

Le protocole STP

STP (Spanning Tree Protocol) est un protocole réseau de couche 2. Il est défini dans la norme IEEE 802.1D.

STP apporte une solution au problème posé par la présence de boucles dans les réseaux commutés de type Ethernet.

C'est quoi le problème et comment STP résout ce problème ?

Pour aller d'un point à un autre, un réseau commuté de type Ethernet doit avoir au moins deux chemins : Un chemin principal et un chemin alternatif au cas où il y a une coupure du chemin principal ou une panne d'un switch.

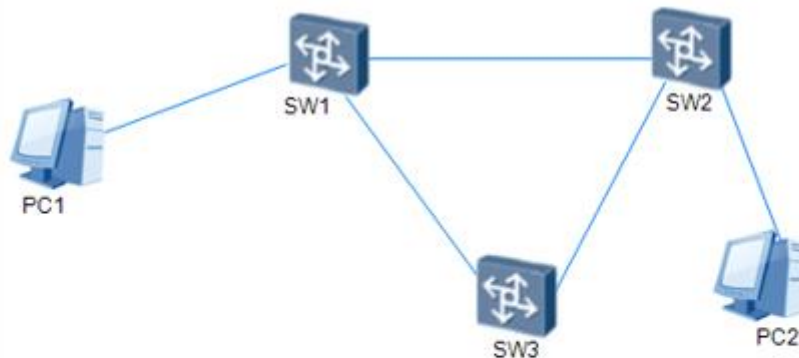


Figure 1 : Réseau commuté redondant

PC 1 et PC2 sont des stations terminales ; SW1, SW2 et SW3 sont des switches

Soit le réseau de la figure 1. Il est clair que la machine PC1 peut communiquer avec la machine PC2 en empruntant soit :

- Le chemin SW1-SW2,
- Le chemin SW1-SW3-SW2.

On parle de redondance. En cas de défaillance du meilleur chemin, la communication est basculée sur le chemin dit alternatif.

Mais il y a un problème : Dans le cas d'un réseau commuté de type Ethernet, cette redondance crée des boucles qui génèrent des tempêtes de diffusion (broadcast storm) .Ces boucles saturent le réseau et finissent par le paralyser complètement.

Considérons la topologie réseau de la figure 2. Au départ les switches n'ont pas encore rempli leurs tables CAM.

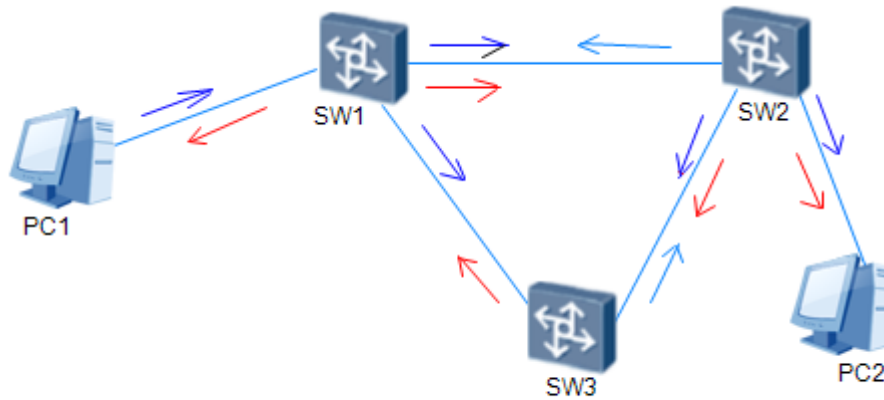


Figure 2 : Broadcast storm dans un réseau commuté

Supposons que La machine PC1 envoie une trame à la machine PC2.

- PC1 a besoin de connaître l'adresse MAC de PC2,
- Pour connaître l'adresse MAC de PC2, une requête ARP est envoyée sur l'adresse MAC de broadcast ff.ff.ff.ff.ff,
- Le switch SW1 voit la trame et la fait sortir de ses deux ports vers les switchs SW2 et SW3,
- Le switch SW2 envoie la trame au switch SW3 et à la station PC2,
- Le switch SW3 envoie la trame aux switchs SW1 et SW2
- Etc...

Il s'agit d'une boucle qui ne finit jamais et qui génère des tempêtes de diffusion ou broadcast storms en anglais.

Conclusion : Nous voyons que d'un côté, nous avons besoin de redondance dans un réseau et de l'autre cette redondance crée des boucles qui génèrent des tempêtes de diffusion (broadcast storm) qui saturent le réseau.

Solution : STP nous permet d'avoir un réseau redondant et sans boucles.

Comment fait le protocole STP pour empêcher les boucles de se produire dans un réseau commuté redondant ?

Le protocole STP utilise des trames de données spéciales appelées BPDU (Bridge Protocol Data Units). Pour permettre aux switchs d'avoir une trace des changements sur le réseau, des BPDU sont échangées toutes les deux secondes sur l'adresse multicast 01:80:C2:00:00:00.

Les informations contenues dans les BPDU sont utilisées pour activer ou désactiver les ports selon la topologie réseau requise. Lorsqu'un switch ou un pont est connecté au réseau, il commence par envoyer des BPDU pour déterminer la topologie du réseau, avant de pouvoir transférer des données. Les BPDU sont envoyés sur l'adresse multicast 01:80:C2:00:00:00.

On distingue trois types de BPDU :

- CBPDU (Configuration BPDU) : Utilisé pour le calcul du spanning tree,

- TCN BPDU : utilisé pour annoncer les changements topologiques,
- TCA BPDU : utilisé pour l'acquittement de changement de notification de la Topologie

Le protocole STP procède en quatre phases :

- Election du commutateur racine (Root Bridge ou RB),
- Détermination du port racine (Root Port ou RP) sur chaque switch,
- Détermination du port désigné (Designated Port ou DP) sur chaque segment,
- Blocage des autres ports

Election du commutateur racine (Root Bridge ou RB)

Le Root Bridge est choisi selon le Bridge identifier (BID). Le BID d'un switch est constitué de l'adresse MAC et de la priorité de ce switch. La priorité est un nombre codé sur 12 bits (soit une valeur comprise entre 0 et 65535). La priorité est paramétrable par l'administrateur réseau. Par défaut, elle est égale à 32768 (ou 0x8000 en hexadécimal). **Le switch avec la priorité la plus basse est élu Root Bridge, et en cas d'égalité, c'est le switch dont l'adresse MAC est la plus basse qui est élu Root Bridge.**

Remarque : Généralement, l'administrateur réseau fixe la priorité de telle sorte que le switch le plus près possible du cœur de réseau soit élu et qu'un autre switch devienne Root Switch en cas de défaillance du Root Bridge principal.

Détermination du port racine (Root Port ou RP) sur chaque commutateur

Chaque un des switches restants détermine parmi ses ports actifs un Root Port. Le Root Port est celui qui possède la distance la plus courte (le coût ou cost le moins élevé) vers le Root Bridge. Le coût dépend de la bande passante de chaque lien. Le tableau 1 donne la valeur de quelques liens en fonction du débit.

Débit	Coût
10 Mbit/s	100
100 Mbit/s	19
1 Gbit/s	4
10 Gbit/s	2

Tableau 1: Débit/ Coût

En cas d'égalité, c'est le port ayant le port ID le plus faible qui sera élu. Donc, l'élection d'un Root Port est effectuée en tenant compte des champs path cost et Port ID d'un paquet BPDU. En cas d'égalité, c'est le port ayant le port ID le plus faible qui sera élu.

Détermination du port désigné (Designated Port ou DP) sur chaque segment

Pour chaque segment réseau reliant des switchs, un DP (Designated Port) est ensuite déterminé. Le port désigné est le port relié au segment qui mène le plus directement à la racine (somme totale des coûts des différents segments traversés est la plus petite).

Blocage des autres ports

Les ports qui ne sont ni RP, ni DP sont bloqués (BP : Blocked Port). Un port bloqué peut recevoir des paquets BPDU mais ne peut pas en émettre.

Changements de topologie

Lorsqu'un lien est coupé ou qu'un switch est hors service, l'algorithme est exécuté à nouveau et une nouvelle topologie est mise en place.

Etats des ports d'un switch

Les ports des commutateurs où STP est actif sont dans l'un des états suivants :

- **Listening** : le switch écoute les BPDU et détermine la topologie réseau,
- **Learning** : le commutateur construit une table faisant correspondre les adresses MAC aux numéros des ports,
- **Forwarding** : opération normale, le port reçoit et envoie des données,
- **Blocking** : Aucune donnée n'est ni envoyée ni reçue mais le port peut passer en mode forwarding si un autre lien tombe,
- **Disabled** : Le port est désactivé (l'administrateur réseau peut manuellement désactiver un port).

Quand un périphérique (un PC, une imprimante, un serveur etc...) est connecté au réseau, son port se mettra automatiquement d'abord en mode listening puis en mode learning et ensuite en mode forwarding.

Forward delay

Le forward delay est le délai de transition entre les modes listening vers learning et learning vers forwarding. Ce délai est fixé par le Root Bridge et vaut quinze secondes par défaut.

STP a connu des évolutions notamment pour accélérer le temps de convergence. L'IEEE a publié le document 802.1w pour décrire le RSTP (Rapid STP) qui accélère la convergence du protocole STP après un changement de topologie. Alors que le temps de convergence est de trente à cinquante secondes pour le STP classique, RSTP est capable de converger en 6 secondes.