

3

Architecture des réseaux locaux

Le système de câblage étant maintenant prêt à l'usage, on peut alors commencer à y installer un réseau.

La conception d'une architecture réseau est élaborée en fonction du nombre de postes de travail à connecter. On peut considérer les cas suivants :

- petit réseau : moins de 200 postes dans un même bâtiment ;
- réseau moyen : de 200 à 800 postes dans un même bâtiment ;
- gros réseau : plus de 800 postes dans un même bâtiment ;
- 1^{ère} variante : plusieurs bâtiments contenant un nombre varié de postes de travail ;
- 2^{ème} variante : plusieurs sites contenant un nombre varié de postes de travail.

Bien qu'arbitraires, ces bornes correspondent à des ordres de grandeur et à des sauts technologiques. En effet, plus le nombre de postes est important, plus il faut répondre à un certain nombre de contraintes et d'exigences qui n'apparaissent qu'avec la complexité du réseau.

L'architecture réseau est bien sûr liée au système de câblage, mais celui-ci doit avoir été conçu pour en limiter les contraintes, c'est-à-dire s'adapter à toutes les situations. Le chapitre précédent a aidé à œuvrer en ce sens. Tous les systèmes de câblage sont en étoile, de même que la topologie ; les équipements actifs seront donc positionnés dans les locaux techniques. Ils serviront à connecter les postes de travail aux serveurs.

Dans le cas le plus simple, la conception d'une architecture consiste à choisir et à positionner les équipements actifs, puis à les connecter entre eux en utilisant le câblage.

Les choix de base

Quel type de réseau choisir ?

La création d'un réseau local nécessite de faire des choix, et tout d'abord celui du type : Ethernet, Token-Ring ou un autre ?

Le premier est le plus répandu et le moins cher, alors que, pour le deuxième c'est l'inverse. Le choix ira donc de préférence au premier. L'intérêt de Token-Ring est surtout sa compatibilité avec les équipements grands systèmes IBM (3090, etc.). Si vous avez des contrôleurs 3174, des terminaux 3270 ou des PC devant se connecter en émulation 3270 à un système central IBM, le choix se portera naturellement vers ce type de réseau local. Mais ce n'est pas une obligation, car la plupart des équipements IBM acceptent désormais des cartes Ethernet. Cela étant, les postes de travail peuvent toujours être connectés à un réseau Ethernet, et les systèmes IBM à des réseaux Token-Ring.

D'autres solutions sont envisageables pour construire un réseau local, mais elles sont nettement plus chères. Il s'agit, par exemple, d'ATM (*Asynchronous Transfer Mode*). Ce type de réseau est surtout destiné à d'autres usages que nous verrons au chapitre 11.

Pour toutes ces raisons, le choix d'Ethernet s'impose.

Concernant la topologie, la plus pratique est celle de l'étoile : tous les systèmes de câblage sont, on l'a vu, fondés sur ce principe.

Quel débit retenir ?

La décision suivante concerne le débit du réseau, c'est-à-dire la vitesse de transmission des trames Ethernet, encore appelée bande passante. La norme Ethernet est déclinée en plusieurs variantes : 10 Mbit/s (norme 10bT), 100 Mbit/s (norme 100bT) et 1 Gbit/s (norme 1000bT). De par son coût et son caractère innovateur, le Gigabit Ethernet est réservé aux liaisons entre les équipements de concentration et aux serveurs.

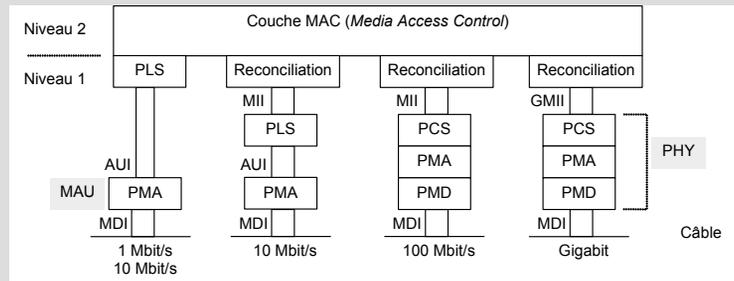
Le choix du débit se fera donc en fonction des coûts, plutôt 10/100 Mbit/s pour les PC et 100/1000 Mbit/s pour les serveurs.

Débit	Utilisation
10 Mbit/s	Postes de travail bureautique
100 Mbit/s	Postes de travail multimédias et serveurs
1 Gbit/s	Pour connecter les équipements réseaux entre eux ainsi que les gros serveurs

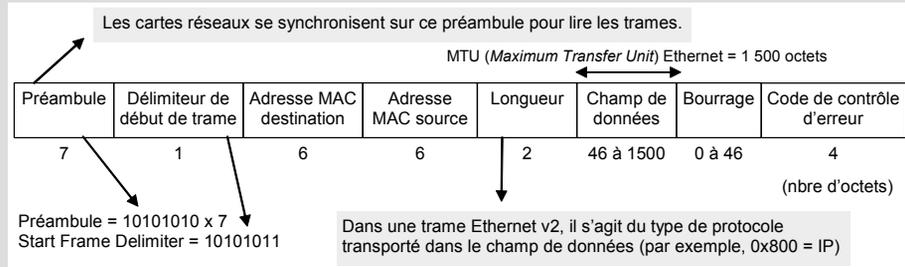
Côté poste de travail, la plupart des cartes réseau fonctionnent à 10 et 100 Mbit/s, et sont au même prix que les cartes 10 Mbit/s. En outre, la plupart des commutateurs offrent des ports à détection automatique de vitesse (port *autosense*).

LE POINT SUR ETHERNET (IEEE 802.3)

La norme 802.3 décrit un protocole de niveau 1 (couche physique) – c'est-à-dire la manière de générer les signaux sur le câble – ainsi que le protocole de niveau 2 (couche liaison). Les spécifications, édictées par l'IEEE (*Institute of Electrical and Electronics Engineers*), reposent sur le modèle suivant :



Au niveau 2, la couche **MAC** (*Medium Access Control*) a pour rôle de transporter les données entre deux points d'un segment Ethernet. Pour ce faire, elle encapsule les données dans une **trame MAC**, identifie émetteurs et destinataires par des **adresses MAC** et assure que les données sont transmises sans erreur, grâce à un code de contrôle.



Chaque carte réseau est identifiée par une adresse MAC unique sur 6 octets (par exemple, 01-00-0A-FB-5D-52). Toute trame MAC émise comporte les adresses de l'émetteur et du destinataire. La carte réseau ne prend en compte que les trames MAC qui lui est destinée. Exception à cette règle : une carte peut émettre une trame spéciale, appelée trame de diffusion (ou trame de **broadcast**), qui sera lue par toutes les autres cartes (adresse destination dont tous les bits sont à 1 — FF-FF-FF-FF-FF-FF).

La couche MAC est également responsable de gérer l'accès au segment Ethernet quand il est partagé par plusieurs nœuds. Lorsque deux stations émettent simultanément, la somme des deux signaux dépasse un seuil qui est interprété comme étant une **collision**. Dans un tel cas, chaque station arrête d'émettre et attend un temps aléatoire avant d'essayer de réémettre leur trame. Cette méthode de gestion de la contention est appelée **CSMA/CD** (*Carrier Sense Multiple Access / Collision Detection*).

...

LE POINT SUR ETHERNET (SUITE)

Afin d'être le plus possible indépendant des méthodes et des supports de transmission, le modèle de référence de la norme IEEE 802.3 repose sur plusieurs sous-couches physiques :

- La sous-couche **réconciliation**, qui a pour rôle d'adapter les bits manipulés par la couche MAC avec les signaux utilisés par les couches physiques sous-jacentes.
- **PLS** (*Physical Layer Signaling*), qui interprète les bits véhiculés sur le câble (données, détection de collision, auto-négociation, etc.).
- **PCS** (*Physical Coding sublayer*), qui code les groupes de bits en symboles, c'est-à-dire en des signaux. Par exemple, le codage 4B/5B consiste à coder 4 bits en un groupe de 5 signaux.
- **PMA** (*Physical Medium Attachment*), qui transmet et reçoit les signaux véhiculés sur le câble en bande de base (par exemple, NRZI). La PMA, associée à un débit, est activée ou désactivée en fonction de la négociation de débit.
- **PMD** (*Physical Medium Dependant*), qui désigne l'électronique qui génère les signaux électriques ou optiques.

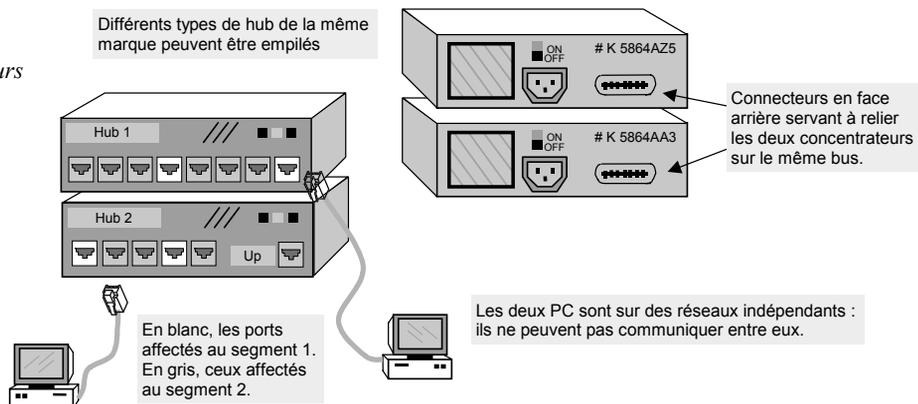
Les interfaces **AUI** (*Attachment Unit Interface*), **MII** (*Media Independent Interface*) et **GMII** (*Gigabit MII*), qui sont optionnelles, assurent la liaison entre la couche MAC (logiciel hébergé dans des composants électroniques) et les composants gérant les signaux (PHY) de la carte réseau. Concrètement, elles se présentent sous la forme de connecteurs, pouvant être étendus par un cordon de quelques centimètres, et raccordant la couche MAC à un **transceiver** ou MAU (*Medium Attachment Unit*), qui intègrent donc les fonctions de la couche PHY. Les GBIC de Cisco sont des exemples de tels *transceivers*.

Quel format d'équipement ?

Le marché offre le choix d'équipements seuls (*stand alone*), empilables (*stackable*), ou en châssis : des cartes sont insérées dans des emplacements prévus à cet effet (*slots*). Pour compliquer le choix, il existe également des *stackables* avec des *slots* d'extension qui permettent d'ajouter un ou deux ports, dans le but de chaîner l'équipement à un autre.

Figure 3-1.

Les possibilités des concentrateurs empilables.



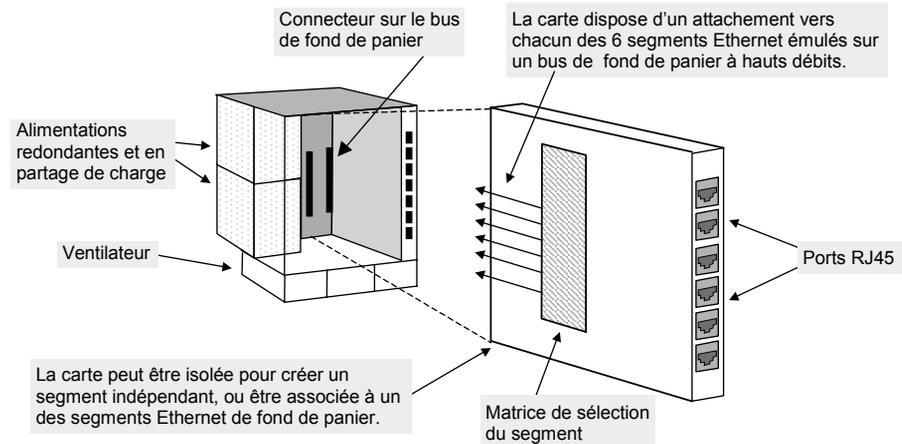
Les modèles *stand alone* visent le marché d'entrée de gamme ; ils sont parfaits pour créer un premier réseau local tel que celui du premier chapitre.

Les modèles empilables sont envisageables dès qu'il y a des possibilités d'extension. Par exemple, votre société dispose de cinquante postes de travail, et vous commencez par en connecter dix dans un premier temps. Vous pouvez alors acheter un hub 12 ports, puis un autre 12 ports plus tard.

Un concentrateur, ou un commutateur, comprend un nombre limité de ports (généralement 8, 16, 24 ou 32). Or, la plupart du temps, un local technique concentre beaucoup plus de postes de travail (jusqu'à plusieurs centaines). La solution est alors de chaîner les *stackables* entre eux *via* un bus spécial, dédié à cet effet ; il s'agit d'un câble externe reliant les équipements entre eux pour n'en faire qu'une unité logique. Attention, le bus étant propriétaire, seuls les équipements d'un même constructeur pourront être chaînés entre eux, généralement de 5 à 8 au maximum.

Une autre solution consiste à installer des châssis, certes plus chers, mais qui offrent de plus grandes capacités d'accueil. Ces équipements permettent de créer plusieurs segments indépendants à des débits différents.

Figure 3-2.
*Principes
et fonctionnalités
d'un châssis.*



Il est possible d'insérer différents types de cartes dans un châssis : concentrateur Ethernet, commutateur Ethernet, carte Token-Ring, FDDI, ATM, etc. Il est également possible de combiner les débits (10, 100 et 1 000 Mbit/s) sur des segments séparés. Mais, attention, il n'est en aucun cas possible de mélanger des débits sur un même segment Ethernet. De même, les réseaux Token-Ring, FDDI et ATM créés seront indépendants.

LES DIFFÉRENTS RÉSEAUX ÉTHERNET

Il existe actuellement trois déclinaisons d'Ethernet normalisées par l'IEEE (*Institute of Electrical and Electronics Engineers*) : le **10bT** à 10 Mbit/s (norme 802.3), le **100bTx** alias Fast Ethernet à 100 Mbit/s (norme 802.3u) et le **1000bT** alias Gigabit Ethernet (norme 802.3ab).

Le premier chiffre qualifie le débit du réseau Ethernet, la lettre " b " signifie un codage des signaux en bande de base (ex : Manchester) et la lettre " T " représente " *Twisted Pair* ", ce qui signifie que le réseau Ethernet fonctionne sur un câblage en cuivre paires torsadées.

Il existe également les mêmes déclinaisons fonctionnant sur fibre optique : **10bF**, **100bFx** et **1000bX** (norme 802.3z). Parmi cette dernière, on distingue le **1000bSX** (S pour *short wavelength*) opérant à 850 nm sur fibre optique multimode et le **1000bLX** (L pour *long wavelength*) opérant à 1 300 nm sur les fibres multimode et monomode.

Ethernet	Codage en ligne	Codage complet	Paires cuivre utilisées	Fréquence du signal / Longueur d'onde	Distance maximale
10bT	Manchester	---	1,2 / 3,6	10 MHz ($\pm 15\%$)	100 m
10bF (FL et FB)	Manchester	---	---	850 nm	2 000 m
100bTx	NRZI	4B/5B	1,2 / 7,8	62,5 MHz	100 m
100bFx	NRZI	4B/5B	---	850 nm	2 000 m
1000bT	PAM5	8B1Q4	Toutes	125 MHz	100 m
1000bSX	NRZ	8B/10B	---	850 et 1 300 nm	220 à 550 m*
1000bLX	NRZ	8B/10B	---	1 300 et 1 550 nm	550 à 5 000 m*

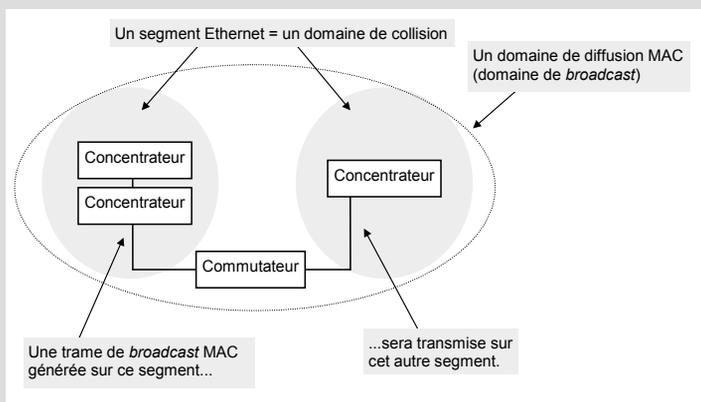
* Selon le type de fibre optique et la qualité des équipements (voir chapitre 12).

Ces variantes utilisent toutes le même principe d'accès au support de transmission (détection de collision en half duplex), les mêmes trames Ethernet et le même adressage MAC (*Medium Access Control*).

Il existe deux types d'équipements : les **concentrateurs** (*hubs*), qui partagent un segment Ethernet entre plusieurs ports, et les **commutateurs** (*switch*), qui créent un segment Ethernet par port.

Certains concentrateurs peuvent être partagés en plusieurs segments indépendants, tandis que le commutateur est capable d'**interconnecter** des segments qui sont physiquement indépendants.

Un **domaine de collision** comprend un ou plusieurs concentrateurs Ethernet ; une trame émise sur un port est régénérée sur tous les autres ports des concentrateurs chaînés (*via* un cordon de brassage) ou empilés (*via* un bus externe par un câble spécifique). Le commutateur ne laisse pas passer les collisions, mais une trame de *broadcast* émise sur un port sera régénérée sur tous les autres ports.



Un PC connecté à un port du commutateur est seul sur le segment Ethernet : aucune collision n'est donc possible, et il dispose de toute la bande passante. Inversement, tous les PC connectés à un concentrateur partagent la même bande passante (10, 100 ou 1000 Mbit/s) et peuvent émettre des trames en même temps, d'où une probabilité de plus en plus importante de collision qui croît avec le nombre de PC.

Le tableau suivant présente quelques éléments de comparaison.

Critère	Empilable	Châssis
Évolutivité	Ajout d'éléments empilables limité à cinq environ	Ajout de cartes limité par le nombre de slots
Segmentation	Limitée à un ou deux segments	Plusieurs segments par port ou par groupe de ports
Capacité de traitement	Bus externe limité à quelques centaines de Mbit/s	Bus de fond de panier de 100 Mbit/s à plusieurs Gbit/s
Alimentation	Une par élément ou, rarement, une pour tous	Une à trois pour l'ensemble du châssis
Redondance d'alimentation	Pas tout le temps	Oui
Création de plusieurs segments indépendants	Oui	Oui
Utilisation	Moins de cent postes en Ethernet, en Token-Ring ou en ATM	Plus de cent postes ; combinaison Ethernet, Token-Ring et ATM possible

En définitive, l'empilable sera choisi pour des faibles densités (moins de cent postes par local technique) ; les châssis seront privilégiés dans les autres cas. D'une manière générale, plus le réseau concentre de postes de travail, plus la fiabilité des équipements doit être importante.

Les châssis seront donc choisis là où le besoin en bande passante est élevé et où la fiabilité est primordiale, c'est-à-dire à des points de concentration stratégiques du réseau.

Concentrateur ou commutateur ?

L'autre décision à prendre consiste à choisir entre les concentrateurs et les commutateurs. Les premiers se contentent de générer le signal, alors que les seconds permettent de créer un segment par port. Ces derniers sont bien sûr plus chers.

Pour une utilisation bureautique du réseau (traitement de texte, comptabilité, base de données, connexion à un serveur, etc.), les concentrateurs suffisent pour connecter les postes de travail car il y a peu de trafic entre eux. Pour améliorer les performances, on peut jouer sur la vitesse (10 ou 100 Mbit/s).

L'utilisation des commutateurs s'envisage dans plusieurs cas de figures :

- lorsqu'on emploie des applications multimédias (voix et vidéo) générant des débits importants et nécessitant des temps de réponse courts ;
- d'une manière générale, lorsque le flux réseau est important et que les temps de réponse sont mauvais ;
- pour interconnecter plusieurs segments Ethernet.

QU'EST-CE QU'UN COMMUTATEUR ?

Un commutateur (*switch*) est un équipement qui offre une bande passante dédiée pour chaque port (10, 100 ou 1 000 Mbit/s par port) alors que le concentrateur partage la bande passante entre tous ses ports. Cela revient à créer un segment Ethernet par port.

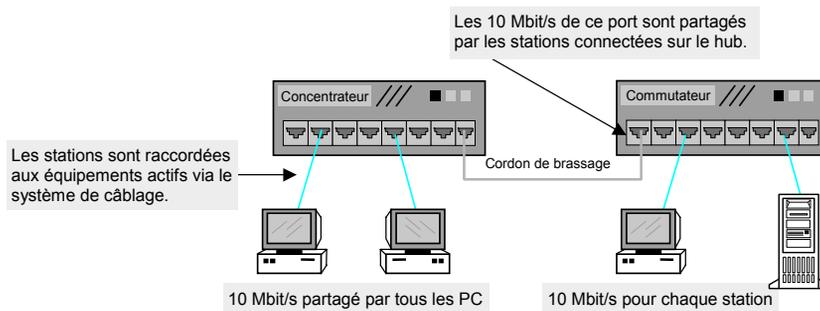
On distingue les *switches cut-through* (*on the fly*, à la volée) et les *switches store and forward* (les plus courants aujourd'hui). Les premiers se contentent de régénérer la trame (même les trames erronées et les collisions), tandis que les seconds la stockent en mémoire avant de la régénérer. La méthode *adaptive cut-through* combine les deux principes : dès qu'une trame est en erreur, le commutateur bascule en mode *store and forward* pendant un certain temps. La méthode *fragment free*, la plus performante, lit les 64 premiers octets avant de décider du mode de transmission.

Chaque port du commutateur correspond à un segment Ethernet, c'est-à-dire à un domaine de collision (voir encadré "Le point sur Ethernet").

Si vous constatez un nombre élevé de collisions lié à une charge réseau importante, vous pouvez, dans un premier temps, segmenter le réseau (c'est-à-dire le couper en deux). Dans ce cas, les PC situés sur un segment ne pourront plus communiquer avec ceux situés sur l'autre.

Au lieu d'acheter un second concentrateur, l'achat d'un commutateur résoudra le problème : les postes seront répartis sur les deux équipements, et les segments ajoutés seront interconnectés.

Figure 3-3.
Concentrateur
et commutateur.



Les stations consommant le plus de bande passante seront de préférence connectées au commutateur. C'est le cas des serveurs qui concentrent toutes les connexions des utilisateurs.

Les commutateurs apportent, par ailleurs, de nouvelles fonctionnalités et ils peuvent être préférés aux concentrateurs rien que pour cela, indépendamment de tout problème de charge réseau.

Étant donné qu'un commutateur crée un segment par port, il est possible de combiner les débits (10, 100 et 1 000 Mbit/s) au sein de la même boîte. Il est à noter que seul le mode *store and forward* le permet tandis que le mode *cut-through*, quant à lui, ne bénéficie pas de

LE POINT SUR L'ETHERNET FULL DUPLEX (IEEE 802.3x)

Le mode full duplex conserve le débit nominal de 10, 100 ou 1 000 Mbit/s, mais sépare les canaux émission et transmission, et donc, élimine le besoin de la détection de collision CSMA/CD employée par l'Ethernet classique (en half duplex).

Le mode full duplex n'est possible :

- que sur des liaisons point à point entre un PC et un commutateur ou entre deux commutateurs ;
- qu'avec des câbles qui séparent physiquement les canaux émission et réception, c'est-à-dire le cuivre en paires torsadées et la fibre optique, mais pas les câbles coaxiaux ;
- qu'avec les commutateurs qui sont seuls capables de stocker les données à envoyer lorsque le canal émission est occupé.

L'absence de détection de collision CSMA/CD a une conséquence importante : **la limitation de longueur des câbles imposée par le délai de propagation des collisions disparaît**. Seule l'atténuation du signal, et donc la performance du câble, limite désormais la distance entre deux équipements : en pratique, on atteint des valeurs comprises entre 150 et 200 m au lieu des 100 m habituels, et plusieurs centaines de kilomètres en fibre optique.

cette possibilité car les trames sont commutées à la volée (elles ne sont pas stockées en mémoire).

On trouve sur le marché les formules suivantes :

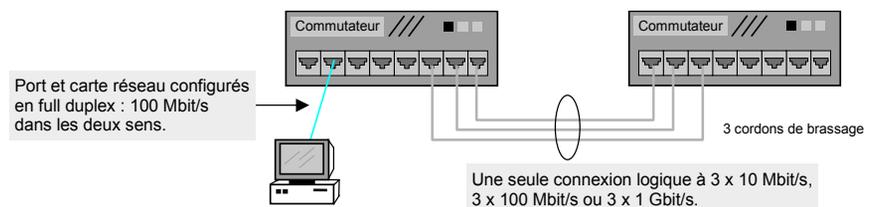
- ports à détection automatique de vitesse (*port autosense*, 10/100 Mbit/s) ;
- port uplink à 100 Mbit/s ou 1 Gbit/s.

👉 Les cartes réseau autosense 10/100 posent des problèmes avec les commutateurs qui sont également autosense, car il n'y a pas de négociation de débit ; chacun essaie de se caler sur la vitesse de l'autre. Il est donc recommandé de désactiver cette fonction au niveau de la carte et de fixer la vitesse manuellement (configuration à l'aide de Windows).

Les commutateurs permettent également d'augmenter les débits de plusieurs manières :

- avec le mode full duplex entre un PC et un port du commutateur ou entre deux commutateurs ;
- en agrégeant plusieurs ports full duplex du commutateur (technique du *port trunking*) pour le relier à un autre commutateur.

Figure 3-4.
Augmentation
des débits
grâce
aux commutateurs.



Le mode full duplex ne permet pas d'augmenter la vitesse de transmission, mais consiste à séparer les canaux émission et réception. Le débit global est ainsi augmenté, et peut, théoriquement, être multiplié par deux. Dans la pratique, ce mode est donc intéressant pour les serveurs et les liaisons intercommutateurs qui peuvent avoir à traiter un flux simultané dans les deux sens.

La carte réseau doit également supporter le mode full duplex.

À l'inverse, l'agrégation consiste à créer une seule liaison logique constituée de plusieurs liaisons physiques. Le débit obtenu est alors égal à la somme des débits des ports agrégés. Cette technique, normalisée 802.3ad, est utilisée pour interconnecter deux commutateurs généralement de même marque. Certaines cartes réseau supportent ce mode de fonctionnement.

Fonctionnalité	Concentrateur	Commutateur
Segment	Un seul pour le concentrateur	Un par port
Segmentation	(Optionnel) Réseaux indépendants (c'est-à-dire isolés, ne pouvant pas communiquer entre eux)	Segments interconnectés par la matrice de commutation
Mélange des débits	Non	Oui pour le mode <i>store and forward</i>
Port Uplink	Pour chaîner les hubs ; 4 maximum en cascade	Pour chaîner les commutateurs entre eux
Administrable	SNMP en option	SNMP + RMON en option
Autres fonctionnalités	Partitionnement des ports sur erreur	Détection automatique 10/100 Mbit/s Full duplex Agrégation de ports VLAN

Les VLAN (*Virtual LAN*) seront étudiés au chapitre 12. Pour l'instant, nous n'en avons pas besoin.

L'architecture

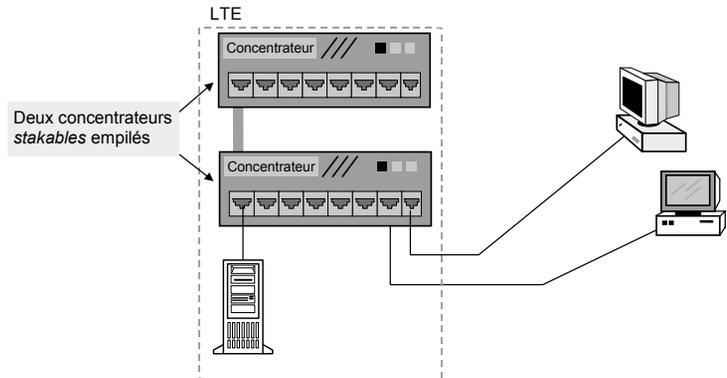
Mise en place d'un réseau local d'étage

Ayant en tête toutes les possibilités des équipements à notre disposition, la conception d'une architecture réseau simple consiste à assembler les concentrateurs et les commutateurs en exploitant au mieux les capacités du câblage.

Partons d'un cas simple : une cinquantaine de PC situés au même étage d'un immeuble quelconque. Les utilisateurs ont juste besoin d'échanger des données entre eux et de partager des applications (un serveur web, une base de données, des traitements de texte, etc.).

Le schéma suivant décrit l'architecture de base qui en résulte.

Figure 3-5.
Réseau local
sur un étage.



C'est on ne peut plus simple :

- En vertu de ce qui a été dit aux sections précédentes, deux concentrateurs empilables 10bT ont été installés et chaînés entre eux sur un bus externe.
- Chaque poste est raccordé à un port d'un concentrateur *via* un système de câblage tel que décrit au chapitre précédent : un câblage en cuivre à paires torsadées (avec, bien sûr, des prises RJ45) centré en étoile autour d'un local technique.
- Les serveurs sont ici situés dans le local technique, et chacun d'entre eux est raccordé directement à un port d'un concentrateur *via* un cordon de brassage RJ45/RJ45.

Il n'y a pas de question à se poser.

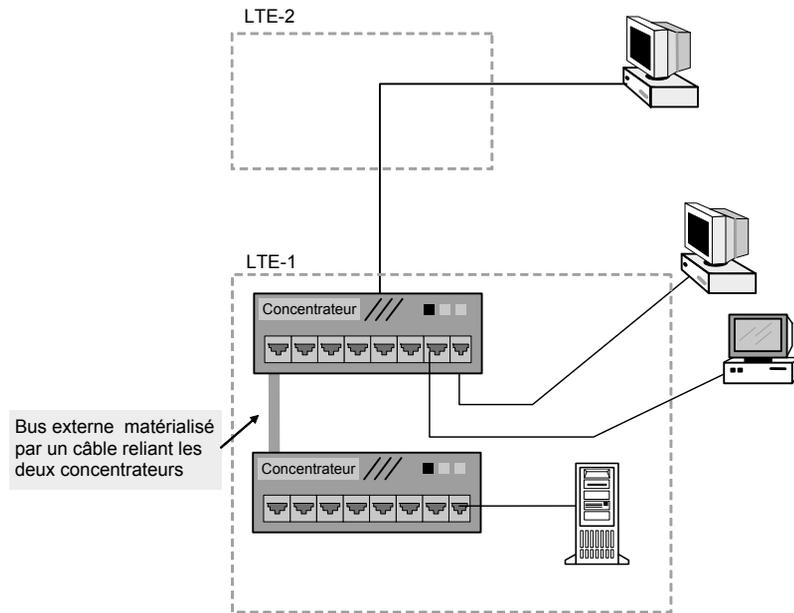
Extension du réseau d'étage

Maintenant, des utilisateurs, situés à l'étage au-dessus, ont les mêmes besoins. Pas de problème : il y a moins de dix PC, ce qui nous permet d'utiliser les rocades en cuivre existantes, et de les connecter directement aux concentrateurs qui sont déjà installés.

QUELS CORDONS DE BRASSAGE ?

Pour connecter un PC à un hub, un cordon de brassage **droit** (de même nature que le câblage en cuivre à paires torsadées) doit être utilisé. Pour connecter deux hubs entre eux, un cordon **croisé** (paires émission et réception) doit être utilisé, sauf si le port uplink est utilisé.

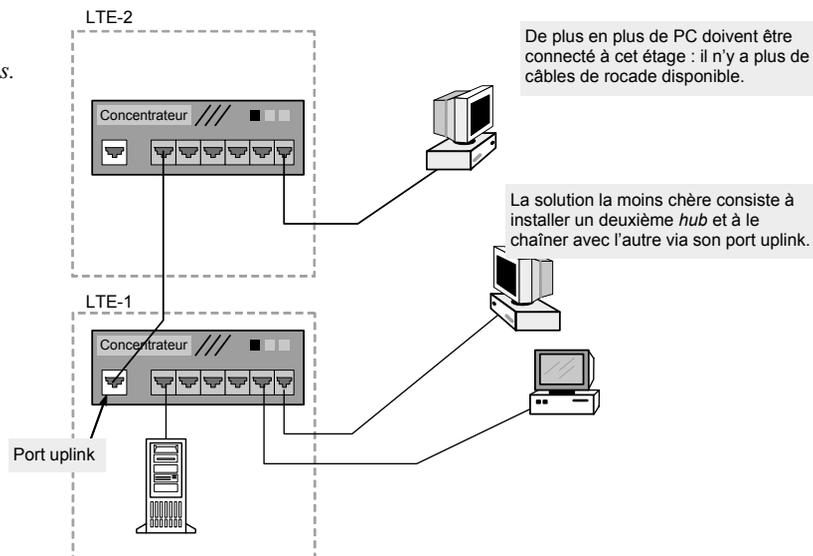
Figure 3-6.
Extension d'un réseau local
sur plusieurs étages.



S'il y a davantage de PC, ou, d'une manière générale, s'il n'y a plus de câbles de roudage en nombre suffisant (en général, entre 6 et 12), il suffit de créer un autre réseau sur le même modèle. Le problème est maintenant de connecter les réseaux construits sur les deux étages pour que tout le monde puisse accéder aux mêmes données et aux mêmes applications.

La solution la plus simple consiste à connecter les deux concentrateurs en cascade, *via* un câble de roudage en cuivre.

Figure 3-7.
Un réseau local
étendu sur plusieurs étages.



Conception d'un réseau d'immeuble

Maintenant, la situation se corse un peu : le réseau est un succès, il y a de plus en plus de demandes de connexions, la société utilise de plus en plus l'informatique. Il faut maintenant créer des réseaux à chaque étage. La solution précédente, consistant à chaîner les concentrateurs n'est plus applicable, car on est limité par le nombre de cascades possible. De plus, au-delà de cent postes connectés, le réseau Ethernet deviendrait saturé (du fait du nombre de collisions qui augmente avec le nombre d'utilisateurs) et, en définitive, les temps de réponse seraient trop grands.

L'installation de boîtes doit maintenant faire place à une plus grande réflexion, c'est-à-dire à un travail d'architecture. Imaginons donc que nous ayons trois cents utilisateurs répartis sur une demi-douzaine d'étages, soit en moyenne cinquante postes par étage, plus les imprimantes et les serveurs. On se retrouve avec soixante à soixante-dix connexions par étage.

Comme précédemment, on peut installer une pile de concentrateurs à chaque étage pour créer un réseau local d'étage. Là encore, il faut se poser de nouveau les mêmes questions :

- Quel débit : 10bT, 100bT ou Gigabit ?
- Quelle technologie : concentrateurs ou commutateurs ?
- Un seul segment Ethernet ou plusieurs ?

Les réponses à ces questions dépendent avant tout du trafic prévisionnel, des perspectives d'évolution et des performances mises en balance par rapport au coût. Si l'immeuble comprend trente autres étages, on peut supposer qu'il faudra tôt ou tard étendre le réseau. Si, en revanche, l'immeuble n'en comprend que six, on sait que la configuration sera figée pour un bon moment.

L'architecture doit donc être conçue pour couvrir les besoins futurs et non seulement ceux du moment. Elle doit donc être évolutive, c'est-à-dire être bâtie sur des équipements que l'on pourra récupérer (recycler pour d'autres usages).

Mise en place d'un réseau fédérateur

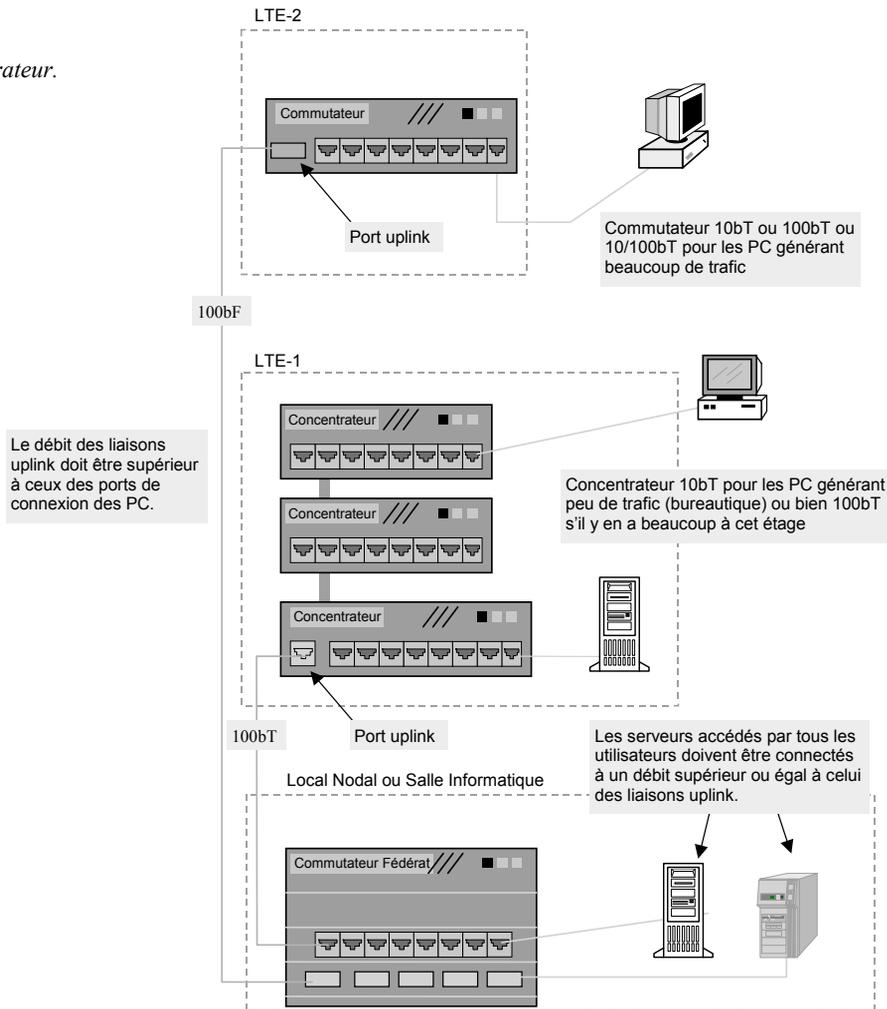
Si les besoins en trafic sont importants (applications multimédias, applications, voix, données et visioconférences), on peut envisager des commutateurs 10bT à tous les étages. Dans la plupart des cas, un débit de 10 Mbit/s suffira, mais la différence de coût étant minime, on peut envisager le 100bT.

La question est maintenant de savoir comment connecter les réseaux locaux entre eux. La solution repose sur la création d'un réseau fédérateur (*backbone*).

On peut imaginer un câble Ethernet (topologie en bus), ou FDDI (anneau), qui parcourt tous les étages et auquel on connecte les concentrateurs. C'est une solution peu évolutive, car le débit est limité à la technologie utilisée (100 Mbit/s pour FDDI). En outre, le Gigabit Ethernet et l'ATM ne sont pas prévus pour une topologie en bus ou en anneau. C'est également une solution peu sûre : le câble étant un élément passif, il n'y a aucun moyen de le superviser à distance.

L'architecture couramment utilisée est de type *collapse backbone* (littéralement, réseau fédérateur effondré). Le principe consiste à concentrer le backbone en un seul point : au lieu d'avoir un réseau qui parcourt tous les étages, le backbone est réalisé dans un commutateur unique. Cela revient à créer une architecture en étoile à deux niveaux, un premier concentrant les PC à chaque étage et un second concentrant les équipements d'étage en un point central, en général la salle informatique ou un local nodal dédié aux équipements réseau.

Figure 3-8.
Conception
d'un réseau fédérateur.



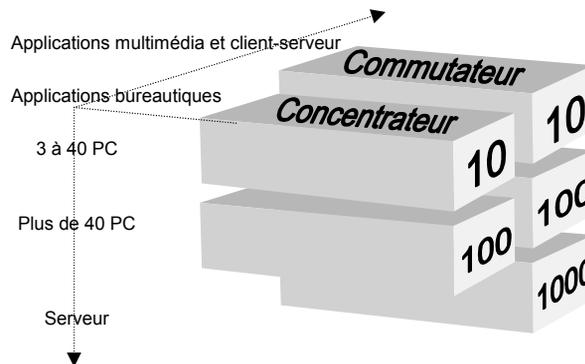
Pour un réseau de taille moyenne (de 200 à 800 utilisateurs), l'équipement central doit être de grande capacité en termes d'accueil et de performances. Le choix se portera donc sur un châssis qui offre une matrice de commutation à haut débit. Le réseau fédérateur n'est alors pas limité à 10 Mbit/s, mais à 100 Mbit/s par étage et à la capacité de la matrice de commutation du commutateur central, généralement plusieurs gigabits.

Quel débit et quelle technologie ?

Le choix du débit du réseau fédérateur dépend de celui utilisé par les PC.

Si les PC sont connectés à un...	Le débit des liens uplink vers le commutateur central doit être au moins égal à...
Concentrateur à 10 Mbit/s	10 Mbit/s
Concentrateur à 100 Mbit/s	100 Mbit/s
Commutateur 10 Mbit/s	100 Mbit/s
Commutateur 100 Mbit/s	1 Gbit/s

Le choix du débit des PC dépend, quant à lui, du volume de trafic généré et du type d'application (du flux bureautique au flux multimédia). Mais le passage du concentrateur au commutateur évite ou retarde l'augmentation du débit, ce qui permet de conserver les cartes réseau existantes dans les PC.



Indépendamment de la charge réseau et du nombre de PC, certaines combinaisons sont plus appropriées que d'autres.

COMMENT FONCTIONNE UN COMMUTATEUR ?

Un commutateur permet d'**interconnecter** plusieurs segments Ethernet. Sur chacun de ses ports, on peut raccorder un concentrateur (plusieurs PC partagent alors la bande passante sur ce port) ou un seul PC (technique de la **microsegmentation**).

Afin de limiter le trafic réseau inutile, les trames Ethernet échangées entre les PC d'un même segment ne sont pas transmises sur les autres segments gérés par le commutateur (voir encart "Le point sur Ethernet").

Pour savoir quelles sont les trames qui doivent sortir du segment et celles qui ne le doivent pas, le commutateur regarde toutes les adresses MAC sources des trames qui entrent sur chacun de ses ports, et les enregistre dans ses **tables d'adresses** (il existe une table par port). C'est le mécanisme d'**apprentissage**. Si un PC est déplacé d'un port à l'autre ou d'un commutateur à l'autre, son adresse MAC peut se retrouver dans deux tables et créer ainsi des conflits. Pour éviter cela, les adresses sont effacées de la table au bout de 15 à 30 secondes.

Les commutateurs d'entrée de gamme ne disposent que de très peu de mémoire et ne peuvent apprendre qu'une, deux ou quatre adresses MAC par port. Ils sont plutôt dédiés à la microsegmentation. Les commutateurs fédérateurs doivent en revanche disposer de beaucoup de mémoire et être capables d'enregistrer plusieurs milliers d'adresses MAC, car ils fédèrent tous les flux interréseaux locaux (c'est-à-dire intersegment Ethernet).

La **mémoire tampon** doit également être suffisamment importante pour permettre l'adaptation des débits (entre 10 et 100 Mbit/s, et surtout entre 10/100 et 1 Gbit/s).

L'interconnexion des segments est réalisée par une **matrice de commutation** à haut débit capable de supporter la somme des débits des ports (8×10 Mbit/s, par exemple). Là encore, les commutateurs fédérateurs doivent comprendre une matrice de commutation très puissante (généralement des ASIC et des processeurs RISC).

Type de trafic	Réseau d'étage	Fédérateur
Applications bureautiques (Word, Excel, bases de données)	Concentrateurs 10bT	Commutateurs 10/100bT
Client-serveur et un peu de multimédia (voix et vidéo)	Commutateurs 10/100bT	Commutateurs 100bT
Applications multimédias intensives	Commutateurs 100bT	Commutateurs Gigabit

Le coût est un autre critère de décision, sans doute le plus important. Lors du choix d'une technologie (concentration ou commutation) et du débit, il faut tenir compte du nombre de cartes réseau pour les PC ainsi que du nombre d'équipements.

Équipement actif 8 ports	Coût en Euros HT
Carte 10bT	De 38 à 120
Cartes 100bT et 10/100bT	De 75 à 185
Concentrateur 10bT	De 75 à 770
Concentrateur 100bT (ou 10/100bT)	De 385 à 1 000
Commutateur 10bT	De 770 à 2 300
Commutateur 100bT (ou 10/100bT)	De 2 300 à 3 000
Concentrateur 1000bT	N'existe pas
Commutateur 1000bT	De 7 600 à 13 800

Les écarts de prix sont dus à des différences dans les fonctionnalités proposées (concentrateur administrable ou non, empilable ou non, avec ou sans slot d'extension, etc.). Un exemple en a été donné au premier chapitre concernant les concentrateurs 10bT. Le choix de la fibre optique fait également monter les prix.

Suivre l'évolution des besoins

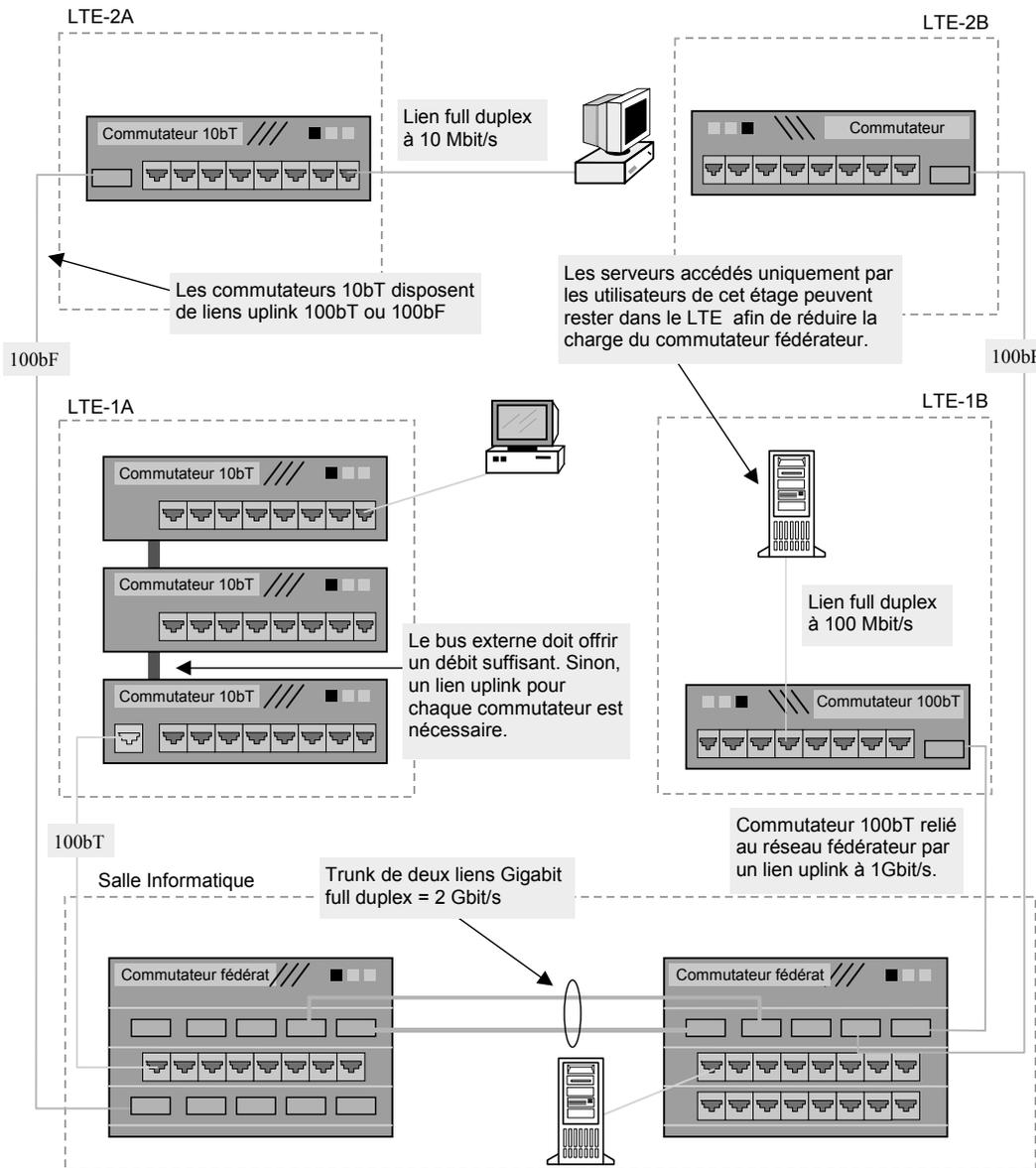
La direction a maintenant décidé de regrouper plusieurs sites sur le nôtre et de louer l'autre aile de l'immeuble. Le nombre d'utilisateurs va ainsi passer de 300 à 500. Ils seront répartis sur 10 étages, soit une moyenne de 70 connexions par étage en comptant les imprimantes, les doubles connexions, etc.

Par ailleurs, les utilisateurs de bases de données client-serveur se plaignent des mauvais temps de réponse. Il s'agit de gros consommateurs de bande passante (gestion électronique de documents, applications décisionnelles de type datawarehouse, etc.) qui perturbe les autres trafics. Les utilisateurs qui se connectent sur des serveurs Unix se plaignent également de ralentissements brusques lors de l'affichage des écrans et du déplacement du curseur à l'écran.

Pour faire face à cette montée en charge, nous avons plusieurs solutions, que nous pouvons combiner.

En premier lieu, les commutateurs peuvent être généralisés. Cela permet de segmenter le réseau afin qu'un flux important ne viennent en perturber un autre. Ensuite, les débits peuvent être augmentés à 100 Mbit/s pour les plus gros consommateurs. Afin de limiter le coût du réseau, les autres utilisateurs peuvent rester connectés à 10 Mbit/s.

Figure 3-9.
Utilisation des capacités des commutateurs
pour monter en charge.



Au niveau du réseau fédérateur, la capacité d'accueil doit être augmentée. La sécurité de fonctionnement doit également être améliorée, car la panne de l'équipement central entraînerait l'arrêt de tout le réseau. Le ou les commutateurs empilables doivent donc faire place à un châssis plus performant (en termes de bande passante de la matrice de commutation) et qui offre des alimentations redondantes.

L'ajout d'un second commutateur, identique au premier et donc interchangeable, renforce encore la fiabilité du réseau fédérateur.

Afin de ne pas créer un engorgement entre les deux commutateurs, un lien à très haut débit doit être créé. Cela est réalisé en agrégeant plusieurs ports Gigabit (technique du *port trunking*).

Les serveurs très sollicités et certains postes de travail consommateurs de bande passante peuvent voir leurs performances d'accès réseau accrues par des connexions full duplex.

Assurer la continuité de service

Près de 500 utilisateurs sont connectés. Le réseau devient un élément stratégique du système d'information. En effet, sans lui, les utilisateurs ne peuvent plus travailler : plus de connexion au système central, plus d'accès aux bases de données, plus de messagerie, etc. Bref, il est nécessaire de construire une architecture redondante pour parer aux éventuelles pannes. Celles-ci ne sont pas un mythe. En voici quelques-unes qui arrivent fréquemment :

- panne d'alimentation ;
- coupure d'un câble entre les commutateurs (erreur de manipulation lors du brassage dans les locaux techniques) ;
- débranchement d'un câble, d'un convertisseur ou de tout autre petit boîtier : eh oui, un *transceiver* ou un connecteur mal enfoncés dans la prise tendent non pas à se remettre en place d'eux-mêmes (dommage), mais plutôt à tomber sous l'effet des vibrations et de la pesanteur...

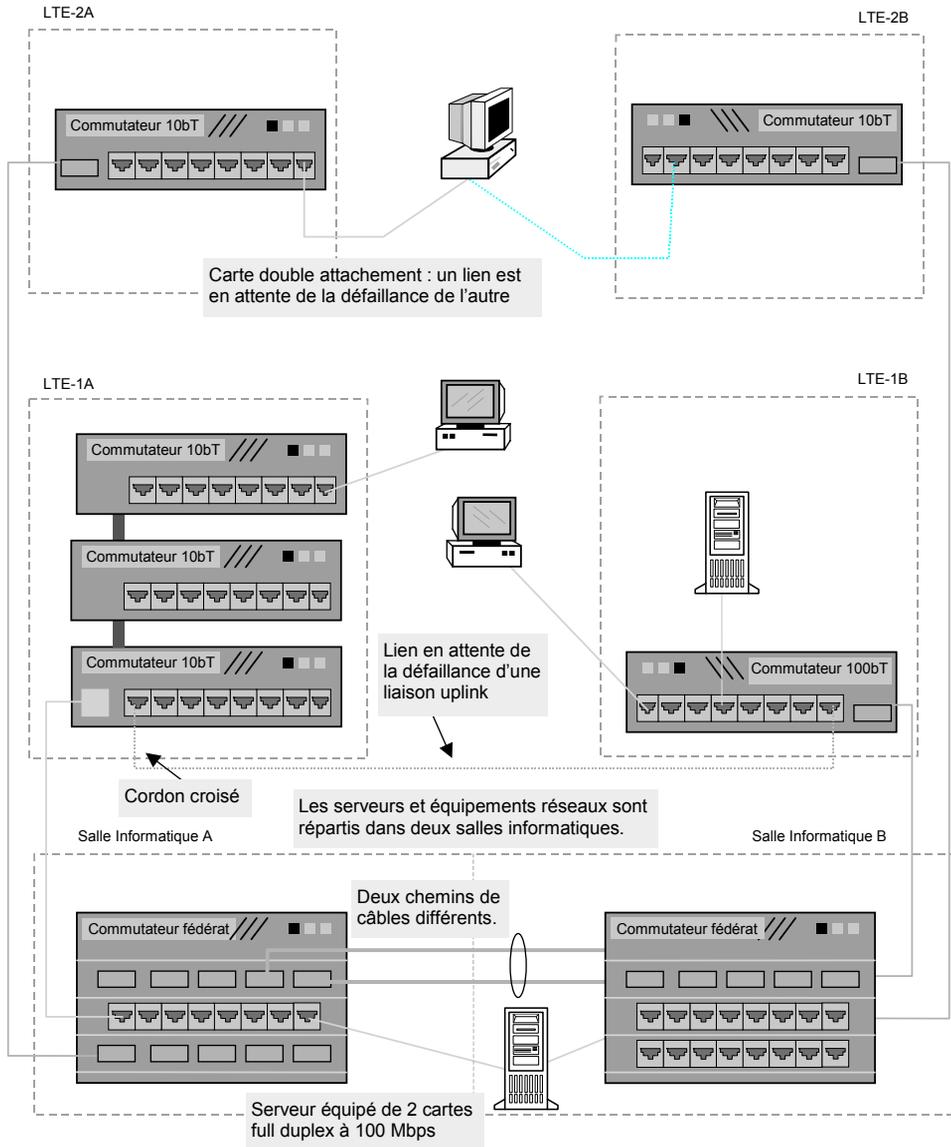
L'incendie ou l'inondation sont plus rares, mais, quand ils se produisent, il n'est pas question de reconstruire un système de câblage et un réseau avant quelques heures, voire plusieurs jours.

La mise en place d'un réseau redondant repose d'une part sur un système de câblage qui offre l'alternative de plusieurs cheminements, d'autre part sur la mise en place d'équipements de secours.

La redondance est généralement limitée au réseau fédérateur, mais, dans des cas critiques (salles de marché, processus industriels en flux tendus, etc.), il peut être nécessaire de l'étendre jusqu'au poste de travail.

On peut alors prévoir d'équiper ces derniers en carte deux ports : quand le premier perd le contact avec le port du concentrateur, il bascule sur le second, connecté à un autre concentrateur situé dans un autre local technique. Côté serveur, on peut envisager la même solution ou encore prévoir deux cartes ou plus, ce qui offre l'intérêt du partage de charge en fonctionnement normal.

Figure 3-10.
Cas
d'un réseau
redondant.



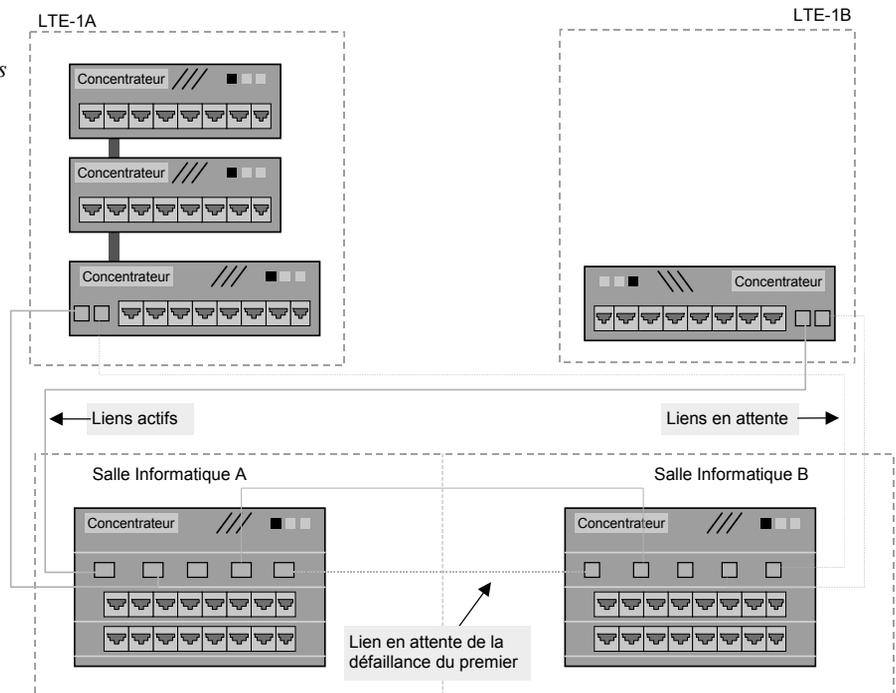
Les équipements fédérateurs sont ici des commutateurs reliés entre eux pour former un seul réseau fédérateur. Cette liaison doit être de débit élevé afin d'absorber tous les flux issus des étages et allant vers un serveur connecté au commutateur fédérateur opposé. On peut envisager ici du Gigabit Ethernet. Les liaisons entre les commutateurs fédérateurs et les équipements d'étage peuvent être réalisées en cuivre ou en fibre optique, le choix dépendant de la distance.

L'agrégation de liens offre ici une protection supplémentaire car chaque lien passe par deux gaines techniques différentes : en cas de perte d'une des liaisons (coupure du câble, incendie, etc.), l'autre reste opérationnelle et permet un fonctionnement en mode dégradé à 1 Gbit/s au lieu de 2 Gbit/s.

Il est à noter que cette architecture redondante n'est possible que parce que les équipements sont des commutateurs. Ceux-ci échangent des informations sur la topologie du réseau et déterminent les routes actives et celles qui ne le sont pas. Le protocole utilisé s'appelle le **spanning tree**.

La même architecture avec des concentrateurs serait impossible car le segment Ethernet ferait une boucle, ce qui est interdit par la norme IEEE. Pour assurer le même niveau de redondance, une configuration équivalente avec des concentrateurs serait donc beaucoup plus complexe et nécessiterait de doubler le nombre de câbles. De plus, seule l'utilisation de la fibre optique permet ce type de redondance.

Figure 3-11.
*Redondance
avec des concentrateurs
Ethernet.*

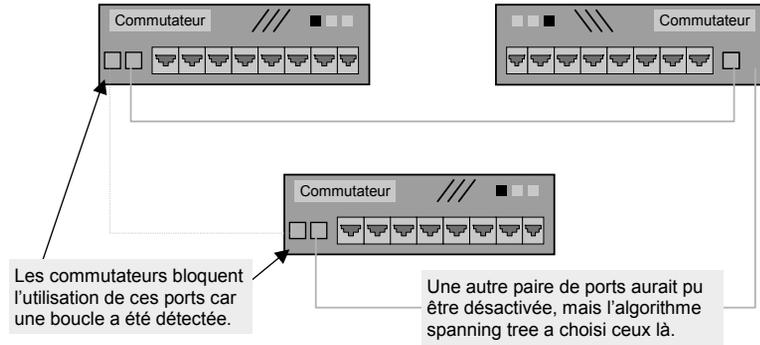


Les liaisons en pointillé sont inactives en fonctionnement normal. En cas de rupture de la liaison principale, le concentrateur d'étage active la seconde liaison en quelques microsecondes.

La liaison entre les deux concentrateurs fédérateurs doit être doublée. En effet, si elle était perdue, on se retrouverait avec deux réseaux isolés ayant le même numéro de réseau IP.

Inversement, avec le spanning tree, les liens sont tous actifs, mais une des routes est invalidée par les commutateurs. En cas de rupture d'un lien, une nouvelle route sera calculée, généralement au bout de trente secondes.

Figure 3-12.
La redondance avec le spanning tree.



L'échange de trames Ethernet

Un réseau local tel qu'Ethernet est constitué de concentrateurs et de commutateurs, deux équipements au comportement bien différent.

Caractéristique	Ethernet partagé	Ethernet commuté
Équipement	Concentrateur	Commutateur
Segment	Un segment partagé par tous les ports	Un segment par port
Architecture matérielle	Composants électroniques ; un ou plusieurs bus Ethernet	ASIC et processeur RISC, matrice de commutation
Architecture logicielle	Petit logiciel pour des options de configuration et pour l'administration SNMP	Logiciel plus complexe, mais traitement essentiellement matériel
Algorithme de routage	Aucun	Spanning tree
Traitement des trames	Aucun traitement (transparent pour les trames Ethernet)	Filtrage et transmission (<i>forward</i>) des trames ; apprentissage des adresses MAC
Couche réseau	Niveau 1 = couche physique	Niveaux 1 et 2 = couche physique et liaison
Fonction	Connecter plusieurs PC au sein d'un segment	Interconnecter plusieurs segments

Échange de trames sur un segment Ethernet

Commençons par étudier le fonctionnement au sein d'un segment Ethernet partagé (la trame circule sur un seul segment).

La norme Ethernet spécifie l'utilisation d'adresses physiques liées aux cartes réseau : les adresses MAC.

QU'EST-CE QUE LE SPANNING TREE ?

Lorsqu'un commutateur reçoit une trame, il recherche son adresse de destination dans ses tables d'adresses MAC (une par port). Si elle ne s'y trouve pas, il la transmet sur tous ses ports. Sinon, il l'envoie uniquement sur le port identifié. C'est le mécanisme de **forwarding**.

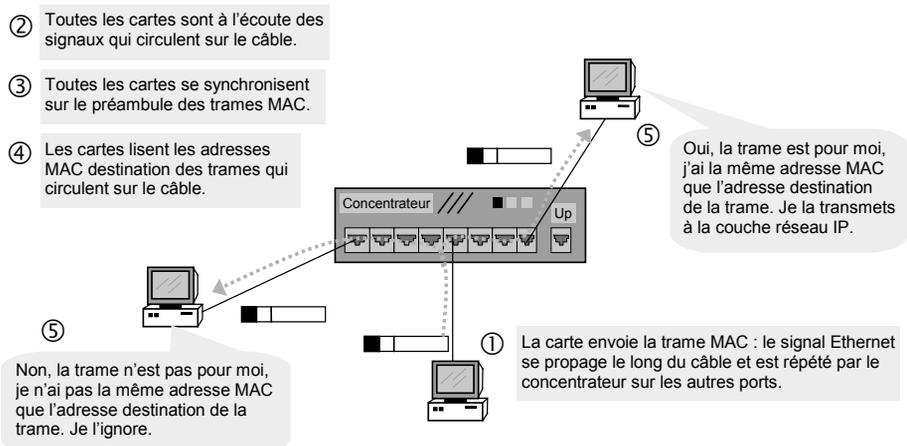
S'il existe plusieurs chemins, par exemple entre un commutateur d'étage et deux fédérateurs, la trame risque de boucler, car chacun des commutateurs transmet la trame (*forward*).

Le spanning tree est un **protocole de routage de niveau 2** associé à un algorithme qui permet d'éviter que les trames bouclent dans le réseau. Les commutateurs s'échangent des trames *spanning tree* et calculent une route en invalidant les chemins multiples susceptibles de créer des boucles au sein du segment Ethernet.

La carte recevant une trame Ethernet ne la prendra en compte que si l'adresse MAC de destination de la trame est identique à celle qui est inscrite dans sa mémoire. La seule exception à cette règle concerne les adresses de broadcast et, éventuellement, les adresses multicast.

En résumé, les cartes sont programmées pour accepter les trames qui leur sont destinées, plus toutes les trames de broadcast ainsi que les trames multicast qui ont été configurées.

Figure 3-13.
*Échange
des trames
Ethernet.*



Les trames qui sont acceptées sont remises au protocole de niveau 3, qui correspond à l'identifiant trouvé dans le champ « Type », soit 0x0800 pour IP. Celles qui ne sont pas destinées à la station sont ignorées.

LES ADRESSES MAC (IEEE 802.3)

Une adresse MAC est constituée de 6 octets et s'écrit sous forme hexadécimale séparé par des tirets, comme par exemple 02-60-8C-EB-25-D2. L'adresse de broadcast est FF-FF-FF-FF-FF-FF.

OUI (*Organizationally Unique Identifier*) = Partie de l'adresse affectée par l'IEEE à un fabricant de carte : 02608C pour 3com, 00000C pour Cisco, etc.

Partie de l'adresse affectée par le fabricant de la carte : de 000001 à FFFFFE

I/G	U/L	Adresse du constructeur sur 22 bits	Sous-adresse sur 24 bits
-----	-----	-------------------------------------	--------------------------

← sens de transmission des bits

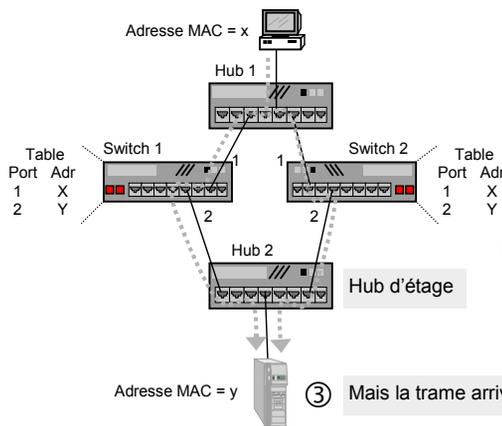
Bit I/G 0 = Adresse unicast (individuelle)
 1 = Adresse multicast (groupe)
 Bit U/L 0 = Adresse Universelle attribuée par l'IEEE
 1 = Adresse Locale

Le schéma montre l'adresse dans l'ordre de transmission des bits, alors qu'en représentation hexadécimale, l'ordre des bits au sein de chaque octet est inversé (cf. exemple au début du chapitre 15). Ainsi, le premier octet " 02 " indique que le bit U/L est à " 1 ". Le site standards.ieee.org/search.html permet de rechercher les OUI affectés par l'IEEE.

Échange de trames entre différents segments Ethernet

On l'a vu, le commutateur est un équipement permettant de segmenter le réseau Ethernet et d'interconnecter ces différents segments. Pour décider si la trame doit passer d'un segment à l'autre, il se base sur son adresse MAC qu'il compare avec celles se trouvant dans ses tables d'adresses en mémoire.

On peut envisager différentes situations dans lesquelles il existe plusieurs chemins possibles entre deux stations.



① Le PC émet une trame.

② Chaque commutateur transmet la trame.

④ De plus, le hub 2 régénère la trame sur tous ses ports. Chaque commutateur reçoit sur son port n° 2 la trame émise par l'autre commutateur : celle-ci est bloquée car l'adresse MAC source « X » se trouve déjà sur un autre port (le n° 1).

③ Mais la trame arrive en deux exemplaires : erreur !

Il faut noter que, si le serveur n'a jamais émis de trame, son adresse MAC « Y » n'est pas encore connue des commutateurs. Ces derniers transmettent alors la trame sur tous les autres ports, dont le port n° 2. De plus, la même situation se produit pour toutes les trames de broadcast et de multicast.

Pour éviter ces problèmes, il faut qu'une des deux routes soit interdite. C'est là qu'intervient le **spanning tree**. Le but de ce protocole est de définir une route unique vers un commutateur désigné racine en se basant sur des coûts et des priorités.

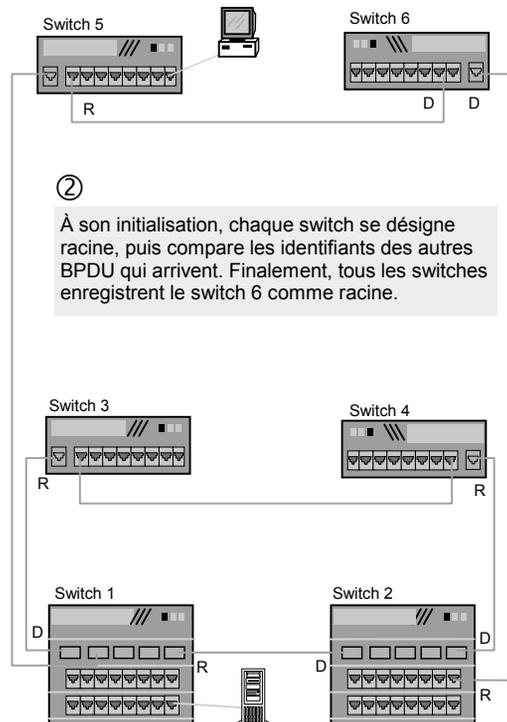
Compte tenu des processus d'élection, il est important de bien paramétrer les coûts (exprimés en nombre de sauts et/ou dépendant du débit du port) ainsi que les priorités. Les valeurs par défaut (fixées en usine) peuvent en effet aboutir à choisir des chemins qui ne sont pas les meilleurs. Prenons le cas de notre réseau local redondant.

Figure 3-14.
*Conséquence
d'un spanning tree
mal paramétré.*

① Tous les switches ont été installés avec les valeurs par défaut (valeurs usine). Ils ont les mêmes coûts et les mêmes priorités.

③ Résultat : le flux entre la station et le serveur transite par deux commutateurs au lieu d'être direct.

R = port racine
D = port désigné
En pointillé, les liens invalidés : les liens ont été bloqués par le spanning tree.



LE POINT SUR LE SPANNING TREE (IEEE 802.1D)

L'algorithme du spanning tree consiste à construire un arbre définissant un chemin unique entre un commutateur et sa racine.

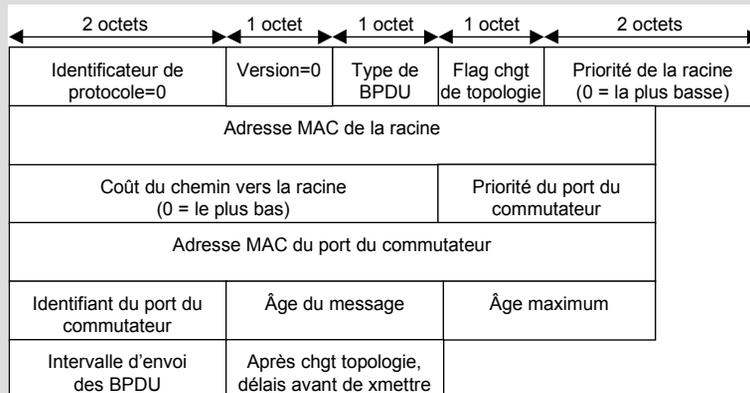
Lors de la construction de l'arbre (suite à un changement de topologie), chaque commutateur émet un **BPDU** (*Bridge Protocol Data Unit*) de configuration sur tous ses ports. Inversement, il retransmet tous les BPDU (éventuellement en les modifiant) qui lui arrivent, et ainsi de suite jusqu'à ce que les BPDU échangés contiennent tous la même valeur.

La première étape de ce processus consiste à élire un **commutateur racine** : c'est celui dont la **priorité** est la plus basse ou, en cas d'égalité, celui dont l'adresse MAC est la plus basse.

Ensuite, chaque commutateur détermine le **port racine** — celui par lequel un BPDU émis par la racine arrive. S'il y en a plusieurs, le port choisi est celui qui a le **coût** de chemin vers la racine le plus bas. Le coût est déterminé par la somme des coûts des ports situés entre le commutateur et la racine. En cas d'égalité, le port choisi est celui qui a la priorité la plus basse ; en cas de nouvelle égalité, c'est celui qui a l'adresse MAC la plus basse.

Enfin, sur chaque segment Ethernet, le commutateur dont le port racine a le coût de chemin vers la racine le plus bas est élu **commutateur désigné**. En cas d'égalité, c'est celui qui a la priorité la plus basse et, en cas de nouvelle égalité, celui qui a l'adresse MAC la plus basse.

En définitive, sur chaque segment Ethernet, un seul chemin vers le commutateur racine sera calculé. Les commutateurs désactivent tous leurs ports qui ne sont ni racines ni désignés.



Afin de détecter les changements de topologie (apparition ou disparition d'un commutateur), la racine envoie régulièrement (toutes les deux secondes par défaut) un BPDU d'annonce (comprenant seulement les trois premiers octets) sur tous ses ports. Les commutateurs transmettent ce BPDU sur leurs ports désignés.

Les BPDU sont envoyés dans des trames Ethernet multicast d'adresse MAC 01:80:C2:00:00:10 transportant les trames LLC identifiées par le SAP 0x42.

Tous les commutateurs ont la même priorité (par exemple 32 768) et le même coût sur chaque port de même débit (par exemple 19 pour les ports à 100 Mbit/s). Le switch 6 a été désigné racine parce que son identifiant (priorité + adresse MAC) était le plus bas. Le même processus de sélection a déterminé les routes menant vers la racine uniquement en se fon-

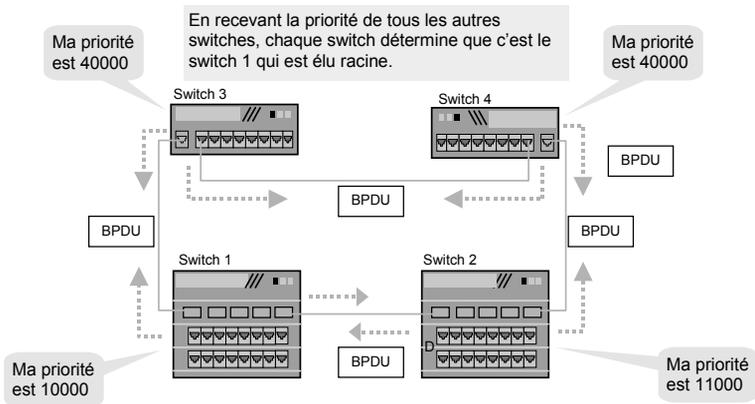
nant sur les valeurs des adresses MAC, puisque toutes les autres valeurs (priorité et coût) sont identiques.

Résultat, certains flux ne sont pas optimisés et peuvent dégrader les performances.

Reprenons les différentes phases de calcul de l'arbre spanning tree. Les commutateurs désignent la racine. Afin d'optimiser les flux, il est préférable que ce soient les commutateurs fédérateurs qui assurent ce rôle. Leur priorité doit donc être abaissée par rapport aux commutateurs d'étage. Étant donné qu'il s'agit de Catalyst 5000, la commande est la suivante :

```
Console> (enable)set spantree priority 10000
VLAN 1 bridge priority set to 10000.
```

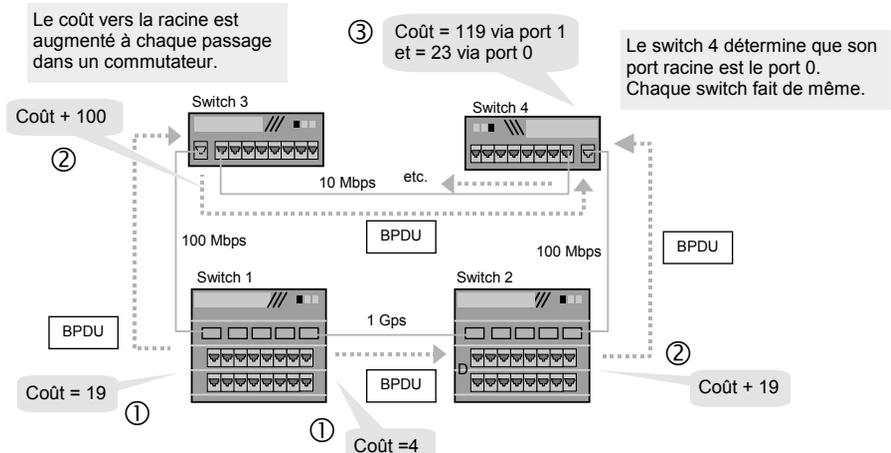
Figure 3-15.
Élection du commutateur racine.



Chaque commutateur choisit ensuite son port racine, celui dont le coût de chemin vers la racine est le plus bas. Sur un Catalyst, le coût de chaque port dépend de son débit : 4 pour 1 Gbit/s, 19 pour 100 Mbit/s, et 100 pour 10 Mbit/s. La commande suivante permet de changer la valeur par défaut :

```
Console> (enable)set spantree portcost 1/1 4
```

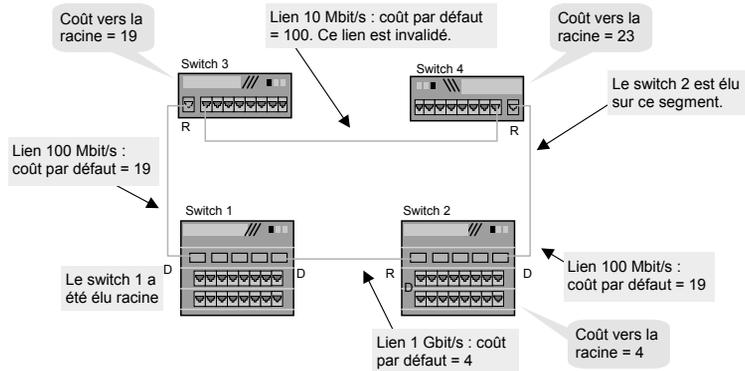
Figure 3-16.
Choix des ports racines.



Sur chaque segment Ethernet, le commutateur désigné est celui dont le port racine a le coût le plus bas. En cas d'égalité, la priorité détermine ce coût. Étant donné que tous les ports ayant un même débit ont le même coût et la même priorité par défaut, le choix s'effectuera en fonction de l'adresse MAC. Pour éviter les mauvaises surprises, il est possible d'abaisser la priorité d'un port pour être sûr qu'il soit désigné en cas de routes multiples :

```
set spantree portpri 1/1 32
```

Figure 3-17.
Élection des commutateurs désignés.



En définitive, les chemins redondants ne sont pas utilisés, et le partage de la charge entre plusieurs routes n'est pas possible.

Les mêmes BPDU sont envoyés sur tous les ports, même là où il n'y a qu'un PC connecté. Le spanning tree peut donc être désactivé sur ces ports, ce qui présente l'avantage de diminuer (un peu) le trafic et d'éviter que des ajouts sauvages de commutateurs (qui seraient connectés sur ces ports) ne viennent perturber votre réseau.

```
set spantree disable
```

Le commutateur racine émet régulièrement (toutes les deux secondes par défaut) des BPDU pour maintenir l'état du spanning tree. Si le réseau est stable (peu d'incidents et de changements), il est possible d'augmenter cette valeur afin de diminuer le trafic

```
set spantree hello 5
```

On peut s'assurer que, sur le switch 1, le spanning tree s'est stabilisé dans une bonne configuration.

```
Console> (enable) show spantree
VLAN 1
Spanning tree enabled
Designated Root          00-1f-00-40-0b-eb-25-d2
Designated Root Priority  45
Designated Root Cost     0
Designated Root Port     1/1
```

```

Root Max Age 20 sec   Hello Time 2 sec   Forward Delay 20 sec
Bridge ID MAC ADDR      00-40-0b-eb-25-d2
Bridge ID Priority      45
Bridge Max Age 20 sec   Hello Time 2 sec   Forward Delay 20 sec

```

Port	Vlan	Port-State	Cost	Priority	Fast-Start
1/1	1	forwarding	4	32	disabled
1/2	1	forwarding	19	32	disabled
2/1	1	forwarding	19	32	disabled
2/2	1	not-connected	19	32	disabled
2/3	1	not-connected	19	32	disabled
3/1	1	not-connected	100	32	disabled
3/2	1	forwarding	100	32	disabled

Le processus de création de l'arbre spanning tree peut durer plusieurs dizaines de secondes. Ce temps est, en réalité, proportionnel au nombre de commutateurs.

Pendant cette phase, aucun commutateur ne traite de trame au cours des 15 premières secondes (valeur par défaut) ; le réseau s'arrête donc de fonctionner chaque fois qu'un commutateur est allumé ou éteint quelque part dans le réseau.

La phase d'apprentissage est encore plus longue lorsque tous les commutateurs s'initialisent en même temps (suite à une panne de courant, par exemple). En effet, les BPDU sont reçus par plusieurs ports dont l'un peut être élu racine, puis invalidé par la suite si un commutateur situé en aval a invalidé sa route. Le temps de stabilisation de l'arbre peut ainsi atteindre plusieurs minutes.

Dans certains cas, notamment sur les commutateurs fédérateurs, il peut être intéressant de diminuer ce temps, surtout si l'architecture réseau est conçue sans aucune boucle.

```
set spantree fwddelay 5
```

Inversement, si ce temps est trop court par rapport au délai de construction de l'arbre, des trames peuvent commencer à circuler et potentiellement être dupliquées dans le cas de routes multiples. Il vaut alors mieux augmenter le paramètre « forward delay » au-delà des 15 secondes par défaut.

La meilleure solution consiste à activer plus rapidement les ports qui ne sont pas concernés par le *spanning tree*, c'est-à-dire ceux sur lesquels est connectée une seule station.

```
set spantree portfast 1/2 enable
```

On peut remarquer que l'échange de BPDU et l'élection d'un commutateur désigné impliquent que chaque port du commutateur soit identifié par une adresse MAC (comme une carte réseau).

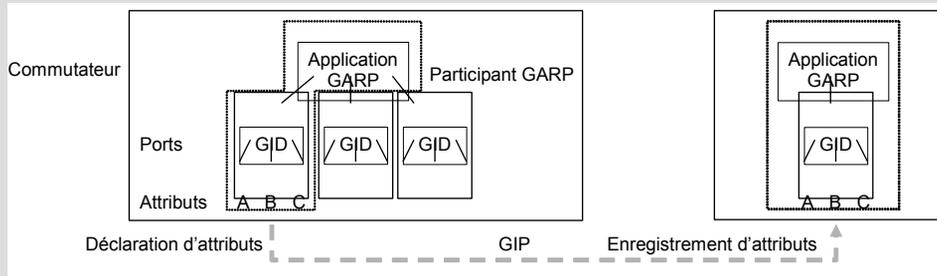
L'alternative à cette optimisation repose sur la nouvelle version du spanning tree, appelée **RSTP** (*Rapid Spanning Tree Protocol*) et issue du groupe de travail IEEE **802.1w**. Le protocole RSTP est compatible avec le spanning tree et sera, à terme, intégré dans le standard 802.1d.

LE POINT SUR GARP (IEEE 802.1d)

La norme 802.d spécifie, en plus du *spanning tree*, le protocole **GARP** (*Generic Attribute Registration Protocol*), qui permet aux commutateurs d'échanger des informations de gestion. Chaque commutateur peut ainsi déclarer aux autres commutateurs appartenir à un groupe possédant certains attributs.

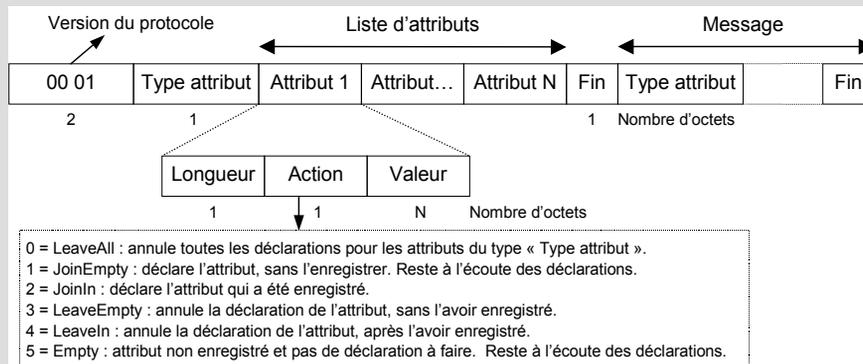
La même norme spécifie une application de GARP à l'enregistrement des groupes multicast, protocole appelé **GMRP** (*GARP Multicast Registration Protocol*). La norme 802.q utilise également GARP pour la propagation dynamique des VLAN entre commutateurs : le protocole est appelé **GVRP** (*GARP VLAN registration Protocol*).

Un commutateur fonctionne avec plusieurs participants GARP, un par application et par port du commutateur. Un tel participant est constitué de l'application et d'une instance du GID (*GARP Information Declaration*), qui gère, pour chaque attribut utilisé par l'application, l'état du participant (membre, observateur...) associé à l'état des attributs (enregistré, déclaré...). Les GID utilisent le protocole GIP (*GARP Information Propagation*) pour échanger ces attributs.



À l'aide de ces composants, un port recevant la déclaration d'un attribut l'enregistre et le déclare (i.e. le propage) aux autres ports. Inversement, l'annulation d'un attribut est propagée vers un port si tous les autres ports ont annulé l'enregistrement de cet attribut.

Le PDU (*Protocol Data Unit*) utilisé par GIP est constitué de messages contenant des listes d'attributs regroupés par types.



Les PDU GARP sont transportés dans des trames LLC identifiées par le SAP 0x42. Les applications GARP sont identifiées par des adresses MAC qui leur sont réservées : 01-80-C2-00-00-20 pour GMRP et 01-80-C2-00-00-21 pour GVRP.