 STP 模式运行，即使它可以在 RSTP 模式下更有效地与其唯一的相邻网桥 B 一起工作。这是因为 A 无法识别出网桥 C 已从网段移除。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。

当端口在 802.1D 兼容模式时，也能处理拓扑变化通知 (TCN) BPDU 和 BPDU 与 TC 或 TCA 位集。

## 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 能完成在一个适当地配置的网络的快速收敛，有时按几百毫秒的顺序。如果管理员正确标识和设置了点对点链路和边缘端口，则经典的 802.1D 计时器（例如转发延迟和 max\_age）可仅用作备份，且不是必要的。此外，如果不存在与传统网桥的交互，则计时器也不是必要的。

## Related Information

- [配置在运行 CatOS 的 Catalyst 系列交换机的 MST \(802.1s\)/RSTP \(802.1w\)](#)
- [了解和配置 Cisco 上行链路快速的功能](#)
- [LAN 交换技术支持](#)
- [LAN 产品支持](#)
- [工具和资源](#)
- [Technical Support & Documentation - Cisco Systems](#)