Gestion et Administration

des commutateurs Cisco

I. Les concepts de base

Les commutateurs en fonction de base travaillent (comme les ponts) au niveau de la couche 2 de l'OSI (Open System Interconnexion) nommée aussi Liaison de données. Le commutateur est le centre de la topologie étoile. La couche 2 a pour unité de transmission la trame.
A la différence du concentrateur, qui travaille au niveau de la couche 1 de l'OSI (unité de transmission le bit), les commutateurs ou les ponts ont la capacité d'analyser le trafic et ainsi de posséder une connaissance des adresses MAC (Medium Access Control) et construire des tables.

1. Fonctionnement

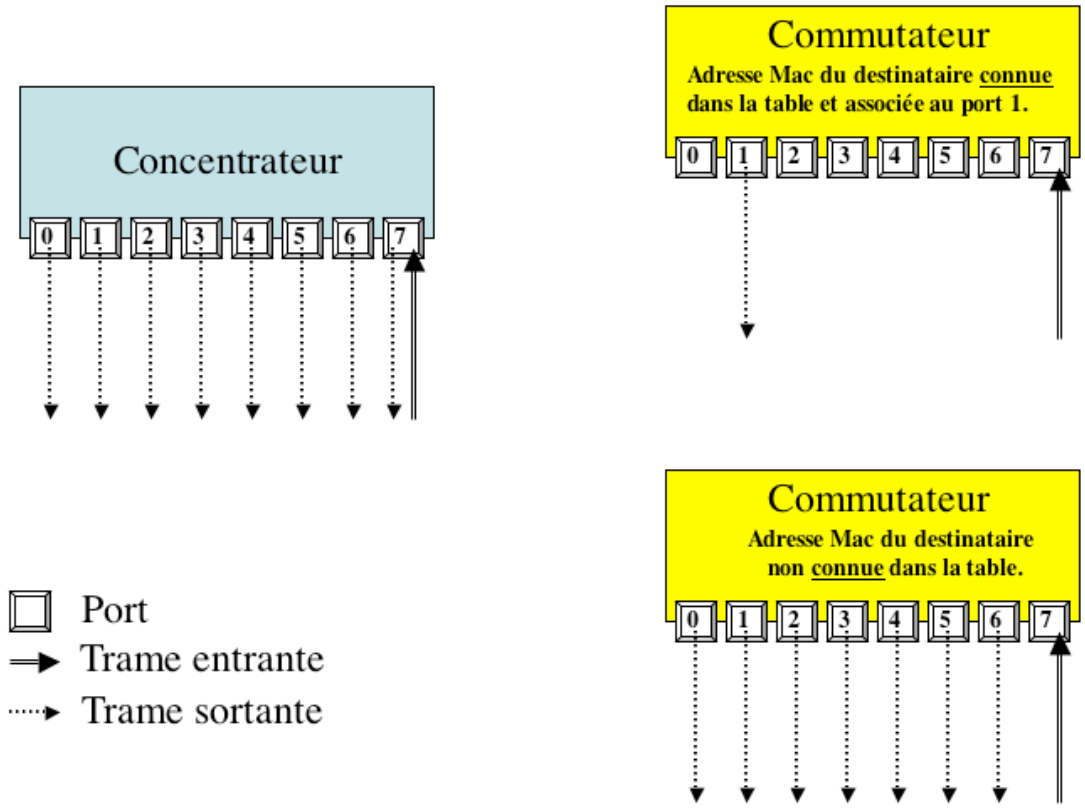
a.Les particularités du commutateur

Le commutateur possède 3 particularités essentielles

- la capacité d'apprendre les adresses MAC des matériels attachés à ses ports.
- La capacité de diriger la trame vers le bon port si l'adresse MAC destinatrice est référencée dans sa table.
- La capacité de détecter et d'éviter les bouclages ou redondances grâce au protocole spanning-tree.

Grâce à ces particularités, le commutateur diminue de manière importante l'utilisation de la bande passante par rapport à l'utilisation d'une topologie bus ou étoile associée à un concentrateur (on parle d'émulation de bus dans le cas de concentrateur).
En effet, leur capacité à « dresser » un état des adresses MAC par port au sein d'une table d'adressage MAC (aussi appelé « content-addressable memory : CAM ») leurs permet, via l'écoute du trafic entrant

de diriger les trames vers un port unique dans le cas ou cette adresse est référencée dans la table. Si l'adresse MAC n'est pas connue, il envoie la trame vers tous les ports hormis celui d'où provient la trame (port entrant).
Là où le concentrateur inonde tous les ports, le commutateur choisit au mieux le bon port. Ceci implique une diminution importante du « bruit » sur le réseaux.



b. Les modes de transmission

Afin de transmettre les trames, le commutateur propose 3 modes implémentés selon les modèles de commutateurs.

- Store&Forward

Une analyse poussée de la trame entrante est effectuée. La trame est reçue en totalité avant d'être retransmise. Les adresses MAC destinatrices et sources contenues dans l'entête ethernet sont lues. Dans ce cas, un contrôle de redondance est effectué et, si un filtrage est mis en place, il est appliqué à la trame. La trame est transmise si les contrôles sont bons.

Ce système implique un temps de latence dépendant de la taille de la trame. En effet celle-ci est reçue en totalité avant son analyse et son éventuelle transmission.

- Cut-through

Seule l'adresse MAC destinatrice est lue dans l'entête ethernet de la trame. Dès que cette adresse est récupérée, la transmission de a trame se fait et ce même si celle-ci n'est pas encore reçue en totalité. Malgré tout, dans certains modèles de commutateurs, une analyse identique au mode store&forward peut être en plus activée. Par contre, si la trame est erronée, elle sera tout de même envoyée. Il est malgré tout possible, dans le cas où il y a trop d'erreurs, de basculer en mode store&forward manuellement ou de manière automatique.
Dans ce cas le mode cut-through allie rapidité et gestion des erreurs.

- Fragment-free

Ce mode propose d'analyser seulement les 64 premiers octets de la trame. Ce chiffre n'est bien sûr pas innocent, il représente la taille minimum d'une trame ethernet. En effet, lors d'une collision, une « trame de collision » est générée (appelée aussi fragment), sa taille est inférieure à 64 octets. Ceci permet au commutateur de les détecter et ainsi de les détruire, à contrario du mode cut-through.
Du fait de l'analyse un peu plus poussée, ce mode est plus lent que le cut-through.

c. L'apprentissage des adresses MAC

Le fonctionnement du commutateur est basé sur sa capacité à apprendre les adresses MAC des matériels connectés à ses ports.
Cet apprentissage permet la création et la mise à jour d'une table contenant des couples adresses MAC-port.
Selon le type de commutateur, elle peut posséder 1024, 8192 (catalyst 2950) … entrées.
Au démarrage du commutateur, cette table est vide.
Comment se passent alors l'envoi des trames et l'apprentissage des adresses ?
Lorsque le commutateur doit envoyer une trame et qu'il ne possède pas dans sa tabme la correspondance adresse MAC-port, il est obligé d'envoyer la trame sur tous les ports sauf le port entrant (d'où provient la trame).
Au démarrage, c'est aussi ce qui se déroule. On parle alors de « flooding ».

Prenons le cas d'un commutateur store&forward. Le commutateur va mettre à jour, en mémoire, sa table de couples port/adresses Mac à chaque passage d'une trame entrante. Il récupère l'adresse mac source (et non l'adresse Mac destination) puis ajoute ou met à jour une entrée dans la table (port entrant/adresse Mac source). Cette entrée n'est pas éternellement présente dans la table. En effet , lors de changement de carte réseau, de port de connexion sur le commutateur, de débranchement, d'arrêt de matériel … il faut que la table soit mise à jour pour qu'elle prenne en compte ces modifications.
Ceci est permis grâce à une durée maximale de présence dans la table d'une de ces associations (port/adresse Mac).
Cette durée est appelée « **time-age** ».

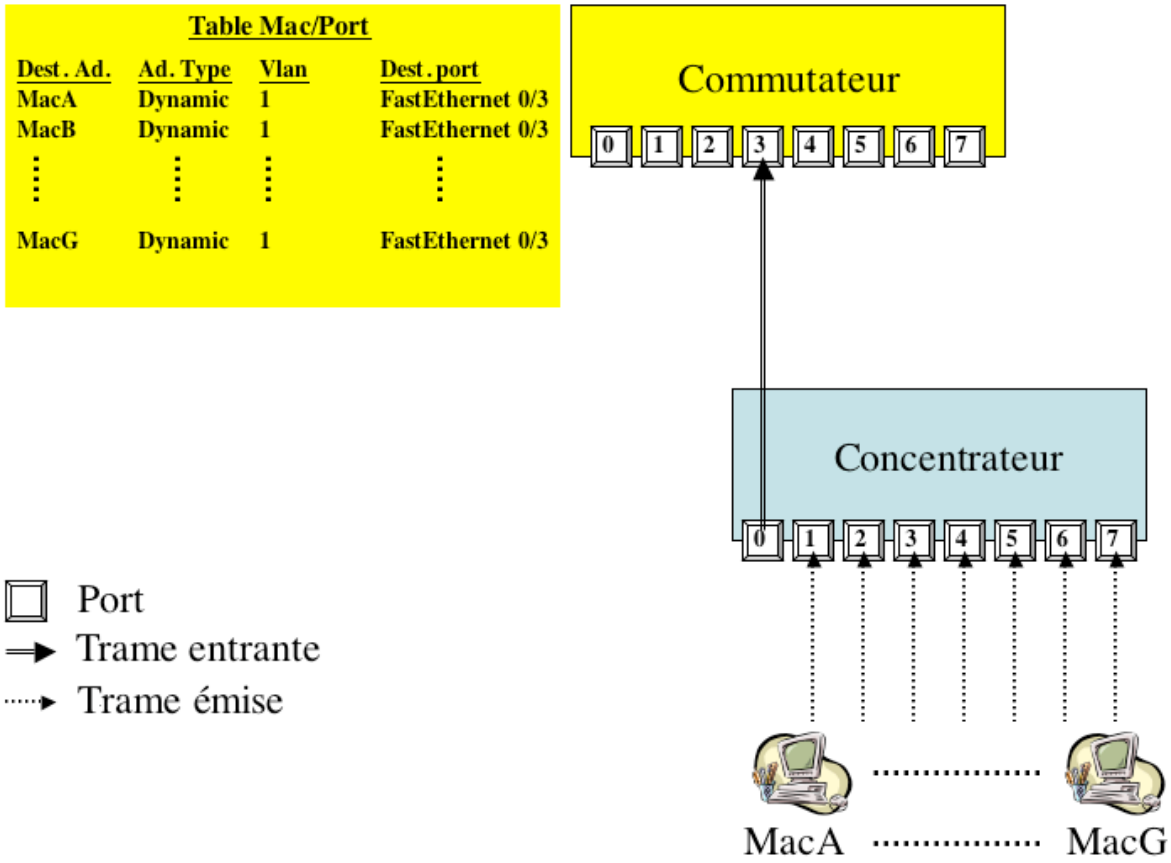
- Si un matériel n'émet pas une trame avant que cette durée ne soit atteinte, l'entrée est supprimée de la table.
- Par contre, si une émission a lieu avant le terme de cette durée, le « time-age » est alors remis à zéro.

Le commutateur ne se souvient donc que des matériels les plus actifs. Si un matériel est changé de port et qu'il émet, son entrée sera alors mise à jour en prenant en compte le nouveau port d'attache de ce matériel.
Le « time-age » est paramétrable. Par défaut il est de 5 minutes sur le 2950.
On retrouve ce « time-age » au niveau des tables de routages des routeurs.

Remarque
Si la table est pleine, les communications vers des adresses inconnues sont envoyées vers tous les ports du commutateur y compris le port entrant.

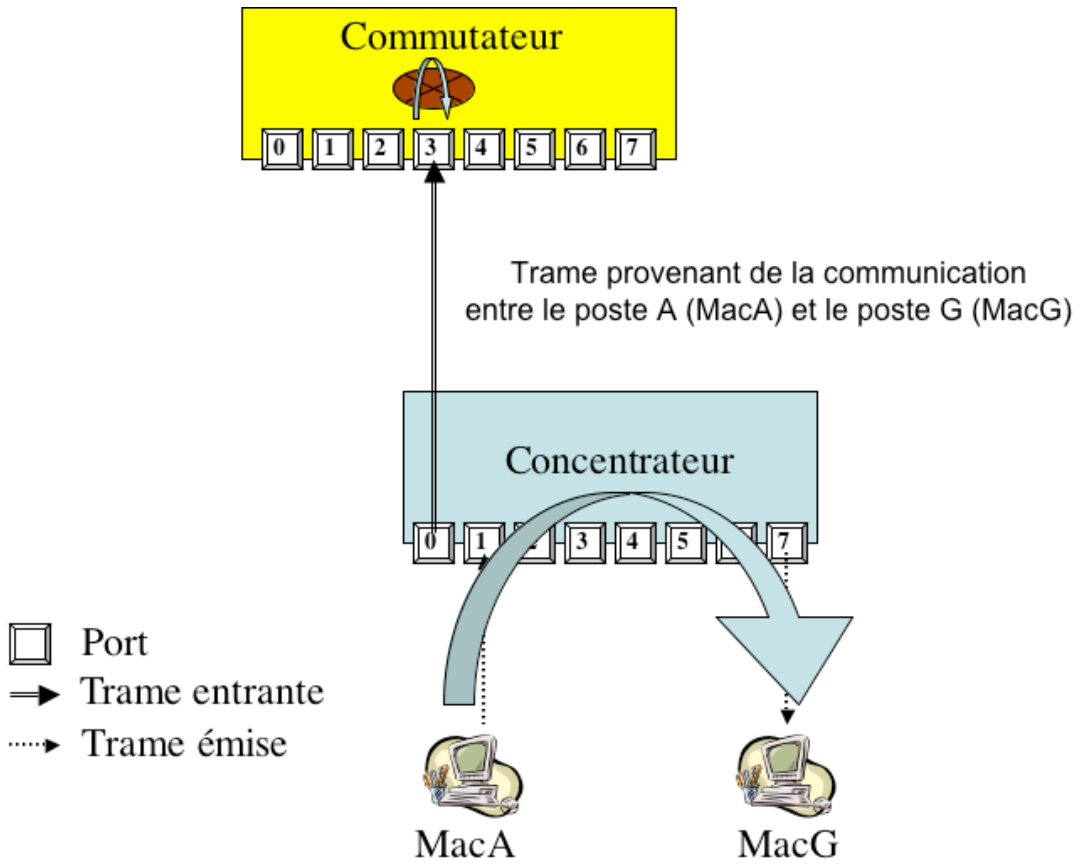
Rappel
Lors de l'émission d'une trame, il y a commutation vers le bon port si l'adresse Mac destinatrice est connue dans la table. Il y a en même temps regénération de l'entrée associée à l'adresse Mac source dans cette table.

Si a un port, plusieurs matériels sont reliés (via un concentrateur, un commutateur …), la table génère une entrée par matériel émetteur.



d. Commutation particulière

- Si dans votre topologie réseau, un concentrateur est connecté à un port d'un commutateur et que deux matériels attachés au concentrateur communiquent entre eux, le commutateur rejettera la trame en provenance du concentrateur. En effet, le commutateur dans ce cas devrait envoyer la trame vers le port entrant ce qui implique une redondance ou une erreur de chemin et par conséquent, une destruction de la trame.



- Les trames de broadcast ou de multicast sont des cas à part.

Ces modes permettent de communiquer avec tous les matériels d'un réseau.
Par conséquent, dans ces deux cas, le commutateur effectue un « flooding ».

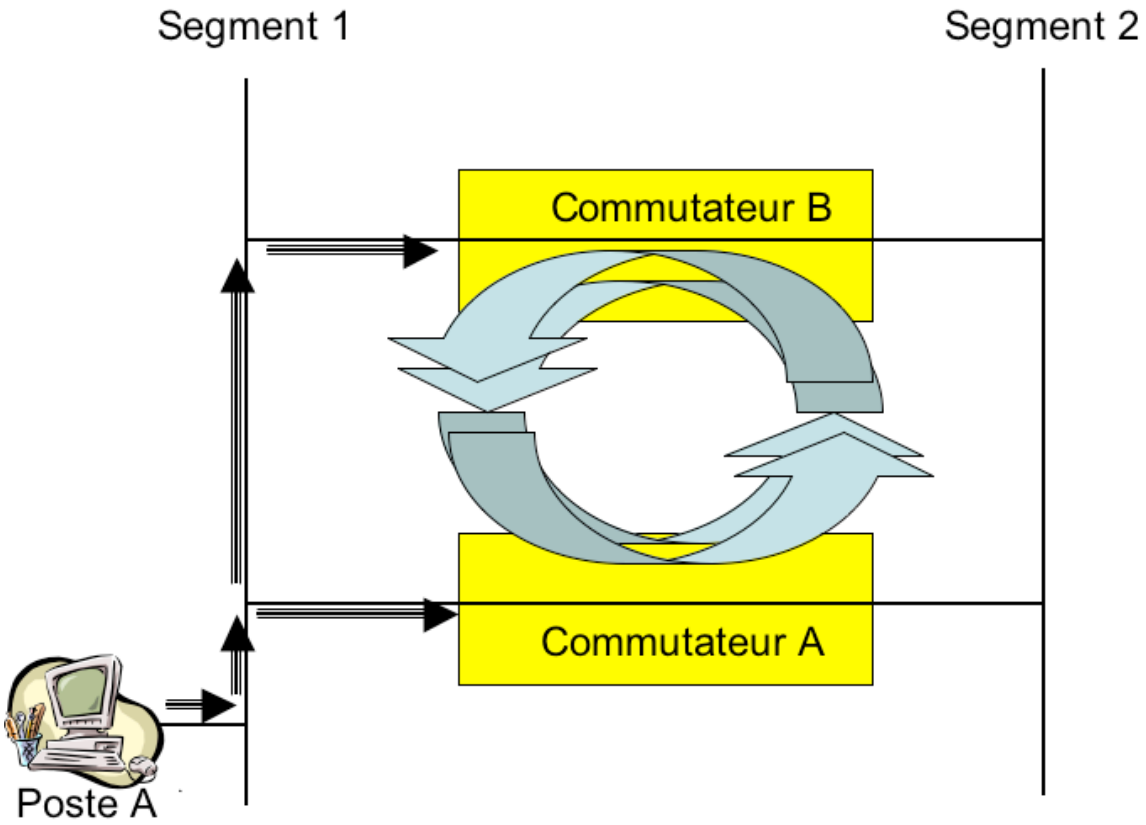
2. Le Spanning Tree

L'un des problèmes lors de ma définition d'une topologie réseau est la redondance. C'est-à-dire, la possibilité pour une trame d'atteindre le destinataire en empruntant plusieurs chemins. Cette possibilité augmente la fiabilité du réseau en évitant la paralysie des transmissions en cas de panne d'une des matériels d'interconnexion.
Par contre, cela peut entraîner des cascades de problèmes. La mise en place du spanning tree permet de les éviter.

a. Les problèmes

- La tempête de broadcast

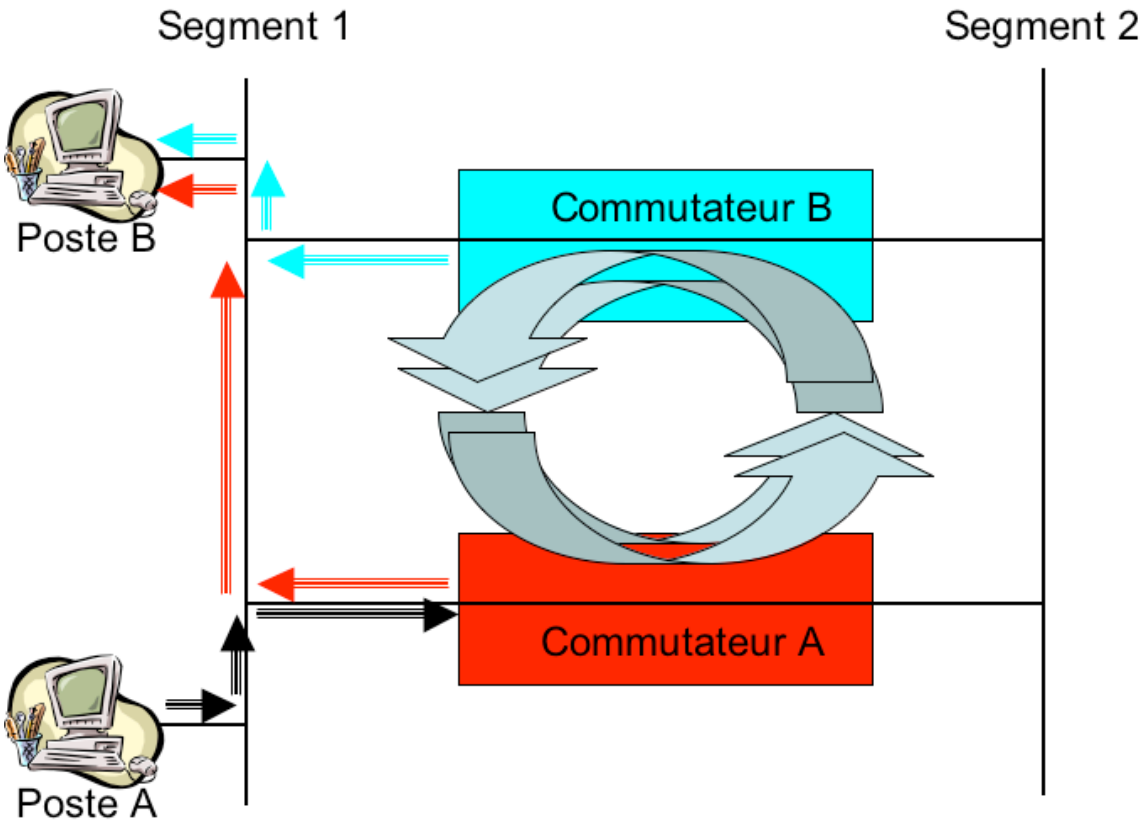
Une trame de broadcast est envoyée vers le commutateur A par le poste A.



Ce commutateur A renvoie alors la trame vers tous les ports (sauf le port entrant) y compris en direction du commutateur B. Le commutateur B va dupliquer et renvoyer la trame vers tous les ports (sauf entrant) et donc vers le commutateur A et ainsi de suite. La trame va être dupliquée et ce schéma de communication va recommencer en boucle.

- Réception multiple d'une trame

Dans le schéma qui suit, il est tout a fait possible que le matériel B (Poste B) reçoive via le commutateur A ET le commutateur B la même trame en provenance du poste A. Les matériels ne sont pas capables de gérer ce type de communication d'où une perturbation importante des communications.



- Instabilité de la table des adresses Mac

Dans le cas où, la trame est dupliquée et transite du poste A vers le commutateur A en passant par son port 0 et si dans le même temps, cette trame est envoyée du commutateur B vers le commutateur A sur son port 1. Alors, la table des adresses Mac du commutateur A va constamment se remettre à jour en indiquant une fois sur 2 (lors de l'envoi par le commutateur B) un port erroné ou du moins non optimum.

b. La solution : le spanning tree

Pour éviter ces phénomènes, il est nécessaire de mettre en place un processus de suppression des boucles de façon logique : le spanning tree (STP : Spanning Tre Protocol).
Ce protocole travail au niveau 2 de la couche OSI et est implémenté généralement sur les commutateurs, ponts et routeur (matériels pontés).
Historiquement créé par DEC (Digital Equipment Corporation), l'algorithme est ensuite normalisé par l'IEEE sous la dénomination 802.1d.
Ce protocole permet d'éviter les boucles en bloquant certains ports sur les matériels précités et en les débloquant en cas de panne.

c. Le fonctionnement du spanning tree

Ce protocole se base sur l'identification du « pont » principal (matériel ayant une fonction de port : commutateur et ou routeur) et du calcul du meilleur coût pour aller d'un point A du réseau à un point B.

- Détermination du pont principal ou « root bridge »

Ce matériel voit tous ses ports définis comme non bloqués (designated ports). Ils peuvent donc faire transiter les trames.

Remarque

Le port ayant le meilleur coût (valeur minimum) n'est pas forcément celui qui est le plus direct en terme de sauts (nombre de matériels par lesquels la trame doit transiter pour arriver à destination). L'élément pris en compte est la bande passante. Le chemin le plus rapide est donc privilégié. La norme de l'IEEE définit les coûts en fonction de la vitesse.

Débit	Coût
10Mbits/s	100
100Mbits/s	19
1Gbits/s	4
10Gbits/s	2

Le « root bridge » est déterminé en prenant la valeur la plus basse sur 8 octets. Cette valeur est calculée à partir de 2 octets pour la priorité (à définir sur la matériel sinon il y a un valeur par défaut) et des 6 octets de l'adresse Mac du matériel.
En clair entre 2 commutateurs ayant la même priorité, le « root bridge » est celui qui possède l'adresse Mac la plus petite.
Les matériels s'échangent ces informations via des trames multicats (toutes les 2 secondes par défaut chez Cisco). Ces trames se nomment BPDU (Bridge Protocol Data Unit)

- Une fois le « root bridge » déterminé, les autres ponts ou matériels pontés, vont définir pour chacun de leurs ports, le chemin de moindre coût pour atteindre le « root bridge ». Ces matériels sont appelés « non root bridge ».

Chaque port peut prendre deux états finaux :

- ✓ Blocking, les trames ne transitent pas par ce port
- ✓ Forwarding, les trames transitent par ce port

Entre ces 2 états, il existe 2 autres états intermédiaires :

- ✓ Listening, le port écoute et récupère les BPDU
- ✓ Learning, le port apprend et met à jour la table des adresses Mac

Pour mieux comprendre, voyons le processus d'un « non root bridge » de son démarrage à son activation.

- Au démarrage

Il se désigne comme « root bridge » par défaut. Tous les ports sont bloqués puis passent en mode listening.
Il y a alors échange avec les autres ponts de BPDU (2 secondes de latence). Il s'aperçoit alors qu'il est en fait un « non root bridge ». Il lui faut donc définir le coût de ses ports vers le « root bridge ».
Les ports « élus » deviennent des « designated ports » après 15 secondes et passent alors en mode learning. Les autres ports passent en mode blocking.

Dans l'état learning, le port construit la table Mac. Il reçoit des trames mais ne les fait pas transiter.
Cet état dure 15 secondes. Le port passe alors en mode forwarding et peut enfin faire transiter les trames des utilisateurs.

Remarque
Tant que le port n'est pas en mode forwarding, aucune trame n'est transmise. Durant la phase learning,le port reçoit les trames des utilisateurs pour générer la table Mac mais les trames ne sont pas transmises. Ce système est extrêmement important car il diminue voir supprime l'effet de flooding.

- Mise à jour

Il existe un élément appelé « max-age » qui correspond à une timer de 20 secondes qui permet de déterminer si le « root bridge » est en panne.
En effet, si pendant ce max-age, acun BPDU n'est arrivé, le « non root bridge » estime que le « root bridge » est en panne.
Il recalcule alors le spanning tree en passant tous ses ports (y compris les port dans l'état blocking) en mode listening.
Au bout de 30 secondes, le spanning tree est à nouveau opérationnel.
.
Remarque
On parle de convergence lorsque tous les ports de tous les commutateurs sont dans l'un des 2 états finaux.
Plus le temps de convergence est faible mieux c'est.
Afin d'améliorer ce temps de convergence, l'IEEE a défini le RSTP (Rapid Spanning Tree Protocol) sous la norme 802.1w.
Ce protocole permet une transition rapide de l'état blocking à forwarding en cas de recalcule.