

# 瞭解多重生成樹通訊協定(802.1s)

## 目錄

---

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[背景資訊](#)

[使用MST的位置](#)

[PVST+使用案例](#)

[標準802.1q使用案例](#)

[MST案例](#)

[MST區域](#)

[MST配置和MST區域](#)

[區域邊界](#)

[MST例項](#)

[IST例項](#)

[MSTI](#)

[常見錯誤配置](#)

[IST例項在所有埠 \( 中繼或接入 \) 上都處於活動狀態](#)

[對映至同一例項的兩個VLAN會阻塞相同的埠](#)

[MST區域與外部環境之間的互動](#)

[建議的配置](#)

[替代組態 \( 不建議使用 \)](#)

[配置無效](#)

[遷移策略](#)

[結論](#)

[相關資訊](#)

---

## 簡介

本檔案介紹多重跨距樹狀目錄通訊協定(802.1s)的功能及組態。

## 必要條件

### 需求

思科建議您瞭解以下主題：

- 熟悉快速STP (RSTP) (802.1w)

## 採用元件

本文件所述內容不限於特定軟體和硬體版本。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

## 背景資訊

多生成樹(MST)是IEEE標準，源自於Cisco專有的多例項生成樹協定(MISTP)實施。下表顯示各種Catalyst交換器中的MST支援：

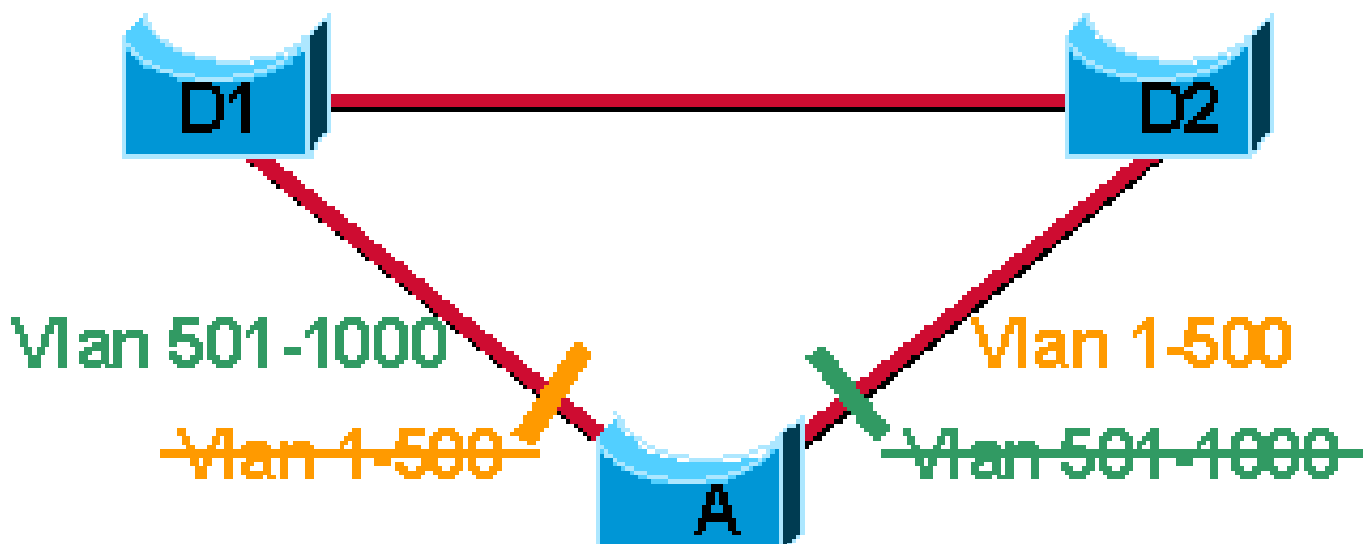
Catalyst平台	具有RSTP的MST
Catalyst 2900 XL和3500 XL	不可用
Catalyst 2950和3550	Cisco IOS® 12.1(9)EA1
Catalyst 2955	所有Cisco IOS版本
Catalyst 2948G-L3和4908G-L3	不可用
Catalyst 4000和4500(Cisco IOS)	12.1(12c)EW
Catalyst 5000和5500	不可用
Catalyst 6000和6500(Cisco IOS)	12.1(11b)EX、12.1(13)E、12.2(14)SX
Catalyst 8500	不可用

有關RSTP (802.1w)的詳細資訊，請參閱[瞭解快速生成樹協定\(802.1w\)](#)。

## 使用MST的位置

此圖顯示一種常見設計，其特徵是接入交換機A的1000個VLAN冗餘連線到兩台分佈層交換機D1和D2。

在此設定中，使用者連線到交換機A，網路管理員通常根據偶數或奇數VLAN或任何被認為合適的其它方案，在接入交換機上行鏈路上實現負載均衡。



具有1000個VLAN的接入交換機A冗餘連線到交換機D1和D2

以下部分是此設定中使用不同型別的STP的示例：

## PVST+使用案例

在Cisco Per-VLAN Spanning Tree (PVST+)環境中，會調整生成樹引數，以便在每個上行鏈路中繼上轉發一半的VLAN。

為了輕鬆地實現此目的，請選擇網橋D1作為VLAN 501至1000的根網橋，網橋D2作為VLAN 1至500的根網橋。下列陳述式適用於此組態：

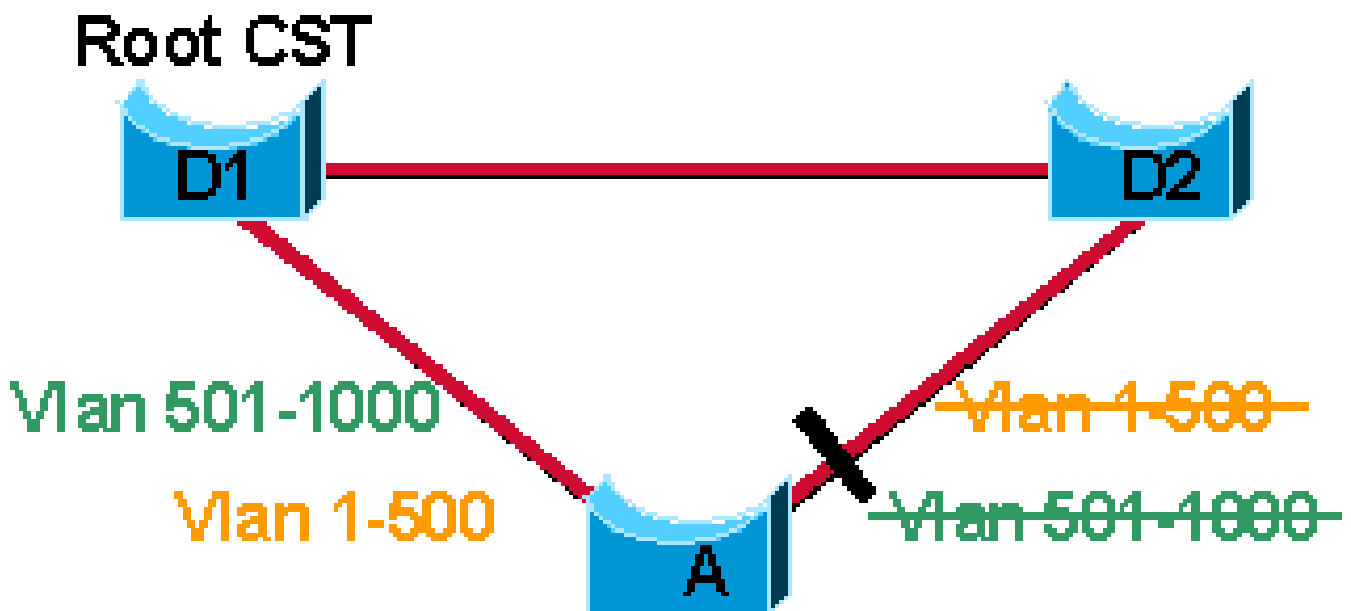
- 在這種情況下，會產生最佳的負載平衡。
- 每個VLAN維護一個生成樹例項，這意味著只為兩個不同的最終邏輯拓撲維護1000個例項。

這嚴重浪費了網路中所有交換機的CPU週期(除了每個例項傳送自己的網橋協定資料單元(BPDU)所用的頻寬之外)。

## 標準802.1q使用案例

原始IEEE 802.1q標準定義的不只是中繼。此標準定義了一個通用生成樹(CST)，它僅假定整個橋接網路只有一個生成樹例項，而無論VLAN的數量如何。


如果CST應用於下圖中的拓撲，則結果將與下圖類似：



應用於網路的通用生成樹(CST)

在運行CST的網路中，以下說法正確：

- 無法進行負載均衡；需要阻止所有VLAN的一個上行鏈路。
- CPU被省略；僅需要計算一個例項。

 注意：Cisco實施增強了802.1q以支援一個PVST。此功能的行為與本示例中的PVST完全相同。Cisco每VLAN BPDU透過純802.1q網橋進行隧道傳輸。

## MST案例

MST (IEEE 802.1s)結合了PVST+和802.1q的最佳方面。

其思想是可以將多個VLAN對映到數量較少的生成樹例項，因為大多數網路只需要幾個邏輯拓撲。

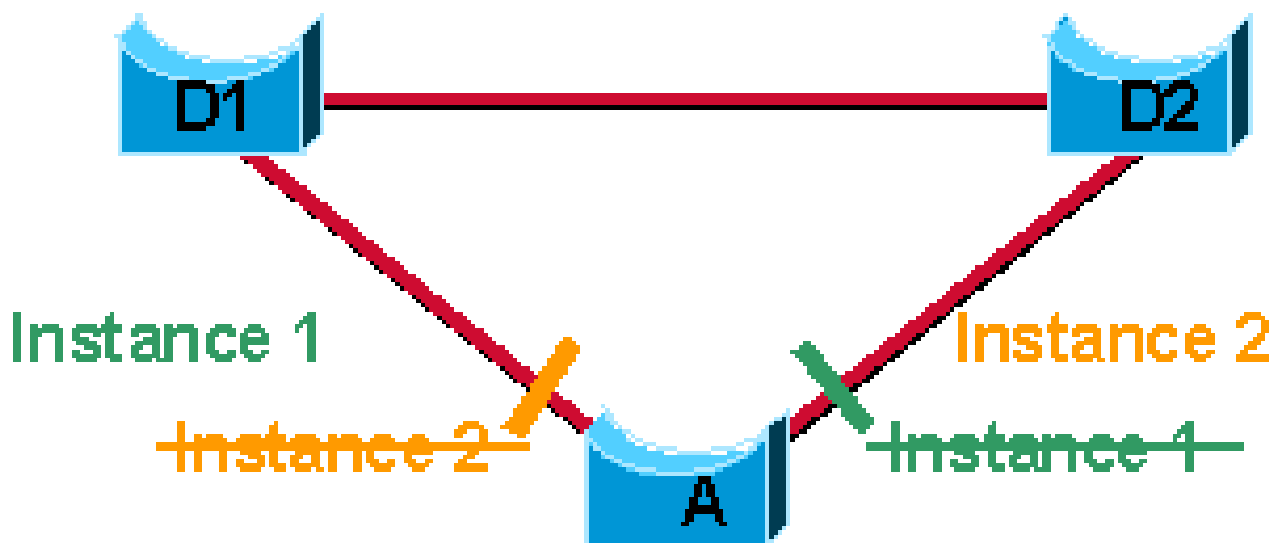
在第一個圖中所描述的拓撲中，只有兩種不同的最終邏輯拓撲，因此實際上只需要兩個生成樹例項。

無需運行1000個例項。如果將1000個VLAN的一半對映到不同的生成樹例項（如圖所示），則以下說法正確：

- 由於一半的VLAN連線到一個單獨的例項，因此仍可實現所需的負載均衡方案。
- 由於只計算兩個例項，因此CPU可倖免。

## Root Instance 1

## Root Instance 2



將1000個VLAN的一半對映到不同的生成樹例項

從技術角度來看，MST是最佳解決方案。從終端使用者的角度來看，遷移到MST的主要缺點是：

- 該協定比通常的生成樹更複雜，並且需要對員工進行額外的培訓。
- 與傳統網橋的互動可能是一項挑戰。有關詳細資訊，請參閱本文檔的[MST區域和外界之間的互動](#)部分。

## MST區域

如上所述，MST引入的主要增強功能是可以將多個VLAN對映到單個生成樹例項。

這就提出了以下問題：如何確定哪個VLAN與哪個例項相關聯；更準確地說，如何標籤BPDU，以便接收裝置可以辨識例項以及每個裝置應用的VLAN。

在802.1q標準情況下，該問題無關緊要，因為所有例項都對映到唯一的例項。在PVST+實施中，關聯為：

- 不同的VLAN會傳送各自例項的BPDU（每個VLAN一個BPDU）。

為了解決此問題，Cisco MSTP為每個例項傳送了一個BPDU，其中包含該BPDU負責的VLAN清單。

如果因為錯誤而導致兩台交換機配置不正確，並且與同一例項相關聯的VLAN範圍也不同，則協定很難從此情況中正確恢復。

IEEE 802.1s委員會採用更簡單的方法，引入了MST區域。將一個區域等同於邊界網關協定(BGP)自治系統，它是置於公共行政下的一組交換機。

## MST配置和MST區域

網路中運行MST的每台交換機都有一個MST配置，該配置包括以下三個屬性：


1. 字母數字配置名稱（32位元組）
2. 配置修訂版號（兩個位元組）
3. 一個4096元素表，用於將機箱上支援的每個潛在4096 VLAN與給定例項相關聯

要成為通用MST區域的一部分，一組交換機必須共用相同的配置屬性。由網路管理員負責將組態正確傳播到整個區域。

目前，只有透過指令行介面(CLI)或簡易網路管理通訊協定(SNMP)才能執行此步驟。

也可以設想其他方法，因為IEEE規範沒有明確提及如何完成該步驟。

---

 注意：如果由於某種原因，一個或多個配置屬性的兩個交換機不同，則這些交換機屬於不同區域。有關詳細資訊，請參閱下一節區域邊界。

---

## 區域邊界

為了確保一致的VLAN到例項對映，協定必須能夠準確確定區域的邊界。

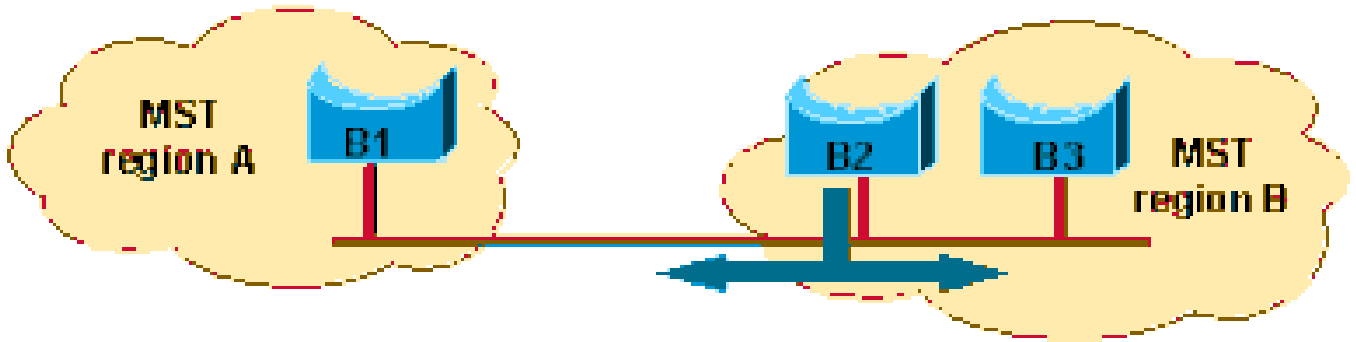
為此，BPDU中包含了該區域的特徵。在BPDU中不會傳輸精確的VLAN到例項對映，因為交換機只需要知道它們是否與鄰居位於同一區域。

因此，僅傳送VLAN到例項對映表的摘要，以及修訂版號和名稱。

交換機收到BPDU後，將提取摘要（透過數學函式從VLAN到例項的對映表中衍生出的數值），並將此摘要與自己的計算摘要進行比較。

如果摘要不同，則接收BPDU的埠位於區域邊界。

一般而言，如果某埠網段上的指定網橋位於不同區域，或者收到舊版802.1d BPDU，則該埠位於區域邊界。在此圖中，B1上的連線埠位於區域A的邊界，而B2和B3上的連線埠位於區域B的內部：



## MST例項

根據IEEE 802.1s規範，MST網橋必須至少能夠處理以下兩個例項：

- 一個內部生成樹(IST)
- 一個或多個多生成樹例項(MSTI)

由於802.1s實際上處於試行標準階段，因此術語不斷演變。這些名稱可能會在802.1s的最終發行版本中變更。思科實施支援16個例項：一個IST ( 例項0 ) 和15個MSTI。

## IST例項

為了清楚地瞭解IST例項的作用，請記住MST源自IEEE。因此，MST必須能夠與基於802.1q的網路互動，因為802.1q是另一個IEEE標準。對於802.1q，橋接網路僅實施一個生成樹(CST)。IST例項只是在MST區域內擴展CST的RSTP例項。

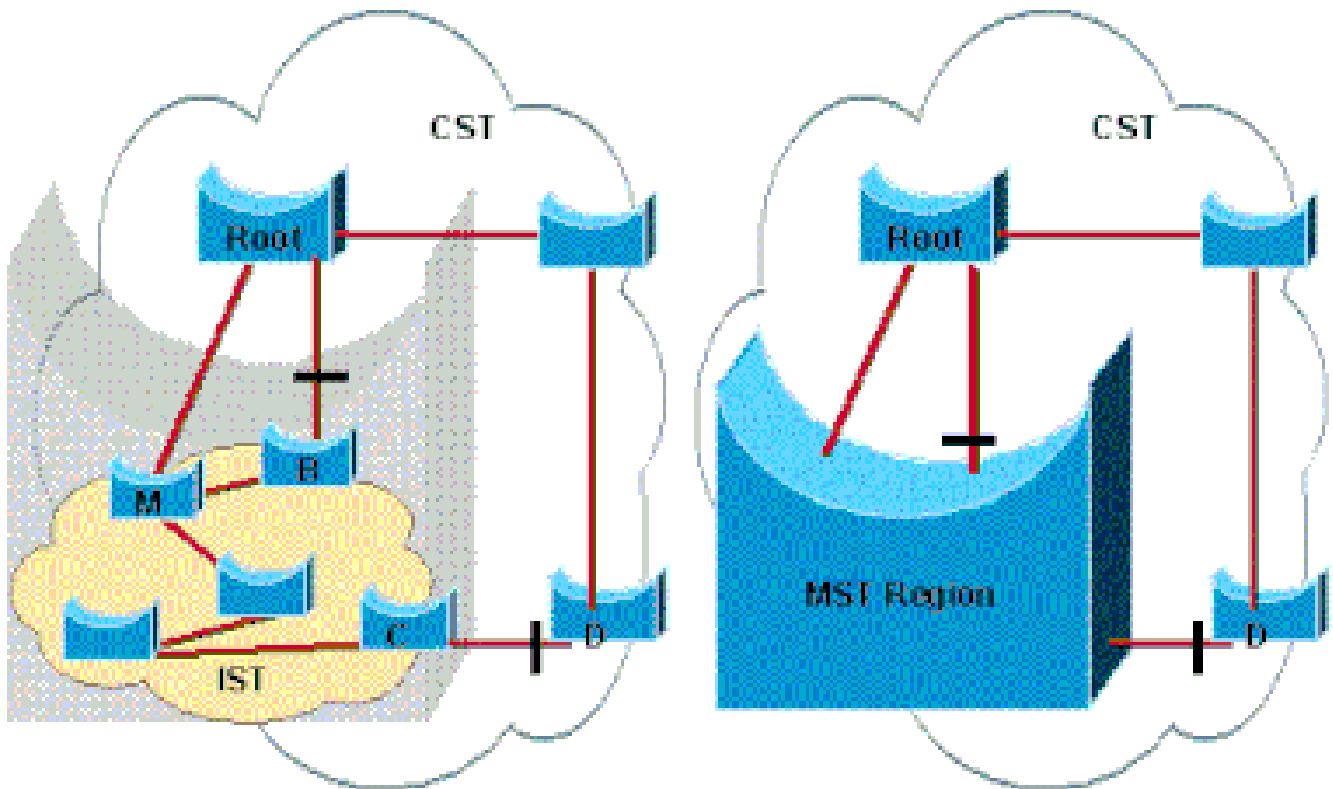
IST例項接收BPDU並將其傳送到CST。IST可以將整個MST區域表示為通往外部世界的CST虛擬網橋。

這兩個圖在功能上是等價的。注意不同阻塞埠的位置。在典型的橋接網路中，您會看到交換機M和B之間的阻塞埠。

您預計在MST區域中間的某個阻塞埠會破壞第二個環路，而不是在D上阻塞。

但是，由於IST的原因，整個區域顯示為運行單個生成樹(CST)的一個虛擬網橋。這使我們能夠瞭解虛擬網橋阻塞了B上的備用埠。

此外，該虛擬網橋位於C到D網段上，導致交換機D阻塞其埠。



使該區域顯示為虛擬CST網橋的確切機制不在本文檔的討論範圍之內，但已在IEEE 802.1s規範中進行了詳細說明。

但是，如果您記住了MST區域的虛擬網橋屬性，則與外界的互動將更容易理解。

## MSTI

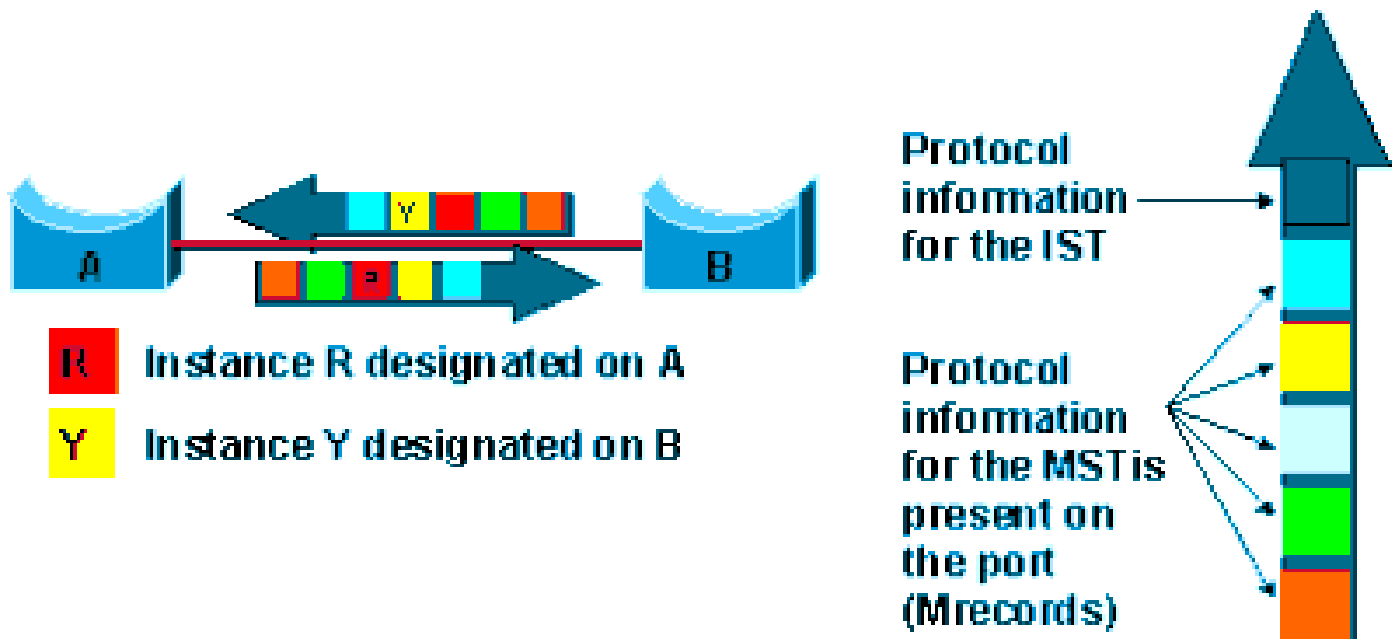
MSTI是僅存在於某個區域內的簡單RSTP例項。預設情況下，這些例項會自動運行RSTP，無需任何額外的配置工作。


與IST不同，MSTI從不與區域外部互動。請記住，MST僅在區域外部運行一個生成樹，因此除IST例項外，區域內部的常規例項沒有外部對應例項。

此外，MSTI不會將BPDU傳送到區域之外，只有IST會傳送。

MSTI不會傳送獨立的單個BPDU。在MST區域內，網橋交換IST的MST BPDU可視為正常RSTP BPDU，也包含每個MSTI的附加資訊。

此圖顯示MST區域內交換機A和B之間的BPDU交換。每台交換機只傳送一個BPDU，但每個埠上存在的每個MSTI都包含一個MRecord。



 注意：在此圖中，請注意MST BDPDU攜帶的第一個資訊欄位包含有關IST的資料。這表示IST（例項0）始終存在於MST區域內的任何地方。但是，網路管理員不必將VLAN對映到例項0上，因此不必擔心。

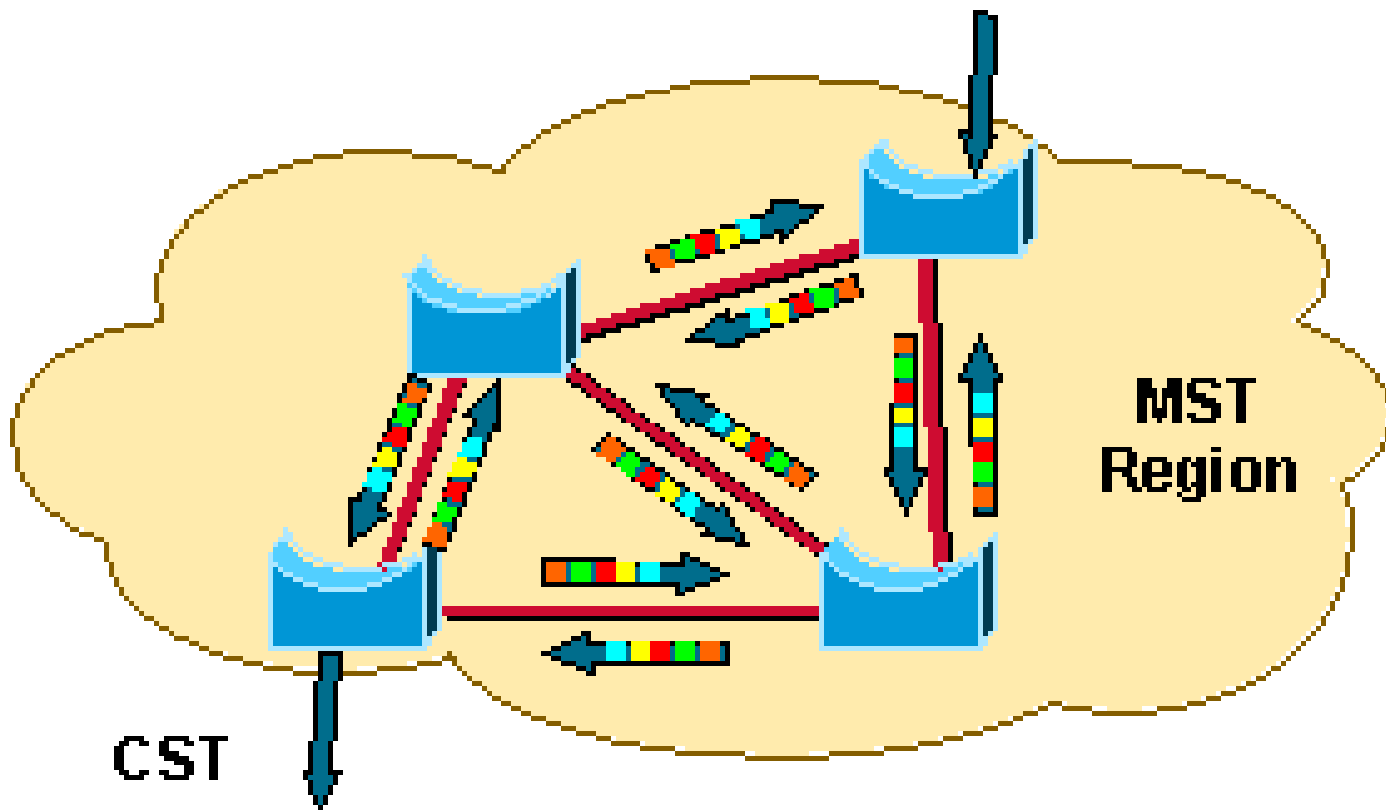
與常規融合生成樹拓撲不同，鏈路的兩端可以同時傳送和接收BPDU。

這是因為如圖所示，每個網橋都可以指定給一個或多個例項，並且需要傳輸BPDU。

一旦在埠上指定了單個MST例項，就會傳送包含所有例項(IST+ MSTI)資訊的BPDU。

此處顯示的圖演示了在MST區域內部和外部傳送的MST BDPU：





在MST區域內部和外部傳送MST BDPUs

MRecord包含相應例項所需的足夠資訊（主要是根網橋和傳送方網橋優先順序引數），用於計算其最終拓撲。

MRecord不需要任何與計時器相關的引數，例如通常在常規IEEE 802.1d或802.1q CST BPDU中找到的呼叫時間、轉發延遲和最大老化時間。

MST區域中唯一使用這些引數的例項是IST；Hello時間確定BPDU的傳送頻率，轉發延遲引數主要用於無法進行快速轉換時（請記住，快速轉換不會發生在共用鏈路上）。

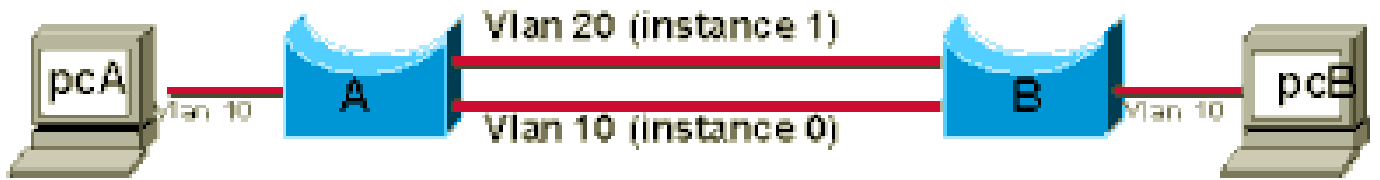
由於MSTI依賴IST傳輸其資訊，因此MSTI不需要這些計時器。

## 常見錯誤配置

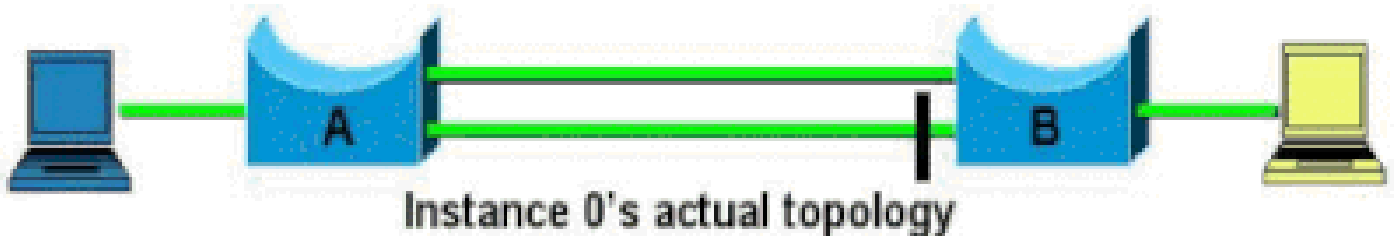
例項和VLAN之間的獨立性是一個新概念，意味著您必須仔細規劃配置。[所有埠（中繼或接入）上的IST例項都是活動的](#)部分說明了一些常見的缺陷以及如何避免它們。

## IST例項在所有埠（中繼或接入）上都處於活動狀態

此圖顯示交換機A和B與位於不同VLAN中的接入埠相連。VLAN 10和VLAN 20對映到不同的例項。VLAN 10對映到例項0，而VLAN 20對映到例項1。



此配置導致pcA無法將幀傳送到pcB。show命令表明交換機B要阻塞VLAN 10中指向交換機A的鏈路，如下圖所示：



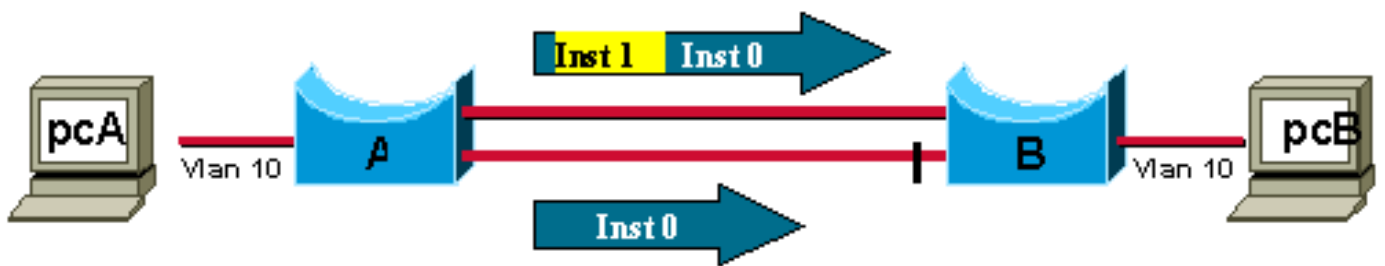
在這樣一個沒有明顯環路的簡單拓撲中，這怎麼可能？

此問題的解釋是，MST資訊僅與一個BPDU (IST BPDU)一起傳輸，而不考慮內部例項的數量。單個例項不會傳送單個BPDU。

當交換機A和交換機B交換VLAN 20的STP資訊時，交換機會傳送一個帶有例項1的MRecord的IST BPDU，因為這是VLAN 20的對映位置。

但是，因為這是IST BPDU，所以此BPDU也包含例項0的資訊。這表示IST例項在MST區域中的所有埠上都處於活動狀態，無論這些埠是否承載對映到IST例項的VLAN。

此圖顯示IST例項的邏輯拓撲：



交換機B從交換機A接收兩個BPDU，例如0（每個埠一個）。很明顯，交換機B必須阻塞其埠之一，以避免環路。

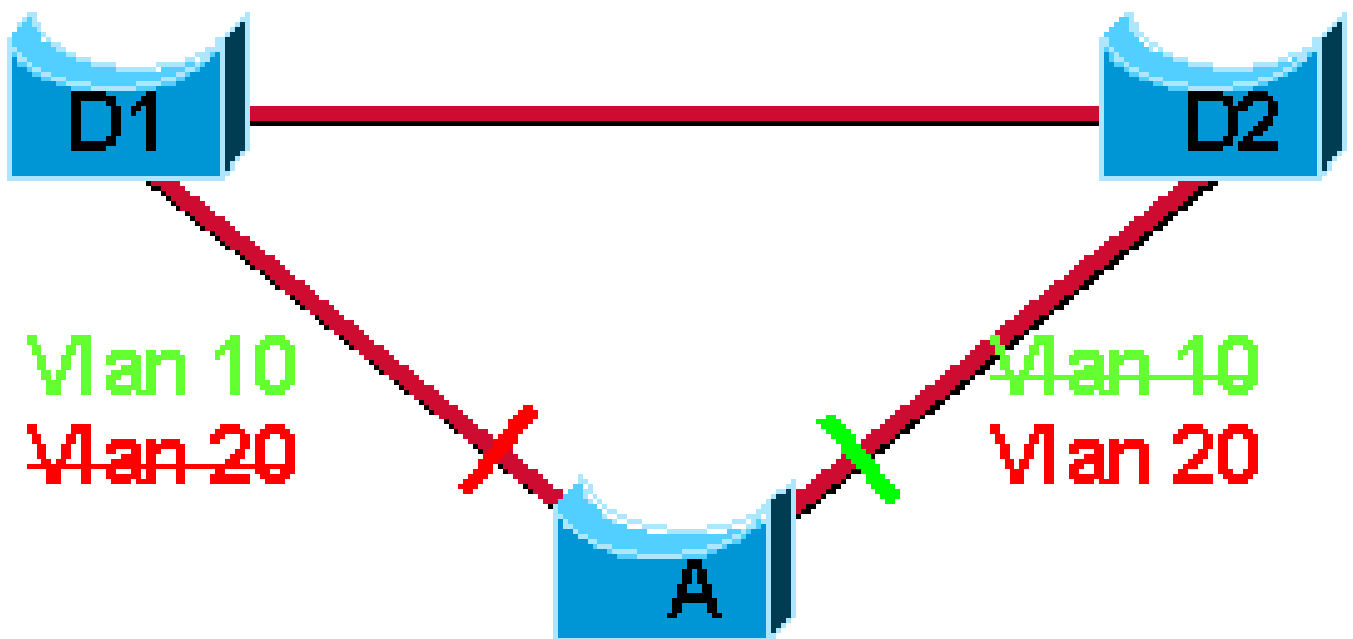
首選解決方案是對VLAN 10使用一個例項，對VLAN 20使用另一個例項，以避免將VLAN對映到IST例項。

另一種方法是在所有鏈路上承載對映到IST的VLAN（在兩個埠上允許VLAN 10，如下圖所示）。

## 對映至同一例項的兩個VLAN會阻塞相同的埠

請記住，VLAN不再表示生成樹例項。該拓撲由例項確定，而不考慮對映到該例項的VLAN。

下圖顯示了一個問題，該問題是在所有埠（中繼或接入）上的IST例項都是活動的部分中討論的問題的另一種形式：



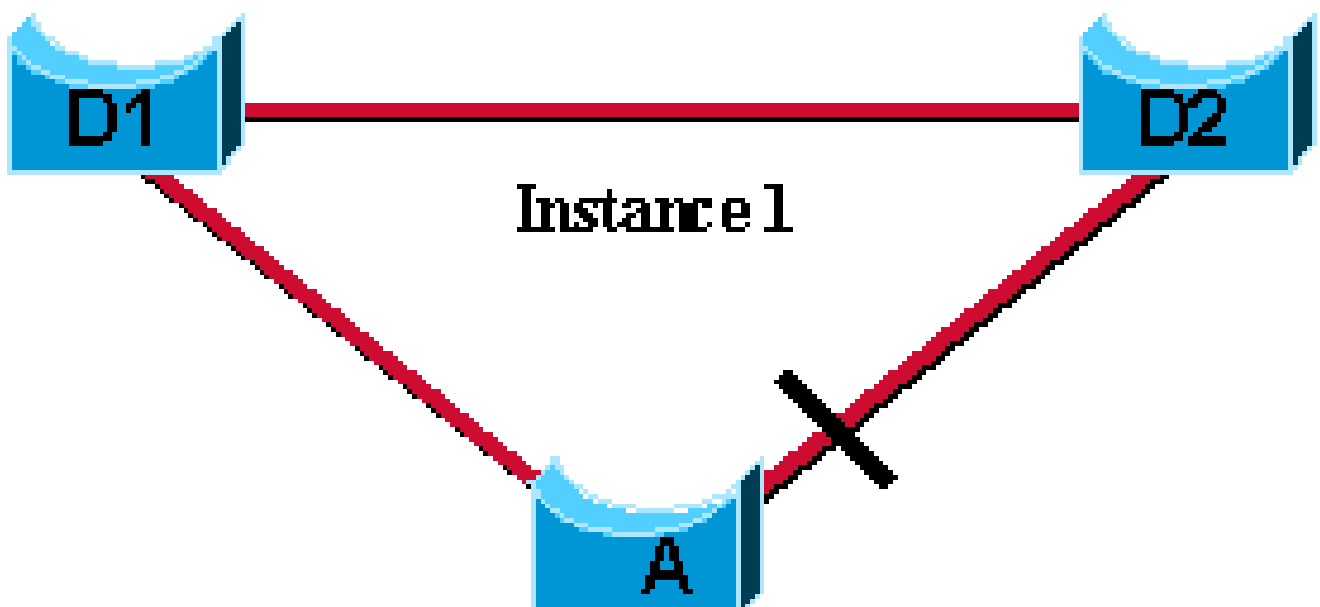
拓撲由例項確定，無論對映至該例項的VLAN如何

假設VLAN 10和20都對映到同一例項（例項1）。

網路管理員想要手動修剪一個上行鏈路上的VLAN 10和另一個上行鏈路上的VLAN 20，以限制從交換機A到分佈層交換機D1和D2的上行鏈路中繼流量（嘗試實現上圖所述的拓撲）。

完成此操作後不久，網路管理員發現VLAN 20中的使用者已失去與網路的連線。

這是一個典型的配置錯誤問題。VLAN 10和20均對映到例項1，這意味著兩個VLAN只有一個邏輯拓撲。無法實現負載分擔，如下所示：



典型配置錯誤問題

由於手動修剪，VLAN 20僅允許在阻塞的埠上使用，這可以解釋連線丟失。為了實現負載均衡，網路管理員必須將VLAN 10和20對映到兩個不同的例項。

為避免此問題，可以使用一個簡單的規則，即絕不手動修剪中繼上的VLAN。如果您決定從中繼中刪除某些VLAN，請同時刪除對映到給定例項的所有VLAN。

切勿從中繼中刪除單個VLAN，也不要刪除對映到同一例項的所有VLAN。

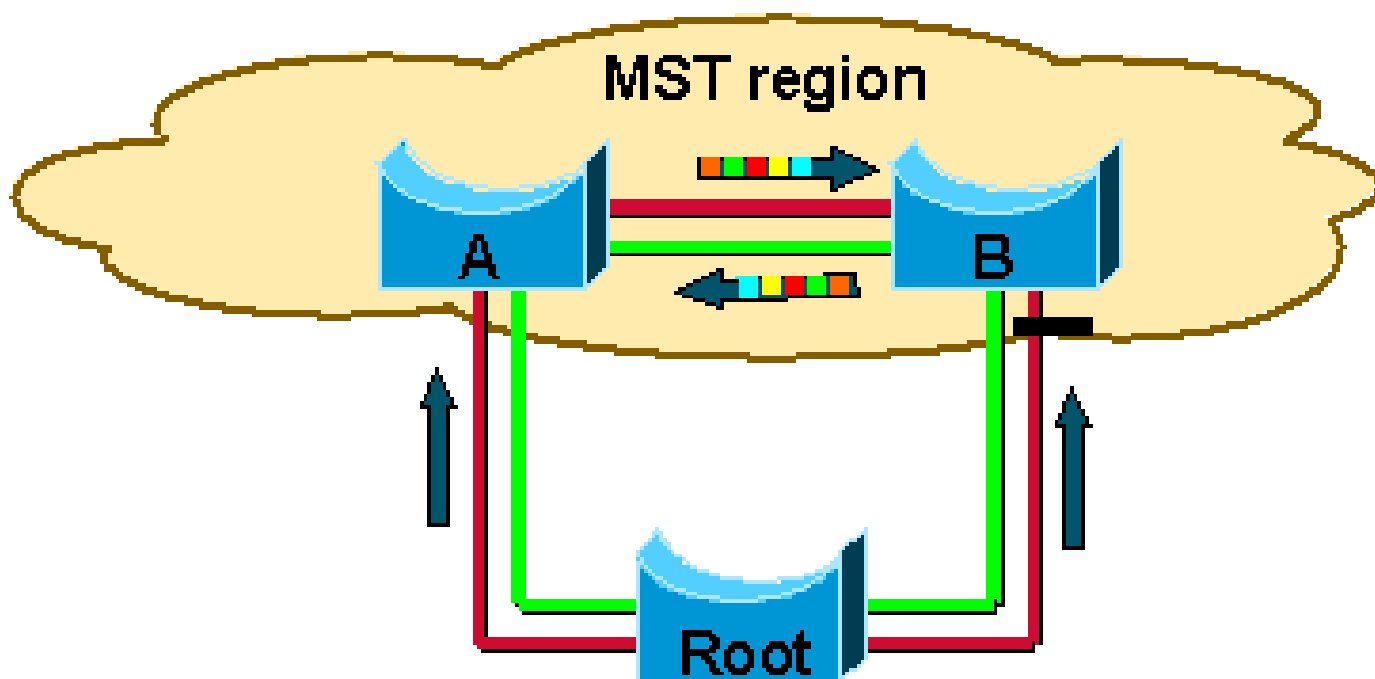
## MST區域與外部環境之間的互動

遷移到MST網路後，管理員可能需要處理MST和傳統協定之間的互操作性問題。

MST可與標準802.1q CST網路無縫操作；但是，由於具有單個生成樹限制，只有少數網路基於802.1q標準。

Cisco在宣佈支援802.1q的同時發佈了PVST+。Cisco還在MST和PVST+之間提供高效而簡單的相容機制。此機制將在本文檔後面部分進行介紹。

MST區域的第一個屬性是，在邊界埠上，沒有MSTI BPDU被傳送，只有IST BPDU被傳送。內部例項(MSTI)始終自動匹配邊界埠上的IST拓撲，如下圖所示：



內部例項(MSTI)始終自動匹配邊界埠上的IST拓撲

在此圖中，假設VLAN 10到50對映到綠色例項，而綠色例項僅是內部例項(MSTI)。

紅色連結代表IST，因此也代表CST。VLAN 10到50在拓撲中處處可用。

綠色例項的BPDU不會從MST區域傳送出去。

這並不意味著在VLAN 10到50中存在環路。MSTI在邊界埠跟蹤IST，交換機B上的邊界埠還會阻止

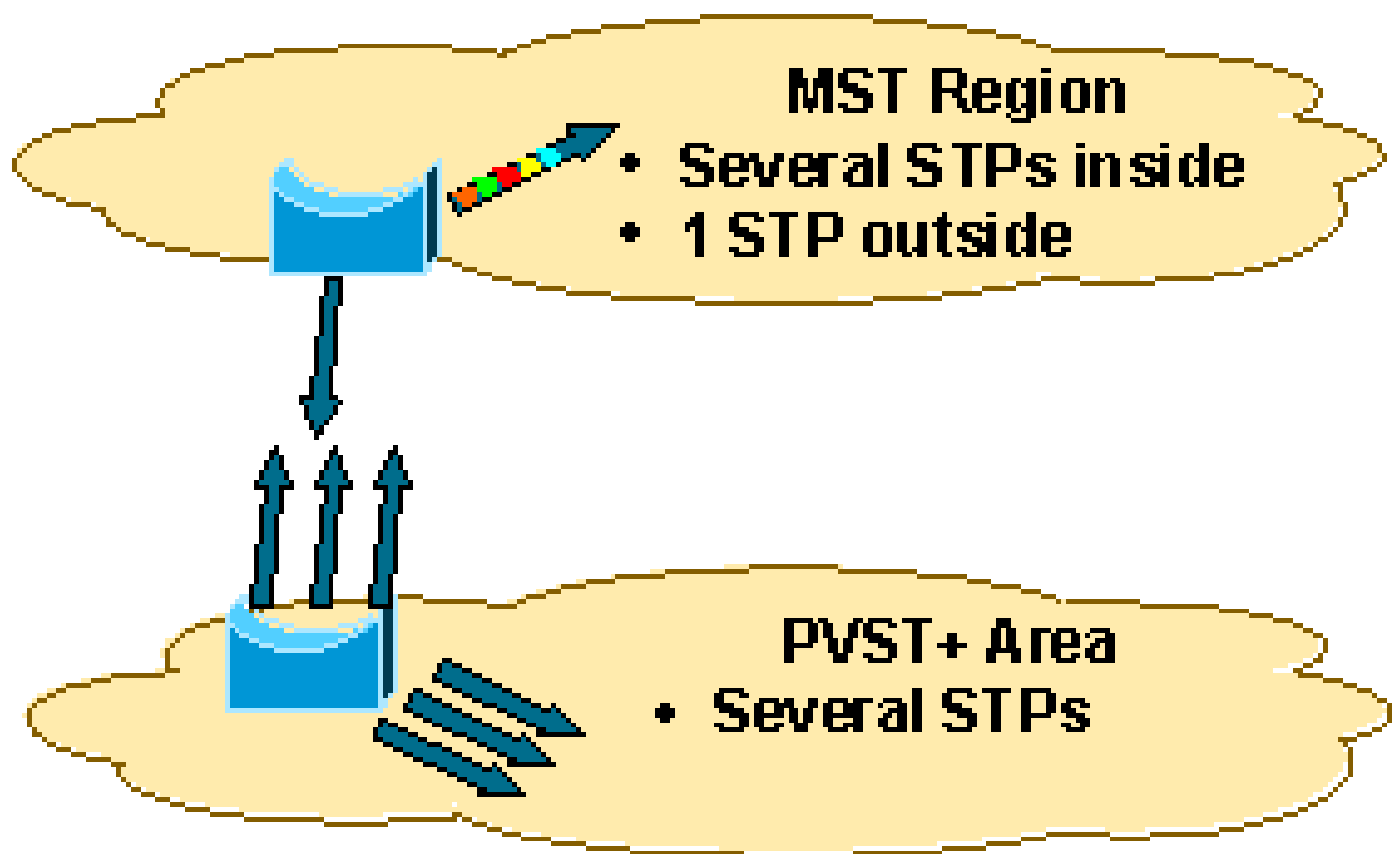
綠色例項的流量。

運行MST的交換機能夠自動檢測邊界上的PVST+鄰居。這些交換機能夠檢測到例項的中繼埠的不同VLAN上收到多個BPDU。

此圖顯示了一個互操作性問題。MST區域只與該區域外的一個生成樹(CST)互動。

但是，PVST+網橋對每個VLAN運行一個生成樹演算法(STA)，因此，每兩秒在每個VLAN上傳送一個BPDU。

邊界MST網橋預計不會收到那麼多BPDU。MST網橋期望接收或傳送一個，這取決於網橋是否為CST的根。



MST橋接器預期會接收或傳送一個

思科開發了一種機制來解決下圖中顯示的問題。一個可能性可能包括將PVST+網橋傳送的額外BPDU透過隧道傳輸到MST區域。

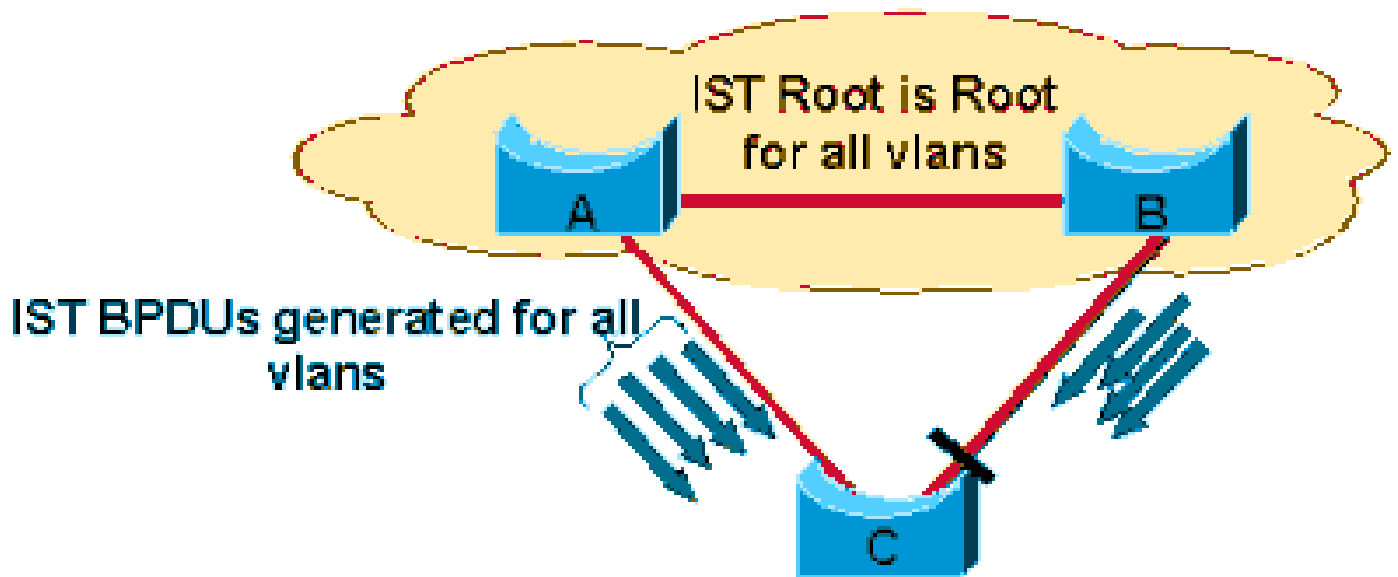
但是，此解決方案在MISTP中首次實施時過於複雜且可能存在危險。一種更簡單的方法被創造出來。

MST區域在所有VLAN上複製IST BPDU以模擬PVST+鄰居。此解決方案包含一些本文檔中討論的限制。

## 建議的配置

由於MST區域現在在邊界上的每個VLAN上複製IST BPDU，因此每個PVST+例項會偵聽到來自IST根的BPDU（這意味著根位於MST區域內）。

建議IST根的優先順序高於網路中的任何其他網橋，以便IST根成為所有不同PVST+例項的根，如下圖所示：



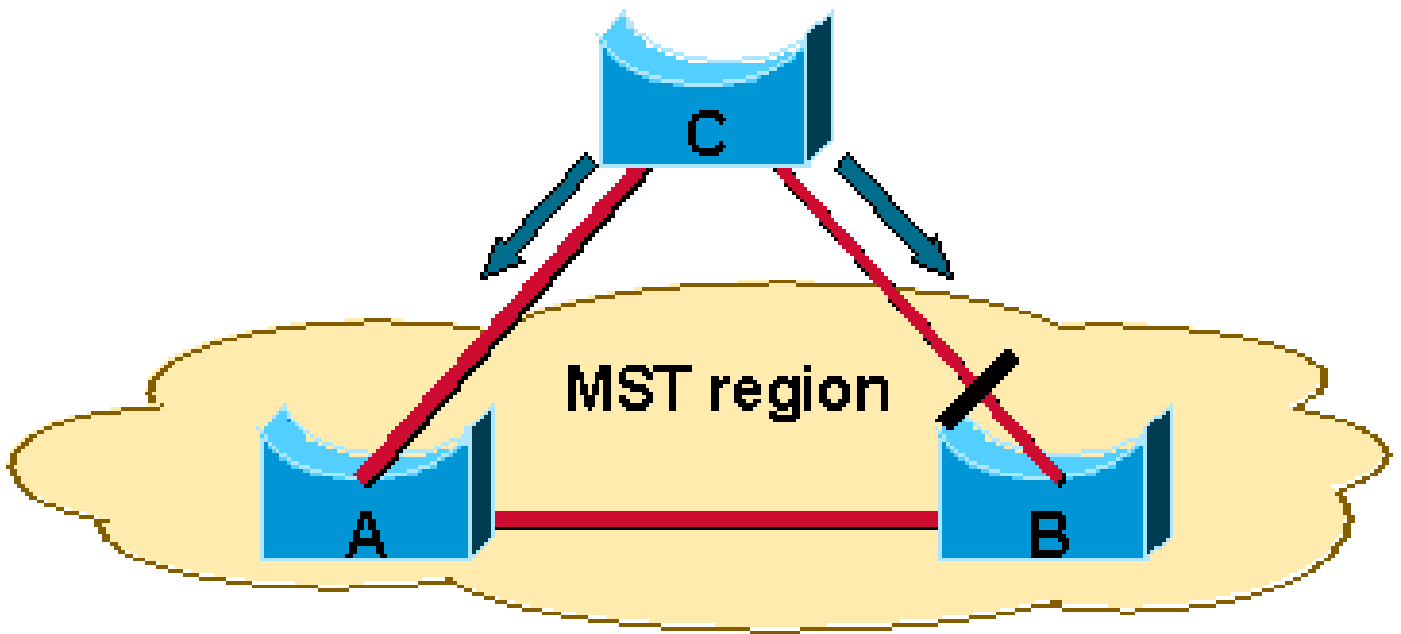
在此圖中，交換機C是冗餘連線到MST區域的PVST+。IST根是交換機C上存在的所有PVST+例項的根。

因此，交換機C會阻塞其上行鏈路之一，以防止出現環路。在此特定情況下，PVST+和MST區域之間的互動是最佳的，因為：

- 交換機C的上行鏈路埠成本可以調節，以實現上行鏈路埠上不同VLAN的負載均衡（因為交換機C在每個VLAN上運行一個生成樹，所以此交換機能夠選擇每個VLAN上的上行鏈路埠塊）。
- 可在交換機C上使用UplinkFast，以便在出現上行鏈路故障時實現快速收斂。

### 替代組態（不建議使用）

另一種可能是讓IST區域成為絕對沒有PVST+例項的根。這意味著所有PVST+例項都具有比IST例項更好的根，如下圖所示：



所有PVST+例項都有比IST例項更好的根

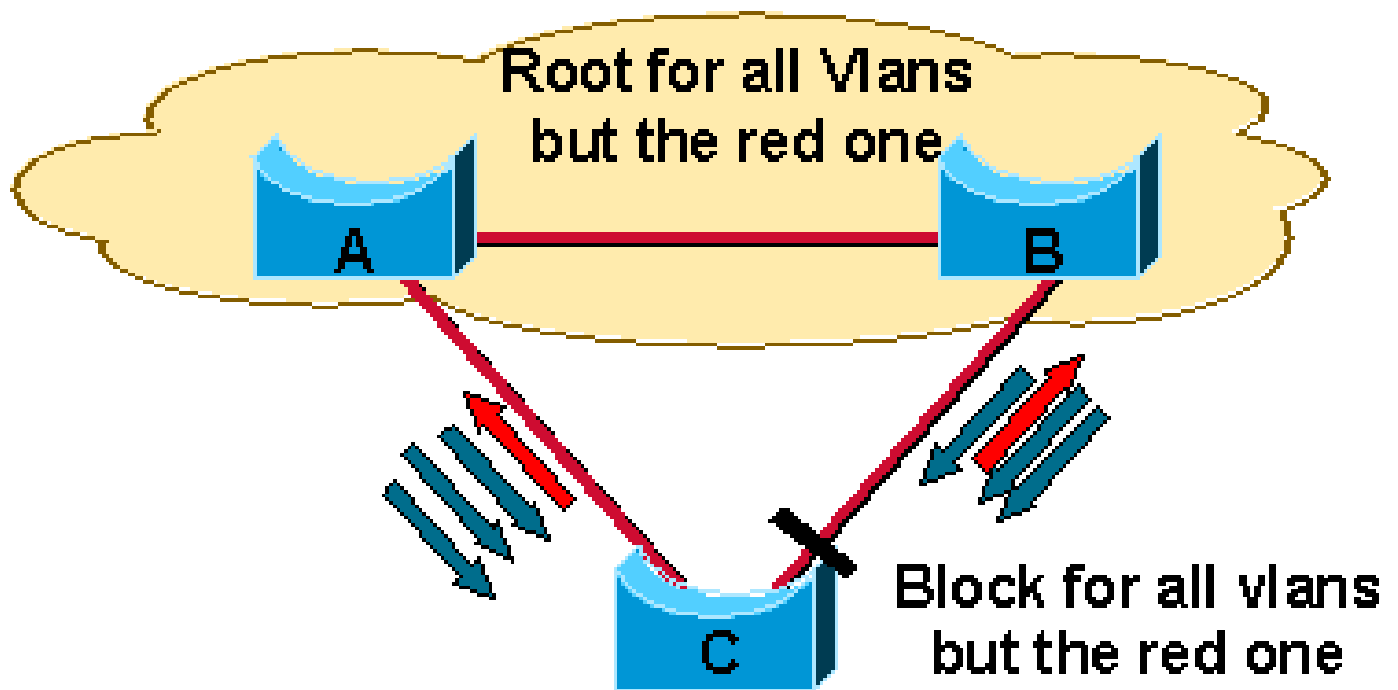
此案例對應於PVST+核心和MST接入層或分佈層，這種情況並不常見。如果在區域之外建立根網橋，則與之前建議的配置相比，存在以下缺點：

- MST區域只運行一個與外界互動的生成樹例項。這基本上意味著邊界埠只能阻塞或轉發所有VLAN。換句話說，在通向交換機C的區域中的兩個上行鏈路之間不可能實現負載均衡。交換機B上用於該例項的上行鏈路可能阻塞所有VLAN，而交換機A可能轉發所有VLAN。
- 此配置仍允許區域內快速收斂。如果交換機A上的上行鏈路發生故障，則需要快速切換到另一台交換機上的上行鏈路。雖然沒有詳細討論IST在區域內的行為方式，以使整個MST區域類似於CST網橋，但可以想象，跨區域的切換從未像跨單個網橋的切換那樣高效。

## 配置無效

雖然PVST+模擬機制在MST和PVST+之間提供了簡單無縫的互操作性，但此機制意味著除上述兩個配置之外的所有配置均無效。要成功進行MST和PVST+互動，必須遵循以下基本規則：

1. 如果MST網橋是根，則此網橋必須是所有VLAN的根。
2. 如果PVST+網橋是根，則此網橋必須是所有VLAN的根（包括CST，當CST運行PVST+時，CST始終在VLAN 1上運行，而不考慮本地VLAN）。
3. 如果MST網橋是CST的根網橋，而PVST+網橋是一個或多個其他VLAN的根網橋，則模擬將失敗並生成錯誤消息。失敗的模擬將邊界埠置於根不一致模式。



失敗的模擬將邊界埠置於根不一致模式

在此圖中，MST區域中的網橋A是除一個之外（紅色VLAN）的所有三個PVST+例項的根。網橋C是紅色VLAN的根網橋。

假設在紅色VLAN上建立的環路（其中網橋C是根）被網橋B阻塞。這意味著網橋B指定給除紅色的VLAN以外的所有VLAN。

MST區域無法執行此操作。由於MST區域只與外部世界運行一個生成樹，因此邊界埠只能阻塞或轉發所有VLAN。

因此，當網橋B在其邊界埠上檢測到更好的BPDU時，網橋會呼叫BPDU防護來阻塞此埠。埠處於根不一致模式。

完全相同的機制也會導致網橋A阻塞其邊界埠。連線丟失；但是，即使存在此類錯誤配置，仍會保留無環拓撲。

 注意：一旦邊界埠出現根不一致錯誤，請檢查PVST+網橋是否嘗試成為某些VLAN的根。

## 遷移策略

遷移到802.1s/w的第一步是正確辨識點對點和邊緣埠。確保需要快速轉換的所有交換機到交換機鏈路都是全雙工。

邊緣埠是透過PortFast功能定義的。請仔細確定交換網路中需要多少例項，並牢記例項會轉換為邏輯拓撲。

確定要對映到這些例項的VLAN，然後仔細為每個例項選擇根和備用根。



選擇網路中的所有交換機可以共用的配置名稱和修訂版號。

思科建議您將儘可能多的交換機放置在單個區域中；將網路分段為單獨的區域並不有利。

避免將任何VLAN對映到例項0。首先遷移核心。將STP型別更改為MST，然後按照您的方式進入接入交換機。

MST可以基於每個埠與運行PVST+的傳統網橋互動，因此，如果清楚地瞭解互動，混合這兩種型別的網橋就不是問題。

請始終嘗試將CST和IST的根保留在區域內部。如果透過中繼與PVST+網橋互動，請確保MST網橋是該中繼上允許的所有VLAN的根網橋。

有關配置示例，請參閱：

- [將生成樹從PVST+遷移到MST的配置示例](#)

## 結論

交換網路必須滿足嚴格的穩定性、恢復能力和高可用性要求。

隨著IP語音(VoIP)和IP影片等新技術的出現，圍繞鏈路或元件故障的快速收斂不再是理想特徵：必須實現快速收斂。

但是，直到最近，冗餘交換網路仍不得不依賴相對緩慢的802.1d STP來實現這些目標。這經常成為一項挑戰網路管理員的任務。

只需調整協定計時器，但常常會損害網路的運行狀況，從而縮短協定幾秒鐘的時間。

Cisco發佈了許多802.1d STP增強功能，如UplinkFast、BackboneFast和PortFast。這些功能為更快的生成樹收斂鋪平了道路。

思科還在MISTP的開發中解決了大型第2層(L2)網路的可擴充性問題。IEEE最近決定將大多數這些概念納入兩個標準：802.1w (RSTP)和802.1s (MST)。

透過實施這些新協定，您可在幾百毫秒的時間內實現收斂，同時還可擴展到數千個VLAN。

思科仍然是行業中的領導者，提供這兩種協定以及專有增強功能，以便促進傳統網橋的遷移以及與傳統網橋的互操作性。

## 相關資訊

- [瞭解快速跨距樹狀目錄通訊協定 \(802.1w\)](#)
- [LAN 交換技術支援](#)
- [思科技術支援與下載](#)

## 關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。