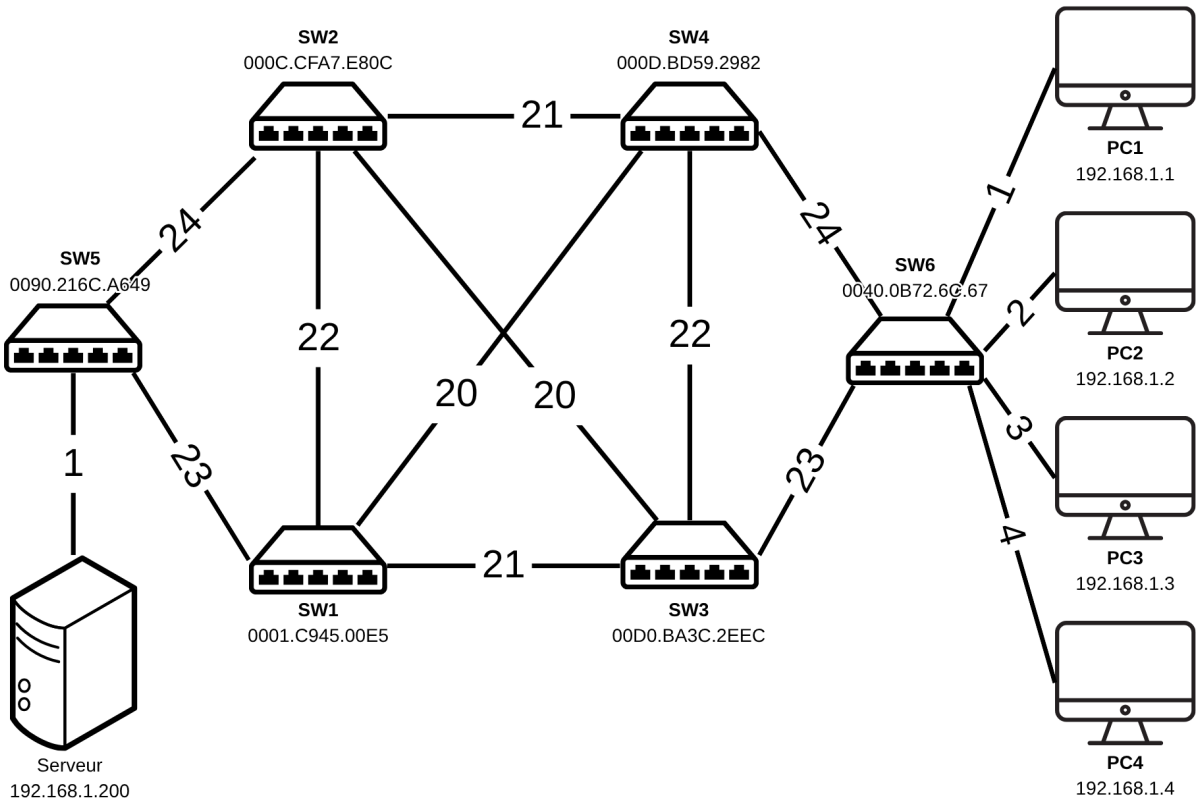
TP - Spanning Tree Protocol

Contexte

L'entreprise dispose d'un réseau conséquent. Elle a décidé de mettre en place des liens redondants entre les différents locaux techniques.

Les liens redondants permettent de limiter l'impact d'une panne de commutateur ou d'une coupure de liaison (lien défectueux, lien supprimé par erreur, etc.). Si la redondance est suffisante, l'impact se limitera aux seuls postes reliés directement au commutateur. Pour réduire encore plus cet impact, les serveurs sont parfois reliés à 2 commutateurs (double attachement).

Par simplification pour la compréhension de cet exolab, seule une partie des actifs et des postes est représentée, chaque actif étant situé dans un local technique différent. Les liaisons entre les locaux techniques ne dépassent pas 80m.



Vous trouverez un fichier Packet Tracer sur Moodle, nommé `Maquette_STP_init.pkt`. Cette maquette va servir de base pour la découverte du protocole Spanning-Tree. Une description du rôle et du fonctionnement du protocole STP est donnée en annexe.

Le numéro indiqué sur les liens correspond au port utilisé sur chaque commutateur. Par simplification, le même port est utilisé de chaque côté du lien pour les liaisons entre commutateurs.

Étape 1 - Observer le bon fonctionnement des communications

Ouvrir la maquette Packet Tracer.

Nous allons déjà observer le fonctionnement du protocole Spanning-Tree dans son fonctionnement par défaut sur les commutateurs CISCO.

Q1. Indiquer quels sont les commutateurs traversés lors d'une communication entre PC1 et Serveur.

Pour vous aider, consultez les annexes suivantes : définition, boucles et principe.

Étape 2 - Observer le résultat de la configuration

Q2. Expliquer s'il est possible de savoir, en observant les ports actifs après convergence du protocole STP, quel est le commutateur dit "Pont racine", autrement dit celui qui constitue la racine de l'arbre de recouvrement.

Pour information Spanning-Tree peut se traduire par arbre de recouvrement, voir dans l'annexe.

Q3. Observer et analyser la configuration STP des commutateurs en :

a) Vérifier la configuration de chaque commutateur à partir de la commande show spanning-tree.

Vous pouvez vous aider des annexes : rôles et décryptage.

Effectuer une copie d'écran et expliquer brièvement le résultat obtenu pour chaque commutateur.

SW1	SW2
Screen du switch	Screen du switch
Explication	Explication
SW3	SW4
Screen du switch	Screen du switch
Explication	Explication

b) Indiquant, d'après les informations obtenues, quel est le commutateur qui joue le rôle de pont racine.

Consultez cette annexe.

Enregistrez votre fichier au nom `Maquette_STP_etape12.pkt`.

Étape 3 - Observer la « reconvergence » et donc la tolérance aux pannes dans plusieurs cas de figure

Dupliquez votre fichier pour chaque scénario au nom `Maquette_STP_etape3_sc1.pkt`

Scénario 1 : Observer la « reconvergence » suite à la coupure d'un lien

Couper une liaison active entre 2 commutateurs (ports actifs sur les 2 commutateurs), par exemple SW1-SW4, et observer le recalcul du protocole Spanning-Tree, aboutissant à un nouvel arbre de recouvrement.

Remarque : le temps de « reconvergence » dure entre 30 et 50 secondes.

Comme expliqué en annexe, il existe une version plus rapide du protocole STP (PVST sous CISCO) : le RSTP (rapid-PVST sous CISCO), qui permet une reconvergence en 6s.

Pour cet exolab, l'utilisation du protocole STP de base (PVST activé par défaut sur les commutateurs CISCO) est volontaire afin de mieux observer son fonctionnement. En production, la version la plus performante de ce protocole doit être privilégiée.

Sur les équipements CISCO, l'activation du protocole rapid-PVST se fait à partir de la commande : spanning-tree mode rapid-pvst

Q4. Indiquer quel chemin emprunte désormais une communication entre PC1 et Serveur.

Rebranchez le lien coupé pour revenir à l'état initial.

Scénario 2 : Observer la « reconvergence » suite au passage en Gigabits de certains liens

Les liaisons sont toutes de type FastEthernet (donc 100 Mbits/s). Imaginons que les liaisons entre SW1 et SW2 et entre SW2 et SW4 soient en GigabitEthernet. On peut supposer que la liaison directe SW1-SW4, actuellement active, pourrait être désormais plus coûteuse qu'une liaison indirecte passant par SW2.

Q5. Donner le coût actuel de la liaison SW1-SW4.

Tenter ensuite l'expérience en procédant comme suit :

- Remplacer la liaison SW2-SW4 FastEthernet (utilisant les ports Fa0/21 des commutateurs) par une liaison Gigabits utilisant les ports Gi0/1 des commutateurs.
- Remplacer la liaison SW1-SW2 FastEthernet (utilisant les ports Fa0/22 des commutateurs) par une liaison Gigabits utilisant les ports Gi0/2 des commutateurs.

Attendre que l'algorithme parvienne à converger, suite à ces modifications.

Q6. Indiquer les modifications observées.

Q7. Calculer, d'après les nouveaux débits affectés à chaque liaison, les coûts de :

- la liaison directe SW1-SW4
- la liaison indirecte passant par SW2.

Q8. En déduire si le nouvel arbre de recouvrement est cohérent par rapport à ces calculs.

Scénario 3 : Observer la « reconvergence » suite au changement de priorité sur un commutateur

On souhaite forcer l'élection de SW2 plutôt que SW1 comme "Pont racine".

Le choix explicite d'un commutateur comme pont racine peut être motivé par différentes raisons : son emplacement stratégique pour que les chemins soient optimaux, ainsi que les caractéristiques du commutateur : fond de panier, débit des ports, modules fibre optique, etc …

Par exemple on peut choisir volontairement le commutateur coeur de réseau équipé de ports en 10 Gbits/s Fibre comme pont racine. Ce commutateur sera relié de manière redondante aux commutateurs de distribution qui desserviront à leur tour les commutateurs d'accès auxquels sont les postes clients en 1 Gigabits/s ou 100 Mbits/s.

Pour forcer l'élection de SW2 comme « pont racine », on doit modifier la priorité de ce commutateur, en la fixant à une valeur inférieure à la priorité fixée par défaut (32768).

Remarque : comme expliqué en annexe, sur les équipements CISCO ce n'est pas le protocole STP standard qui est utilisé mais le protocole PVST (Per VLAN Spanning Tree) même si seul le VLAN 1 existe. La modification de priorité doit donc être réalisée par VLAN.

Par exemple, pour le VLAN 1, la syntaxe est : spanning-tree vlan 1 priority <nouvelle-priorité>

La priorité se définit comme un multiple de 4096 comme l'indique le message en cas d'erreur de valeur :

```
SW2(config)#spanning-tree vlan 1 priority 10000
% Bridge Priority must be in increments of 4096.
% Allowed values are :
   0      4096   8192   12288  16384  20480  28672
   32768  36864  40960  45056  53248  57344  6144
```

Pour cette expérimentation qui fait suite au scénario 2, on conservera les liaisons qui ont été modifiées en Gigabits.

A faire 🚀

Modifier la valeur de la priorité sur SW2 en 8192 pour le seul vlan utilisé ici (vlan 1).

Vérifier la prise en compte de cette valeur en affichant les propriétés Spanning-tree.

Q9. Vérifier que SW2 devient bien le pont racine suite à ce changement de priorité et indiquer les changements au niveau des liaisons.

Scénario 4 : Observer la « reconvergence » suite à la panne d'un commutateur

Plusieurs cas de figure pourraient être étudiés :

- panne du commutateur racine actuel (SW2),
- panne d'un commutateur dont tous les ports sont en forward (SW1),

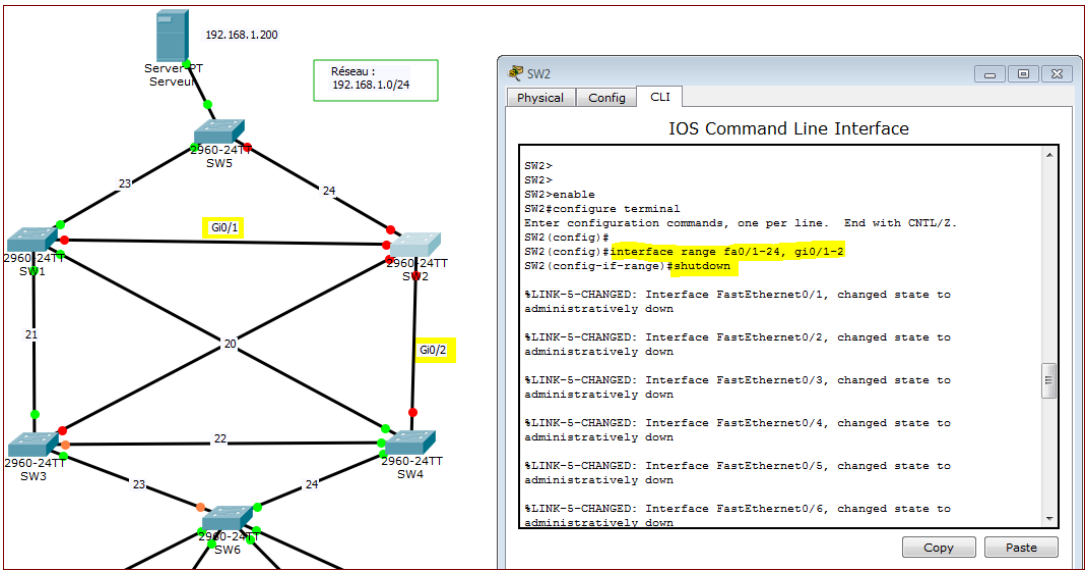
- panne du commutateur relié aux postes.
  Ce dernier cas ne serait pas très intéressant puisqu'on ne pourrait plus tester les communications.

Le premier cas va entrainer obligatoirement une modification de l'arbre de recouvrement, puisque cette panne va provoquer une réélection. Nous allons donc traiter ce cas.

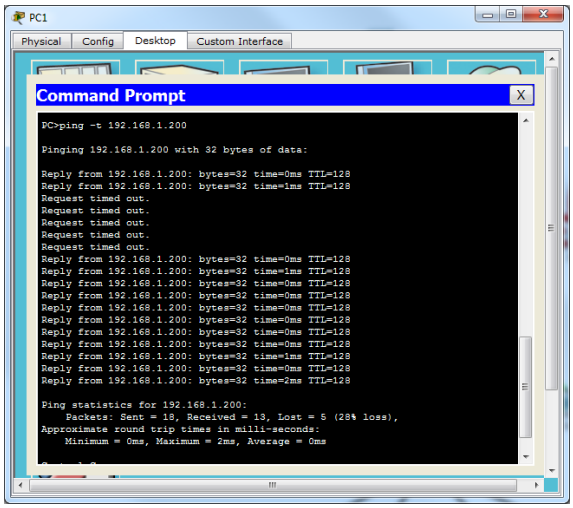
Pour éviter une suppression du commutateur et des liens, et puisqu'il n'est pas possible de couper l'alimentation d'un commutateur sous Packet Tracer, on va se contenter de désactiver tous les ports, ce qui sera plus ou moins équivalent à une panne du commutateur.

A faire 🚀

Lancer un ping permanent, par exemple de PC1 vers Serveur, pour pouvoir observer la perte de communication temporaire pendant le recalcul Spanning-Tree. Sur SW2, désactiver l'ensemble des ports. Pour désactiver les ports, il suffit d'utiliser la commande `shutdown` et la commande `no shutdown` pour les réactiver.



Les ports de SW2 sont bien entendu désactivés, et par conséquent les ports situés à l'autre extrémité de chaque liaison de SW2, comme le montre les points rouges ci-dessus.

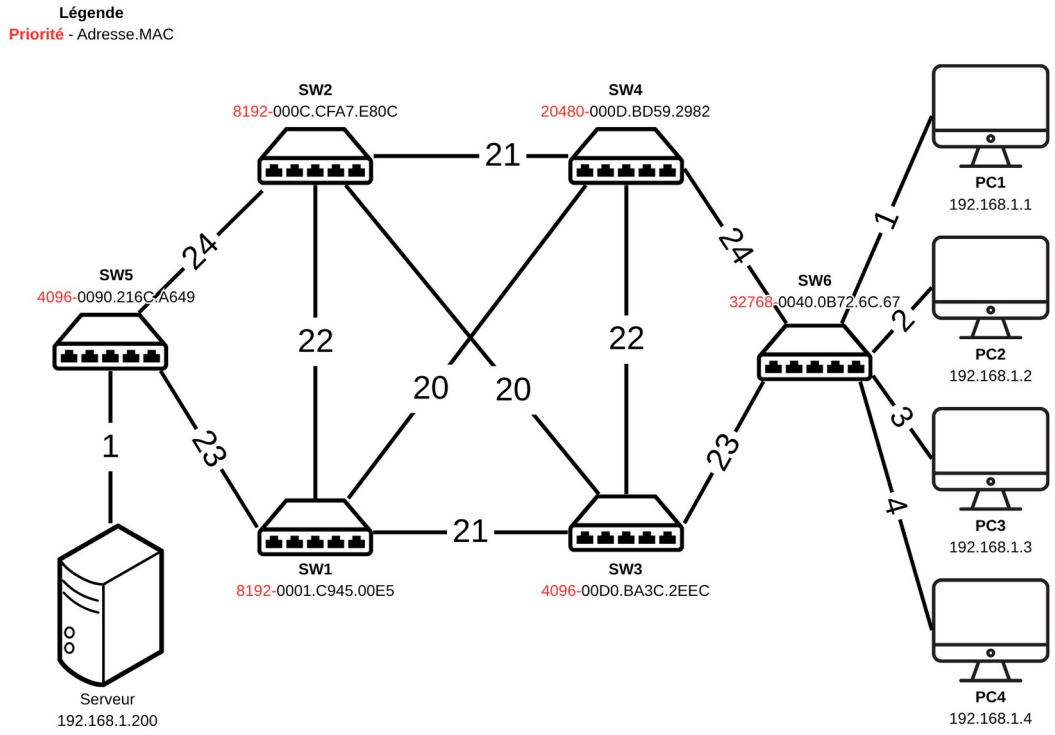


Après quelques instants, un autre arbre de recouvrement est construit et permet à nouveau une communication entre PC1 et Serveur :

Q10. Indiquer, en justifiant, quel est le nouveau pont racine.

Étape 4 - Mise en application

Partons du schéma suivant :



Tous les ports utilisés sur les commutateurs sont en FastEthernet (100 Mbits/s).

La priorité par défaut est 32768.

Remarque : La définition des priorités est ici un « cas d'école ». En général, on se contente d'affecter une priorité au(x) seul(s) commutateur(s) que l'on souhaite privilégier dans l'élection.

A faire 🚀

Déterminer l'arbre de recouvrement en tenant compte des principales règles STP :

1. Élection du pont racine

   Le commutateur ayant le Bridge ID [Priorité-MAC Adresse] le plus bas est élu pont racine. A égalité de priorité (ce qui est le cas par défaut – priorité de 32768 définie sur tous les commutateurs), le commutateur ayant la plus petite adresse MAC est élu pont racine.
2. Détermination des liens actifs

   Chaque commutateur doit être relié au pont racine avec le coût le plus faible. Par défaut le coût correspond à une valeur par défaut pour chaque débit : 19 pour 100 Mbits/s
3. Désactivation des liens redondants

   Les liens alternatifs sont désactivés. Le port bloqué est celui correspondant à au Bridge ID le plus grand en cas de coût identique depuis le pont racine.

Vérifier votre réponse en ouvrant le fichier `Maquette-STP-etape4.pkt` : les priorités indiquées sur le schéma ont été définies sur chaque commutateur.

Annexe - Le protocole STP en bref

Définition

Le spanning-tree est un protocole réseau permettant de définir une topologie sans boucle, dans un réseau local composé de commutateurs avec des liens redondants.

Le protocole STP est défini dans la norme IEEE 802.1D

Les principales versions

L'explication détaillée des différences entre les versions n'entre pas dans le cadre de cet exolab.

Pour plus d'informations voir par exemple Wikipédia ou d'autres sites comme goffinet.org.

STP [IEEE 802.1D]	Spanning-Tree Protocol : protocole de base
PVST	Per Vlan Spanning-Tree : quand plusieurs VLAN existent dans un réseau commute, STP peut fonctionner de manière indépendante sur chacun des VLAN. C'est le mode par défaut sur les commutateurs CISCO.
RSTP [IEEE 802.1w]	Rapid Spanning Tree Protocol : la convergence est plus rapide en cas de changement de topologie (6 secondes au lieu de 30 à 50 secondes).
MSTP	Multiple Spanning Tree Protocol : inclus dans la norme IEEE 802.1Q. Il existe une instance de RSTP par groupe de VLAN. Il est possible par conséquent de mieux utiliser les liaisons redondantes pour effectuer de la répartition de charges sur les liens.

Les boucles dans un LAN

En cas de lien redondant, comme dans le schéma ci-dessous, si un poste (par ex. P1) envoie une trame de diffusion, SW2 va la diffuser sur tous ses ports, donc vers SW2 (en bleu) et SW3 (en vert). De même SW2 va la rediffuser vers PC2 et SW3, puis SW3 vers PC3 et SW1, et la boucle est bouclée La diffusion pourrait même continuer (en orange).

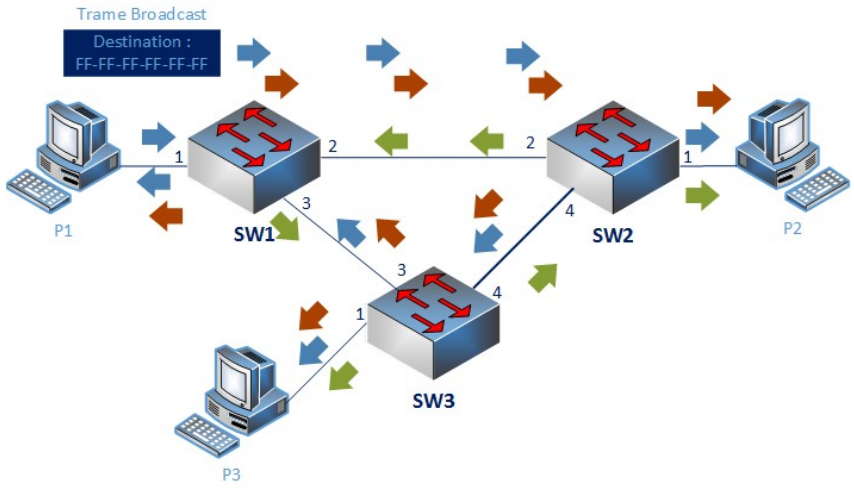


Figure 1 - Sans STP, la boucle provoquera une tempête de Broadcast

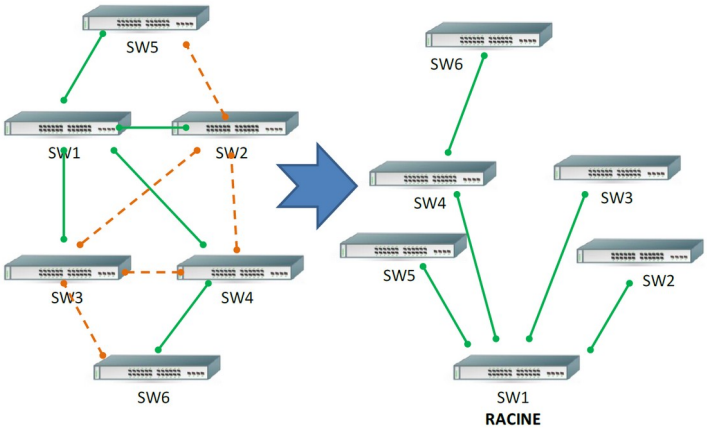
Cette situation s'appelle une tempête de broadcast (Broadcast Storm) et provoque une instabilité des tables d'adresse MAC et une saturation de la bande passante et des ressources du commutateur, puisque par exemple SW3 recevra la trame de broadcast provenant de P1 sur son port 3, mais aussi sur son port 4 : il associera donc probablement l'adresse MAC de P1 d'abord au port 3, puis au port 4, alors que le poste n'a pas changé de place.

Par ailleurs, le commutateur ne peut pas savoir quand une trame broadcast a atteint tous les destinataires. Le principe est de la réémettre sur tous les ports actifs (sauf celui d'origine). Cette « boucle sans fin » provoque au final une indisponibilité du réseau.

Le principe de STP

Le principe de base est relativement simple : il faut créer un arbre qui permette de relier l'ensemble des équipements sans boucle, et autant que possible avec un coût moindre pour atteindre chacun des équipements depuis la racine.

On parle d'arbre de recouvrement, car pour éviter les boucles, on obtient bien un arbre de liens actifs qui permet de relier tous les commutateurs et donc tous les postes. Par ailleurs, tout arbre possède bien une racine, et les branches se développent uniquement au-dessus de celle-ci.



Une fois les liens désactivés pour casser les boucles, on obtient bien un arbre ; pour s'en convaincre, il faut faire en sorte que le commutateur qui n'a que des liens actifs se retrouve en bas.

Un des commutateurs constitue bien la « racine » ; les autres commutateurs sont reliés par un seul chemin au commutateur racine.

Il faut donc :

1. Élire le commutateur racine appelé « Root Bridge » ou « Pont Racine ».
2. Déterminer le chemin de « coût » le plus faible pour chaque équipement vers chaque équipement depuis ce « pont racine ». On parle bien de coût le plus faible, et pas forcément de chemin le plus court, car les liens ne sont pas forcément équivalents (en fonction du débit qui les caractérise).

L’élection du « Pont Racine » (ou « Root Bridge »)

Dans une infrastructure STP, un des commutateurs est élu « pont racine » et représente ainsi le point de départ de l’algorithme. C’est à partir de ce dernier que seront prises toutes les décisions de modifier ou non l’état d’un port d’un commutateur.

Les commutateurs sont identifiés par un Bridge ID unique. Il est constitué par une priorité (valeur sur 12 bits) associée à l’adresse MAC du commutateur. Par défaut la priorité est de 32768 sur tous les commutateurs. Le commutateur dont le Bridge ID est le plus petit remporte l’élection.

Au démarrage de l’algorithme chaque commutateur s’annonce comme le « Pont Racine», mais s’il reçoit une information d’un commutateur adjacent signalant un « Pont Racine » de Bridge ID plus faible que celui qu’il connaît actuellement, il corrige en le remplaçant par celui de Bridge ID plus faible.

Pendant l’élection, tous les commutateurs émettent des messages BPDUs (Bridge Protocol Data Units, trames contenant les informations STP) et seulement des messages BPDUs (les autres trames éventuellement envoyées par des postes ne sont pas transmises).

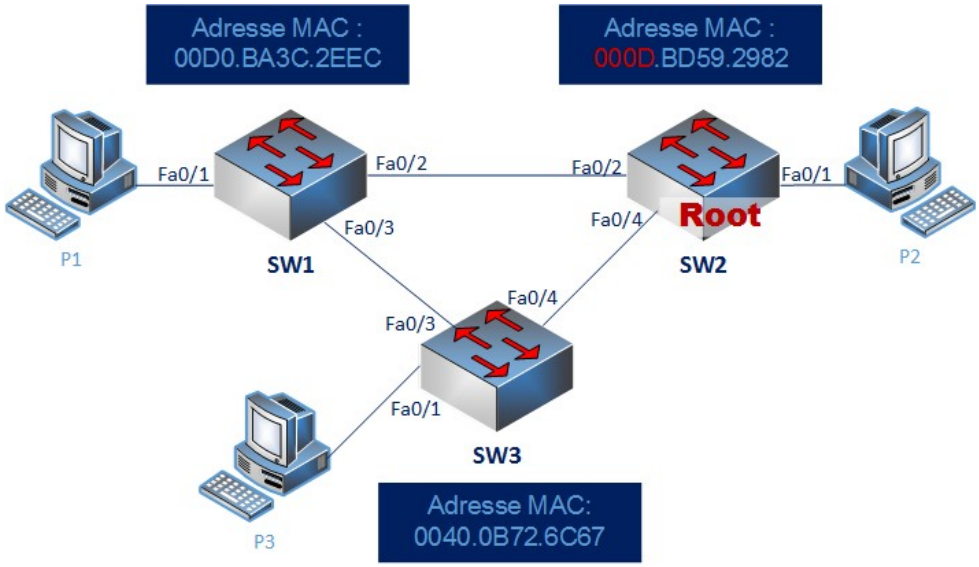


Figure 2 - SW2, qui a le plus petit ID, est élu "Pont Racine".

Détermination des rôles des ports

En fonction des coûts associés à chaque liaison et de l'emplacement du pont Racine, l'algorithme STP va définir le rôle de chaque port.

Rôle des ports STP		**Coûts standards des liens**	
Port **root**	Meilleur chemin vers le "root bridge"	Ethernet	10 Mbits/s ⇒ 100
Port **désigné**	Port non root en "forwarding"	Fast-Ethernet	100 Mbits/s ⇒ 19
Port **bloqué**	Port ni root, ni désigné	Gigabit-Ethernet	1 Gbits/s ⇒ 4
		TenGigabit-Ethernet	10 Gbits/s ⇒ 2

Dans l'exemple ci-dessus, si on considère que les ports fonctionnent tous en 100 Mbits/s, il est facile de déterminer, une fois le pont racine identifié, quel sera le lien coupé.
En effet SW1 et SW3 sont reliés directement à SW2 par un lien à 100 Mbits/s, donc avec un coût de 19, alors qu'indirectement, SW1 et SW3 seraient à un coût de 2 x 19 = 38.

C'est donc le lien entre SW1 et SW3 qui sera coupé. Reste à savoir quel port sera bloqué.

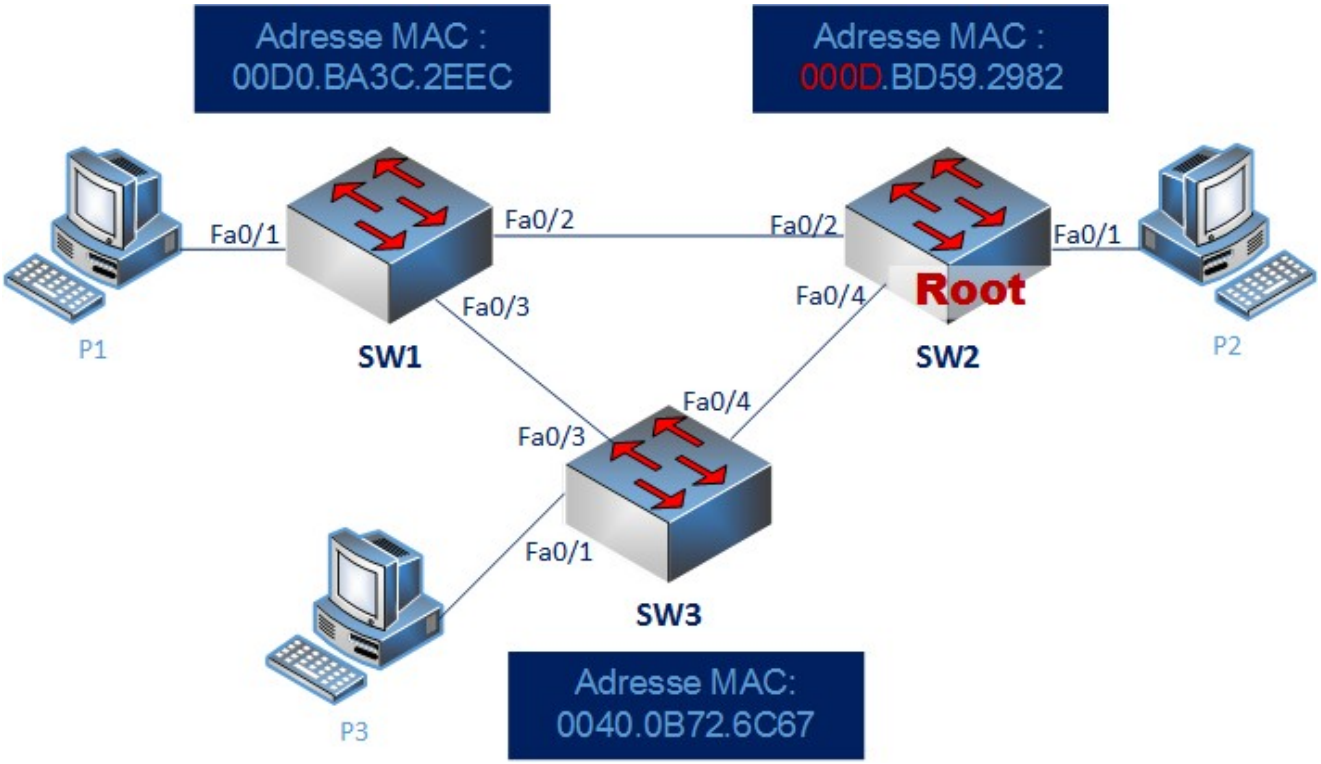
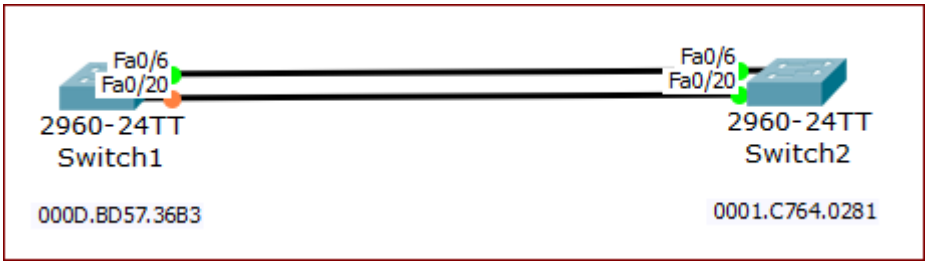


Figure 3 - Le lien entre SW1 et SW3 est coupé par le protocole STP

Cas particulier

En cas d'égalité de valeur pour l'ID de commutateur (Bridge ID), c'est le numéro de port qui permettra de déterminer le lien qui sera désactivé par Spanning-Tree.

Ce cas se produit notamment quand une double liaison redondante existe entre deux commutateurs



Dans ce cas de figure :

- L'adresse MAC la plus faible est celle de Switch2, donc Switch2 est élu "Pont racine"
- Le nom du port Fa0/6 est plus petit que Fa0/20, c'est donc Fa0/20 qui sera bloqué.

Les informations visibles dans la configuration spanning-tree le confirment :

```
Switch-1#sh spanning-tree
VLAN0001
  Spanning tree enabled protocol ieee
  Root ID     Priority    32769
              Address     0001.C764.0281
              Cost        19
              Port        6(FastEthernet0/6)
              Hello Time  2 sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID   Priority    32769  (priority 32768 sys-id-ext 1)
              Address     000D.BD57.36B3
              Hello Time  2 sec  Max Age 20 sec  Forward Delay 15 sec
              Aging Time  20

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Fa0/20           Altn BLK 19        128.20   P2p
Fa0/6            Root FWD 19        128.6    P2p
```

Prise en compte des débits

Le coût d'un lien est fonction du débit associé à ce lien. Imaginons que les liens qui relient SW3 à SW1 et SW2 soient en Gigabits. Le lien coupé ne sera plus le même :

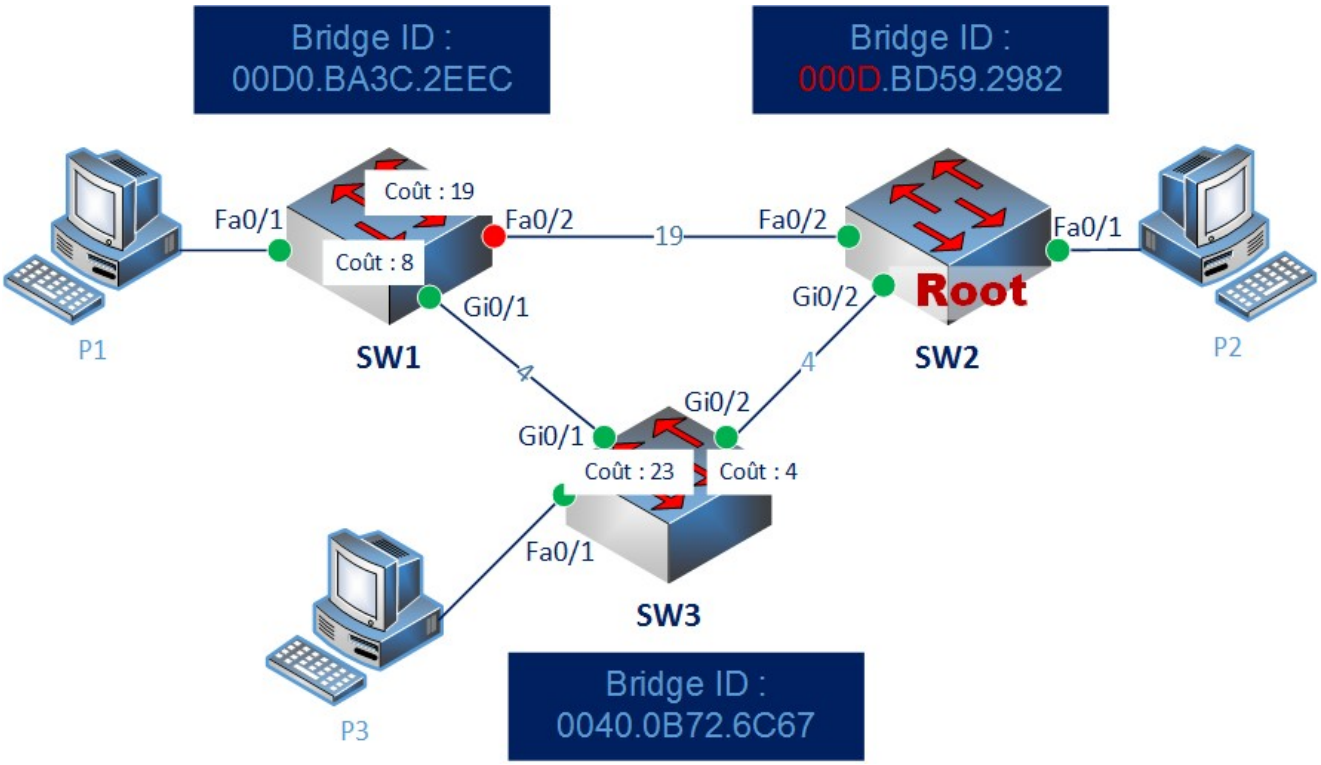


Figure 4 - Calcul de coût en fonction des débits

En effet, le lien direct en 100 Mbits/s qui relie directement SW1 au pont racine SW2 correspond à un coût de 19, alors que la liaison indirecte passant par SW2, avec des liens en Gbits/s totalise un coût de 4+4 = 8.

C'est donc le lien entre SW1 et SW2 qui sera désactivé dans ce cas de figure.

Priorisation d'un commutateur

Pour des raisons diverses on peut vouloir forcer l'élection d'un commutateur plutôt qu'un autre en tant que pont racine. Pour ce faire, il est possible de définir une priorité.

Par défaut la priorité est de 32768, mais elle peut être n'importe quel multiple de 4096 (sur 2 octets), donc de 0 à 61 440.

L'ID du commutateur n'est pas son adresse MAC seulement, mais la priorité associée au commutateur, suivi de son adresse MAC. Dans ce contexte, le commutateur ayant la plus faible priorité sera élu "pont racine", et en cas d'égalité sur la priorité, la plus petite adresse MAC emportera l'élection.

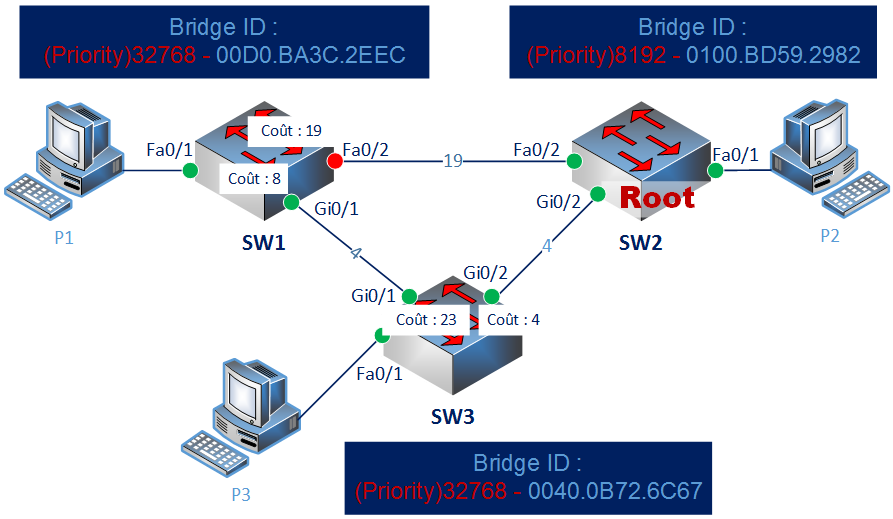


Figure 5 - Élection du pont racine en cas de priorité

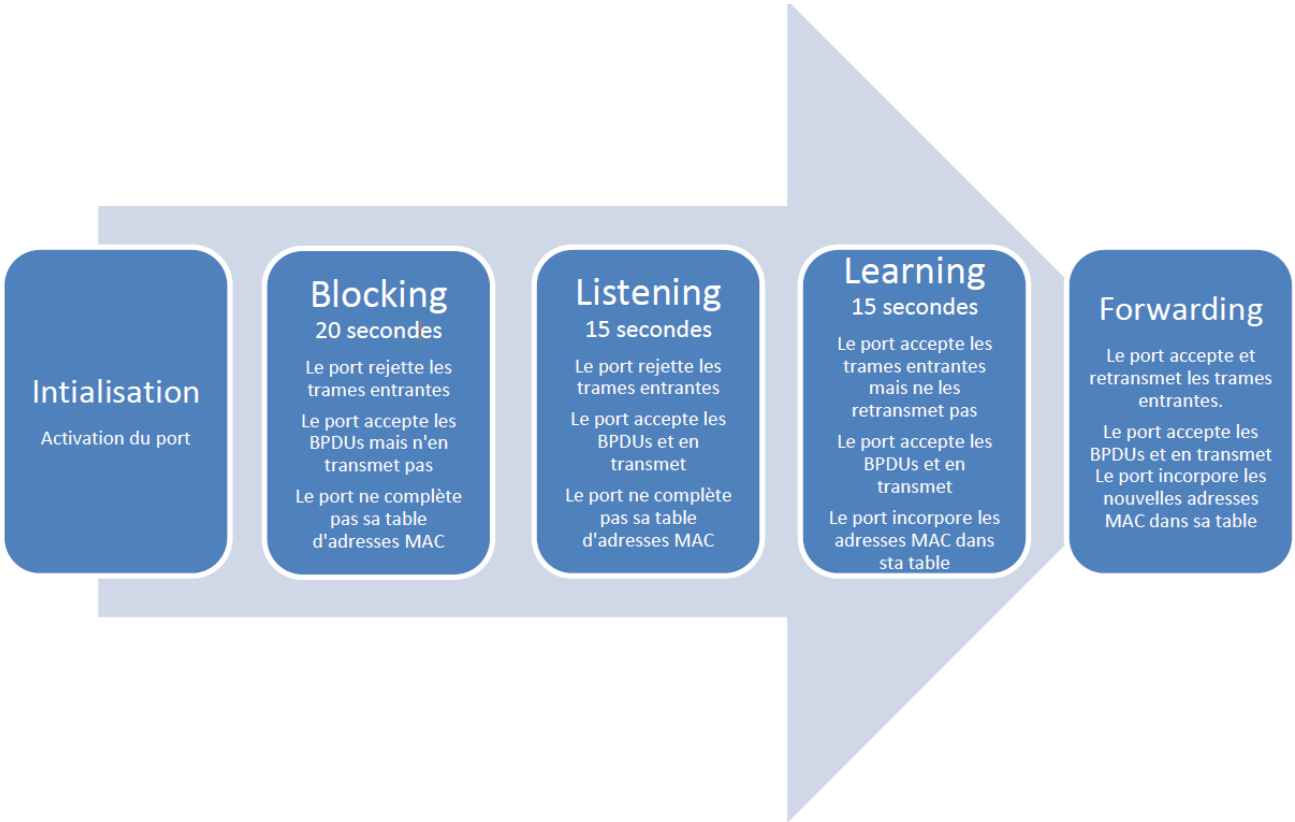
Dans le cas de figure ci-dessus, SW2 est élu "Pont racine", non pas parce qu'il a la plus petite adresse MAC, mais parce qu'on lui a affecté la plus petite valeur de priorité.

Sur les commutateurs CISCO, on fonctionne par défaut en PVST (Per Vlan Spanning-Tree). Il faut donc préciser le ou les vlans sur le(s)quel(s) la priorité doit être affectée.

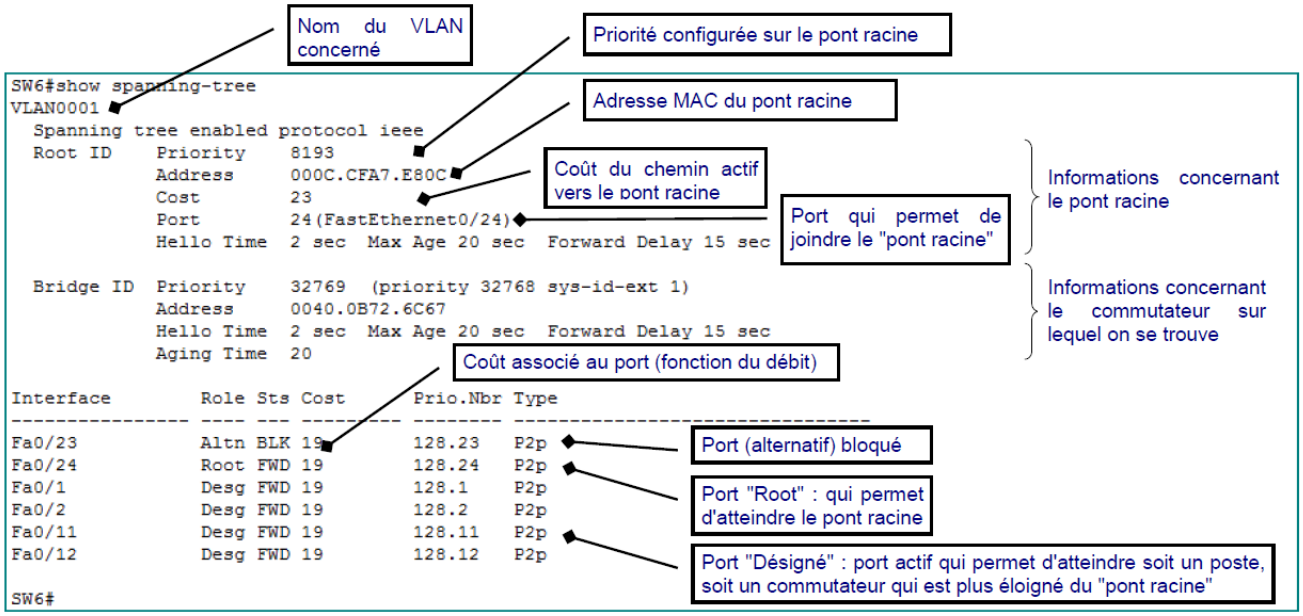
Exemple : spanning-tree vlan 1 priority 4096

Progression de l'état d'un port (suite à un branchement)

Voici en résumé, comment un port s'active et passe en mode "Forwarding" lorsqu'un câble est branché sur ce port :



Décryptage des informations spanning-tree sur un commutateur



Quelques remarques complémentaires :

- Dans le cas de figure ci-dessus, un seul vlan, le vlan par défaut, est configuré.
- Les priorités affichées ne sont pas celles configurées : l'identifiant de vlan est inclus dans la priorité.

  Exemple : 32 769 correspond bien à la priorité 32 768 pour le vlan 1 comme indiqué dans la priorité du Bridge ID (Priority 32768 sys-id-est 1).

Sources d'informations utilisées pour construire ce support – ou pour aller plus loin

CISCOMADESIMPLE.BE (notamment la fiche récapitulative sur STP)

STP – Spanning Tree Protocol

Spanning Tree – Théorie | Networklab

Cisco Nexus 5000 Series NX-OS Software Configuration Guide - Configuring Multiple Spanning Tree [Cisco Nexus 5000 Series Switches] - Cisco