Haute disponibilité : Présentation et étude du niveau 2

Importance et définitions

Le rôle de l'administrateur est de veiller au bon fonctionnement du système d'information. Pour cela il devra avoir recours à des principes tels que la haute disponibilité pour la tolérance de panne. Ceci peut se traduire par l'utilisation d'un second lien physique ou d'un équipement redondant prêt à prendre le relai en cas de panne du premier. Celui-ci peut être sous la forme d'un cable, d'une seconde ligne internet, d'un second routeur, d'un second firewall, …

L'administrateur doit également assurer les performances du réseau. Nous pourrons ainsi aller un peu plus loin dans la haute disponibilité à partir du moment où notre second équipement ne reste pas passif, mais devient actif. Notre réseau sera donc disponible, mais aussi on pourra faire de la répartition de charge (Load Balancing) entre différents équipements actifs à un instant t.

Notre réseau gagnera ainsi en performance, en capacité, et en haute disponibilité.

On peut classer les différents mécanismes de haute disponibilité selon qu'ils assurent une redondance de niveau 2 s'ils sont mis en place sur un commutateur (dans ce cas cette redondance désignera un ou des liens supplémentaires physique entre plusieurs commutateurs ou entre un commutateur et un équipement du réseau), ou une redondance de niveau 3 s'ils interviennent au niveau de la couche IP

Au niveau 2 nous trouvons des protocoles de haute disponibilités tels que :

- Le STP (spanning tree)
- Port Channel (EtherChannel)
- PAgP et LACP
- Stacking de commutateur

Au niveau 3 :

- FHRP
- HSRP
- VRPP

I. Le Spanning Tree

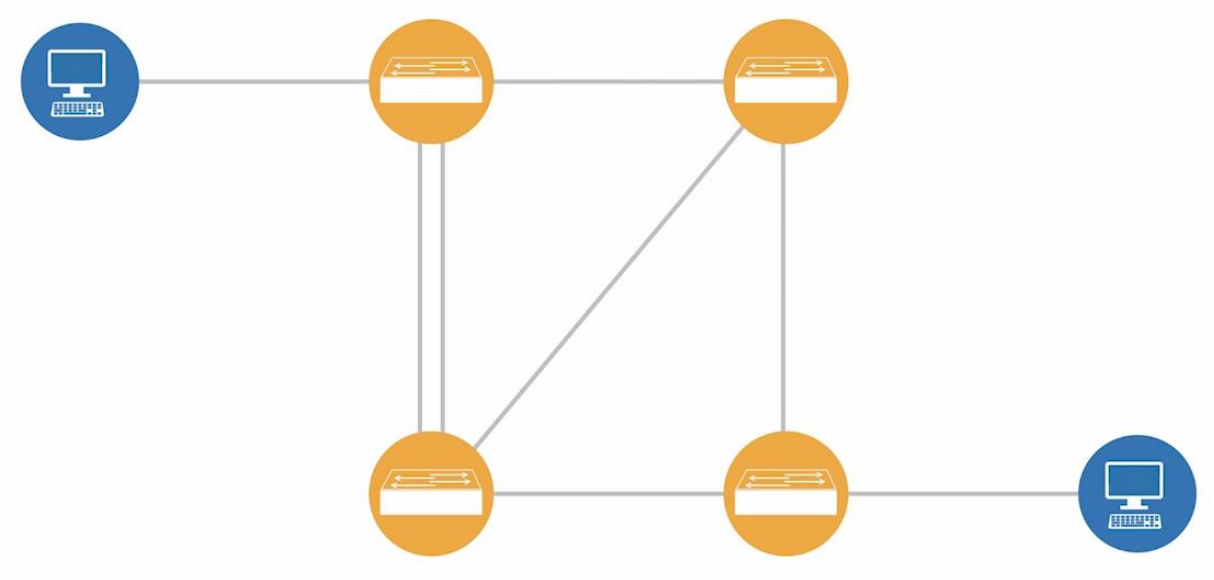
1. Présentation

La mission première du spanning tree est de détecter et de couper automatiquement les boucles au niveau ethernet dans un réseau

On utilise souvent le terme spanning tree pour désigner en fait, plusieurs protocoles qui ont pour origine le spanning tree 802.1D initial mais sur lequel les constructeurs ont effectué leurs propres améliorations.

On peut citer par exemple le RSTP, rapid spanning tree, ou chez Cisco le rps tv plus qui permet notamment, de prendre en compte les VLAN.

Le spanning tree va permettre également la mise en place d'une double liaison entre switches à des fins de tolérance de panne d'un switch ou d'un lien.



En fait, avec ces systèmes, on détourne un peu le but premier du spanning tree qui était celui d'empêcher les boucles étant donné que cet algorithme est recalculé à chaque changement dans la topologie, donc une coupure de lien, une coupure de switch, c'est un changement et ça permet en fait du coup de mettre en place des architectures avec une redondance assurée automatiquement.

2. Fonctionnement

Pour comprendre le spanning tree et ensuite bien le mettre en place sur le switch Cisco, il est nécessaire de le comprendre. Cet algorithme va permettre la détection de boucles et bloquer un ou plusieurs ports sur un ou plusieurs switches de l'infrastructure. L'algorithme ne dépend pas uniquement d'un seul switch mais c'est toute l'infrastructure qui va devoir être prise en compte.

Les switches vont devoir dialoguer, se mettre d'accord sur les ports à bloquer ou pas, mais avant cela, l'élection d'un switch dit racine ou root bridge dans l'infrastructure va être réalisée.

L'algorithme de spanning tree appelé STA désigne un commutateur unique comme pont racine. Et il l'utilise ensuite comme point de référence pour le calcul de tous les chemins possibles.

On parle de bridge root. Tous les commutateurs associés au protocole spanning tree, échangent des trames appelées des BPDU pour identifier le commutateur doté de l'ID de pont le plus faible sur le réseau. Ce dernier devient automatiquement le pont racine pour la suite du calcul de l'algorithme.

L'identifiant est composé de la façon suivante :

PRIORITÉ FIXÉE	ID SYSTÈME ÉTENDU	ADRESSE MAC
4 bits	12 bits	48 bits

La priorité qui est paramétrable, l'ID système étendu n'était pas obligatoire dans les premières versions qui ne géraient pas les VLAN. Maintenant, il est indispensable et si on veut gérer les VLAN, il faudra que cette priorité contienne le VLAN ID.

Enfin si on imagine des switches configurés par défaut et sans VLAN, donc, VLAN par défaut, le VLAN 1, ce qui est très souvent le cas, on se basera alors sur l'adresse Mac de l'équipement. Le switch contenant l'adresse Mac avec le plus de bit de poids fort à 0 sera le root bridge.

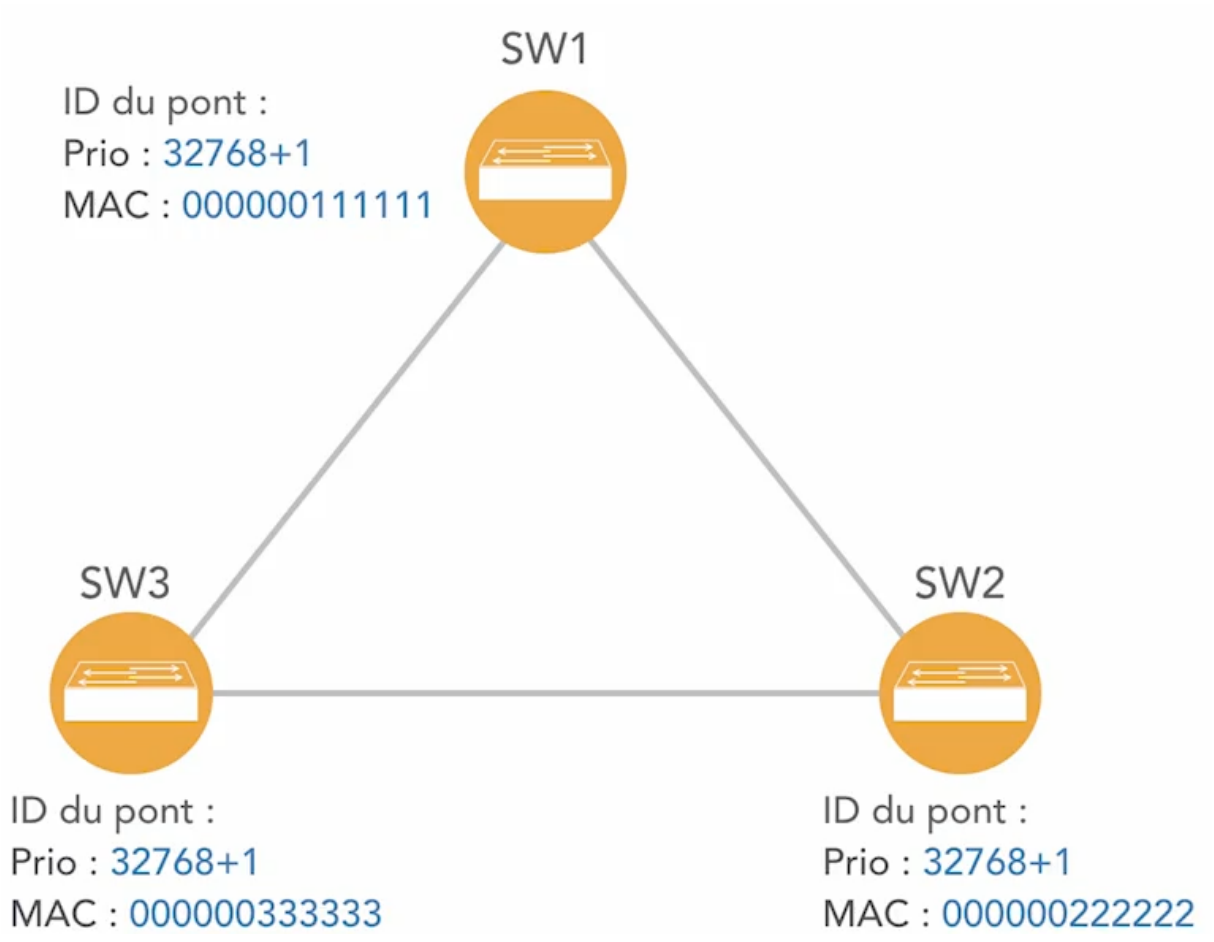
Il faut bien comprendre que quand on parlera de priorités, on prendra en fait en compte les 16 premiers bits. On ne modifiera jamais l'adresse Mac du switch en fait.

Dans le spanning tree, les switches s'envoient des trames BPDU avec cet identifiant à intervalle de temps régulier.

Exemple : chacun de ces switches ont pour défaut une priorité à 32 769. Alors en fait, la priorité par défaut est égale à 32768 et comme nous sommes dans le VLAN 1, la priorité est 32 769.

On prend en compte l'intégralité de l'identifiant.

Les switches vont s'envoyer chacun des trames BPDU avec leurs identifiants en se proclamant root.



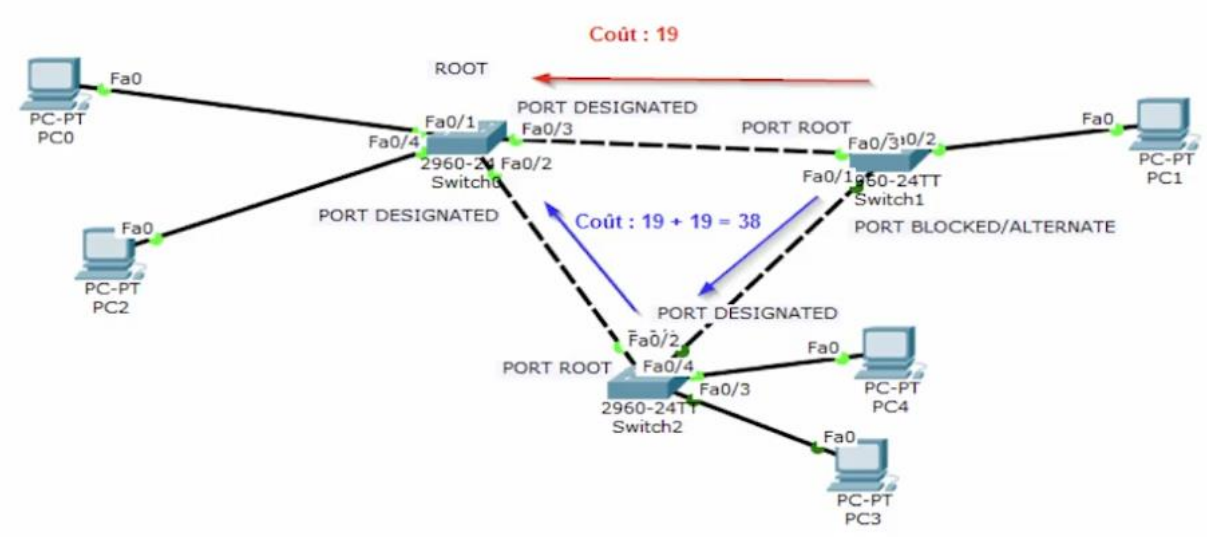
Quel est le switch root ? et pourquoi ?

Une fois le switch racine défini, l'algorithme a calculé le chemin le plus court pour y parvenir à partir de chaque switch. Chaque switch utilise l'algorithme STA pour identifier les ports devant être bloqués. Pendant que l'algorithme détermine les meilleurs chemins pour accéder au switch racine, depuis l'ensemble des ports de commutation de l'infrastructure, le réseau est paralysé et les échanges ne sont pas encore possibles. L'algorithme STA prend en compte le coût des chemins comme celui des ports lorsqu'il détermine quel port bloquer. Le coût de la route est calculé à l'aide des valeurs de coût de port qui sont en fait fonction de la vitesse du port. La somme des valeurs des coûts de ports, détermine le coût du chemin global vers le pont racine. Si plusieurs chemins sont disponibles, l'algorithme STA choisit le chemin doté du coût de chemin le plus faible.

Dans le tableau suivant, qui fait le lien en fait entre la vitesse de liaison et le coût du port par défaut, on remarque que plus la liaison a une vitesse rapide, plus le coût du port est bas et plus du coût, le chemin sera prioritaire.

VITESSE DE LIAISON	COÛT
10 Gbit/s	2
1 Gbit/s	4
100 Mbit/s	19
10 Mbit/s	100

On peut fixer dans tous les cas manuellement si on le souhaite, le coût du port avec la commande ***spanning tree cost*** et la valeur.



Dans ce schéma-là, le switch en haut à gauche est le switch qui a été élu root. Il est relié au switch 1 ici par un lien fast internet avec donc un coût de 19. Et il est relié aussi au switch 2 avec un coût de 19. Et le switch root est aussi relié par l'intermédiaire de ce switch 2 au premier switch avec un coût de 19, plus 19 soit 38. Donc le meilleur chemin sera celui du haut.

Lorsque l'algorithme STA a déterminé quels sont les meilleurs chemins possibles pour chaque commutateur, on va attribuer un rôle à chacun des ports de l'infrastructure. Sur un switch donné, on aura donc des ports racines (Root Port RP). Il s'agit des ports de commutation les plus proches du port racine, des ports désignés (Designated Port DP), il s'agit de tous les ports non racine qui sont autorisés à acheminer du trafic sur le réseau, tous les ports d'un switch root par exemple sont des

ports désignés. Enfin, on aura des ports bloqués (Blocking Port BLK) ou en mode alternate, il s'agit du port qui a été désigné comme devant être bloqué.

À noter que la notion de port alternate est une notion qui apparaît lorsqu'on met en place le spanning tree sur Cisco avec le protocole par défaut. Lorsque sur un lien entre deux switches qui ne sont pas root, le coût du chemin est le même, donc les ports ont la même vitesse, le port qui sera bloqué sera le port du switch avec l'identifiant, le bridge ID, le plus élevé. On pourra alors agir sur la priorité du port directement qui est une autre variable, pour choisir le port bloqué et le port qui sera désigné sur le segment. Par défaut la priorité d'un port est de 128.x, x étant le numéro physique du port. Le port fa0/15 aura donc une priorité 128.15. Le port avec la priorité la plus faible sera donc prioritaire.

3. Les différents protocoles basés sur STP

Le RSTP, « Rapid Spanning Tree », est, comme son nom l'indique, plus rapide que le Spanning Tree initial. Il le remplace tout en conservant une rétro-compatibilité. La majeure partie de la terminologie du protocole 802.1 d initial perdure, et les principaux paramètres restent inchangés. La rapidité est notamment possible par la limitation de la diffusion des BPDUS aux voisins directs. La convergence est aussi plus rapide car un switch peut diffuser des trames sur un port dit « alternate », sans attendre que l'algorithme ait convergé dans toute l'infrastructure. Par contre, le RSTP normalisé 802.1w ne sait gérer en l'état qu'une seule instance de VLAN.

Chez Cisco et uniquement chez Cisco existe le protocole PVST, PVST+. En fait, c'est ce protocole qui est utilisé par défaut et c'est autour de ce-dernier que nous avons tout à l'heure observé le fonctionnement de l'algorithme. Il n'est pas très différent du Spanning Tree initial mais sait gérer en fait plusieurs instances de Spanning Tree dans l'infrastructure : une par VLAN.

Avec le STP initial, si un lien est bloqué, c'est tous les VLANs qui le sont. Or une boucle réseau est valable pour un domaine de broadcast donné. Et donc pour un VLAN donné il fallait donc trouver un algorithme qui permette de gérer les VLAN. Le PVST+ saura le faire.

Il existe également le PVRST+ chez Cisco, qui reprend en fait les avantages du RSTP mais qui est basé sur PVST, donc avec prise en compte de la gestion des multi VLAN.