

STP Part 3: 802.1D Port States & Convergence

IEEE 802.1D 埠狀態遷移、計時器運作、拓撲收斂與 TCN 機制全解析

主題: Spanning Tree Protocol (STP 3 - Port States and Converge) 標準規範: IEEE 802.1D

核心模組: 5 大埠狀態 / STP Timers / 拓撲收斂 (30s vs 50s) / TCN BPDU & Aging 刷新

1. IEEE 802.1D STP 五大埠狀態 (Port States)

在經典 IEEE 802.1D 生成樹協定中，為了防止暫態環路 (Transient Loops) 以及因拓撲變更瞬間導致廣播風暴，交換機介面在啟用後不會立即進入轉發 (Forwarding) 狀態，而是必須循序漸進地通過多個狀態階段。

五大標準埠狀態詳解

- **Disabled (禁用狀態)**：介面處於管理性關閉 (Administratively Down, shutdown) 或實體故障狀態。此狀態下介面既不發送也不接收任何數據訊框或 BPDU。
- **Blocking (阻塞狀態)**：介面完成初始化後預設進入的狀態，或是選舉失敗未成為 RP / DP 的非指定埠。此狀態**只接收 (Receive Only) BPDU**，不發送 BPDU、不轉發用戶數據、不學習 MAC 地址，負責持續監聽鄰居交換機的健康狀態。
- **Listening (監聽狀態)**：介面被選為 Root Port 或 Designated Port 後進入此狀態。此時介面**允許雙向收發 BPDU**，以建立並確認活性樹狀拓撲。若在該狀態期間失去了 RP 或 DP 資格，會立即退回 Blocking 狀態。此階段**不學習 MAC 地址**，亦不轉發用戶數據。
- **Learning (學習狀態)**：監聽期滿 (Forward Delay 15s) 後進入學習狀態。此狀態持續雙向收發 BPDU，同時**開始監聽流量並學習來源 MAC 地址寫入 MAC Address Table**，但依然**不轉發普通數據訊框**，以避免未建表前引發全網 Flooding。
- **Forwarding (轉發狀態)**：學習期滿 (Forward Delay 15s) 後正式進入轉發狀態。此時介面具備完整功能：**轉發/接收用戶數據、持續學習 MAC 地址，並收發 BPDU**。

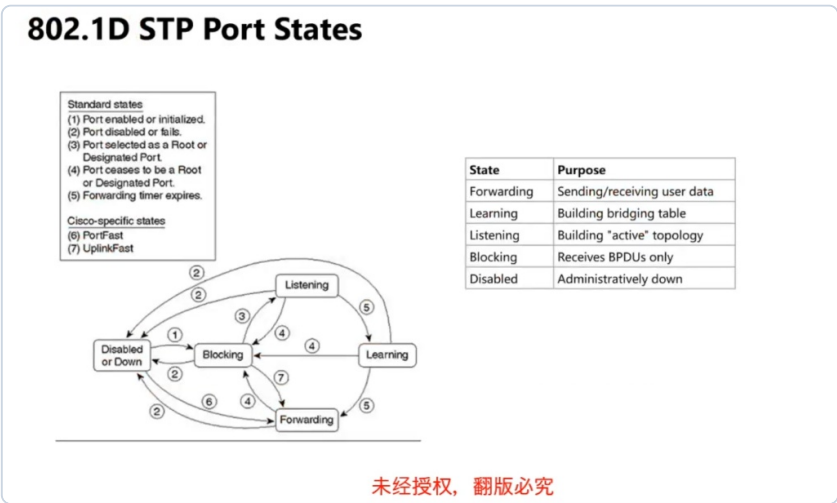


圖 1：802.1D STP 五大埠狀態定義與收發行為特徵

埠狀態功能對照矩陣

埠狀態 (Port State)	接收 BPDU	發送 BPDU	學習 MAC 地址	轉發用戶數據訊框	主要目的 / 角色功能
Disabled	No	No	No	No	介面關閉 (Administratively Down)
Blocking	Yes	No	No	No	阻斷環路，僅監聽 BPDU (Receives BPDUs only)

埠狀態 (Port State)	接收 BPDU	發送 BPDU	學習 MAC 地址	轉發用戶數據 訊框	主要目的 / 角色功能
Listening	Yes	Yes	No	No	計算並建立活性拓撲 (Building "Active" Topology)
Learning	Yes	Yes	Yes	No	預填轉發表，學習 MAC (Building Bridging Table)
Forwarding	Yes	Yes	Yes	Yes	正常轉發用戶數據 (Sending/Receiving User Data)

2. 埠狀態遷移機制與 Cisco 擴充增強技術

交換機介面在不同事件觸發下，會於五大狀態間進行轉換。下圖詳細標示了狀態遷移流程以及各項觸發條件：

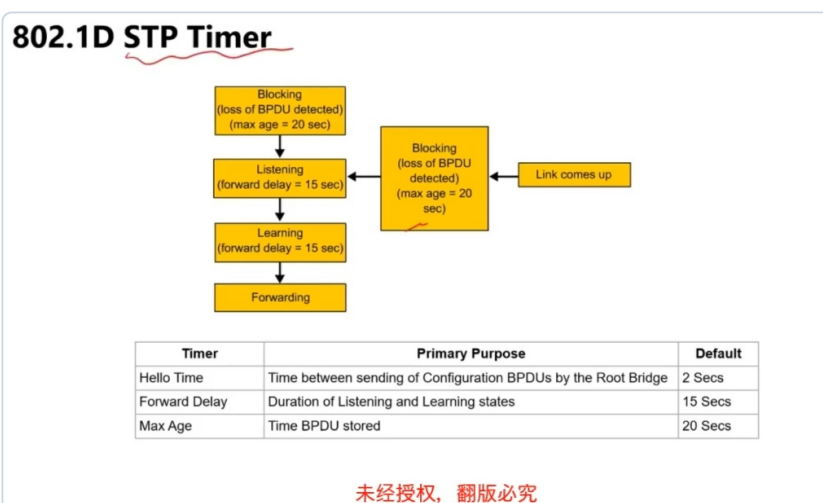


圖 2：802.1D 狀態轉換狀態機 (State Machine) 與 Cisco 專屬加速技術遷移路徑

標準狀態遷移事件說明 (Standard Transition Events)：

- **(1) Port enabled or initialized:** 介面啟用或交換機開機初始化，直接從 Disabled 進入 Blocking 狀態。
- **(2) Port disabled or fails:** 介面被管理員關閉或鏈路發生故障，任何狀態皆立即退回 Disabled 狀態。
- **(3) Port selected as RP or DP:** 介面在 STP 選舉中贏得 Root Port 或 Designated Port 角色，由 Blocking 進入 Listening 狀態。
- **(4) Port ceases to be RP or DP:** 介面在重新計算時失去 RP/DP 資格，無論處於 Listening、Learning 或 Forwarding，均立即退回 Blocking 狀態。
- **(5) Forwarding timer expires:** 轉發計時器 (Forward Delay = 15s) 到期，驅動介面從 Listening 進入 Learning，再等 15s 後進入 Forwarding。

Cisco 專屬狀態遷移技術 (Cisco-Specific States)：

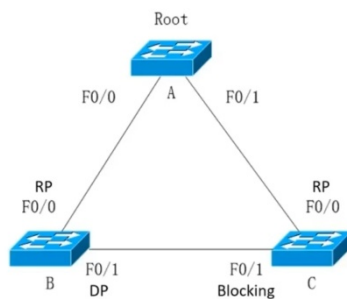
- **(6) PortFast (路徑 6)：** 配置於連接終端主機 (PC/Server) 的 Access 埠。介面在啟用後**直接由 Disabled 跨越 Blocking/Listening/Learning，瞬間進入 Forwarding 狀態 (0 秒收斂)**，大幅避免 DHCP 逾時與終端連線延遲。
- **(7) UplinkFast (路徑 7)：** 配置於接入層交換機。當直接上行鏈路 (Root Port) 斷線時，**備援 Blocking 埠無需等待 30 秒，直接跳轉進入 Forwarding 狀態**，實現次秒級切換。

3. 802.1D STP 三大核心計時器 (STP Timers)

STP 的拓撲計算與狀態轉換完全依賴三大標準計時器。所有的計時器參數均由 **Root Bridge 統一決定**，並透過 Configuration BPDU 宣告給全網所有下游交換機遵循。

計時器 (Timer)	預設數值	主要用途與功能描述
Hello Time	2 秒	Root Bridge 發送 Configuration BPDU 的時間間隔 (Time between sending of Configuration BPDUs)。
Forward Delay	15 秒	介面在 Listening 狀態與 Learning 狀態各自停留的維持時間 (Listening 15s + Learning 15s = 30s)。
Max Age	20 秒	BPDU 儲存的最大有效壽命。若連續超過 Max Age 未收到上游發送的 BPDU，交換機判定該 BPDU 逾期失效 (Loss of BPDU detected)。

802.1D STP Converge



50 second case:

- Link between A and B down
- C F0/1 starts Max Age Timer for 20 second
- C F0/1 goes into Listening state for 15 second
- C F0/1 goes into Learning state for 15 second
- STP Converge (Forwarding state)

30 second case:

- Link between A and C down
- C lost RP, F0/1 goes into Listening state for 15 second (bypass Max Age Timer)
- C F0/1 goes into Learning for 15 second
- STP Converge (Forwarding state)

未经授权，翻版必究

圖 3：STP 計時器與狀態前進時間流程圖 (Max Age 20s → Listening 15s → Learning 15s)

4. 802.1D STP 收斂時間實戰解析：30 秒 vs 50 秒

當網路發生故障時，STP 重新計算並完成收斂所需的時間，取決於是**直接鏈路故障 (Direct Failure)** 還是**間接鏈路故障 (Indirect Failure)**。

5 Minutes Down Time

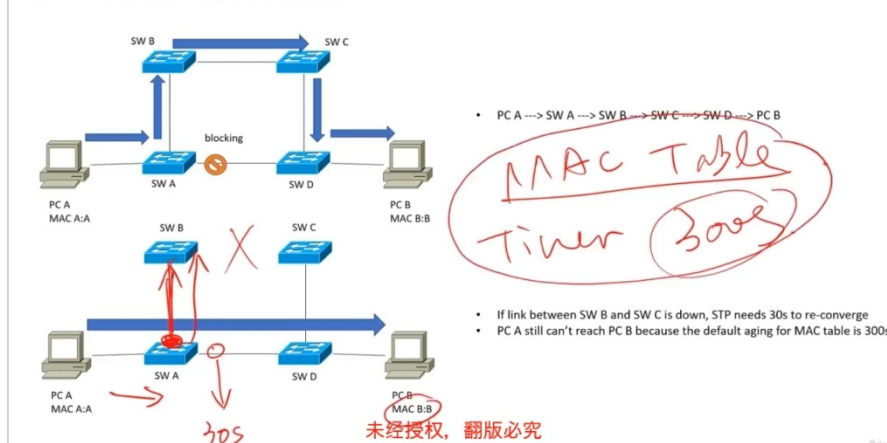


圖 4：STP 收斂實例：間接故障 50 秒 vs 直接故障 30 秒分析圖

情境：Switch A 與 Switch B 之間的鏈路中斷 (Link between A and B down)。

- ### 案例二：30 秒收斂 (直接故障 / Direct Link Failure)

情境： Switch A 與 Switch C 之間的直接鏈路中斷 (Link between A and C down)。

- ## 5. 拓撲變更與 5 分鐘斷網危機 (5 Minutes Down Time Problem)

STP Topology Change Mechanism

- As soon as a bridge detects a change in the topology of the network (a link that goes down or goes to forwarding), it advertises the event to the whole bridged network.
- Every bridge is then notified and reduces the aging time to forward_delay (15 seconds by default) for a certain period of time (max_age + forward_delay, 20+15=35 seconds by default).
- It is more beneficial to reduce the aging time instead of clearing the table because currently active hosts, that effectively transmit traffic, are not cleared from the table.
- The notification is through TCN BPDU.
- Definition of "Topology Change":
 - When a port that was **forwarding is going down (blocking for instance)**
 - When a port **transitions to forwarding** and the bridge has a designated port. (This means that the bridge is not standalone.)
- Two steps to send notification to all switches in the network
 - The bridge notifies the root bridge of the spanning tree
 - The root bridge "broadcasts" the information into the whole network



圖 5：MAC 表 300 秒老化時間引發的 5 分鐘流量黑洞問題

1. 正常狀態下，PC A 與 PC B 之間的流量經過 SW A → SW B → SW C → SW D 傳送。此時 SW A 的 MAC 表記錄「PC B 的 MAC 地址位於朝向 SW B 的上行介面」。
2. 當 SW B 與 SW C 之間的鏈路中斷時，STP 經過 30 秒重新收斂，解除 SW A 與 SW D 之間的 Blocking 狀態。

3. **危機產生**：雖然 STP 拓撲已打通新路徑，但 SW A 記憶體中的 MAC 表依然殘留舊記錄 (有效時間 300 秒)。因此當 PC A 發送單播流量給 PC B 時，SW A 依然死板地將訊框丟往已經斷線的 SW B，導致**長達 5 分鐘流量完全無法送達 PC B**！

6. STP 拓撲變更通知機制 (TCN Mechanism)

為徹底解決 300 秒 MAC 老化引發的斷網問題，IEEE 802.1D 設計了專屬的 **Topology Change Notification (TCN)** 機制。

Topology Change 的嚴格定義

交換機在以下情況判定發生了拓撲變更 (Topology Change)：

- 當原本處於 **Forwarding** 狀態的介面變為 **Down** 或 **Blocking** 狀態。
- 當介面成功轉換進入 **Forwarding** 狀態，且交換機至少擁有一個 **Designated Port** (非孤立 Standalone 節點)。

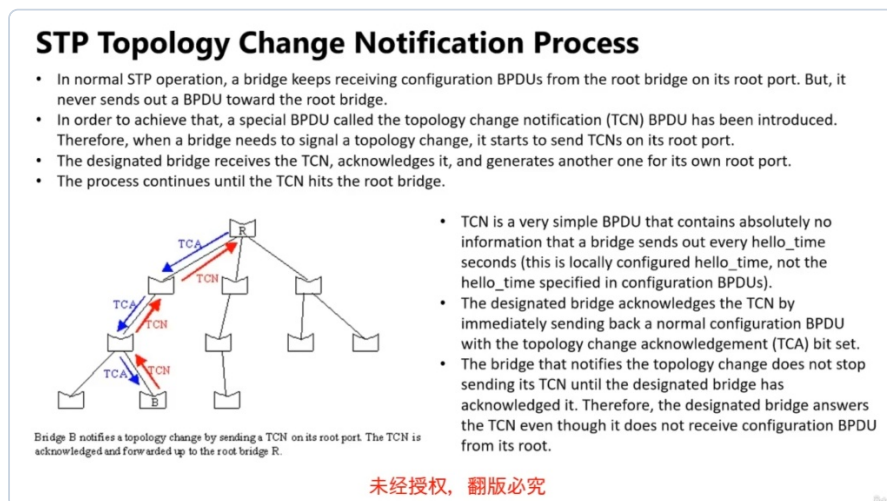


圖 6：TCN 兩階段通報流程：下游往根橋發送 TCN，根橋全網廣播 TC 旗標

TCN 的核心解法：縮短 MAC 老化時間為 Forward Delay (15 秒)

當拓撲變更發生時，交換機不會粗暴地直接清空 (Clear/Flush) 整個 MAC 地址表，因為這會造成全網廣播泛洪。相反地，**STP 會通知全網所有交換機，將 MAC 老化時間從預設的 300 秒暫時縮短為 Forward Delay (15 秒)**，維持期間為 $\text{Max Age} + \text{Forward Delay} = 35$ 秒。這使得無效的舊路徑 MAC 記錄在 15 秒內迅速自然老化，而持續活躍通信的主機則不會受到影響。

7. TCN 封包交互詳細流程 (Two-Step Notification Process)

在 STP 正常運作下，非根交換機只會從 Root Port 接收來自 Root Bridge 的 BPDU，絕不會向上游主動發送 Configuration BPDU。因此拓撲變更通報採用嚴謹的兩階段流程：

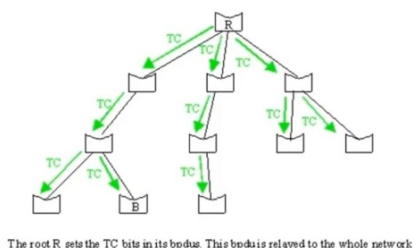
第一階段：由下而上通報根橋接器 (TCN & TCA)

- 發起 TCN**：偵測到拓撲變更的交換機 (Bridge B)，透過其 **Root Port** 向上游 **Designated Bridge** 發送 **TCN BPDU**。TCN 封包非常精簡，僅包含 Protocol ID、Version 與 Type (0x80)，不帶任何參數。

2. **逐跳發送與 TCA 確認**：交換機每隔 Hello Time（2秒）持續發送 TCN。上游指定橋接器收到後，會在下一次發送的 Configuration BPDU 中將 **Topology Change Acknowledgment (TCA) 位元設為 1** 回應給下游，通知下游停止發送 TCN。
3. **向上接力轉發**：上游指定橋接器接力透過自己的 Root Port 繼續向上發送 TCN，直到 TCN 正式抵達 **Root Bridge (R)** 為止。

STP Topology Change Notification Process

- Once the root is aware that there has been a topology change event in the network, it starts to send out its configuration BPDUs with the topology change (TC) bit set.
- These BPDUs are relayed by every bridge in the network with this bit set. As a result all bridges become aware of the topology change situation and it can reduce its aging time to forward_delay.
- As a result all bridges become aware of the topology change situation and it can reduce its aging time to forward_delay.



未经授权，翻版必究

圖 7：第一階段：下游交換機沿著樹狀路徑向上通報 TCN，上游以 TCA 確認

第二階段：由根橋接器全網廣播 TC 旗標 (TC Bit Broadcast)

1. **Root Bridge 廣播 TC**：Root Bridge 得知全網拓撲變更後，開始在向下發送的 Configuration BPDU 中將 **Topology Change (TC) 旗標位元設為 1**。
2. **全網擴散**：全網所有下游交換機收到帶有 TC=1 的 BPDU 後，立即將自身的 **MAC Table 老化時間由 300 秒縮短為 15 秒 (Forward Delay)**。
3. **廣播持續時間**：Root Bridge 會持續發送帶有 TC 旗標的 BPDU 長達 $\text{Max Age} + \text{Forward Delay} = 20s + 15s = 35 \text{ 秒}$ ，確保全網所有交換機皆完成 MAC 表老化刷新。

STP Topology Change Case Study

If the link between Switch A and Switch C goes down, the network must re-converge into a stable STP topology. The sequence of events as follows.

1. Switch C detected a link down on its port g1/0/1; Switch A detected a link down on its port g1/0/2.
2. Switch C failed to send Topology Change Notification(TCN) to root bridge because its root port is down.
3. Switch A send TCN out its port g1/0/1. This is received and replied by other switches among the way.
4. Switch B and Switch C receive TCN message.
5. Switch C received TCN message on its port g1/0/2, which was Blocking states previously, now it becomes the new root port.

未经授权，翻版必究

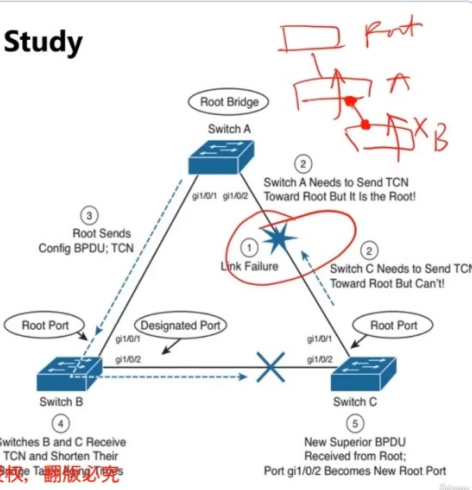


圖 8：第二階段：Root Bridge 全網廣播 TC 旗標，全網交換機縮短 MAC 老化時間為 15 秒

8. 拓撲變更實戰個案分析與 TCN 泛洪效應 (TCN Flood)

STP 拓撲變更個案推演 (Case Study)

參照下圖經典三台交換機拓撲（Switch A 為 Root Bridge）：

STP Topology Change Case Study

If the link between Switch A and Switch C goes down, the network must re-converge into a stable STP topology. The sequence of events as follows.

1. Switch C detected a link down on its port g1/0/1; Switch A detected a link down on its port g1/0/2.
2. Switch C failed to send Topology Change Notification(TCN) to root bridge because its root port is down.
3. Switch A send TCN out its port g1/0/1. This is received and replied by other switches among the way.
4. Switch B and Switch C receive TCN message.
5. Switch C received TCN message on its port g1/0/2, which was Blocking states previously, now it becomes the new root port.

未经授权，翻版必究

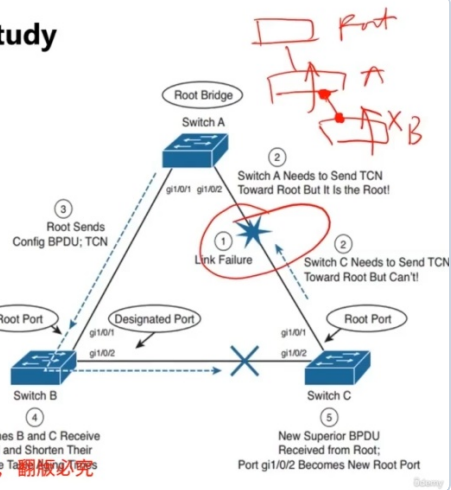


圖 9：Switch A 與 Switch C 鏈路斷線時的 TCN 處理與狀態轉換推演

事件推演順序步驟：

1. **(1) 鏈路故障：** Switch A 與 Switch C 之間的鏈路中斷 (Link Failure)。Switch C 偵測到 g1/0/1 Down，Switch A 偵測到 g1/0/2 Down。
2. **(2) 嘗試發送 TCN：** Switch C 想要發送 TCN 通知 Root Bridge，但由於其 Root Port 已經斷線，**無法直接向上通報！**
3. **(3) Switch A 發起通報：** 同時身為 Root Bridge 的 Switch A 偵測到自己原本 Forwarding 的介面斷線，直接透過 g1/0/1 發送帶有 TCN/TC 旗標的 BPDU 給 Switch B。
4. **(4) 接收並縮短 MAC 老化：** Switch B 與 Switch C 接收到來自 Root 的訊息，立即將 MAC 表老化時間縮短為 15 秒。
5. **(5) 狀態遷移完成：** Switch C 收到來自 Root 經由 Switch B 轉發的 Superior BPDU，原先處於 Blocking 的 g1/0/2 歷經 Listening (15s) 與 Learning (15s)，正式成為新的 **Root Port**，全網恢復連通！

TCN 泛洪帶來的副作用 (TCN Flood Side Effect)

網路中主機越多，主機開關機引發鏈路 Up/Down 造成 Topology Change 的機率越高。當 TCN 頻繁觸發時，會引發 **TCN Flood 效應**：

TCN Flood

The more hosts are in the network, the higher are the probabilities of getting a topology change.

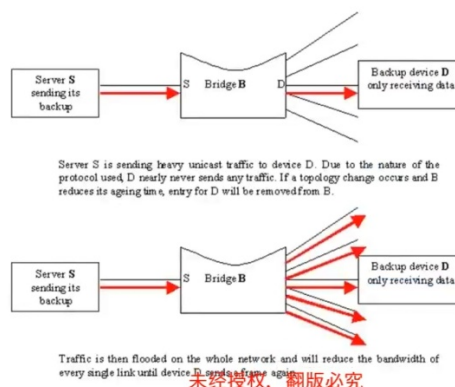


圖 10：頻繁 TCN 導致靜默設備 (Backup Device D) MAC 條目遭提早清除引發單播泛洪

TCN 泛洪的影響與防範：

- **靜默設備流量泛洪：**如圖 10 所示，若 Server S 持續單向傳送大量備份數據給備份設備 D，而 D 鮮少主動發送流量。當其他無關鏈路觸發 TCN 導致交換機 MAC 表老化時間縮短為 15 秒時，交換機內 D 的 MAC 條目會被提早清除。交換機因找不到目的 MAC，被迫將高頻寬的備份流量**泛洪 (Flooding) 到全網所有鏈路**，嚴重消耗網路頻寬！
- **最佳解法 (PortFast)：**在所有連接終端電腦、伺服器的介面上啟用 **Cisco PortFast**。啟用 PortFast 的介面在 Up/Down 時**絕不會向網路產生任何 TCN BPDU**，徹底根除 TCN 泛洪！

9. 總結與核心考訊重點 (Key Summary)

STP 3 核心考訊速記：

- **五大狀態收發行為：**Blocking (只收 BPDU)、Listening (雙向收發 BPDU)、Learning (收發 BPDU + 學習 MAC)、Forwarding (轉發數據 + 學習 MAC + 收發 BPDU)、Disabled (介面關閉，完全不運作)。
- **三大計時器預設值：**Hello Time = 2s, Forward Delay = 15s, Max Age = 20s。
- **收斂時間差異：**直接故障 (Direct Failure) Bypass Max Age 耗時 **30 秒**；間接故障 (Indirect Failure) 需等待 Max Age 逾時，耗時 **50 秒**。
- **MAC 老化問題：**預設 MAC 表老化時間為 300 秒 (5 分鐘)，若未及時刷新會導致流量黑洞。
- **TCN 解決之道：**拓撲變更時發送 TCN 至 Root Bridge，由 Root 廣播 TC 旗標，全網暫時將 **MAC 表老化時間縮短為 Forward Delay (15 秒)**。
- **防範 TCN 泛洪：**終端介面務必配置 spanning-tree portfast，避免終端設備開關機產生 TCN 擾亂全網。