Architecture réseau - Isima F2 (3ème année)

Alexandre Guitton

année 2007-2008, version 1.0

Table des matières

Chapitre 1

Introduction

Les réseaux prennent une place de plus en plus importante dans les entreprises. Ils véhiculent en effet des volumes importants de données numériques confidentielles, et ils sont de plus en plus utilisés pour faire transiter les communications téléphoniques.

Les réseaux d'entreprise peuvent couvrir de quelques dizaines à quelques milliers d'équipements, sur des échelles allant d'un étage d'un bâtiment à des complexes industriels dans des pays distants interconnectés.

Dans ce cours, nous allons présenter les architectures de divers types de réseaux d'entreprise. Nous discuterons des réseaux locaux, étendus et métropolitains. Avant de commencer, nous allons cependant rappeler quelques notions importantes.

1.1 Adresses MAC et IP

Définition 1 (Adresse MAC) *L'adresse MAC est un identificateur (supposé unique globalement, et devant être unique sur un même réseau) associé à une carte réseau. Il s'agit d'un identificateur physique, dépendant du constructeur de la carte et du numéro de série de celle-ci.*

Une adresse MAC occupe 48 bits (soit 6 octets). Elle est souvent représentée par six valeurs hexadécimales (chacune représentant un octet) séparées par des points, comme : `12.34.56.78.9A.BC`*. L'adresse MAC de diffusion est :* `FF.FF.FF.FF.FF.FF`*.*

Définition 2 (Adresse IP) *L'adresse IP est un identificateur logique (devant être unique publiquement) associé à une carte réseau.*

Une adresse IPv4 occupe 32 bits (soit 4 octets). Elle est souvent représentée par quatre valeurs décimales (chacune représentant un octet) séparées par des points, comme : `123.45.67.89`*.*

Un routeur possédant au moins deux cartes réseaux, il possèdera donc plusieurs adresses MAC et plusieurs adresses IP différentes.

1.2 Modèles réseaux

Il existe deux principaux modèles réseaux : le modèle historique et le modèle d'Internet. Ces deux modèles présentent quelques différences, mais ils sont essentiels tous les

deux pour la bonne compréhension des mécanismes réseaux.

1.2.1 Modèle OSI

Le modèle OSI (pour *open systems interconnection*, interconnexion de systèmes ouverts) est un modèle à sept couches décrivant les fonctionnalités de communication et d'organisation de tout élément d'un réseau, du niveau physique au niveau logiciel.

Les noms et fonctionnalités des sept couches suivent :

La couche physique touche au support physique. Elle s'assure de la synchronisation, de l'encodage/décodage et de la transmission/réception de bits.

La couche liaison de données gère des communications entre deux machines connectées au travers d'un support physique commun. Ces deux machines communiquent par l'intermédiaire de trames.

La couche réseau concerne des interconnexions de machines distantes. Elle assure une transmission de bout en bout. Les machines communiquent en s'échangeant des paquets.

La couche transport gère des communications de bout en bout entre processus. Les données échangées sont des messages.

La couche session gère des modèles de communications transactionnels, appelés sessions.

La couche présentation se charge du codage des données applicatives pour les présenter à l'application de manière transparente, indépendemment des codages des différents réseaux ou des différentes machines.

La couche application offre un point d'accès aux applications réseaux.

Le modèle OSI présente plusieurs problèmes :

- Il fait l'hypothèse qu'il n'y a qu'un seul protocole par couche. Cette hypothèse n'est pas vérifiée lorsque l'on fait de l'interconnexion de réseaux hétérogènes par exemple.
- Certains protocoles des couches basses offrent des services que le modèle OSI propose dans les couches plus hautes. Le modèle perd de l'efficacité si la redondance n'est pas supprimée. Cela peut être pire : un protocole d'une couche peut nécessiter des services supplémentaires d'un protocole inférieur.

1.2.2 Modèle IP

Le modèle IP (pour *Internet protocol*) est un découpage en couches plus réaliste que le modèle OSI en ce qui concerne la pile de protocoles IP. On le nomme aussi modèle TCP/IP (pour *transport control protocol / Internet protocol*), par abus de langage. Il définit quatre couches :

La couche d'accès réseau fait circuler les informations sur le réseau sous-jacent.

La couche Internet fournit les mécanismes nécessaires à la gestion et à l'acheminement de paquets de données (protocole IP).

La couche transport assure la communication de données entre deux machines distantes (protocoles TCP et UDP (pour *user datagram protocol*)).

La couche application englobe toutes les applications.

1.2.3 Comparaison du modèle OSI et du modèle IP

Les modèles OSI et IP sont très similaires. La figure 1.1 présente la manière dont certaines couches du modèle OSI sont regroupées dans le modèle IP.

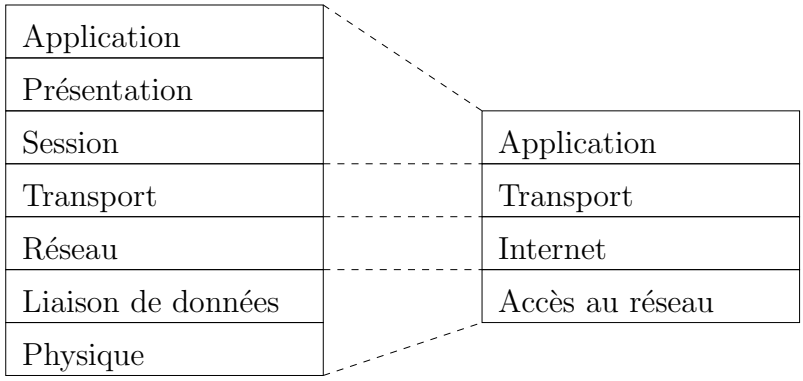

FIG. 1.1 – Comparaison du modèle OSI (à gauche) et du modèle IP (à droite).

1.3 Rappels sur la pile de protocoles IP

La pile de protocoles IP est au cœur de l'Internet. Ici, nous ne ferons que des rappels sur les protocoles de base.

1.3.1 Protocole IP

Le protocole IP (pour *Internet Protocol*) est un protocole de la couche réseau. La version la plus récente du protocole est la version 6 (IPv6), mais celle qui est mondialement utilisée est la version 4 (IPv4).

Le protocole IP offre des services minimaux :
- l'identification d'une machine par l'intermédiaire d'un identifiant public unique que l'on appelle adresse IP,
- l'acheminement des paquets sans garantie (on parle de *best effort*),
- la fragmentation des messages trop gros en plusieurs paquets (on parle encore de fragments).

Le protocole IP n'assure pas de :
- l'intégrité des données[1],
- l'ordre d'arrivée des paquets ou des fragments,
- l'unicité des paquets reçus (certains pouvant être dupliqués),
- la réception des paquets envoyés.

Le protocole IP utilise conjointement un protocole de contrôle nommé ICMP et un protocole de traduction d'adresses nommé ARP. Au-dessus d'IP, on trouve le protocole de transport UDP offrant des services additionels minimaux, et le protocole de transport TCP fournissant une grande fiabilité.

[1]IP possède une somme de contrôle pour se protéger des erreurs de transmission, mais elle n'est pas un obstacle pour un attaquant voulant corrompre les données

1.3.2 Protocole ICMP

Le protocole ICMP (pour *Internet Control Message Protocol*) est le protocole permettant de véhiculer des messages de contrôle ou d'erreur. Les messages de contrôle incluent :

- la demande ou la réponse d'écho (le célèbre `ping`),
- la demande ou la réponse de redirection.

Les messages d'erreur sont :

- la notification d'une destination inaccessible,
- la notification d'une extinction de la source (quand les routeurs ne sont plus en mesure de sauvegarder les paquets reçus par la source),
- la notification de temps d'acheminement dépassé,
- la notification d'entête erroné.

1.3.3 Protocole ARP

Le protocole ARP (pour *Address Resolution Protocol*) est un protocole de la couche réseau effectuant la traduction d'adresses IP en adresses MAC.

Un routeur recevant un paquet IP en extrait une adresse de destination. En examinant sa table de routage[2], le routeur détermine l'adresse IP et l'interface de sortie correspondant au prochain saut vers la destination. Cependant, la trame dans lequel ce paquet sortant va être encapsulée doit posséder l'adresse MAC de l'interface du prochain saut. Si cette adresse MAC n'est pas connue, l'émetteur émet une requête ARP demandant quelle est l'adresse MAC correspondant à l'adresse IP donnée. La machine ayant l'adresse IP répond[3]. Le routeur connait à présent l'adresse MAC du prochain saut et peut envoyer la trame à la bonne adresse. Les correspondances entre adresses MAC et adresses IP sont mémorisées dans un cache.

Voici un exemple de traitement d'un paquet par un routeur. Cet exemple est illustré sur la figure 1.2.

1. Le routeur R reçoit un paquet P.
   - L'adresse MAC destination de P est l'adresse MAC mac_1 de l'interface par laquelle R a reçu P.
   - L'adresse IP destination ip_5 de P est l'adresse IP de la destination finale de P.
2. R traite le paquet et se rend compte que le paquet n'est pas pour lui (car R ne possède pas l'adresse IP ip_5). R consulte alors sa table de routage[4] et détermine que le routeur suivant pour ip_5 est le routeur d'adresse IP ip_3. R détermine aussi quelle est l'interface de sortie pour ce paquet.
3. R ne connait pas l'adresse MAC associée à ip_3 (ce cas est en fait très peu probable ; il apparait pour le premier paquet échangé entre ces deux routeurs). R envoie donc une requête ARP de diffusion demandant qui possède l'adresse IP ip_3.

[2]Plus précisément, le routeur examine sa table d'acheminement, la table de routage étant utilisée pour construire la table d'acheminement.

[3]Toute machine connaissant la réponse à la requête pourrait répondre. C'est d'ailleurs ce que font les proxys ARP

[4]Plus précisément, R consulte sa table d'acheminement. La table de routage est utilisée pour construire la table d'acheminement.

- L'adresse MAC source de la requête ARP est l'adresse de l'interface de sortie de P. Il s'agit de mac_2.
- L'adresse MAC destination de la requête ARP est -1.

4. Le routeur possédant l'adresse ip_3 reçoit la trame ARP et émet une réponse ARP contenant l'adresse MAC correspondant à l'adresse ip_3.
   - L'adresse MAC source de la réponse ARP est l'adresse de l'interface par laquelle la requête ARP a été reçue. Il s'agit de mac_3.
   - L'adresse MAC destination de la réponse ARP est l'adresse MAC source de la requête, c'est-à-dire mac_2.
5. R reçoit la réponse ARP et peut à présent envoyer le paquet P à son voisin.
   - L'adresse MAC source de P est mac_2.
   - L'adresse MAC destination de P est mac_3.
   - L'adresse IP source et destination de P n'a pas changé. Elle reste ip_5.

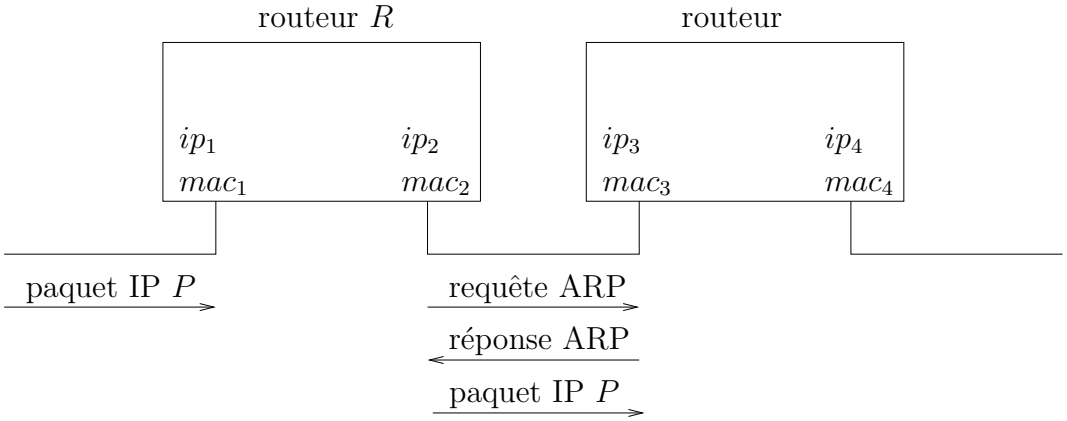

FIG. 1.2 – ARP est utilisé pour connaitre l'adresse MAC d'une machine voisine dont on ne connait que l'adresse IP.

1.3.4 Protocole UDP

Le protocole UDP (pour *User Datagram Protocol*) est un protocole de la couche transport, au-dessus du protocole IP. Par rapport à IP, il ajoute un seul service : le multiplexage de communications par le biais de ports. En d'autres termes, UDP offre aux applications différents points d'accès. Deux applications différentes ne peuvent pas partager un même point d'accès.

Le protocole UDP est un protocole simple. Cela fait de lui un protocole de choix pour les applications temps-réels, audio et vidéo, pouvant tolérer quelques erreurs de transmission.

1.3.5 Protocole TCP

Le protocole TCP (pour *Transmission Control Protocol*) est un protocole de la couche transport, au-dessus du protocole IP. Par rapport à IP, il offre de nombreux services, notamment :

– l'ordonnancement des données,
– la réception d'une et une seule copie de chaque paquet envoyé,
– la gestion de données urgentes (dites hors-flux),
– l'adaptation du débit aux conditions du réseau (ce qui se nomme le contrôle de congestion) et à la mémoire des partenaires (ce qui se nomme le contrôle de flux).

Le protocole TCP offre un service connecté au-dessus d'un service non-connecté tel qu'IP. Le protocole TCP utilise les fonctionnalités suivantes :

– la numérotation de chaque paquet (au moyen d'un *sequence number*),
– un acquittement explicite (quasi-)systématique,
– la séparation logique des deux canaux de communications unidirectionnels.

Nous allons voir brièvement le mécanisme d'établissement de la connexion, le mécanisme de transfert d'information, et le mécanisme de fermeture de la connexion.

Ouverture de la connexion

L'ouverture de la connexion avec le protocole TCP se fait en trois étapes, comme indiqué sur la figure 1.3. Tout d'abord, le client envoie un paquet TCP ayant le bit SYN (pour *synchronisation*) au serveur. Le serveur répond avec un paquet TCP ayant les bits SYN et ACK (pour *acquittement*) au client. Finalement, le client répond avec un paquet TCP ayant simplement le bit ACK. Le numéro de séquence SN et le numéro d'acquittement AN sont incrémentés à tour de rôle.

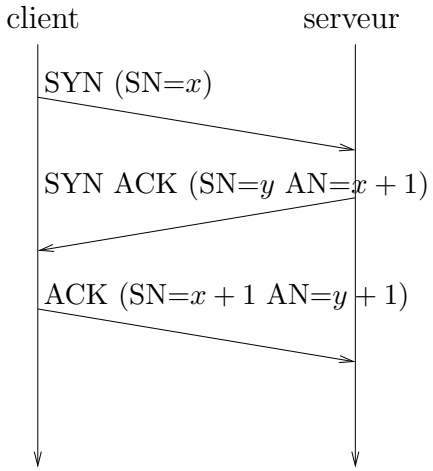

Fig. 1.3 – L'ouverture de la connexion avec le protocole TCP.

Transfert d'information

L'envoi de données se fait en positionnant le bit PSH (pour *push*) dans un paquet. L'autre extrémité de la connexion répond en acquittant le nombre d'octets reçus. Le numéro de séquence augmente de la taille des données reçues. Le schéma de transfert est présenté sur la figure 1.4.

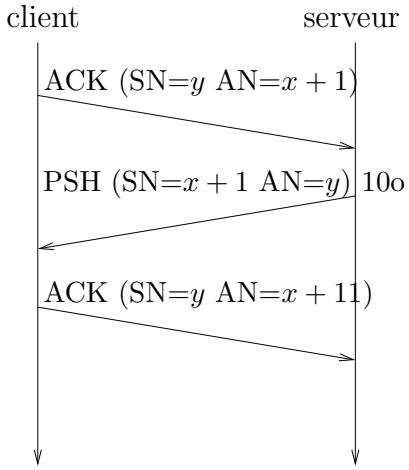

FIG. 1.4 – L'envoi de 10 octets de données avec le protocole TCP.

Fermeture de la connexion

La fermeture de la connexion avec le protocole TCP se fait en quatre étapes, comme indiqué sur la figure 1.5. Cela vient du fait que les deux canaux unidirectionnels peuvent être fermés indépendemment. Pour chaque canal, une extrémité envoie sa demande de fermeture au moyen d'un paquet TCP ayant le bit FIN. L'autre extrémité répond en envoyant un bit paquet TCP ayant le bit ACK.

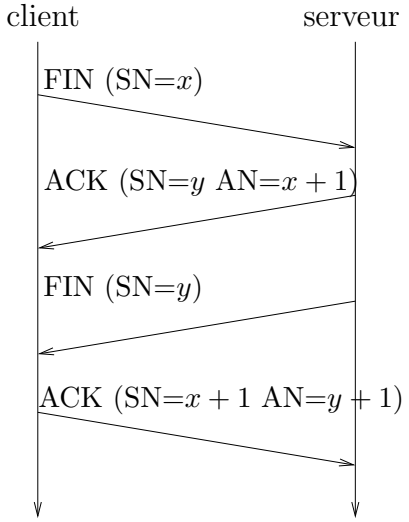

FIG. 1.5 – La fermeture de la connexion avec le protocole TCP.

1.4 Références

Ce cours s'appuie sur les notes de cours de Christophe GOUINAUD, et sur plusieurs ouvrages comme :

- le livre de Jean-Luc MONTAGNIER sur les réseaux d'entreprise [2],
- le livre de Guy PUJOLLE sur les réseaux [3],

– le livre du groupe Gisèle Cizault sur IPv6 [1].

Beaucoup de documentation est tirée d'Internet, et notamment du site `http ://www.wikipedia.com`.

Chapitre 2

Réseaux locaux

Les réseaux locaux sont des réseaux de faible étendue géographique (allant jusqu'à quelques bâtiments sur un même site). On les nomme LAN (pour *local area network*) ou RLE (pour réseau local d'entreprise).

Nous allons étudier dans ce chapitre l'architecture et les protocoles des petits LAN, puis ceux des grands LAN. Nous nous concentrerons sur les LAN filaires, en ignorant les topologies et protocoles des WLAN (pour *wireless LAN*, les LAN sans fils).

2.1 Topologies des petits réseaux locaux

Il existe trois topologies de petits réseaux locaux :

La topologie en bus où toutes les stations se connectent sur un même support physique.

La topologie en étoile où toutes les stations sont disposées sur les extrémités d'une étoile, dont le centre est un dispositif spécial.

La topologie en anneau où toutes les stations sont connectées sur un même anneau logique. Chacune des stations reçoit des données de sa station précédente et émet des données vers sa station suivante.

Ces trois topologies sont représentées sur la figure 2.1.

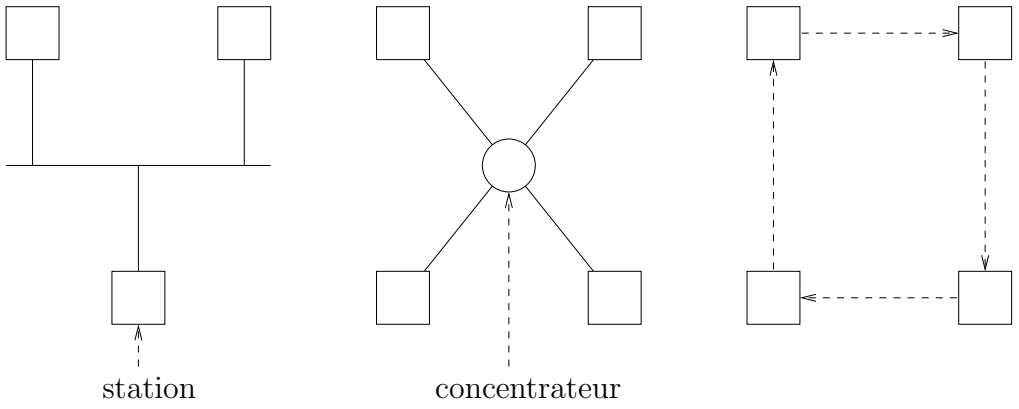

Fig. 2.1 – Les topologies en bus (à gauche), en étoile (au centre) et en anneau (à droite).

2.1.1 Topologie en bus

La topologie en bus se réalise à l'aide de câbles coaxiaux qui sont chainés les uns aux autres par l'intermédiaire de prises BNC en T. Les cartes réseaux des ordinateurs se connectent sur la branche centrale du T par un connecteur BNC, tandis que les deux branches latérales servent à étendre le câble coaxial. Aux deux extrémités du bus, on place un bouchon. La figure 2.2 décrit cette topologie.

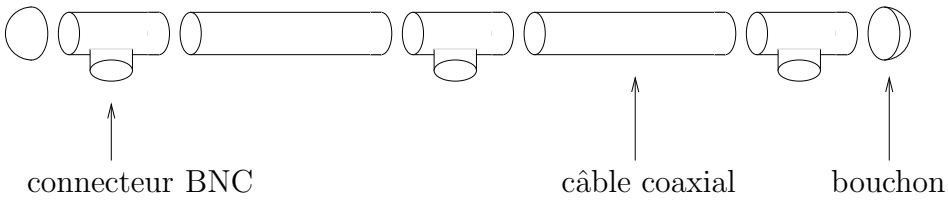

FIG. 2.2 – La topologie en bus

Le câble coaxial possède une âme en cuivre. L'atténuation du signal sur un câble en cuivre est assez forte. Un bus étant dépourvu de répéteurs de signal, sa longueur totale ne peut pas excéder quelques centaines de mètres (entre 90 et 200 mètres).

Sur une topologie en bus, le protocole Ethernet est utilisé.

Définition 3 (Ethernet) *Ethernet est un protocole de la couche 2 définissant une méthode d'accès au support physique partagé. Lorsqu'une station a des données à émettre, elle teste le support physique afin de déterminer si une autre station est en train d'émettre ou non. Si ce n'est pas le cas, la station transmet sa trame, sinon, elle attend un temps aléatoire. Lorsque deux stations émettent en même temps, une collision est détectée par les deux stations. Elles arrêtent la transmission et reprennent toutes les deux après un temps aléatoire. Le temps d'attente aléatoire est choisi dans un intervalle dont la taille double à chaque tentative infructueuse.*

Ethernet est un protocole de type CSMA/CD (pour carrier sense multiple access / collision detection). Plusieurs stations ont accès au medium, et le protocole met en œuvre une détection de la porteuse et des collisions.

Ethernet utilise un TBEB (pour *Truncated Binary Exponential Backoff*) pour calculer le délai avant une autre tentative (aussi appelé *backoff*). Le délai à la n-ème tentative est de la forme 2^n. Quand cette valeur devient trop grande, elle est tronquée.

On parle de câbles Ethernet fins ou 10base2, le 10 indiquant que le débit théorique atteignable est de 10 Mbit/s et le 2 indiquant la longueur maximale de 200 mètres.

Avantages et inconvénients de la topologie en bus. Une topologie en bus ne coûte pas cher, mais elle est difficile à mettre en place, à maintenir (lors de l'ajout ou la suppression de stations), offre un faible débit et ne peut pas s'étendre sur de longues distances. Elle est adaptée à des réseaux de très peu de machines.

2.1.2 Topologie en étoile

La topologie en étoile se fait autour d'un équipement appelé concentrateur (ou *hub*). Les câbles utilisés sont des paires torsadées : quatre paires de cuivre enroulées dans une

gaine protectrice. L'enroulement permet de réduire les interférences électro-magnétiques des fils entre eux. Aux extrémités des paires torsadées, on trouve un connecteur RJ45. La figure 2.3 décrit cette topologie.

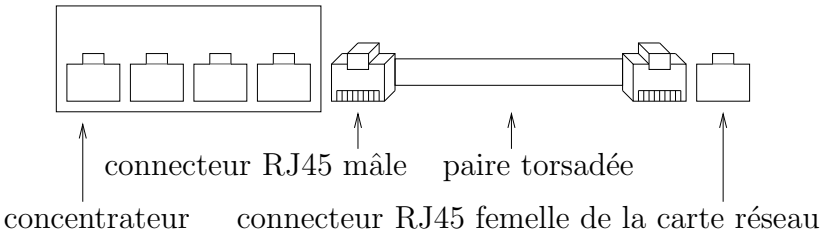

FIG. 2.3 – La topologie en étoile

Pour que les signaux émis par un ordinateur A soient reçus par un ordinateur B, il convient qu'à un moment, les fils soient croisés : les fils d'émission de A doivent être reliés aux fils de réception de B, et les fils de réception de A doivent être reliés aux fils d'émission de B. C'est le concentrateur qui va effectuer ce croisement. Ainsi, tous les câbles entre le concentrateur et les ordinateurs sont des câbles droits (c'est-à-dire non croisés). Il existe des câbles croisés permettant de relier deux ordinateurs (uniquement) directement, sans passer par un concentrateur.

Sur un réseau en étoile, les débits peuvent aller de 10 Mbit/s à 100 Mbit/s ou même 1 Gbit/s, en fonction des équipements. Dans le nom 100baseT, le T signifie torsadé pour rappeler que les paires sont torsadées.

Il existe de nombreux types de câbles. On rencontre fréquemment les catégories 5 (classe D), 5E, 6 (classe E) et 7 (classe F) (avec des débits maximum croissants). La catégorie la plus répandue est la catégorie 5 supportant le 100 Mbit/s et pouvant supporter le 1 Gbit/s, mais il est préférable d'opter pour la catégorie 6 qui supporte obligatoirement le 1 Gbit/s. Les connecteurs de la catégorie 7 ne sont pas encore normalisés. De plus, les câbles peuvent disposer de protections supplémentaires contre les interférences (causées par les courants forts voisins). On distingue :

- les câbles UTP (pour *unshielded twisted pair*), n'offrant pas de protection supplémentaire,
- les câbles FTP (pour *foiled twisted pair*), dont les paires sont entourées d'un écran en aluminium,
- les câbles STP (pour *shielded twisted pair*), dont les paires sont entourées d'un blindage métallique tressé,
- les câbles SFTP (pour *shielded foiled twisted pair*), possédant à la fois un écran et un blindage.

La longueur des câbles ne peut pas excéder 90 mètres. Cette contrainte ne s'applique qu'entre le concentrateur et les ordinateurs, ou entre deux concentrateurs, puisque les concentrateurs agissent comme des répéteurs de signal.

Le concentrateur est un équipement réseau de la couche 2 qui répète les signaux reçus d'une interface vers toutes les autres interfaces. Les signaux sont régénérés par le concentrateur. Le choix d'un concentrateur se fait sur plusieurs critères :

- Le nombre de ports. C'est le critère le plus important. Les concentrateurs ont souvent de 4 à 48 ports.
- La capacité à désactiver les ports dont le signal n'est pas bon (la plupart des concentrateurs possèdent cette faculté).

– La présence d'un port BNC pour l'interconnexion avec un réseau en bus.
– La capacité à être administrable (ou *manageable*) à distance (par l'intermédiaire d'un serveur SNMP, décrit ultérieurement). C'est bien souvent inutile.
– La présence d'un port spécial dit *uplink*. Ce port permet est simplement un port non croisé.
– La capacité à être chainable (ou *stackable*), c'est-à-dire à se connecter à d'autres concentrateurs. C'est bien souvent inutile, car cela se fait par l'intermédiaire d'un câble spécifique et il existe d'autres moyens de chainer les concentrateurs.
– La capacité de segmentation. Elle permet de séparer virtuellement les ports, comme s'ils appartenaient à deux concentrateurs différents. Cela permet de différencier deux réseaux.

Le chainage de concentrateurs se fait souvent par l'intermédiaire du port *uplink*. Pour connecter deux concentrateurs par ce port, on utilise un câble droit allant du port *uplink* de l'un des concentrateurs vers l'un des ports standard de l'autre concentrateur. Le croisement a lieu dans le port standard, mais pas dans le port *uplink*. Les concentrateurs peuvent aussi être chainés par l'intermédiaire de leurs ports standards, en utilisant des câbles croisés cette fois (pour annuler l'un des croisements).

Sur une topologie en bus, le protocole Ethernet est lui aussi utilisé.

Outre les paires torsadées, il est aussi possible d'utiliser des fibres optiques. La décision de passage en fibre optique se fait sur une base de distance (les fibres optiques n'ont pas la limitation de la centaine de mètres des paires torsadées) ou quand les débits doivent être très importants (plusieurs Gbit/s).

Définition 4 (Fibre optique) *Une fibre optique est un câble transportant des signaux optiques. La fibre optique est composée d'un cœur dans lequel la lumière se propage, et d'une gaine protectrice. L'angle d'indice du signal entrant la fibre est choisi de manière à garantir une réflexion totale du signal, ce qui lui permet d'être transmis sur de grandes distantes sans perdre beaucoup d'intensité.*

Les fibres ne permettant qu'à un seul signal d'être transmis sont appelées fibres monomodes. Certaines fibres permettent à plusieurs signaux d'être transmis avec des angles d'incidence différents. Ces fibres sont appelées fibres multimodes.

Il y a toutefois des contraintes de longueur sur les fibres optiques. En effet, à mesure que les signaux optiques la parcourent, ils sont atténués et se dispersent. Pour un débit d'1 Gbit/s, la longueur maximale d'une fibre optique est :
– environ 3 km pour une fibre monomode et une fréquence de 1300 nm,
– 550 m pour une fibre multimode dont le cœur a un diamètre de 50 microns (ce qui se note 50/125), pour des fréquences de 850 nm et de 1300 nm,
– 550 m pour une fibre multimode dont le coeur a un diamètre de 62.5 microns (ce qui se note 62.5/125) à une fréquence de 1300 nm,
– 300 m pour une fibre multimode dont le coeur a un diamètre de 62.5 microns (ce qui se note 62.5/125) à une fréquence de 850 nm.

Celle que l'on trouve le plus fréquemment est la fibre optique 62.5/125 fonctionnant sur 850 nm.

Les avantages des fibres optiques sur les câbles en cuivre sont :
– les grands débits atteignables (de l'ordre de 25 Tbit/s en théorie),

– la faible atténuation des signaux, qui peuvent ainsi être transmis sur de grandes distances,
– la non-perturbation des signaux par les rayonnements électro-magnétiques,
– l'installation possible dans un milieu déflagrant, puisqu'il n'y a pas d'étincelles.

Avantages et inconvénients de la topologie en étoile. La topologie en étoile est la topologie la plus utilisée actuellement. Le nombre important de câbles partant des concentrateurs rend la maintenance difficile si elle n'est pas prévue initialement. Une telle topologie nécessite l'achat d'un nombre relativement important de concentrateurs. Elle peut couvrir une zone relativement grande étant donné que les concentrateurs régénèrent les signaux.

2.1.3 Topologie en anneau

La topologie en anneau est une topologie logique. Physiquement, elle s'appuie sur une topologie en étoile, comme indiqué sur la figure 2.4. Le concentrateur, aussi appelé MAU (pour *media attach unit*) ou MSAU (pour *multistation access unit*), permet de connecter plusieurs stations. Les MAU peuvent être chainés par l'intermédiaire de ports spécifiques, les ports *Ring In* et *Ring Out*. C'est le MAU qui fait le bouclage.

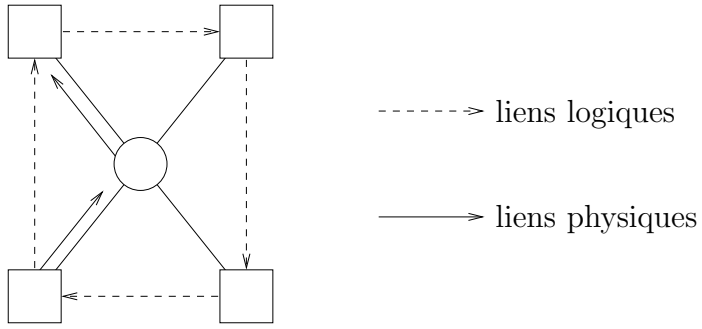

FIG. 2.4 – La topologie logique de l'anneau se base sur une topologie physique en étoile.

Sur une topologie en anneau, c'est le protocole *Token ring* qui est utilisé.

Définition 5 (Token ring) *Token ring (littéralement, anneau à jeton) est un protocole de la couche 2 fonctionnant sur une topologie en anneau. Un jeton virtuel est passé de proche en proche en suivant les liens logiques de l'anneau. Une station ayant des informations à transmettre attend de posséder le jeton, le conserve, émet une trame de données qui parcourt l'anneau. Lorsque la station destinataire reçoit le jeton, elle transforme la trame de données en trame d'acquittements, et renvoie cette trame sur l'anneau. Finalement, la station émettrice reçoit la trame d'acquittement, la détruit et relâche le jeton.*

Le jeton est modélisé par une trame virtuelle. Dans un anneau Token ring, une machine est élue moniteur actif. Elle est en charge de s'assurer qu'il y a toujours un et un seul jeton circulant sur l'anneau. Elle peut supprimer des jetons dupliqués ou régénérer des jetons perdus.

Avantages et inconvénients de la topologie en anneau. La topologie en anneau est physiquement une topologie en étoile. Elle partage donc la plupart des avantages et inconvénients. L'utilisation du protocole *Token Ring* tombe en désuétude et constitue un inconvénient supplémentaire.

2.1.4 Comparaison d'Ethernet et de *Token Ring*

Alors qu'Ethernet est un protocole à accès aléatoire, donc stochastique, *Token Ring* est à accès déterministe. Il a été prouvé que *Token Ring* est plus performant qu'Ethernet. Cependant, l'apparition de l'Ethernet commuté (que nous verrons dans la partie suivante) permet à Ethernet d'obtenir des performances bien supérieures. L'implantation d'Ethernet est aujourd'hui majoritaire, puisqu'environ 85% des LAN utilisent Ethernet, et 15% utilisent *Token Ring*. La tendance est à accentuer cet écart.

2.2 Topologies des grands réseaux locaux

Dans cette partie, nous nous intéressons aux réseaux locaux de plusieurs dizaines d'ordinateurs. Pour que les performances d'un tel réseau soient importantes, il est nécessaire d'utiliser des techniques avancées comme la commutation et l'augmentation du débit. Pour faciliter l'installation et la maintenance, nous verrons comment mettre en place un système de câblage cohérent. Enfin, il convient de s'intéresser à la robustesse du réseau par la mise en place de câbles redondants agissant comme des routes de secours si une partie du réseau tombe en panne.

2.2.1 Commutation

Dans la partie précédente, nous avons vu les concentrateurs, des équipements actifs qui répètent les signaux. Il existe un autre type d'équipement actif pour Ethernet : le commutateur (ou *switch*).

La tâche du commutateur est de répéter une trame qu'il reçoit sur le bon port, plutôt que sur tous les ports. Ils sont en quelque sorte des concentrateurs optimisés.

Définition 6 (Domaine de diffusion) *Un domaine de diffusion est un ensemble de machines atteintes par l'émission d'une trame de diffusion. Le domaine de diffusion d'une machine est l'ensemble de ses voisins au sens de la couche liaison de données.*

Définition 7 (Domaine de collision) *Un domaine de collision est un ensemble de machines qui entrent en compétition (au sens CSMA/CD) lors de l'émission d'une trame.*

Dans un réseau Ethernet ne disposant que de concentrateurs, le domaine de collision est l'ensemble du réseau. Dans un réseau Ethernet commuté, le commutateur sépare le réseau en un domaine de collision par port. Le domaine de diffusion est l'ensemble du réseau dans les deux cas.

Pour savoir à quel port une trame est destinée, le commutateur maintient une table associant des ports à des adresses MAC. À chaque fois que le commutateur reçoit une trame, il stocke l'adresse MAC de l'émetteur et l'associe au port sur lequel la trame a été

reçue. Pour réémettre la trame, le commutateur regarde l'adresse MAC de la destination. Si l'adresse MAC est déjà dans la table, le commutateur vérifie que le port de sortie de la trame soit différent du port d'entrée de la trame, puis si c'est le cas, retransmet la trame sur le port associé. Si l'adresse MAC n'est pas encore présente, le commutateur transmet la trame sur tous les ports, agissant ainsi comme un concentrateur par défaut.

La commutation est différente du routage, notamment au niveau des adresses : les commutateurs utilisent des adresses MAC, tandis que les routeurs utilisent des adresses de la couche réseau (comme les adresses IP).

2.2.2 Augmentation des débits

Le mode standard pour Ethernet est le *half-duplex* : les données peuvent circuler dans les deux sens (de a vers b ou de b vers a) mais pas simultanément. En effet, quand a émet, b le détecte et s'interdit d'émettre.

Il est possible d'utiliser un autre mode, le *full-duplex*, pour dissocier émission et réception et ainsi augmenter les débits. Une autre possibilité pour augmenter les débits est d'agréger plusieurs ports entre deux commutateurs.

Mode *full-duplex*

Le mode *full-duplex* est un mode dans lequel une carte peut simultanément émettre et recevoir des trames. Pour réaliser de l'Ethernet *full-duplex*, il est nécessaire :

- de pouvoir émettre et recevoir simultanément au niveau physique,
- de ne pas effectuer la détection de collision (qui teste le canal de réception lors de l'émission).

Les cartes Ethernet connectant des paires torsadées disposent de circuits différents pour l'émission et pour la réception. La première condition est remplie. De plus, dans un Ethernet commuté basé sur une topologie en étoile, la commutation assure qu'il n'y aura pas de collisions, rendant désué la détection de collisions.

Le mode *full-duplex* permet d'augmenter les débits puisque d'une part le retard inhérent au TBEB qu'utilise Ethernet est limité au strict minimum (il n'y a pas de période de *backoff*) et d'autre part car le canal peut être utilisé dans les deux sens simultanément.

Agrégation de ports

La technique d'agrégation de ports de commutateurs (aussi appelée *port trunking*) permet d'augmenter les débits entre deux commutateurs avec des liens *full-duplex*. Pour cela, il suffit de configurer l'agrégation dans les deux commutateurs et de connecter plusieurs de leurs ports (par des câbles croisés sur des ports banalisés, ou droits si le port *uplink* d'un commutateur est utilisé). Le débit obtenu est égal à la somme des débits des ports agrégés. La figure 2.5 présente l'agrégation de trois ports à 100 Mbit/s, permettant d'obtenir un débit de 300 Mbit/s.

2.2.3 Systèmes de câblages

Un système de câblage est l'ensemble des éléments nécessaires à l'interconnexion de stations. Il regroupe notamment :

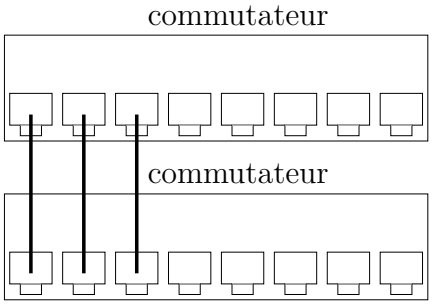

FIG. 2.5 – Trois ports à 100 Mbit/s sont agrégés pour générer un port logique à 300 Mbit/s.

– des câbles,
– des équipements actifs,
– des locaux techniques,
– des chemins de câbles.

Nous avons déjà parlé des câbles et des équipements actifs. Il nous reste à décrire les locaux techniques et les chemins de câbles.

Locaux techniques

Les équipements actifs sont généralement disposés dans des locaux techniques. Ces locaux permettent de concentrer dans une même pièce les câbles connectant les stations aux équipements actifs. Il existe deux types de locaux techniques : les locaux techniques d'étage et les locaux nodaux.

Définition 8 *Local technique d'étage (LTE) Un local technique d'étage est un petit local concentrant les câbles venant des stations et les connectant aux équipements actifs.*

Définition 9 *Local nodal (LN) Un local nodal est un petit local concentrant les câbles venant de tous les locaux techniques d'étage et les connectant à des équipements actifs.*

Généralement, on place un local technique d'étage par étage et on place un local nodal par bâtiment. C'est souvent du local nodal qu'est réalisée l'interconnexion avec d'autres réseaux. Les câbles entre stations et locaux techniques d'étages sont appelés câbles de distributions. Les câbles entre locaux techniques sont appelés rocades.

Le placement des locaux techniques d'étage dépend du nombre de prises dans les équipements actifs ainsi que de la distance maximale entre stations et équipements actifs.

Les figures 2.6, 2.7 et 2.8 présentent respectivement des réseaux locaux d'un étage, de majoritairement un étage et de deux étages. Le choix qui a été fait est de ne pas présenter de local nodal.

Dans les locaux techniques, il est fréquent de disposer de baies de brassage. Ces baies permettent de regrouper les câbles et de faciliter les recâblages. Une baie de brassage comporte plusieurs fermes de brassage. Il existe deux types de fermes :

– les fermes de distribution,
– les fermes de ressources.

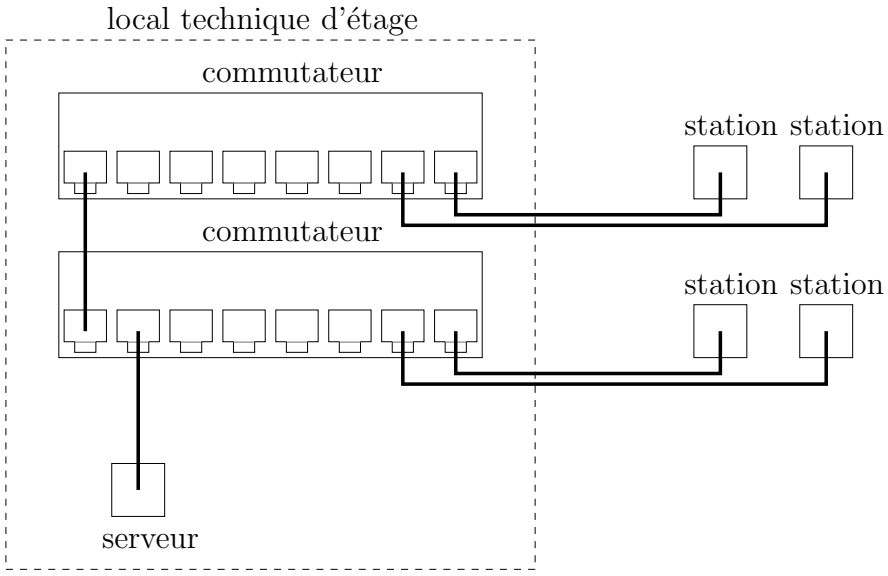

FIG. 2.6 – Réseau d'un étage.

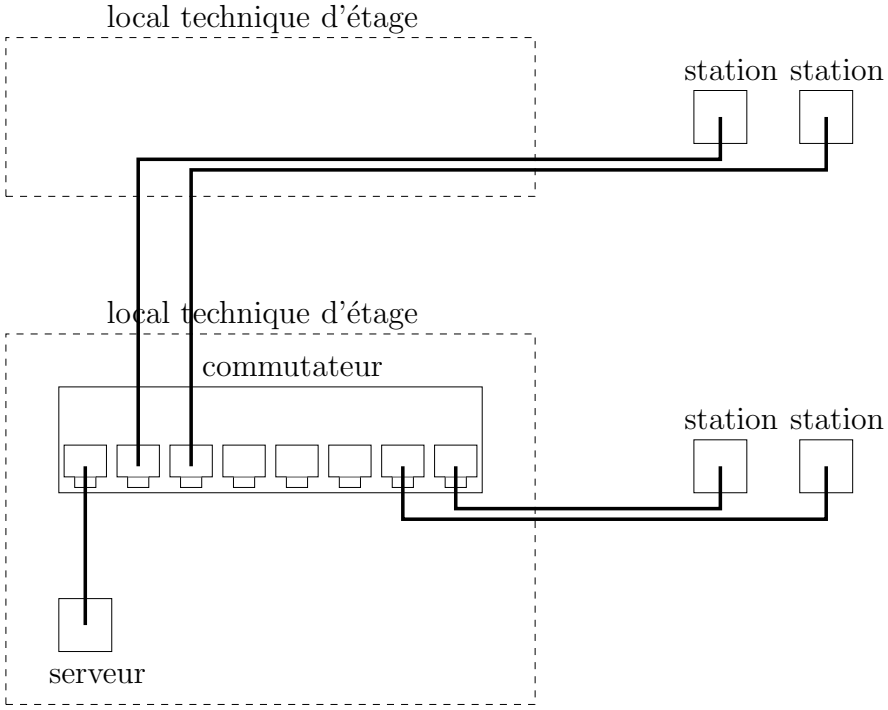

FIG. 2.7 – Réseau majoritairement d'un étage.

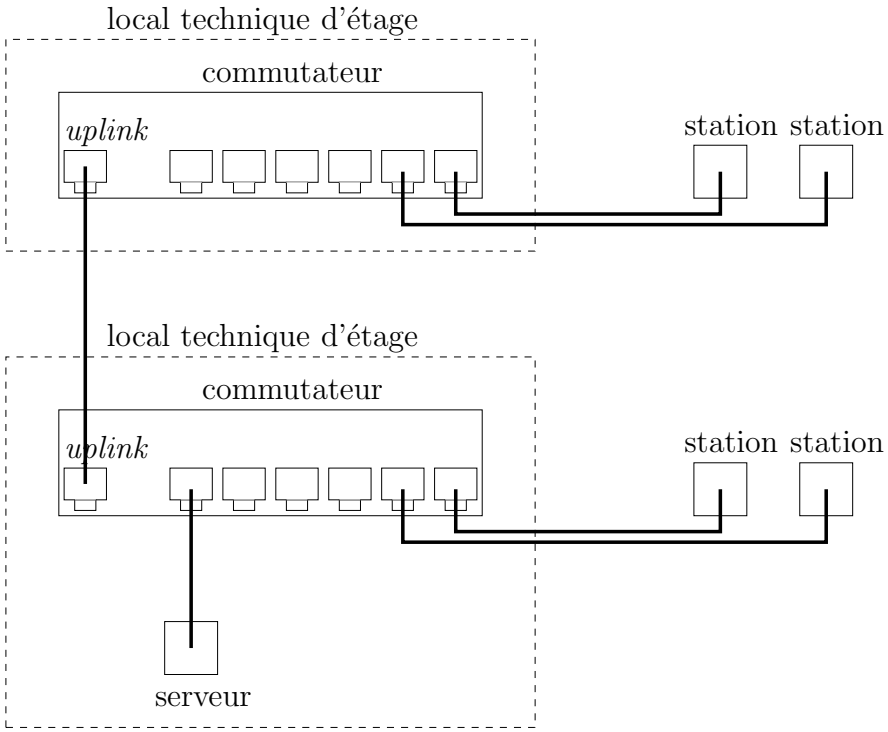

FIG. 2.8 – Réseau de deux étages.

Les fermes de distribution regroupent les câbles partant des LTE pour aller vers les équipements des utilisateurs. Les fermes de ressources sont connectées aux équipements réseau actifs. La figure 2.9 décrit une baie de brassage de manière schématique.

Chemins de câbles

Il existe deux types de chemins de câbles :

- les chemins de câbles principaux,
- les chemins de câbles dans les bureaux.

Les chemins de câbles principaux sont horizontaux (entre LTE et stations) ou verticaux (entre locaux techniques). Ils sont placés de préférence dans des conduits métalliques, le long des couloirs, dans les faux plafonds ou dans les faux planchers. Le cheminement par façade est à éviter, puisqu'il impose le perçage systématique des murs (pour atteindre les bureaux) et qu'il est moins esthétique. La figure 2.10 présente les chemins de câbles principaux et dans les bureaux. Idéalement, les chemins de câbles évitent les zones de perturbations électromagnétiques (comme les éclairages à néons ou les zones de courant fort).

Lors de la conception de chemins de câbles, il est nécessaire de se préoccuper de plusieurs facteurs : le perçage des murs (murs porteurs, murs faisant transiter des canalisations ou des câbles, matériau), la sécurité, l'esthétique.

Les chemins de câbles dans les bureaux sont souvent réalisés dans des goulottes. Elles se trouvent à environ 20 cm du sol. La connexion se fait sur des boitiers VDI (voix, données, images) regroupant des prises RJ45 et des prises de courant.

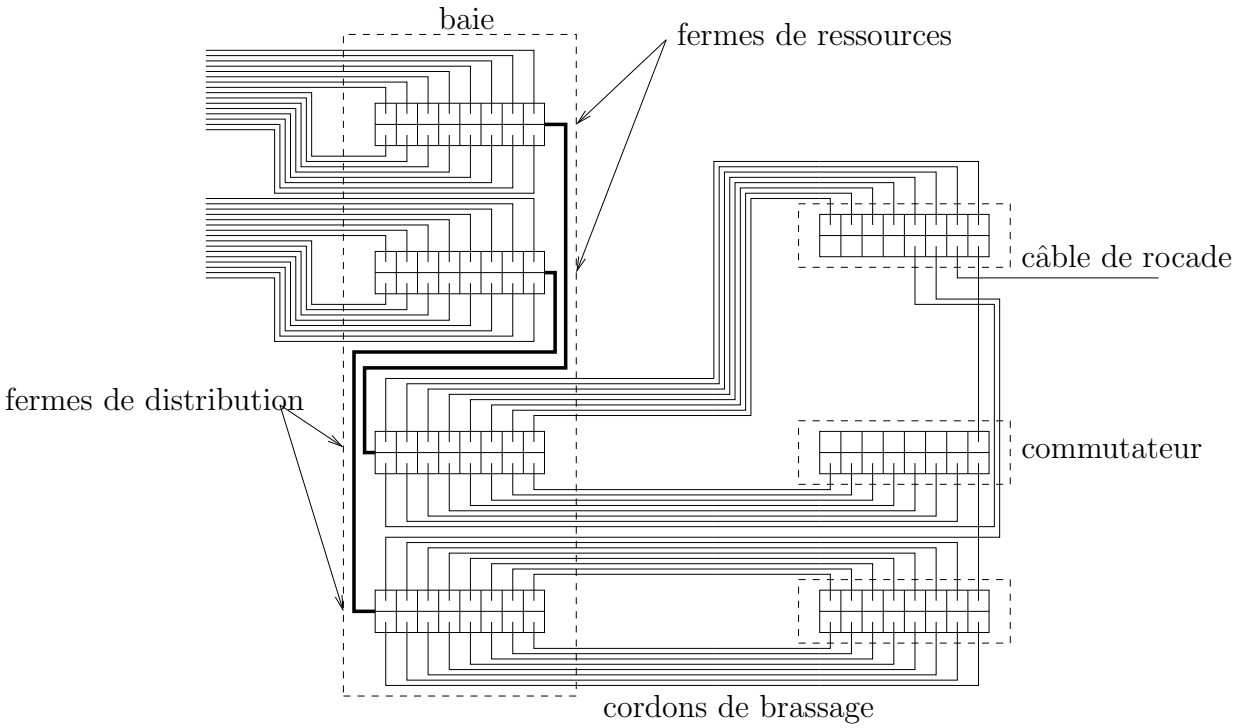

FIG. 2.9 – Les baies de brassage facilitent le câblage et le recâblage.

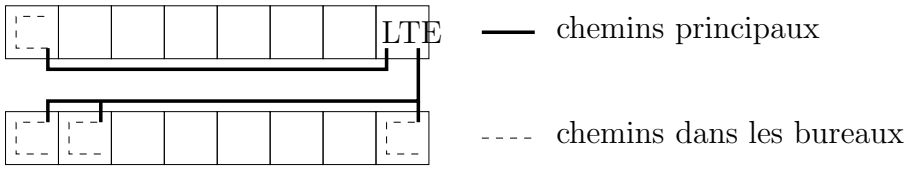

FIG. 2.10 – Les chemins de câbles.

Il est très important de nommer chaque câble, aux deux extrémités. Le nommage doit être clair et concis. En pratique, on utilise un numéro d'étage et un numéro séquentiel. Le nommage doit être fait avant le passage des câbles dans les chemins de câbles.

Courants forts et faibles

Une installation courant fort est une installation distribuant l'énergie électrique aux appareils. L'intensité des courants forts varie de quelques mA à plusieurs kA. Une installation courant faible fournit des voies de communication aux informations. L'intensité des courants faibles est en général de l'ordre de quelques μA. Les installations courants forts génèrent plus d'interférences que les installations courants faibles, et il est nécessaire de la prendre en compte le câblage et les équipements courants forts lors de la réalisation du câblage courant faible.

En plus du réseau électrique et du réseau informatique, il existe deux autres réseaux primordiaux :

- le réseau de terre,
- le réseau de masse.

Le réseau de terre a pour but de protèger les personnes contre les électrocutions. Il relie tout équipement électrique ou conducteur à la terre. Cela inclus entre autres :

- les canalisations métalliques,
- les armatures de béton,
- les chemins de câbles,
- les prises électriques.

Le réseau de masse a pour but de protéger le matériel contre les perturbations électriques, et notamment les surtensions. Il relie les équipements électroniques à la terre. Cela inclus entre autres :

- le matériel informatique,
- le blindage ou l'écran des câbles,
- les fermes de brassage.

Le réseau de masse a généralement deux topologies. La topologie utilisée anciennement était une topologie d'étoile (voir sur la partie (a) de la figure 2.11). Cependant, la topologie maillée (voir sur la partie (b) de la figure 2.11) utilisée de nos jours à l'avantage de mieux résister aux perturbations (qui augmentent avec la longueur des boucles).

Redondance

Pourquoi faire de la redondance ? Si le réseau tombe en panne, les utilisateurs ne peuvent plus travailler. Il est nécessaire de posséder un réseau redondant dans lequel la panne d'un lien (ou d'un équipement) ne remet pas en cause la continuité du service.

La redondance peut être réalisée en dupliquant les câbles, les équipements actifs (concentrateurs, commutateurs ou serveurs) ou les cartes réseaux.

La redondance pose un problème important : si aucune précaution n'est prise, des trames peuvent circuler infiniment dans le réseau, le surcharger et l'effondrer. C'est le cas des réseaux de concentrateurs ayant des boucles (ce qui est interdit par la norme IEEE) ou des trames de diffusion dans les réseaux de commutateurs.

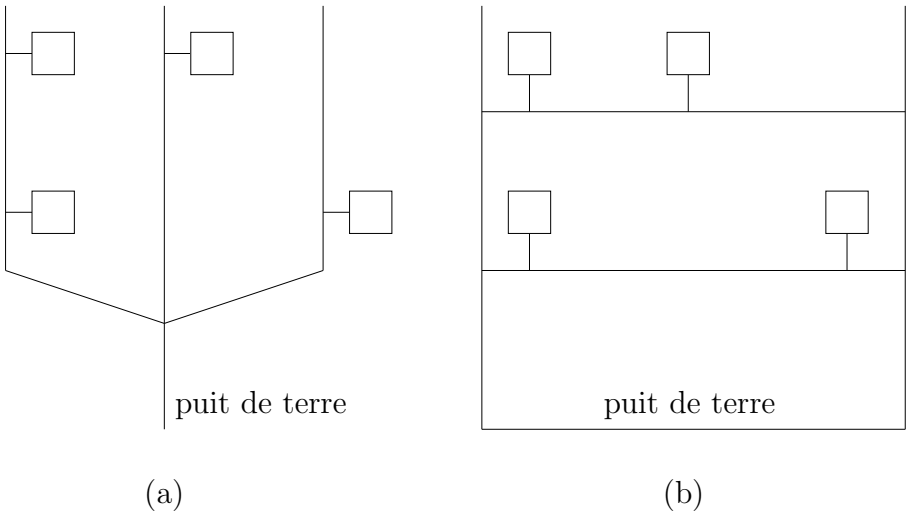

FIG. 2.11 – Le réseau de masse protège les équipements contre les surtensions. L'ancienne topologie en étoile (a) est remplacée par une topologie maillée (b).

Le protocole du *spanning tree*. Le *spanning tree* (qui signifie arbre couvrant) est un protocole qui permet de supprimer logiquement la redondance physique d'un réseau. En présence d'une panne, le réseau peut se reconfigurer afin de maintenir la connectivité. La suppression logique se fait en désactivant des ports. Ces ports ne seront plus utilisé, jusqu'à ce que le protocole les réactive (ce qui arrive lorsque la topologie change).

Le protocole du *spanning tree* est un protocole distribué, c'est-à-dire que chaque commutateur va envoyer des messages à ses voisins et prendre des décisions sans avoir une connaissance complète de la topologie. Cependant, nous allons étudier la construction de l'arbre d'un point de vue global, en ayant la connaissance complète de la topologie. Le protocole du *spanning tree* nous garantit que l'arbre construit de manière distribué et celui construit de manière global sont les mêmes. L'arbre est maintenu par le commutateur racine qui émet des trames de contrôle régulièrement afin de s'assurer que la topologie n'a pas changée.

Tout d'abord, un des commutateurs est élu comme le commutateur racine. L'arbre construit sur le réseau des ponts est un arbre des plus courts chemins. Pour départager deux commutateurs dont les identifiants sont c_1 et c_2, on compare les couples (w_1, c_1) et (w_2, c_2), w_1 et w_2 étant une priorité associée (manuellement) à c_1 et c_2. Dans la suite, nous appelons ID le couple (w, c).

Élection du commutateur racine. Le commutateur racine est celui ayant le plus petit ID.

Poids des liens. Le poids des liens peut être paramétré manuellement, mais par défaut il dépend de la capacité du lien. Un lien de 10 Mbps a un poids de 100, un lien de 100 Mbps un poids de 19 et un lien d'1 Gbps un poids de 4.

Départage. L'arbre des plus courts chemins n'est pas unique : il peut exister plusieurs plus courts chemins entre un commutateur et le commutateur racine. Dans ce cas, il faut

pouvoir départager ces multiples chemins. Les règles sont les suivantes :

- Si un commutateur possède deux plus courts chemins vers le commutateur racine, le port qui est activé est celui correspondant au chemin dont le commutateur suivant a le plus petit ID.
- Quand sur un même segment deux commutateurs permettent d'accéder au commutateur racine par un plus court chemin de même poids, c'est le commutateur ayant le plus petit ID qui est utilisé.
- Quand deux ports d'un même commutateur conduisent à un plus court chemin de même poids, c'est le port de plus petit indice qui est activé.

La figure 2.12 illustre sur un exemple l'activation des ports obtenu par le biais du protocole du *spanning tree*.

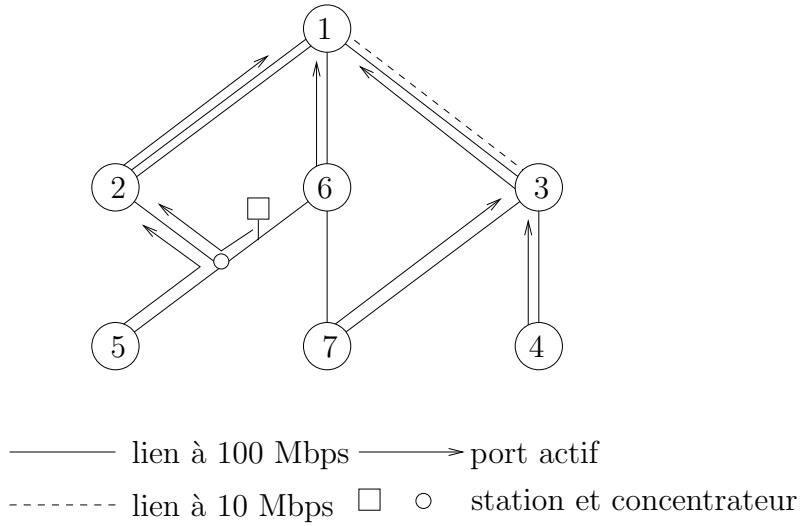

FIG. 2.12 – Le protocole du *spanning tree* désactive les liens induisant de la redondance.

Le processus de création de l'arbre peut durer de quelques dizaines de secondes à quelques minutes, en fonction du nombre de commutateurs. Si tous les commutateurs sont allumés simultanément (suite à une panne de courant par exemple), le processus de création de l'arbre en sera d'autant plus long. Pendant cette période, aucune trame n'est traitée par les commutateurs. Il y a donc une perte de service à chaque fois que la topologie change. Il existe un protocole de *spanning tree* qui converge plus rapidement, le RSTP (pour *rapid spanning tree protocol*).

2.2.4 Cahier des charges et suivi de la recette

Lorsque l'on a un projet de câblage, deux étapes sont très importantes :

- la réalisation d'un cahier des charges, préliminaire aux travaux,
- le suivi de la recette, pendant et après les travaux.

Le cahier des charges permet de spécifier les conditions du projet et de départager les soumissionnaires. Il contient plusieurs parties :

- une description de l'environnement de travail,
- une description de l'existant,
- une description de ce qui est attendu.

La description de l'environnement de travail regroupe notamment :
– les responsabilités techniques de l'entreprise réalisant les travaux,
– le calendrier de l'avancement,
– les plans hygiène, sécurité, qualité,

La description de l'existant contient le cheminement des courants faibles et forts actuels, les plans des bâtiments, etc.

La description de ce qui est attendu inclus :
– la fourniture, la pose, la configuration et le test des composants requis,
– la désignation précise, la quantité et le prix des équipements,
– la documentation que l'entreprise doit fournir (notamment le cahier de tests).

Plus la description est précise, plus on est assuré de la qualité du service qui sera rendu.

Pendant la réalisation, il est nécessaire de suivre la recette de manière régulière. Il est toujours plus facile de modifier certains points pendant les travaux, plutôt qu'après. À l'issue des travaux, il faut s'assurer que l'installation marche. Pour cela, il faut tester chaque câble et chaque prise. Le cahier de test doit donc contenir une fiche de tests par câble notamment. La recette est versée (entièrement) quand tous les tests sont passés. Tout cela est défini dans le cahier des charges.

Donnons deux définitions : le « maître d'ouvrage » est le donneur d'ordre, c'est-à-dire celui qui paye. Le « maître d'œuvre » est le responsable des travaux.

2.3 Protocole IP et architecture

Un réseau local doit garantir une connectivité IP de bout en bout. Pour cela, il faut réunir plusieurs éléments :
– un plan d'adressage définissant comment les adresses sont attribuées,
– la manière dont les paquets IP sont encapsulés dans du Ethernet,
– la manière dont les données sont acheminées,
– les services IP nécessaires à l'utilisation du réseau.

2.3.1 Encapsulation d'IP dans Ethernet

Il existe deux formats de trames Ethernet : le format Ethernet v2 et le format Ethernet 802.3. Le premier correspond à l'encapsulation du paquet IP directement dans la trame Ethernet, le second correspond à l'encapsulation du paquet IP dans un paquet LLC/SNAP, qui est lui-même encapsulé dans la trame Ethernet.

Voici le format de la trame Ethernet v2 :
– préambule : 7 octets,
– délimiteur de début de trame : 1 octet,
– adresse MAC destination : 6 octets,
– adresse MAC source : 6 octets,
– type de protocole : 2 octets,
– paquet IP : 46 à 1500 octets,
– code de contrôle d'erreur : 4 octets.

Pour un paquet IP encapsulé dans une trame Ethernet v2, le champ type de protocole vaut 0x0800 (il vaut 0x0806 pour ARP).

Voici le format de la trame Ethernet 802.3 :

– préambule : 7 octets,
– délimiteur de début de trame : 1 octet,
– adresse MAC destination : 6 octets,
– adresse MAC source : 6 octets,
– longueur : 2 octets,
– entête LLC : 3 octets,
– entête SNAP : 5 octets
– paquet IP : 38 à 1492 octets,
– code de contrôle d'erreur : 4 octets.

L'entête LLC contient le point d'accès au service destination (qui vaut 0xAA), le point d'accès au service source (qui vaut 0xAA) et le type de trame LLC qui vaut toujours 0x03.

L'entête SNAP contient un identifiant unique d'organisation sur 3 octets (qui vaut 0x00) et le type de protocole sur 2 octets. Les valeurs du type de protocole sont similaires à celles d'Ethernet v2.

Pour différencier une trame Ethernet v2 d'une trame Ethernet 802.3, on examine le champ « type de protocole / longueur ». Si la valeur est supérieure à 1500, c'est que c'est une trame Ethernet v2. En effet, la longueur des données comprises dans une trame Ethernet (que ça soit Ethernet v2 ou Ethernet 802.3) est limitée à 1500.

2.3.2 Adressage

L'adressage IP est un adressage logique. À chaque interface (et non pas à chaque machine) correspond une adresse unique globalement[1]. Comme les adresses font 32 bits, l'espace des adresses est constitué de plus de 4 millions d'adresses. Ces adresses sont découpées en cinq classes.

Les adresses contiennent une partie identifiant le réseau et une partie identifiant la machine sur le réseau. Pour savoir combien de bits identifient le réseau, on utilise le masque de sous-réseau (qui a la forme d'une adresse IP). L'identifiant du réseau est obtenu en effectuant un ET binaire entre l'adresse IP et le masque du sous-réseau.

Voici les cinq classes, les adresses et les masques de sous-réseau associées.
– Classe A : le bit le plus significatif est à `0` et le masque de sous-réseau est à `255.0.0.0` (ce qui s'écrit `/8`). Cela met à disposition 126 réseaux dont les adresses vont de `1.0.0.0` à `127.0.0.0`.
– Classe B : les bits les plus significatifs sont à `10` et le masque de sous-réseau est à `255.255.0.0` (ce qui s'écrit `/16`). Cela met à disposition – réseaux dont les adresses vont de `128.0.0.0` à `191.255.0.0`.
– Classe C : les bits les plus significatifs sont à `110` et le masque de sous-réseau est à `255.255.255.0` (ce qui s'écrit `/24`). Cela met à disposition – réseaux dont les adresses vont de `192.0.0.0` à `223.255.255.0`.
– Classe D : les bits les plus significatifs sont à `1110`. Il n'y a pas de masque de sous-réseau. Ces adresses correspondent aux adresses *multicast*. Elles vont de `224.0.0.0` à `239.255.255.255`.
– Classe E : les bits les plus significatifs sont à `11110`. Il n'y a pas de masque de sous-réseau. Ces adresses sont réservées. Elles vont de `240.0.0.0` à `255.255.255.254`.

[1]Pour le moment, nous parlons d'adresses publiques.

Le CIDR (pour *classless inter-domain routing*), utilisé de nos jours, casse le modèle à cinq classes. Il utilise des masques de sous-réseau qui ne sont pas forcément des multiples de 8 bits. Ainsi, on peut avoir un masque de sous-réseau en `255.255.255.240`, ou `/28`, si on ne dispose que d'une dizaine de machines. Les avantages du CIDR par rapport aux classes sont :

- une meilleure allocation des adresses, avec moins de perte, notamment dans le cas des réseaux de classe A,
- une meilleure agrégation des adresses, puisque toutes les adresses d'un sous-réseau ont le même préfixe.

Notons que pour tout sous-réseau, l'adresse de machine n'ayant que des `0` désigne le sous-réseau lui-même, et l'adresse de machine n'ayant que des `1` désigne toutes les machines de ce sous-réseau. Ainsi, un masque de `/28` laisse à disposition $16 - 2 = 14$ adresses de machines.

2.3.3 Acheminement

L'acheminement des paquets IP se fait selon l'algorithme du *longest common prefix match*.

Supposons la table de routage suivante :

Adresse IP	Masque de sous réseau	Interface de sortie	Prochain saut
`192.168.1.0`	`255.255.255.0`	eth0	`192.168.1.1`
`192.168.2.0`	`255.255.255.0`	eth1	`192.168.2.1`
`192.0.0.0`	`255.0.0.0`	eth1	`192.168.2.1`
`default`		eth0	`192.168.1.1`

L'adresse `192.168.1.13` correspond aux lignes 1, 3 et 4, avec des préfixes de longueurs respectives 24, 8 et 0. Le plus long préfixe est le premier. C'est lui qui est choisi. L'adresse `192.1.1.3` correspond aux lignes 3 et 4, avec des préfixes de longueurs respectives 8 et 0. C'est la ligne 3 qui est choisie pour acheminer ces paquets.

2.3.4 Routage

Il existe deux grandes classes de protocoles de routage : les protocoles à vecteur de distances (comme RIP) et les protocoles à états de liens (comme OSPF).

Protocole RIP

Le protocole RIP (pour *Routing Information Protocol*) était le protocole de routage le plus utilisé jusqu'à récemment. Dans RIP, chaque routeur émet périodiquement (toutes les 30 secondes) un paquet contenant un vecteur. Chaque élément de ce vecteur est la distance nécessaire pour atteindre un sous-réseau. Par exemple, supposons que A soit connecté aux réseaux r_1 et r_2, et que B soit connecté aux réseaux r_2 et r_3. Initialement, A avertit B que la distance (A, r_1) vaut 1 et que la distance (A, r_2) vaut 1. B avertit A que la distance (B, r_2) vaut 1 et que la distance (B, r_3) vaut 1. B peut aussi calculer que la distance (B, r_1) est de 2. En recevant ces trois messages, A peut déduire (puis envoyer à B) que la distance (A, r_3) vaut 2.

Étudions un problème du « compte jusqu'à l'infini ». Le scénario est le suivant : A est relié à B qui est relié à C. B annonce régulièrement une distance (B, A) de 1, et C

annonce régulièrement une distance (C, A) de 2. Si le lien de A à B tombe en panne, B perd sa route vers A. Cependant, il reçoit de C la distance (C, A) de 2 indiquant que C peut (ou pense pouvoir) atteindre A en deux sauts. B annonce alors une distance (B, A) valant 3. C met à jour sa distance (C, A) à 4 et l'annonce. B met à jour sa distance à 5, et ainsi de suite, jusqu'à atteindre l'infini.

Le mécanisme du *split horizon* permet d'éviter un problème qui apparaît avec les pannes. Il indique que si C possède une distance (C, A) qu'il obtient par l'intermédiaire de B (B est le saut suivant C pour atteindre A), alors C ne va pas annoncer cette distance à B. B n'en a pas besoin de toutes façons.

Il existe deux autres mécanismes permettant d'éviter (autant que possible) les boucles de routage :

- Le *split horizon with poison update*, dans lequel tous les sous-réseaux sont annoncés d'une part, et d'autre part, les sous-réseaux inaccessibles sont indiqués explicitement.
- La mise à jour déclenchée automatiquement, dans lequel le changement de topologie déclenche immédiatement la mise à jour des vecteurs de distance et leur annonce.

Les protocoles à vecteurs de distances ont plusieurs désavantages :

- Pour éviter les boucles de routage, ils utilisent une distance maximale (qui vaut 15 dans le cas de RIP). Cette distance maximale limite la taille des réseaux qui peuvent être gérés.
- Le temps de convergence, c'est-à-dire le temps nécessaire pour que les tables de routage du réseau soient sans boucles, est lent.
- En cas de panne, un mécanisme spécifique doit être mis en place pour éviter le phénomène qui s'appelle le « compte jusqu'à l'infini » et les boucles de routage,
- la qualité des liens (incluant leur bande passante) est dure à prendre en compte.

Protocole OSPF

Le protocole OSPF (pour *Open Shortest Path First*) est un protocole à états de liens. Il fait partie des protocoles qui remplacent RIP. Dans un tel protocole, chaque routeur doit être en mesure de savoir si un lien est en état de fonctionnement ou pas. Pour cela, les routeurs s'échangent des paquets `Hello`. Ensuite, chaque routeur diffuse sur le réseau l'état de tous ses liens. Chaque routeur connaît donc la topologie complète et est en mesure de calculer des chemins tenant compte de la bande passante par exemple. Le maintien de la topologie n'est pas effectué de manière périodique, mais se déclenche lorsqu'un changement de topologie a lieu. Les états de liens sont échangés par inondation complète du réseau. S'il n'y a pas de changement de topologie, le protocole OSPF ne devrait pas générer beaucoup de traffic (à part les paquets `Hello` qui permettent de savoir si un lien est tombé en panne ou pas).

Pour soulager le réseau, la distribution de la topologie se fait par l'intermédiaire d'un routeur désigné. Ce routeur devient un point critique.

2.3.5 Services IP

Il existe de nombreux services et protocoles qui permettent de s'assurer du bon fonctionnement d'un réseau et de l'administrer. Nous allons en détailler quelques uns.

Outil `ping`

L'outil `ping` (pour *packet Internet groper*, le tâtonneur de paquets Internet) permet(tait) de s'assurer de la connectivité d'une machine à un réseau.

Le `ping` fait l'hypothèse que chaque machine dispose d'un serveur `echo`. À la réception d'un paquet ICMP de type `echo request`, le serveur `echo` répond un paquet ICMP de type `echo reply`. Si l'utilisatrice reçoit l'`echo reply`, c'est que la machine est bien configurée. Dans ce cas, l'utilisatrice peut calculer le temps nécessaire aux paquets pour faire l'aller-retour. Ci-dessous, un exemple décrivant le résultat de la commande `ping` sur l'adresse IP `193.49.118.101` depuis une adresse IP du même sous-réseau. La commande `arp -d` supprime du cache ARP l'adresse donnée en paramètres.

```
alexandre@alex:~$ sudo arp -d 193.49.118.101
alexandre@alex:~$ ping 193.49.118.101
PING 193.49.118.101 (193.49.118.101) 56(84) bytes of data.
64 bytes from 193.49.118.101: icmp_seq=1 ttl=64 time=2.95 ms
64 bytes from 193.49.118.101: icmp_seq=2 ttl=64 time=0.172 ms
64 bytes from 193.49.118.101: icmp_seq=3 ttl=64 time=0.174 ms
64 bytes from 193.49.118.101: icmp_seq=4 ttl=64 time=0.195 ms

--- 193.49.118.101 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 3002ms
rtt min/avg/max/mdev = 0.172/0.874/2.958/1.203 ms
```

Avec la multiplication des *firewalls*, il est de plus en plus rare de pouvoir `ping`er une machine, même si elle est bien configurée.

Outil `traceroute`

L'outil `traceroute` permet d'afficher la route prise par des paquets IP. Il fait l'hypothèse que la route prise par les paquets ne varie pas pendant les quelques secondes pendant lesquelles l'outil est lancé.

L'outil `traceroute` envoie une succession de paquets avec des durées de vie (TTL pour *time to live*) de plus en plus grande. Chaque paquet va arriver à expiration de sa durée de vie avant sa destination. Dans ce cas, le routeur qui détruit le paquet envoie à la source un message ICMP de type `time exceeded`. Seul le dernier paquet, celui qui a une durée de vie suffisamment grande, va être reçu par la machine destination. Ce paquet est envoyé sur un port UDP très grand dans l'espoir que le port ne correspondra à aucun serveur. La machine destination va répondre avec un message ICMP de type `port unreachable` et le programme `traceroute` pourra s'arrêter. La route prise par les paquets à l'aller peut être reconstruite en récupérant et ordonnançant les différents messages `time exceeded` reçus[2]. Un exemple d'exécution de `traceroute` donne :

```
alexandre@alex~: traceroute 209.85.129.99
traceroute to 209.85.129.99, 30 hops max, 40 byte packets
```

[2]Ils peuvent être ordonnancés car le paquet ICMP `time exceeded` contient l'entête IP complet du paquet détruit, ainsi que les huit premiers octets de données.

```
 1  193.49.118.1  0.599 ms  3.243 ms  0.628 ms
 2  195.221.123.252  0.299 ms  0.325 ms  0.317 ms
 3  193.51.186.54  0.614 ms  0.498 ms  0.641 ms
 4  193.51.179.233  8.963 ms  8.770 ms  8.736 ms
 5  193.51.179.13  8.809 ms  8.711 ms  8.617 ms
 6  194.68.129.242  9.199 ms  9.111 ms  9.245 ms
 7  80.231.79.14  9.075 ms  9.319 ms  9.106 ms
 8  80.231.73.5  9.654 ms  9.674 ms  9.695 ms
 9  195.219.228.25  22.312 ms  22.066 ms  22.110 ms
10  195.219.228.2  112.618 ms  46.637 ms  199.751 ms
11  80.231.81.9  22.329 ms  22.062 ms  22.274 ms
12  80.231.82.4  22.152 ms  22.022 ms  22.273 ms
13  80.231.82.146  24.555 ms  24.525 ms  24.364 ms
14  216.239.43.89  24.656 ms  24.753 ms
    72.14.238.118  27.448 ms
15  72.14.232.141  27.223 ms  32.320 ms  27.314 ms
16  72.14.232.243  28.490 ms
    72.14.233.206  32.785 ms
    72.14.232.243  29.283 ms
17  72.14.238.120  28.789 ms
    72.14.238.122  28.714 ms
    72.14.238.120  28.456 ms
18  72.14.232.167  29.060 ms
    209.85.129.99  27.898 ms
    72.14.232.167  29.535 ms
```

On peut se demander pourquoi l'outil `traceroute` n'utilise pas l'option IP d'enregistrement de la route. Il y a deux raisons principales :

- L'option d'enregistrement de la route ne peut stocker qu'un faible nombre d'adresses (9).
- L'option d'enregistrement de la route n'est pas utilisée dans tous les routeurs.

Une différence entre `traceroute` et l'option IP d'enregistrement de la source est que `traceroute` enregistre l'adresse IP de l'interface d'entrée du paquet, et non pas l'adresse IP de l'interface de sortiei du paquet.

Attention, `traceroute` ne suppose pas que les routes sont symétriques, mais seule la route à l'aller est affichée. Pour donner une définition de `traceroute` plus précise, on pourrait dire que c'est un outil permettant d'afficher une route probable pour l'aller d'une source à une destination, à un instant donné.

Protocole DNS

Le DNS (pour *domain name system*) est un annuaire mondial organisé comme une base de données distribuée. Il sert à traduire des noms en adresse IP. Lorsqu'une utilisatrice tape `www.google.com`, l'application transforme grâce à DNS ce nom en adresse IP utilisable.

L'architecture de DNS est constituée de résolveurs (les clients), de serveurs caches et de serveurs de noms. Les serveurs de noms forment une forêt. En haut de chaque arbre,

on trouve les serveurs racines. Une zone DNS est un sous-arbre de l'arbre (en fait de la forêt) DNS administré par une même autorité. Chaque zone possède un serveur primaire et un ou plusieurs serveurs secondaires, autoritaires sur la zone.

L'algorithme de résolution d'un nom est le suivant. Le résolveur émet une requête vers le serveur primaire de sa zone. Si le serveur primaire de la zone connaît l'adresse correspondant au nom, il répond au résolveur. Dans le cas contraire, il interroge un serveur racine. Le serveur racine retourne l'adresse du serveur de deuxième niveau. Le serveur primaire interroge ce serveur suivant, qui retourne soit l'adresse directement, soit l'adresse d'un autre serveur de noms. Ce processus continue jusqu'à ce que l'adresse soit obtenue (ou qu'un serveur puisse garantir que l'adresse n'existe pas).

Les enregistrements DNS s'appellent des RR (pour *resource records*). Il existe plusieurs types de RR, entre autres :

- le type `A` qui stocke des adresses IP,
- le type `CNAME` qui stocke des alias,
- le type `MX` qui indique le nom du serveur mail,
- le type `NS` qui indique l'adresse d'un serveur de noms,
- le type `PTR` qui retourne des noms de domaines.

Il existe des outils interactifs d'interrogation de la base de données DNS, comme l'outil `dig`. Voici deux extraits de son utilisation.

```
alexandre@alex:~$ dig google.com
;; QUESTION SECTION:
;google.com.             IN A
;; ANSWER SECTION:
google.com.       196    IN A     64.233.187.99
google.com.       196    IN A     72.14.207.99
google.com.       196    IN A     64.233.167.99
;; AUTHORITY SECTION:
google.com.       155505 IN NS    ns3.google.com.
google.com.       155505 IN NS    ns4.google.com.
google.com.       155505 IN NS    ns1.google.com.
google.com.       155505 IN NS    ns2.google.com.
;; ADDITIONAL SECTION:
ns1.google.com.   77085  IN A     216.239.32.10
ns2.google.com.   77085  IN A     216.239.34.10
ns3.google.com.   79099  IN A     216.239.36.10
ns4.google.com.   81515  IN A     216.239.38.10

alexandre@alex:~$ dig www.google.com
;; QUESTION SECTION:
;www.google.com.         IN A
;; ANSWER SECTION:
www.google.com.   538263 IN CNAME www.l.google.com.
www.l.google.com. 230    IN A     209.85.129.147
www.l.google.com. 230    IN A     209.85.129.99
www.l.google.com. 230    IN A     209.85.129.104
;; AUTHORITY SECTION:
```

```
l.google.com.     80003  IN NS    a.l.google.com.
l.google.com.     80003  IN NS    b.l.google.com.
;; ADDITIONAL SECTION:
a.l.google.com.   6051   IN A     209.85.139.9
b.l.google.com.   80044  IN A     64.233.179.9
g.l.google.com.   6038   IN A     64.233.167.9
```

Protocole DHCP

Le protocole DHCP (pour *dynamic host configuration protocol*) permet d'attribuer de manière dynamique des adresses IP à des machines clientes, et de leur communiquer des paramètres comme :

- le masque de sous réseau,
- la passerelle par défaut,
- l'adresse des serveurs DNS,
- l'adresse d'autres serveurs (comme le serveur de temps ou le serveur d'impression).

Dans le mode dynamique, lorsqu'un serveur DHCP attribue une adresse IP, il spécifie aussi une durée de contrat. Si le client désire utiliser l'adresse IP pendant plus longtemps, il doit faire une demande de renouvellement du contrat.

Dans le mode automatique, un client obtient toujours la même adresse IP. Dans le mode manuel, c'est le client (et non pas le serveur) qui spécifie son adresse IP. Le protocole DHCP sert alors à informer le serveur DHCP des adresses en cours d'utilisation.

Il n'est pas nécessaire de connaître l'adresse d'un serveur DHCP pour démarrer. Le client doit simplement émettre un paquet de diffusion qui est traité par le serveur DHCP. La réponse du serveur est aussi envoyée en diffusion, mais elle contient l'adresse MAC du client afin qu'il puisse identifier que le message lui est adressé. Le client répond (par diffusion) qu'il désire prendre cette adresse, et le serveur acquitte ce choix, à nouveau par diffusion. À ce moment, le client peut commencer à utiliser l'adresse que le serveur DHCP lui a attribuée.

Si l'hôte et le serveur DHCP ne sont pas sur le même réseau, on peut utiliser un relais DHCP.

Mécanismes NAT et PAT

Le NAT (pour *network address translation*) est un mécanisme permettant de faire correspondre des adresses IP privées et des adresses IP publiques.

Lorsque le nombre d'adresses publiques disponibles est égal au nombre d'adresses privées, la translation d'adresses peut être bijective : chaque adresse IP privée se voit attribuer une adresse IP publique. Le routeur réalisant le NAT conserve une table de correspondance et modifie à la volée :

- les adresses IP des paquets sortants du réseau interne (pour leur donner l'adresse IP publique),
- les adresses IP des paquets entrants dans le réseau interne (pour leur donner l'adresse IP privée),
- les sommes de contrôle IP, UDP et TCP (entre autres).

Les intérêts du NAT dans ce cas sont :

- la possibilité d'avoir un adressage public stable (même si l'adressage privé change),

– la possibilité de cacher vis-à-vis de l'extérieur la structure du réseau privée (notamment par la technique du *masquerading*, où l'adresse IP du routeur est utilisée comme adresse extérieure).

Toutefois, le mécanisme NAT est surtout utilisé en combinaison avec le PAT (pour *port address translation*)[3]. Lorsque le nombre d'adresses publiques disponibles est plus petit que le nombre d'adresses privées, il n'est pas possible d'associer à chaque adresse IP privée une adresse IP publique. On veut toutefois que toutes les machines privées puissent accéder au réseau externe simultanément. La solution du PAT est de faire correspondre à un couple (adresse privée, port interne) un couple (adresse publique, port). La translation d'adresses ne suffit plus, il faut aussi translater les ports pour identifier à laquelle des machines privées correspond l'adresse publique.

Prenons un exemple : une utilisatrice de la machine interne d'IP `10.0.0.1` désire accéder à un serveur HTTP d'une machine externe d'IP `1.2.3.4`. Elle envoie donc une requête TCP (un paquet `SYN`) concernant le quadruplet $(10.0.0.1, 1039, 1.2.3.4, 80)$, 1039 étant un port attribué aléatoirement par la machine de l'utilisatrice. Le routeur réalisant le PAT voit ce paquet et réalise la translation de l'adresse 10.0.0.1 en une autre adresse, par exemple 5.6.7.8. De plus, il réalise la translation du port 1039 en 1402. Puis, il achemine le paquet à la machine $(1, 2, 3, 4)$. Le serveur HTTP reçoit une demande de connexion concernant le quadruplet $(5.6.7.8, 1402, 5.6.7.8, 80)$. Il accepte la connexion (en renvoyant un `SYN ACK`). En recevant cette réponse, le routeur réalisant le PAT tente de retransformer l'adresse `5.6.7.8`. Comme plusieurs machines internes peuvent avoir cette adresse, le routeur ne sait pas par quelle adresse IP il faut remplacer `5.6.7.8`. Cependant, une seule machine correspond au port `1402`, et c'est la machine `10.0.0.1`. Le routeur trouve aussi dans sa table l'ancien port (qui avait été écrasé) et remplace le `1402` par `1039`. La connexion peut être établie.

Faisons quelques remarques :

– Lorsque le NAT est utilisé, la table de correspondance ne fait apparaître que les adresses privées et les adresses publiques.
– Lorsque le NAT et le PAT sont utilisés conjointement, la table de correspondance fait apparaître les couples (adresse privée, port interne) et (adresse publique, port).
– La correspondance est établie à l'initiative d'une machine interne. En effet, si un paquet arrive à la machine faisant le PAT sur un port inconnu, elle ne sait pas à quelle machine interne ce paquet doit être adressé. Le paquet est alors ignoré. Cela pose des problèmes à certains protocoles initiés par les machines externes (comme FTP).

En résumé, le PAT (en combinaison avec le NAT) permet de résoudre le problème de la limitation des adresses IP, mais introduit quelques difficultés au niveau de certains protocoles, comme FTP. Les difficultés viennent du fait que les connexions doivent être initiées par le client d'une part, et que les données FTP peuvent contenir des informations sur les adresses IP (qui sont translatées par le PAT).

Attention à ne pas confondre NAT et DHCP. Les deux peuvent réaliser de l'allocation dynamique d'adresses, mais DHCP alloue des adresses IP à des machines n'en ayant pas *a priori*, tandis que NAT alloue de nouvelles adresses IP (publiques) à des machines ayant déjà des adresses IP (privées). Le mécanisme NAT met en place une table de

[3]Lorsque nous parlerons de PAT, il s'agira toujours de la combinaison de NAT et de PAT. Il existe du PAT seul, sans NAT, mais nous n'en discuterons pas ici.

correspondance, alors que le DHCP n'en utilise pas.

À titre informatif, les adresses IP privées appartiennent à :

- `10.0.0.0` avec pour masque `255.0.0.0`,
- `172.16.0.0` avec pour masque `255.240.0.0`,
- `192.168.0.0` avec pour masque `255.255.0.0`.

Architecture SNMP

L'architecture SNMP permet de gérer les ressources réseau. Elle est constituée de trois éléments :

- La MIB (pour *management information base*) définit une base de données d'informations de gestion.
- La SMI (pour *structure of management information*) définit la structure de la MIB et la syntaxe des éléments enregistrés.
- Le protocole SNMP permet d'accéder aux champs de la MIB des différents éléments du réseau.

Les champs de la MIB correspondent aux valeurs gérées (tant logicielles que matérielles). Ces variables peuvent être récupérées ou modifiées. On peut trouver par exemple la durée des temporisateurs des différents protocoles ou des statistiques d'utilisation.

La SMI définit une hiérarchie de variables. Par exemple, le premier niveau de hiérarchie contient les champs `IP`, `ICMP`, `TCP`, `system`,`interfaces`, et ainsi de suite.

Il existe une station de gestion, le NMS (pour *network management system*), listant dans sa MIB toutes les ressources réseaux. Chaque ressource réseau possède un agent logiciel SNMP et une base de ressources MIB pouvant être accédée.

2.3.6 IPv6

IPv6 (encore appelé IPng pour *IP next generation*) est la nouvelle version d'IP[4]. Elle est destinée à résoudre plusieurs problèmes d'IPv4, notamment :

- l'épuisement des adresses IPv4,
- l'explosion des tables de routage,
- l'exploitation des mêmes adresses de bout en bout.

Le CIDR dans IPv4 est insuffisant pour empêcher l'explosion des tables de routage : d'une part, les adresses IPv4 sont trop petites pour que le routage hiérarchique soit efficace et d'autre part, il y a beaucoup de plages d'adresses déjà allouées de manière non hiérarchique.

De plus, le NAT déroge à la règle du bout en bout. Comme nous l'avons vu, cela pose des problèmes à certaines applications (et notamment à certaines applications de sécurité).

[4]La version du protocole IP est passée directement de 4 à 6 car la valeur 5 correspond à un autre protocole, le protocole *Internet Stream Protocol* (ST). En comparaison à IP, ST offre un mode orienté connexion et une gestion de qualité de service

Champ	Taille	Rôle
version	4	indique que la version d'IP est 4
longueur de l'entête	4	donne la longueur de l'entête IP (options comprises)
type de service	8	spécifie le type de service
longueur	16	spécifie la longueur du paquet
identifiant	16	identifie les fragments d'un même paquet
drapeaux	3	informations de fragmentation
décalage dans le fragment	13	indique la position du fragment
TTL	8	durée de vie du paquet
protocole	8	indique le protocole encapsulé
somme de contrôle	16	somme de contrôle de l'entête
source	32	adresse source du paquet
destination	32	adresse destination du paquet
options	?	options du paquet

TAB. 2.1 – Le format de l'entête IPv4.

Les opposants à IPv6 mentionnent souvent l'espace d'adresses qui est démesuré[5]. Il est cependant probable qu'IPv6 tende à remplacer petit à petit IPv4.

Fonctionnalités d'IPv6

IPv6 ajoute de nouvelles fonctionnalités par rapport à IPv4 :
- l'autoconfiguration des machines sans état,
- des adresses locales pour les liens,
- les jumbogrammes (qui sont des paquets dont la taille pouvant aller jusqu'à 4 Go),
- pas de fragmentation des paquets, et plus de somme de contrôle.

Différences entre IPv4 et IPv6

Les entêtes IPv4 et IPv6 sont différents. Le tableau 2.1 rappelle l'entête IPv4 et le tableau 2.2 donne l'entête IPv6.

Voici quelques autres différences entre IPv4 et IPv6 :
- Au niveau du type d'adresses, IPv4 propose deux types d'adresses : les adresses *unicast* et les adresses *multicast*. IPv6 ajoute les adresses *anycast*[6].
- Au niveau du plan d'adressage, IPv4 offre deux plans d'adressage : le plan d'adressage en classes et le plan d'adressage CIDR. IPv6 permet de gérer de nouveaux plans d'adressages : un plan d'adressage géographique (abandonné), un plan d'adressage dépendant du fournisseur (lui aussi abandonné) ou encore un plan d'adressage hiérarchique.

[5]Avec une allocation d'adresses denses, IPv6 permettrait d'offrir environ 5.10^{28} adresses pour chacun des 6.5 milliards d'individus de la planète. Dit autrement, IPv6 offre environ mille milliards d'adresses par centimètre carré de surface terrestre (océans compris). Cependant, l'idée derrière la taille de l'espace d'adressage de 128 bits est plutôt de simplifier les tables d'acheminement.

[6]Un message envoyé à un groupe *anycast* doit être reçu par un seul élément du groupe, et non pas par tous les éléments du groupe comme c'est le cas en *multicast*.

Champ	Taille	Rôle
version	4	indique que la version d'IP est 6
classe	8	spécifie la classe du service
identificateur	20	identifie le flux (depuis la source)
longueur	16	spécifie la longueur de la partie utile
entête suivant	8	indique l'endroit où se trouve l'entête suivant
nombre de sauts	8	indique la durée de vie du paquet
source	128	adresse source du paquet
destination	128	adresse destination du paquet

TAB. 2.2 – Le format de l'entête IPv6.

– Plusieurs mécanismes optionnels et additionnels dans IPv4 ont été intégré dans IPv6 de base, comme notamment : le *multicast*, la sécurité (avec IPsec) ou encore la mobilité.

Adresses et adressage IPv6

Les adresses IPv6 font 128 bits. Elles sont représentées par une série de nombres hexadécimaux de quatre chiffres séparés par des `:`. Une adresse IPv6 valide est `0123:4567:89AB:CDEF:01`
Une suite de `0:` consécutifs peut être remplacée par un `::`, une seule fois seulement. Ainsi, `FE80:0:0:0:0:0:0:1` peut être écrit `FE80::1`. Similairement, `0:0:0:0:0:0:0:1` s'écrit aussi `::1` et `0:0:0:0:0:0:0:0` s'écrit aussi `::`.

Lorsque les adresses IP sont représentées de manière littérales, par exemple lorsque l'on représente une URL, la notation avec des `:` peut être confondue avec la représentation des numéros de ports. On utilise donc des crochets, comme dans l'exemple suivant : `http://[::1]/`.

Le plan d'adressage agrégé définit trois niveaux de hiérarchie : une partie de l'adresse est réservée pour l'identification publique (sur 48 bits), une partie pour l'identification du site (sur 16 bits) et une partie pour l'interface (sur 64 bits).

Il existe quelques adresses particulières. L'adresse `::` correspond à l'adresse indéterminée (lorsqu'une machine se connecte sur le réseau). L'adresse `::1` correspond au bouclage (le *loopback*).

Dans le DNS, les adresses IPv6 sont indiquées par des champs `AAAA` (les adresses IPv4 sont dénotées par le champ `A`).

Transition et cohabitation

Il existe plusieurs moyens de faire cohabiter IPv4 et IPv6, et de faire la transition de l'un à l'autre. Ainsi, des machines IPv6 devraient pouvoir accéder aux services IPv4 et aux services IPv6.

Une étape dans la cohabitation entre IPv4 et IPv6 est au niveau de la station. À ce niveau, la cohabitation est souvent réalisée par une pile duale (ou *dual-stack*). Une pile duale implémente à la fois IPv4 et IPv6.

Pour faire transiter des paquets IPv6 au dessus d'un réseau IPv4, on utilise souvent des tunnels 6to4. Une manière d'encapsuler des paquets IPv6 dans des paquets IPv4 est d'utiliser des adresses compatibles IPv4. Une adresse compatible IPv4 est une adresse

IPv6 de la forme `::wwxx:yyzz`. Des paquets IPv6 ayant des adresses compatibles IPv4 sont encapsulées dans des paquets IPv4. Ces paquets sont décapsulés à la sortie du réseau IPv4.

Une machine IPv6 dialoguant avec une machine IPv4 peut utiliser des adresses IPv4 mappées. Une adresse IPv4 mappée est une adresse IPv6 représentant une adresse IPv4. Elles s'écrivent sous la forme `::FFFF:wwxx:yyzz` ou `::FFFF:a.b.c.d` (où `a.b.c.d` est l'adresse IP à points, `ww` est la valeur hexadécimale de `a`, `xx` celle de `b`, `yy` celle de `c`, `zz` celle de `d`. L'application IPv6 détecte que l'adresse est une adresse IPv4 mappée et envoie les messages à la couche IPv4. Lorsque la couche IPv4 de la machine IPv6 reçoit des paquets IPv4 pour l'application IPv6, les adresses sont remappées en adresses IPv4 mappées. Des adresses IPv4 mappées ne circulent jamais sur le réseau mais sont utilisées par les applications IPv6.

Historique d'IPv6

Voici quelques dates concernant IPv6 :
- 1994 : naissance du protocole SIPP en remplacement à IPv4,
- 1995 : SIPP est rebaptisé IPv6,
- 1996 : Renater est l'un des premiers réseaux français à utiliser IPv6,
- 1996 : Linux implémente une pile IPv6 en version alpha (dans le noyau `2.1.8`),
- 1997 : AIX d'IBM est la première plateforme commerciale à implémenter IPv6,
- 1998 : Microsoft Research implémente une pile IPv6 expérimentale,
- 2000 : IPv6 est disponible sur les plateformes BSD,
- 2002 : Windows XP SP1 et Windows Server 2003 disposent d'une pile IPv6,
- 2003 : MAC OS X v10.3 supporte IPv6, qui est activé par défaut,
- 2004 : la pile IPv6 de Linux est approuvée (dans le noyau `2.6.10`),
- 2004 : les enregistrements IPv6 indiquant les serveurs de noms de quelques pays (dont la France) apparaissent dans les serveurs DNS racines,
- 2007 : Windows Vista supporte IPv6, qui est activé par défaut.

Chapitre 3

Réseaux métropolitains

Les réseaux métropolitains sont des réseaux dont l'étendue géographique va d'un ensemble de bâtiments à une ville. On les nomme MAN (pour *metropolitan area network*).

Il existe plusieurs supports de transmissions différents pour les MAN, offrant des débits différents. La mise en place d'une interconnexion au moyen de chacun de ces supports nécessite l'installation d'un CPE (pour *customer premises equipement*) et la négociation d'un contrat avec l'opérateur.

3.1 Faibles débits : PPP sur RTC

Lorsque l'interconnexion ne nécessite pas des débits importants, on utilise le protocole PPP sur le support RTC. Nous allons décrire ce qu'est PPP et ce qu'est RTC.

3.1.1 PPP

Le protocole PPP (pour *point-to-point protocol*) est un protocole de la couche liaison de données fonctionnant sur des liaisons point à point. Il est le successeur de HDLC. Il autorise des débits allant jusqu'à 2 Mbit/s. Il utilise trois mécanismes :

- l'encapsulation des données,
- le contrôle de la couche 2 avec le protocole LCP,
- le contrôle de la couche 3 avec un protocole NCP.

L'encapsulation PPP sert à connaître le type de chaque paquet (qui peut être un LCP, un NCP ou données) et à vérifier qu'il n'y a pas eu d'erreur de transmission au moyen d'une somme de contrôle.

Le protocole LCP (pour *link control protocol*) permet une configuration automatique des deux entités aux extrémités de la liaison point à point. Les paramètres configurés incluent la taille maximale des datagrammes échangés sur la liaison. LCP fournit un mécanisme d'authentification, réalisé par le protocole PAP (pour *password authentication protocol*, majoritairement utilisé) ou par le protocole CHAP (pour *challenge-handshake authentication protocol*, plus récent). Une fonctionnalité de LCP est de détecter les liaisons bouclées (par l'intermédiaire d'un numéro magique choisi par chaque extrémité de la liaison).

Finalement, un protocole NCP (pour *network control protocol*) existe par protocole de couche 3 fonctionnant au-dessus de PPP. Pour IP, il s'agit du protocole IPCP (pour

IP control protocol).

3.1.2 RTC

Le RTC (pour *réseau téléphonique commuté*) ou PSTN (pour *public switched telephone network*) désigne le réseau téléphonique analogique[1]. Le RTC est basé sur une boucle sur une paire torsadée entre un commutateur et l'abonné, et sur une architecture arborescente à commutation de circuits. Avec le RTC, les débits vont jusqu'à 64 Kbit/s.

L'architecture RTC est présentée sur la figure 3.1. Les prises sont les prises téléphoniques classiques.

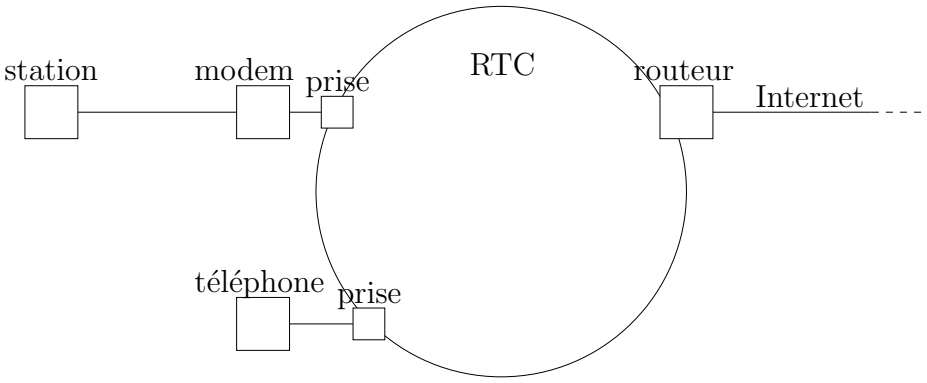

FIG. 3.1 – L'architecture RTC.

3.2 Moyens débits : PPP sur RNIS

Les interconnexions nécessitant des débits moyens est souvent réalisée par l'intermédiaire du support RNIS.

Le RNIS (pour *réseau numérique à intégration de service*) ou ISDN (pour *integrated services digital network*) désigne le réseau téléphonique numérique[2]. Ce protocole est basé sur une architecture à commutation de circuits. RNIS offre des débits allant de 128 Kbit/s à 2 Mbit/s.

Le RNIS offre trois types de canaux : un canal de synchronisation, des canaux de données et un canal de signalisation. Le canal est un bus sur lequel jusqu'à cinq terminaux peuvent se connecter. L'accès au canal se fait au moyen du protocole CSMA-CR (pour *carrier sense multiple access - contention resolution*, que nous ne détaillerons pas).

Selon le type d'accès RNIS choisi, on utilise deux topologies. L'accès de base, correspondant à un plus faible débit, est connecté à un TNR (pour terminal numérique de réseau) qui est connecté au RNIS. L'accès primaire, correspondant à un débit plus important, est connecté à un TNA (pour terminal numérique d'abonné), lui-même connecté à un TNL (pour terminal numérique de ligne). TNR et TNL sont connectés à l'opérateur téléphonique. L'architecture est présentée sur la figure 3.2.

[1]Il fut créé par Alexander Graham BELL pour faire écouter des pièces de théâtre à distance.
[2]Le réseau RNIS de France Télécom est nommé Numéris.

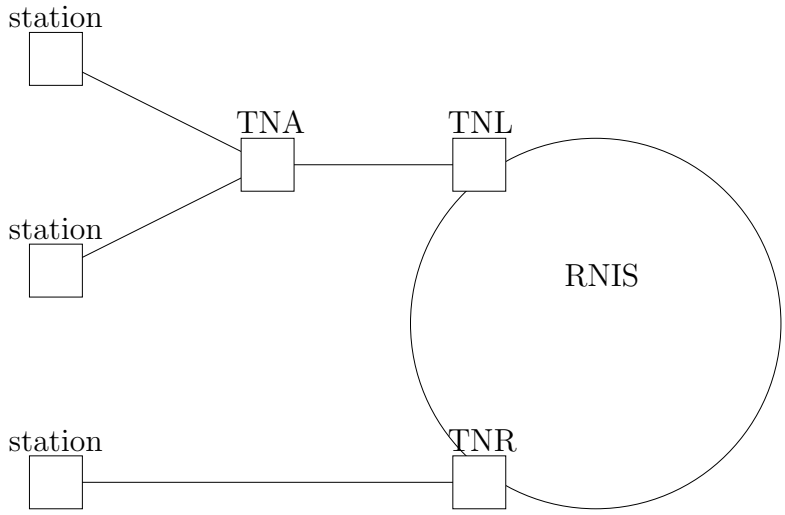

FIG. 3.2 – L'architecture RNIS.

3.3 Hauts débits

Pour faire de l'interconnexion à haut débit, il y a plusieurs possibilités : on peut utiliser des lignes spécialisées, ou les technologies ADSL et HDSL.

3.3.1 PPP sur ligne spécialisée

Les lignes spécialisées sont des liaisons point à point. Comme pour RNIS, elles se basent sur une agrégation de canaux. Les débits que peuvent atteindre les lignes spécialisées peuvent atteindre 34 Mbit/s.

La mise en place d'une ligne spécialisée nécessite l'installation d'un équipement appelé CSU (pour *channel service unit*), comme indiqué sur la figure 3.3.

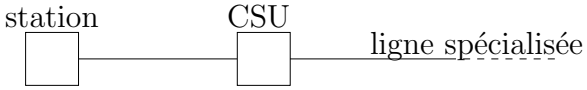

FIG. 3.3 – L'architecture d'une ligne spécialisée.

3.3.2 PPP sur ADSL

L'ADSL (pour *asymetric digital subscriber line*) est une technologie permettant d'offrir des débits asymétriques : le débit de l'extérieur à destination du client (flux descendant) est plus élevé que le débit du client vers l'extérieur (flux montant). Les débits descendants varient en général de 1.5 Mbit/s à 6 Mbit/s, et les débits montants de 512 Kbit/s à 640 Kbit/s.

L'utilisation simultanée d'un accès analogique (pour un téléphone) et d'un accès numérique (pour Internet) nécessite l'utilisation d'un séparateur. L'architecture est présentée sur la figure 3.4.

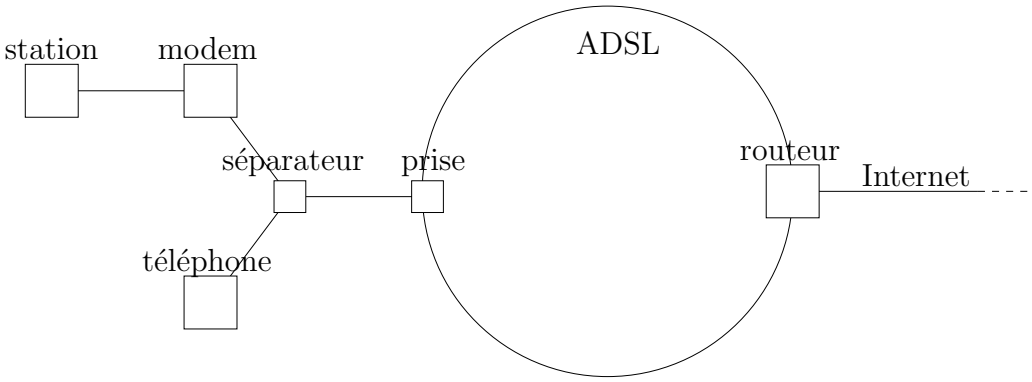

FIG. 3.4 – L'architecture ADSL.

3.3.3 PPP sur HDSL

La technologie HDSL (pour *high bit rate digital subscriber line*) est une technologie antérieure à l'ADSL, plus adaptée aux entreprises qu'aux particuliers. Une liaison HDSL revient moins cher qu'une liaison spécialisée car les répéteurs de signaux peuvent être environ trois fois plus espacés. Les lignes spécialisées tendant à être implémentées avec de l'HDSL. Dans ce cas, c'est PPP qui est utilisé pour transmettre les trames.

3.4 Architecture DMZ

Le réseau Internet est un réseau non protégé, sans contrôle d'accès. Un réseau interne d'entreprise doit donc se protèger contre l'extérieur. Cependant, les entreprises proposent souvent des services aux utilisateurs d'Internet (comme un serveur web ou un serveur FTP par exemple). La mise à disposition de ces machines ne doit pas compromettre le bon fonctionnement du réseau interne de l'entreprise.

On pourrait envisager une solution comme celle présentée sur la figure 3.5. Le réseau interne de l'entreprise est utilisé par les utilisateurs d'Internet pour accéder à la machine faisant tourner le serveur web et à celle faisant tourner le serveur FTP.

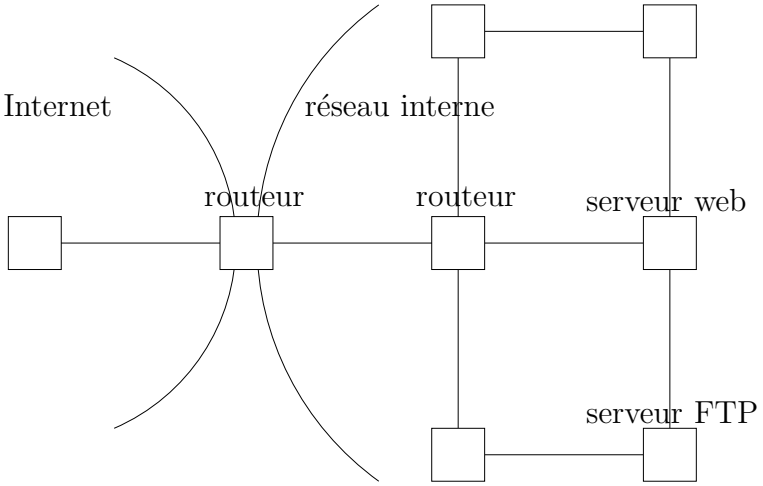

FIG. 3.5 – Architecture sans DMZ. Le contrôle du trafic est difficile.

Il est très difficile de filtrer les communications à l'intérieur d'un tel réseau. Le trafic interne se mélangeant au trafic externe. L'architecture DMZ est une solution à ce problème.

DMZ signifie zone démilitarisée. Il s'agit d'une zone frontière (ou tampon). Dans le vocabulaire réseau, les DMZ se situent entre le réseau externe Internet et le réseau interne d'entreprise. C'est là que l'on place les serveurs qui doivent être accessibles de l'intérieur et de l'extérieur, ainsi qu'un pare feu, comme indiqué sur la figure 3.6.

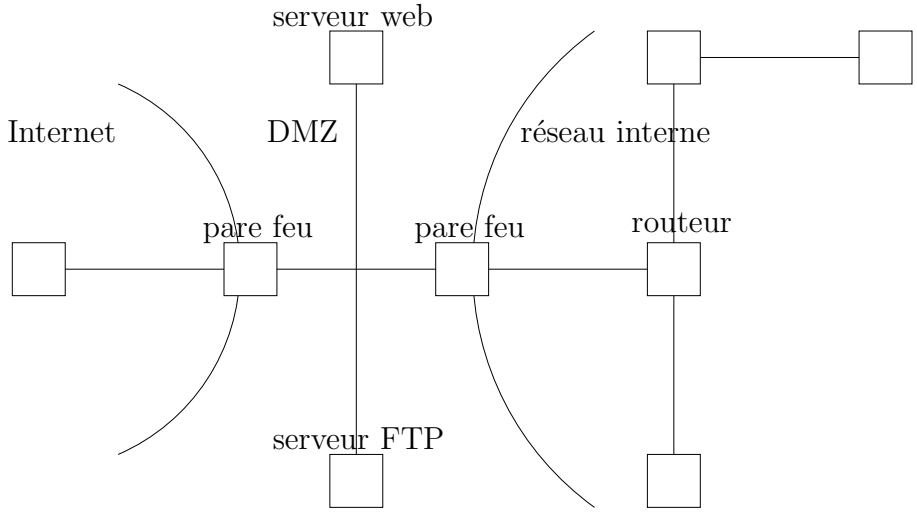

FIG. 3.6 – Architecture avec DMZ. Le contrôle du trafic est facilité.

Le pare feu entre le réseau externe et la DMZ est souvent configuré de la manière suivante :

- le trafic d'Internet vers le réseau interne est bloqué,
- le trafic d'Internet vers la DMZ est autorisé.
- le trafic de la DMZ vers Internet est bloqué,
- le trafic du réseau interne vers Internet est autorisé.

Le pare feu entre le réseau interne et la DMZ est souvent configuré comme suit :

- le trafic du réseau interne vers la DMZ est autorisé,
- le trafic du réseau interne vers l'Internet est autorisé,
- le trafic de la DMZ vers le réseau interne est bloqué,
- le trafic d'Internet vers le réseau interne est bloqué.

Cette architecture existe avec un seul pare feu, ce qui offre moins de sécurité : le réseau interne de l'entreprise est compromis dès que le pare feu est compromis.

Le filtrage des paquets par le pare feu se fait grâce à un mécanisme à état : le pare feu examine l'état des paquets qui passent, et en déduit l'état de la connexion TCP entre les machines. Par exemple, le pare feu autorisera l'établissement d'une connexion de A vers B mais pas de B vers A (si A est dans le réseau interne et B dans le réseau externe).

Chapitre 4

Réseaux étendus

Les réseaux étendus sont des réseaux dont l'étendue géographique dépasse celle d'une ville, pour interconnecter plusieurs villes entre elles, voire plusieurs pays. On les nomme WAN (pour *wide area network*).

4.1 Très hauts débits : PPP sur SONET/SDH

SONET (pour *synchronous optical network*), utilisé aux États-Unis et au Canada, et SDH (pour *synchronous digital hierarchy*), utilisé dans le reste du monde, sont deux standards très proches permettant la communication de grands volumes de données sur des fibres optiques. Ils se placent au niveau de la couche physique. Les débits avec SONET/SDH vont de 51 Mbit/s à plusieurs Gbit/s.

SONET et SDH se basent sur une synchronisation précise de tous les équipements. Au lieu de transmettre une trame séquentiellement, ces deux mécanismes entrelaçent les données de la trame avec son entête. La taille des trames étant fixée, il est possible de démultiplexer en prélevant un octet à intervalle fixe. 8000 trames sont transmises par seconde.

4.2 ATM

ATM (pour *asynchronous transfer mode* est un protocole regroupant des fonctionnalités de la couche liaison de données, réseau et transport. C'est un protocole à relais de cellules. Chaque cellule est une trame de taille fixe (53 octets, séparés en 5 octets d'entête et 48 octets de données). ATM se charge du découpage de données arrivant en paquets de tailles variables en ces cellules de taille fixe. Les cellules non complètes nécessitent du *padding* afin que leur taille fasse bien 53 octets.

L'idée principale derrière le découpage en cellules est de garantir un temps bien défini pour faire la commutation de chaque cellule. Le but de cette technologie est de réduire la gigue (c'est-à-dire la variance) du délai, afin d'améliorer la qualité des communications audio.

ATM fonctionne en mode connecté : avant de pouvoir commuter les cellules, il est nécessaire d'avoir établi un chemin. Chaque chemin est associé à une connexion ayant

une certaine qualité de service. Entre deux entités, il est souvent utile d'établir plusieurs chemins, chacun ayant sa propre qualité de service.

- On appelle VC (pour *virtual channel*) un chemin reliant deux entités du réseau ATM. Chaque VC est identifié par un identifiant dénoté VCI (pour *virtual channel identifier*). La commutation de cellules se fait grâce au VCI contenu dans chaque cellule.
- On appelle VP (pour *virtual path*) une agrégation de VC suivant le même chemin. Chaque VP est identifié par un VPI (pour *virtual path identifier*).

Les types de service fournis par ATM sont regroupés dans des catégories nommées AAL (pour *ATM adaptation layers*). Ces catégories permettent le relais des cellules ATM aux couches supérieures. Entre autres, AAL1 est utilisé pour les applications ayant des débits constants (comme la voix) et AAL5 pour les paquets de données. AAL5 est de loin le plus utilisé (il porte d'ailleurs le nom SEAL, pour *simple and efficient AAL*). AAL5 réduit la partie utile par cellule à 40 octets.

4.3 *Frame relay*

4.4 MPLS

4.5 Alternative du VPN

Bibliographie

[1] G. Cizault. *IPv6 - Théorie et pratique.* O'Reilly, 2 edition, 1999. ISBN-10 : 2841770850.

[2] J.-L. Montagnier. *Réseaux d'entreprise par la pratique.* Eyrolles, 2 edition, Mars 2004. ISBN-10 : 2212112580, ISBN-13 : 978-2212112580.

[3] G. Pujolle. *Les réseaux - Édition 2005.* Eyrolles, 5 edition, Septembre 2004. ISBN-10 : 2212114370.