

Le protocole STP

Spanning-Tree Protocol

Arthur COURT

STP

Spanning Tree Protocol

Table des matières

STP Spanning Tree Protocol	1
I) La redondance	2
A. Qu'est-ce que la redondance ?	2
B. Reproduction sur Packet Tracer	3
II) Le principe du protocole STP	4
A. Le fonctionnement	4
B. Les opérations	4
1. Election du commutateur racine.....	4
2. Configuration des ports.....	5
3. Définition des rôles des ports.....	6
4. Etat des ports	7
C. Les inconvénients de STP	8
III) Commandes.....	8
SOURCES :	10

L'objectif du protocole Spanning Tree est de maintenir un réseau redondant absent de tous problèmes de boucles tout en garantissant une tolérance de pannes.

I) La redondance

A. Qu'est-ce que la redondance ?

La redondance consiste à disposer plusieurs exemplaires d'un même équipement ou d'un même processus afin de :

- Augmenter la capacité totale ou les performances d'un système
- Réduire le nombre de pannes
- Combiner ces deux effets

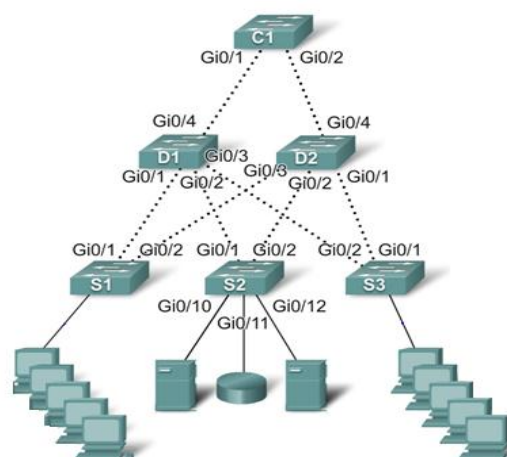
Il existe plusieurs types de redondance :

- **La redondance symétrique** : elle repose sur le principe de dupliquer deux choses semblables à l'identique point par point.
- **La redondance asymétrique** : Elle permet de basculer d'un type de matériel à un autre, il n'est pas forcément identique mais assure les mêmes fonctionnalités avec si possible des performances similaires.
- **La redondance évolutive** : elle est comparable à l'asymétrie mais on isole le système défaillant lors d'une panne pour utiliser une autre partie du système.
- **La redondance modulaire** : c'est une technique qui permet de dévier une panne d'un système sur un autre.

Dans tous les cas on ne parle de redondance seulement si les composants exercent les mêmes fonctions et ce sans dépendre les uns des autres. Leur influence mutuelle se limite en général à se répartir la charge de travail ou des données.

Certains composants effectuent des contrôles sur l'activité de leur voisin afin de se substituer à celui-ci s'ils sont manifestement hors d'usage, ou relancer le service si cela est possible.

Dans le cas de systèmes complexes, on peut dupliquer différents sous-ensembles. On travaille successivement sur chaque sous-ensemble en commençant par ceux jugés les moins fiables ou étant le plus critique. Une fois dupliqué on se concentre sur le prochain sous-ensemble jugé le plus sensible ou fragile et ainsi de suite. On poursuivra ce processus jusqu'à avoir atteint le niveau de capacité, de performance et de fiabilité requis et aussi tant que le surcoût de l'installation est jugé rentable.



Exemple de réseau redondant

Le problème de ce type de configuration, c'est que dans le cas de commutateurs, il y a un risque de saturation du réseau par des trames broadcast à cause des boucles de commutations.

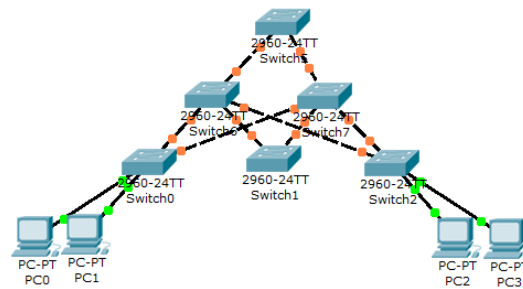
En effet, même si les éléments d'interconnexions sont des commutateurs qui par définition envoient les trames directement sur les bons ports en unicast, il suffit qu'une trame d'adresse MAC FF:FF:FF:FF:FF:FF apparaisse dans le réseau pour que celui-ci vienne à complètement se saturer (la trame étant diffusée sur tous les ports des commutateurs suite à la boucle physique).

Un tel phénomène se nomme **tempête de diffusion** ou **broadcast storm**. Le seul système qui sera capable d'empêcher la propagation d'une telle tempête sera le routeur.

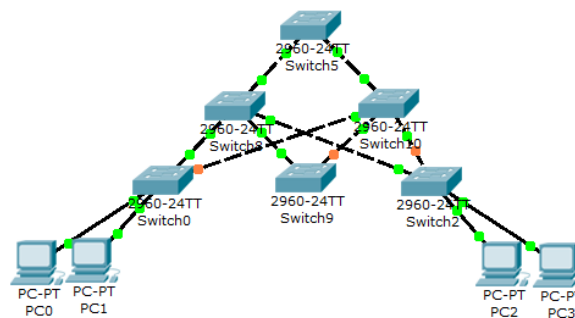
Spanning Tree va permettre de désactiver tout port générant une boucle physique dans un même réseau physique, tout en assurant la redondance des liens dans le cas d'une panne. Le choix des liens se faisant en fonction des segments les plus rapides.

B. Reproduction sur Packet Tracer

Si on s'amuse à reproduire le réseau redondant à l'image un peu plus haut sur Packet Tracer, on verra que certains ports des commutateurs resteront orange à cause du protocole STP pour justement éliminer ces boucles de commutations.



Le réseau lorsque l'on vient de brancher les appareils entre eux : les ports sont entrain de s'initialiser.



Après plusieurs secondes, on observe que 3 ports restent orange, le protocole STP les a bloqué pour éviter les boucles.

Si on désactive le protocole STP (voir plus loin pour les commandes) pour faire en sorte que qu'aucun port ne soit bloqué, et qu'on envoie un ping du PC-0 vers le PC-2 :

```
Packet Tracer PC Command Line 1.0
PC>ping 10.0.0.3

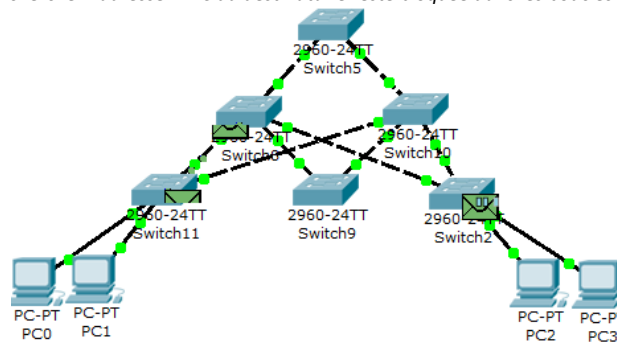
Pinging 10.0.0.3 with 32 bytes of data:

Request timed out.
Request timed out.
Request timed out.
Request timed out.

Ping statistics for 10.0.0.3:
    Packets: Sent = 4, Received = 0, Lost = 4 (100% loss),
```

Le ping n'aboutit pas

Alors que les PC ont été mis dans le même réseau et qu'ils sont reliés par des commutateurs, le ping n'arrive jamais, car la requête ARP qui est envoyée pour chercher l'adresse MAC du destinataire reste bloquée dans les boucles de commutations.



II) Le principe du protocole STP

A. Le fonctionnement

Lorsque le protocole STP est activé, au premier allumage des switch, chacun vérifie si son réseau commuté ne comporte aucune boucle. Si une boucle potentielle est détectée, les switches bloquent certains ports de connexion et en laissent d'autres actifs, afin d'assurer le transfert des trames.

Grâce à STP les Switchs vont définir une arborescence dans le réseau commuté de type étoile étendue. Les Switchs vont vérifier constamment que le réseau ne présente pas de boucle et que chaque port fonctionne correctement.

Afin d'empêcher les boucles de commutation, le protocole STP effectue les opérations suivantes :

- Forcer certaines interfaces à passer en état de veille ou de blocage
- Laisser d'autres interfaces en état de transmission
- Reconfigurer le réseau en activant le chemin de veille approprié, si le chemin de transmission se libère

B. Les opérations

1. Election du commutateur racine

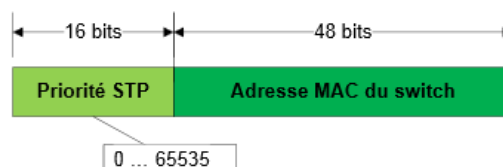
Tout d'abord, STP va procéder à l'élection d'un commutateur ou point racine (Root Bridge). Il constituera le commutateur principal et le point central de la topologie STP.

Chaque réseau ne comporte qu'un seul Root Bridge, qui est choisi en fonction du Bridge ID.

Ce BID est constitué de la valeur de priorité et de l'adresse MAC.

- La valeur par défaut de la priorité s'élève à 32 768. La priorité la plus faible est à 0 et monte de 4096 en 4096.
- Si l'adresse MAC d'un switch est AA:11:BB:22:CC:33, le BID par défaut du Switch sera 32768:AA-11-BB-22-CC-33

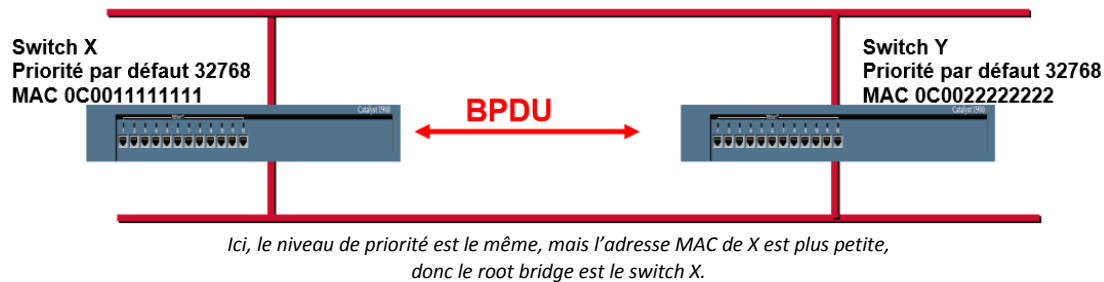
Bridge ID = Bridge priority + MAC address



Le Root bridge est le switch avec le plus petit BID.

Lors d'une égalité de priorité c'est le système qui a l'adresse MAC la plus petite qui emporte l'élection du « root bridge ».

En pratique, lorsqu'un switch est allumé, il pense d'abord qu'il est le commutateur racine. Les switches du réseau vont alors s'échanger à intervalle régulier (toutes les deux secondes) une trame spéciale appelée BPDU Hello (Bridge Protocol Data Unit). Un des champs de la trame BPDU est l'ID, Identifier. Si le Switch2 annonce un BID racine inférieur à celui du Switch1, alors le Switch1 cessera d'annoncer son ID racine et acceptera celui du Switch2. Ainsi, le Switch 2 deviendra le commutateur racine.



Une fois l'élection terminée, seul le Root Bridge envoie des BPDU.

Notes : Idéalement le commutateur racine doit être le switch situé à un point central du réseau. Il est possible de laisser les switch décider qui est le pont racine et le chemin à prendre. Cependant pour une infrastructure mieux conçue il faut décider "qui est qui".

2. Configuration des ports

Par la suite, grâce au Root Bridge et aux BPDU, STP va pouvoir déterminer quels ports seront bloqués ou non. Les trames BPDU envoyées par le Root Bridge vont permettre au réseau de se reconfigurer automatiquement en cas de panne.

On différencie 3 types de ports dans le protocole STP :

- **Les ports racines (root ports)** : Il fournit le chemin avec le coût le plus bas vers le pont racine
- **Les ports désignés (designated ports)** : Ils retransmettent tout le trafic (forwarding state)
- **Les ports bloqués (blocking state)** : Il ne transfère pas de trafic

Tous les ports du pont racine sont des ports désignés.

Sur les switches non racines, ils vont calculer le meilleur chemin (coût le plus bas) vers le switch racine et à partir de ce chemin, élire le port racine correspondant qui sera en forwarding state.

Ce Root Port sera donc choisi en fonction du coût du chemin vers le Root Bridge (**Root Path Cost**), de la taille de la bande passante (la plus grande possible) et du nombre de liens le plus bas.

Chaque type d'interface représente un coût :

Coût standard des interfaces :

Ethernet : 100

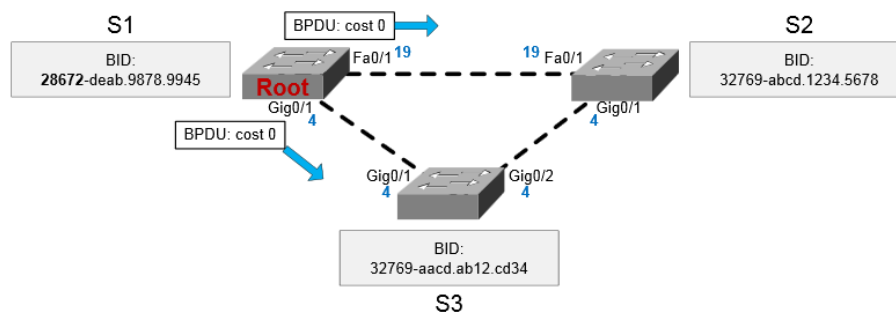
Fast-Ethernet : 19

Gigabit : 4

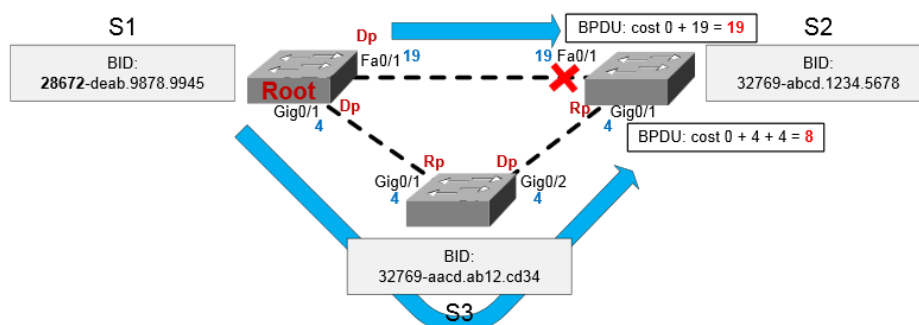
Ethernet channel Gigabit : 3

Ten-Gigabit : 2

3. Définition des rôles des ports



Le Root-Bridge envoie une BPDU dans chaque direction.



A chaque entrée sur une interface, le coût de l'interface est additionné au « Root Path Cost » du BPDU.

Explications :

S2 reçoit deux BPDUs, l'un venant directement de S1, l'autre par le côté de S3.

Celui provenant de S3 a un « Root Path Cost » de 8, inférieur à celui venant de S (19), le chemin passant par S3 est donc le meilleur chemin vers le Root Bridge, l'interface Gig0/1 de S2 est donc un Root port.

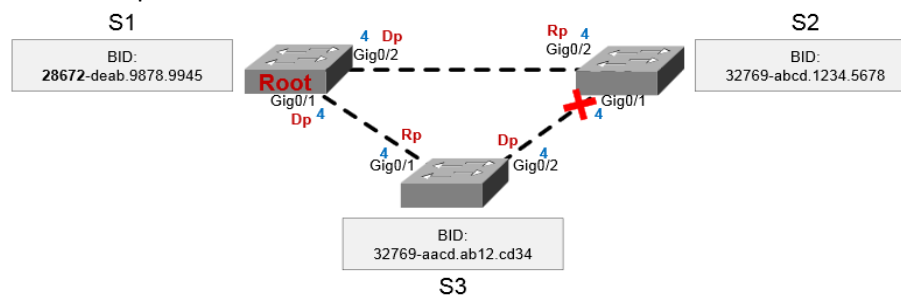
Les ports d'un Root Bridge sont toujours des Designated Ports.

Un port faisant face à un Root port ne peut être qu'un Designated Port (Dp) :

- Gig0/2 sur S3 sera donc un Dp.
- Fa0/1 et Gig0/1 seront donc des Dp.

Pour ouvrir la boucle, il suffit de bloquer un seul port. Dans ce cas, la seule possibilité est Fa0/1 sur S2.

Si le Root Path Cost est égal des deux côtés de la boucle, le Bridge ID est utilisé pour définir le côté du lien où le port sera bloqué.



Adresse MAC S2 : AB-CD-12-34-56-78

Adresse MAC S3 : AA-CD-AB-12-CD-34

Ici, S2 a un BID plus grand que S3 (donc moins bon), le lien entre S3 et S2 sera alors bloqué du côté de S2.

L'adresse MAC de S2 commence par AB et l'adresse MAC de S3 par AA, A (10) est plus petit que B (11) donc S3 est plus petit.

Si ni le « Root Path Cost », ni le BID ne permettent de faire un choix, c'est alors le nom de l'interface qui est utilisé. Le « plus petit » nom d'interface sera le meilleur. (A est plus petit que Z, 1 est plus petit que 2).

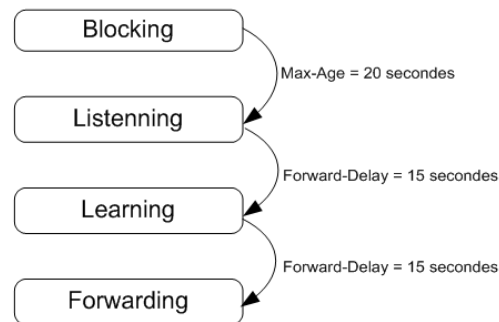


Dans ce cas, S1 est Root Bridge, tous ses ports sont donc Dp. C'est alors du côté de S2 qu'il y aura un port bloqué. Les coûts sont égaux, le BID aussi. C'est donc le nom de l'interface qui va permettre de choisir. Gig0/1 est plus petit que Gig0/2, donc meilleur. Gig0/2 sera donc le port bloqué.

4. Etat des ports

Lorsque l'on branche un nouveau switch, le port passe par plusieurs états avant de s'activer ou non avec STP.

Temps de convergence = 50s :



- Branchement d'un nouveau Switch.
- Activation du port, le protocole STP se met en marche.
- **Max-Age** (20 secondes), temps de sécurité à attendre en cas de changement topologique : le port est **bloqué**, il jette les trames entrantes, accepte les BPDU mais ne les retransmet pas, il ne complète pas sa table d'adresse MAC.
- **Listenning** (15 secondes), temps pour que le commutateur écoute les trames qu'il peut recevoir, sans émettre : il jette les trames entrantes, **accepte les BPDU et les retransmets**, il ne complète pas sa table d'adresse MAC.
- **Learning** (15 secondes), temps pour apprendre les informations de la topologie STP : il accepte les trames entrantes mais ne les retransmet pas, accepte les BPDU et les retransmets, incorpore les nouvelles adresses MAC dans sa table.
- **Forwarding** : le port passe en Forwarding, il accepte et retransmets les trames entrantes, accepte les BPDU et les retransmets, remplit sa table d'adresses MAC.

Le port met donc 50 secondes à s'initialiser.

Lorsqu'une modification topologique est détectée, l'arbre STP est recalculé et le trafic est stoppé.

C. Les inconvénients de STP

STP est conçu pour n'autoriser qu'un seul chemin actif entre deux commutateurs. Lorsque le chemin actif est inaccessible, le protocole sélectionne automatiquement un chemin de secours. La convergence du Spanning-Tree après la perte de liens peut prendre plusieurs secondes et n'est pas en adéquation avec les exigences de disponibilité des applications critiques sensibles à la latence ou à la perte de paquets. De plus, plus la topologie du réseau sera compliquée, plus ce temps de convergence sera important.

Aussi, sa complexité en terme de gestion introduit une lourdeur lors d'intervention sur le réseau tel que l'ajout d'équipements ou encore la mise à jour logicielle de ces derniers.

III) Commandes

- Voir le statut du protocole STP :

Switch# show spanning-tree vlan 1

```
VLAN0001
Spanning tree enabled protocol ieee
Root ID    Priority    32769
           Address    0001.630D.B678
           This bridge is the root
           Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID  Priority    32769 (priority 32768 sys-id-ext 1)
           Address    0001.630D.B678
           Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
           Aging Time 20

Interface Role Sts Cost Prio.Nbr Type
-----
Fa0/1     Desg FWD 19   128.1  P2p
Fa0/2     Desg FWD 19   128.2  P2p
```

On voit dans la partie Root ID, tout ce qui concerne le commutateur racine du réseau :

- Sa priorité (par défaut 32769)
- Son adresse MAC (VLAN 1)
- "This bridge is the root", ce switch est donc le Root Bridge.
- Les paramètres pour les BPDU

Ensuite on peut voir les ports avec notamment leur rôle (donc désignés) et leur coût (19).

Comme le switch est le commutateur racine, la partie Root ID et Bridge ID est la même.

Pour un switch non racine on aurait ça :

```
VLAN0001
Spanning tree enabled protocol ieee
Root ID    Priority    32769
           Address    0001.630D.B678
           Cost        19
           Port        2 (FastEthernet0/2)
           Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec

Bridge ID  Priority    32769 (priority 32768 sys-id-ext 1)
           Address    00E0.B033.7AB6
           Hello Time 2 sec Max Age 20 sec Forward Delay 15 sec
           Aging Time 20

Interface Role Sts Cost Prio.Nbr Type
-----
Fa0/1     Altn BLK 19   128.1  P2p
Fa0/2     Root FWD 19   128.2  P2p
```

On voit dans la partie Root ID, tout ce qui concerne le commutateur racine du réseau, soit les mêmes informations que le switch précédent, avec le coût du chemin vers lui, et par quelle interface (ici 19 et in fa/02).

Ensuite, on voit les informations du switch lui-même, on voit bien que son adresse MAC est plus grande que l'autre.

Concernant les ports, on voit qu'un a le rôle Root, donc le chemin le plus rapide vers le Root ID, et l'autre a le rôle Altn pour alternate qui signifie qu'il a été mis en veille (bloqué).

- Pour sélectionner soi-même un commutateur Root :
Switch(config)# spanning-tree vlan *vlan-id* priority *priority* (= multiple de 4096)
- Changer le coût d'un port en mode access :
Switch(config-if)# spanning tree cost *cost*
- Changer le coût d'un port en mode trunk :
Switch(config-if)# spanning tree vlan *vlan-id* cost *cost*

- **Pour changer la priorité d'un port en mode access ou trunk**, ce sont les mêmes commandes que ci-dessus en remplaçant `cost cost` par `priority priority`.
- **Désactiver STP :**
`Switch(config)# no spanning tree vlan vlan-id`
- **Changer les paramètres de délais :**
`Switch(config)# spanning-tree vlan vlan-id max-age seconds`
`Switch(config)# spanning-tree vlan vlan-id forward-time seconds`
`Switch(config)# spanning-tree vlan vlan-id hello-time seconds`
- **Passer en Rapid Spanning Tree (permet de faire passer le temps de convergence à 6 secondes maximum, ce qui le rend beaucoup plus opérationnel que STP).**
`Switch(config)# spanning-tree mode rapid-pvst`

SOURCES :

<http://home.nordnet.fr/~ericleleu/cours/nfe107/Haute%20Disponibilite.pdf>

<http://vgicquiau.blogspot.fr/2012/05/stp-spanning-tree-protocol.html>

http://fr.wikipedia.org/wiki/Redondance_des_mat%C3%A9riels

http://cisco.goffinet.org/s3/spanning_tree#.U6P_ubHrAQs