

STP 1 - Layer 2 Loop Issue

二層網路冗餘、網路環路危害（廣播風暴/未知單播）與生成樹協定（STP）防範機制深度解析

1. Layer 2 冗餘的起源與需求 (Layer 2 Redundancy)

在區域網路（LAN）設計中，可靠性是核心指標之一。傳統的單路徑網路架構雖然簡單，但存在單點故障（Single Point of Failure, SPOF）的巨大風險。

Layer 2 Redundancy

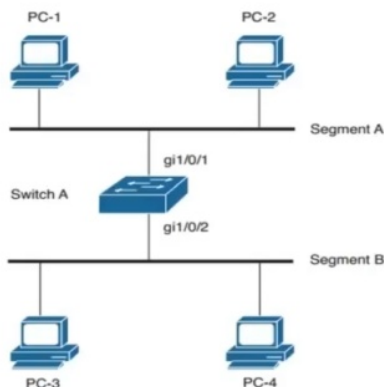


Figure 1: Transparent Bridging with a Switch

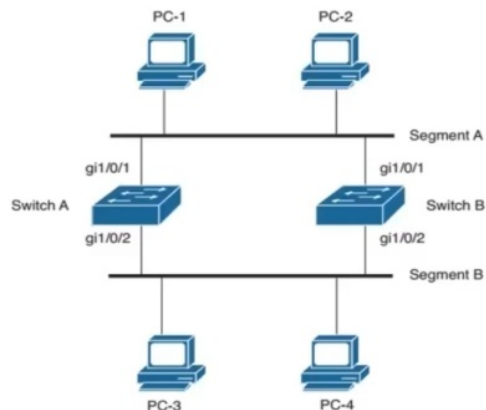


Figure 2: Redundant Bridging with Two Switches

Figure 1 shows a simple two-port switch forwarding frames between devices. However, this network design offers no additional links or paths for redundancy if the switch or one of its links fails.

Figure 2 shows a second switch been added to provide redundancy. One single switch can fail without causing end-to-end connectivity fail.

未经授权，翻版必究

圖 1：單一交換機非冗餘架構 (Figure 1) 與 雙交換機冗餘架構 (Figure 2)

💡 冗餘架構對比：

- **Figure 1 (Transparent Bridging with a Switch):** 左圖僅有一台 Switch A 連接 Segment A 與 Segment B。當 Switch A 發生故障或其鏈路中斷時，PC-1/PC-2 與 PC-3/PC-4 之間的通訊將完全中斷。
- **Figure 2 (Redundant Bridging with Two Switches):** 右圖引入第二台交換機 Switch B 提供冗餘 (Redundancy)。當任意一台交換機故障時，備用交換機可維持端到端的連通性。

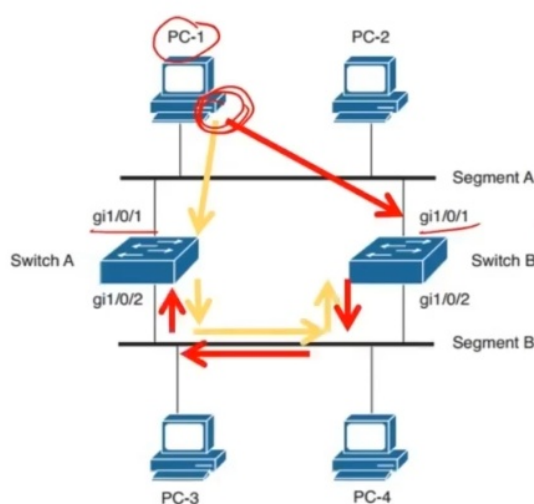
⚠️ 冗餘引發的核心代價：網路環路 (Layer 2 Loop)

雖然冗餘提高了可靠性，但若直接在二層（Layer 2）環境中建立多條實體路徑，且沒有特殊協議干預，將會導致致命的**二層環路 (Layer 2 Loop)**。

2. 二層環路問題一：廣播風暴 (Broadcast Storm)

乙太網路（Ethernet）二層報頭中沒有 TTL（Time to Live，生存時間）機制（TTL 僅存在於 IP 三層報頭）。因此，廣播訊框一旦在環路中產生，將無休止地循環轉發，形成「廣播風暴」。

Layer 2 Loop - Broadcast Storm



Now, consider PC1 is sending **broadcast packet**

1. Switch A and B will receive broadcast frame from gi1/0/1
2. Switch A and B will forward the frame out from gi1/0/2
3. Switch A and B will receive broadcast frame from gi1/0/2
4. Switch A and B will forward the frame out from gi1/0/1

未经授权，翻版必究

圖 2：PC-1 發送廣播數據包引發廣播風暴過程示意圖

🔄 廣播風暴持續循環步驟分析：

- 1 **PC-1 發送廣播訊框**：PC-1 發起廣播數據包（例如 ARP Request）。Switch A 與 Switch B 均在 `gi1/0/1` 埠收到該訊框。
- 2 **二層泛洪轉發**：Switch A 與 Switch B 依照廣播處理原則，將訊框由各自的 `gi1/0/2` 埠泛洪（Flooding）輸出至 Segment B。
- 3 **交叉接收與再轉發**：Switch A 的 `gi1/0/2` 收到 Switch B 剛發出的廣播訊框，Switch B 的 `gi1/0/2` 也收到 Switch A 剛發出的廣播訊框。
- 4 **無窮盡循環**：兩台交換機再次將廣播訊框從 `gi1/0/1` 埠泛洪回 Segment A，導致數據包在環路中無窮無盡地迴圈，瞬間耗盡所有頻寬與 CPU 資源。

3. 二層環路問題二：未知單播與 MAC 位址表震盪 (Unknown Unicast & MAC Flapping)

除了廣播風暴，當交換機收到**未知單播**（Unknown Unicast）訊框時，同樣會進行泛洪，進而導致 MAC 位址轉發表（MAC Address Table）發生持續錯亂與震盪。

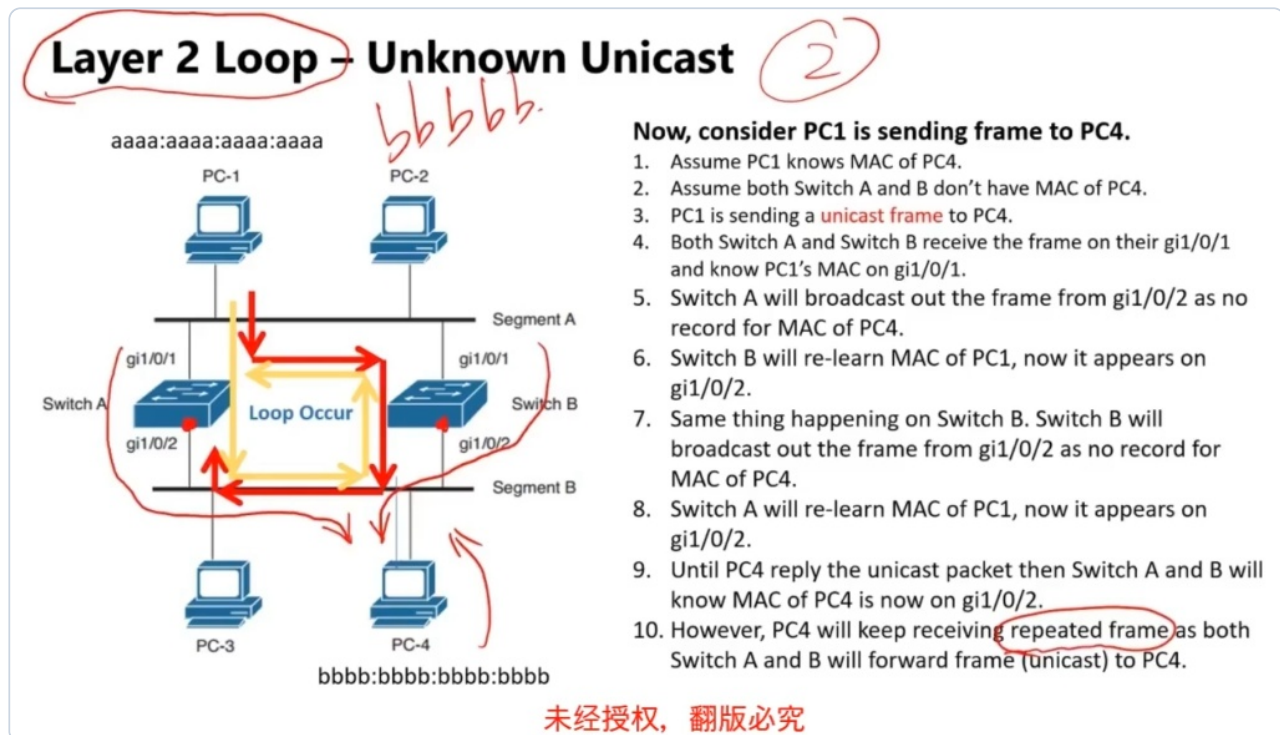


圖 3：未知單播導致 MAC 位址震盪與重複訊框接收

未知單播轉發與 MAC 震盪流程：

- 1 **初始條件**：PC-1 (MAC: aaaa:aaaa:aaaa:aaaa) 欲發送單播訊框給 PC-4 (MAC: bbbb:bbbb:bbbb:bbbb)。PC-1 知道 PC-4 的 MAC 位址，但 Switch A 與 Switch B 的 MAC 位址表中均無 PC-4 的紀錄。
- 2 **學習與泛洪**：Switch A 與 Switch B 從 gi1/0/1 收到訊框，學習到 PC-1 位於 gi1/0/1 。由於不知道 PC-4 在何處，兩台交換機皆將訊框由 gi1/0/2 泛洪輸出。
- 3 **MAC 位址表覆蓋（震盪）**：Switch A 由 gi1/0/2 收到來自 Switch B 轉發的 PC-1 訊框，誤以為 PC-1 轉移到了 gi1/0/2 ，因而更新 MAC 位址表。同理，Switch B 也將 PC-1 的埠更新為 gi1/0/2 。
- 4 **重複訊框接收（Repeated Frames）**：PC-4 會持續收到兩台交換機重複轉發的相同訊框。在 PC-4 回應單播包之前，MAC 表持續翻轉，整個網路處於極度不穩定狀態。

4. 生成樹協定 (STP) 的導入與防環機制

為了兼顧「網路冗餘」與「避免環路」，IEEE 制定了 **生成樹協定 (Spanning Tree Protocol, STP / IEEE 802.1D)**。

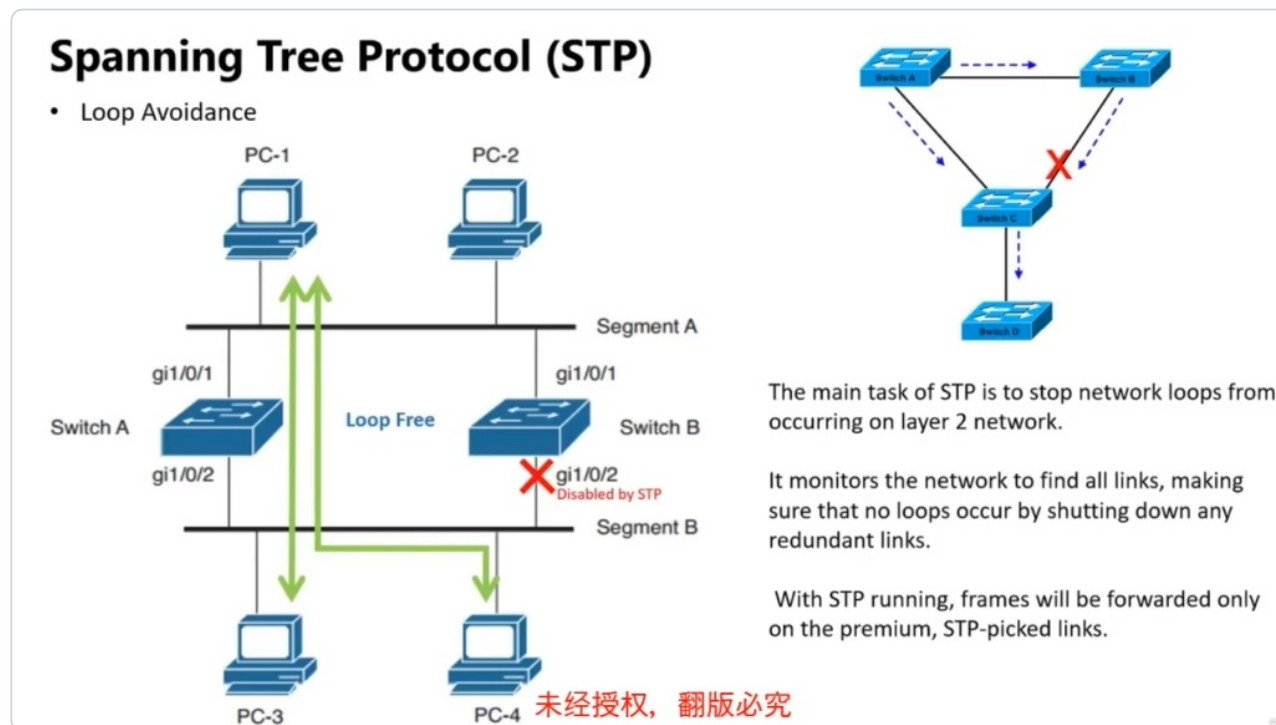


圖 4：STP 阻斷冗餘埠以實現無環路拓撲 (Loop-Free Topology)

STP 的核心工作機制：

- **防止二層環路 (Loop Avoidance)**：STP 的主要任務是防止二層網路中發生網路環路。
- **動態監控與邏輯修剪**：STP 會持續監控整個拓撲的所有鏈路，通過邏輯上關閉 (**Disabled / Blocked**) 冗餘鏈路上的特定端口 (如圖中 Switch B 的 `gi1/0/2` 埠)，打破物理環路。
- **最優路徑轉發**：在 STP 運行下，數據訊框僅會在 STP 選定的最優、無環路鏈路上進行轉發 (Loop Free Path)。
- **自動故障轉移 (Failover)**：當活動鏈路發生故障時，STP 會自動重新計算並啟用原本被阻塞的備用端口，恢復網路通訊。

5. 橋接協定資料單元 (Bridge Protocol Data Unit, BPDU)

交換機之間通過交換一種特殊的二層數據訊框——**BPDU** 來互相溝通、計算生成樹拓撲並達成網路共識。

Bridge Protocol Data Units (BPDU)

Bridge Protocol Data Units (BPDUs) are frames that contain information about the STP, there are TWO types of BPDU:

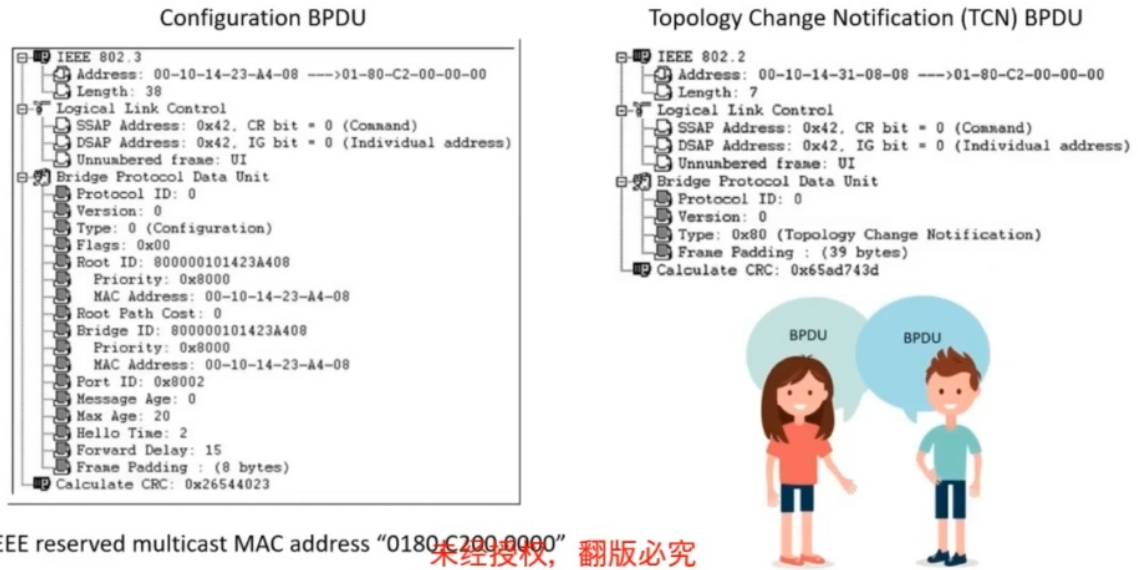


圖 5：Configuration BPDU 與 TCN BPDU 的 Wireshark 封包結構與 Multicast MAC

BPDU 的兩大類型與關鍵欄位：

BPDU 主要分為兩種型態：

- Configuration BPDU (組態 BPDU)：**用于根橋 (Root Bridge) 選舉及生成樹拓撲計算。內含關鍵參數：
 - Root ID：目前網路中根橋的識別碼。
 - Bridge ID：發送該 BPDU 的交換機識別碼 (包含 Priority + MAC)。
 - Port ID：發送該 BPDU 的端口識別碼。
 - Hello Time (預設 2 秒)、Max Age (預設 20 秒)、Forward Delay (預設 15 秒)。
- Topology Change Notification (TCN) BPDU：**當網路拓撲發生變更 (如鏈路 Up/Down) 時，由下游交換機向上游發送，用以通知全網更新 MAC 位址表老化時間。

BPDU 專屬多播 MAC 位址：

交換機如何辨識收到的訊框是否為 STP BPDU？

BPDU 使用 IEEE 預留的特定組播/多播 (Multicast) MAC 位址：**01-80-C2-00-00-00**。當二層交換機接收到目的 MAC 為此位址的訊框時，會直接送交 CPU 的 STP 引擎處理，而不會進行普通數據的泛洪。

6. STP 核心重點彙整

分析項目	無 STP 護航（單純冗餘）	啟用 STP 護航
拓撲狀態	存在實體與邏輯環路	邏輯修剪為樹狀無環拓撲 (Loop-Free)
廣播訊框處理	引發廣播風暴，耗盡 CPU/頻寬	正常泛洪，無無窮迴圈問題
MAC 位址表	頻繁覆蓋震盪 (MAC Flapping)	穩定紀錄，精確轉發
溝通協定	無機制	BPDU (Multicast MAC: 01-80-C2-00-00-00)