RÉSEAUX LOCAUX

SUPPORT DE COURS
E3i, 2001-2002
Université de Tours

Michel Crucianu

École d'Ingénieurs en Informatique pour l'Industrie
64, avenue Jean Portalis
37200 TOURS

Table des matières

1. Réseaux téléinformatiques et modèle OSI — rappel

Intérêt des réseaux téléinformatiques :

1° Le partage des ressources matérielle et logicielles, des données.
2° La fiabilité du système d'information.
3° L'augmentation graduelle des ressources matérielles et logicielles.
4° La communication entre utilisateurs distants et/ou applications distantes.
5° La collaboration entre utilisateurs distants (*groupware*, par exemple Lotus Notes).

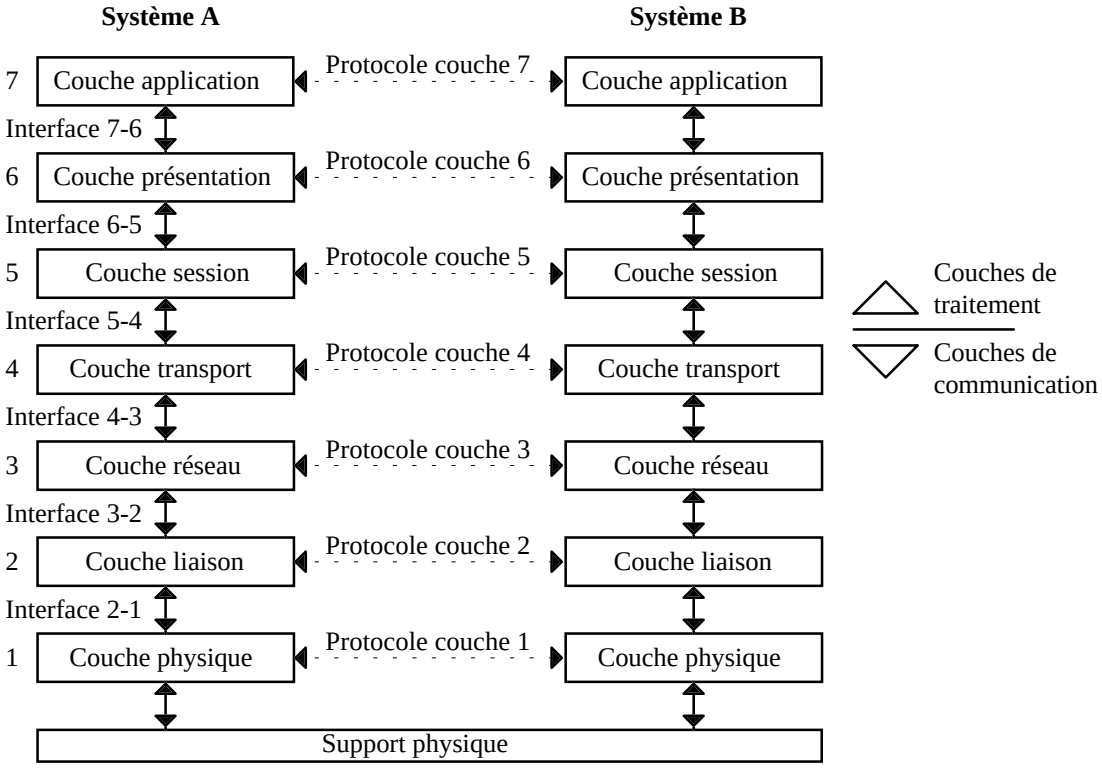
Types de réseaux en fonction de l'aire desservie :

1° Réseaux locaux (*Local Area Networks, LAN*) : 10 m ÷ 1 km.
2° Réseaux métropolitains (*Metropolitan Area Networks, MAN*) : 1 km ÷ 100 km.
3° Réseaux très longue distance (*Wide Area Networks, WAN*) : 100 km ÷ 10 000 km.

Pourquoi un modèle en couches ?

1° Facilité de développement et de modification : une couche (un protocole) peut être modifiée de façon indépendante tant que l'interface avec les deux couches adjacentes reste inchangée.
2° Intéropérabilité : une même couche de niveau *n+1* peut utiliser les services de couches de niveau *n* très différentes à condition que l'interface *n/n+1* soit la même.

Architecture OSI :



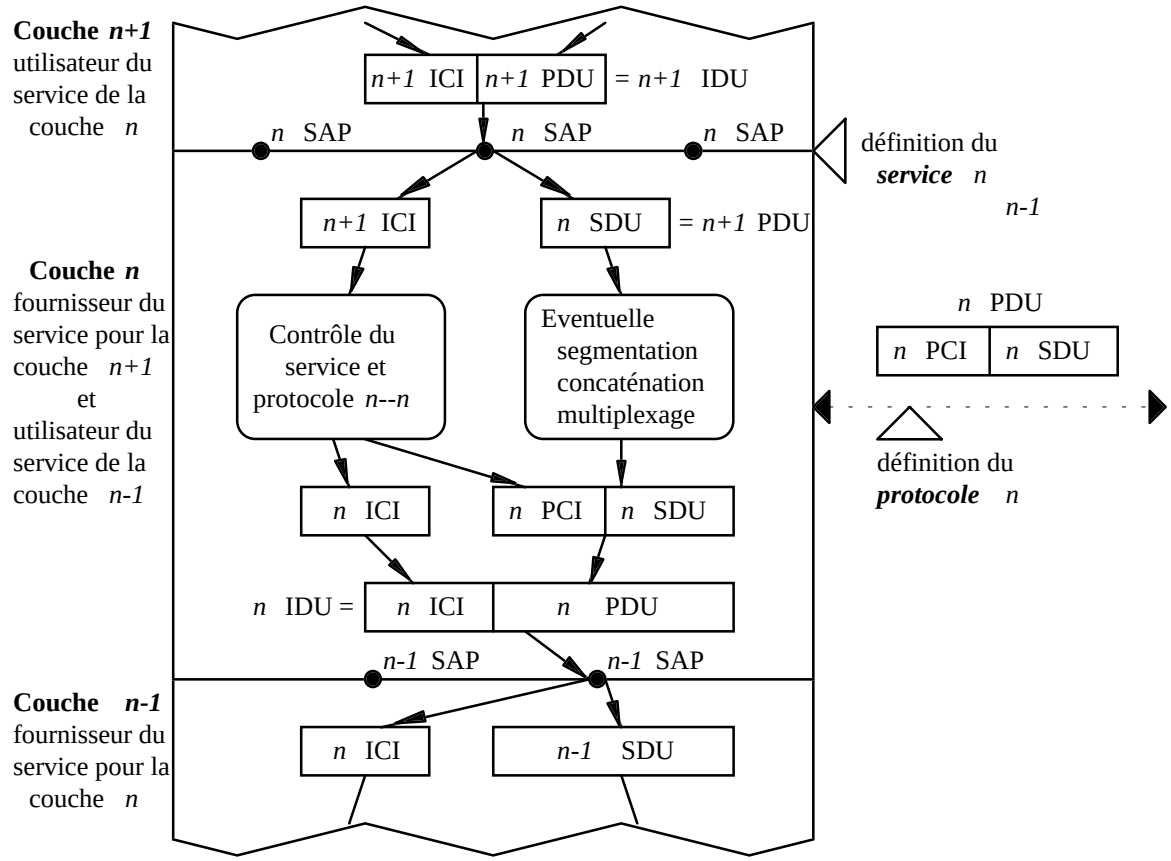
Termes employés : *interface* entre couches adjacentes, *protocole* entre processus *pairs* (de même niveau).

Spécifique de chaque couche OSI :

1° Physique : transmission des bits sur un support physique déterminé.
2° Liaison de données : liaison fiable point à point.
3° Réseau : acheminement des messages.
4° Transport : transport fiable de bout en bout.
5° Session : gestion du dialogue et synchronisation.
6° Présentation : syntaxe de transfert, compression, cryptage[1].
7° Application : services application génériques (terminal virtuel, transfert de fichiers, messagerie).

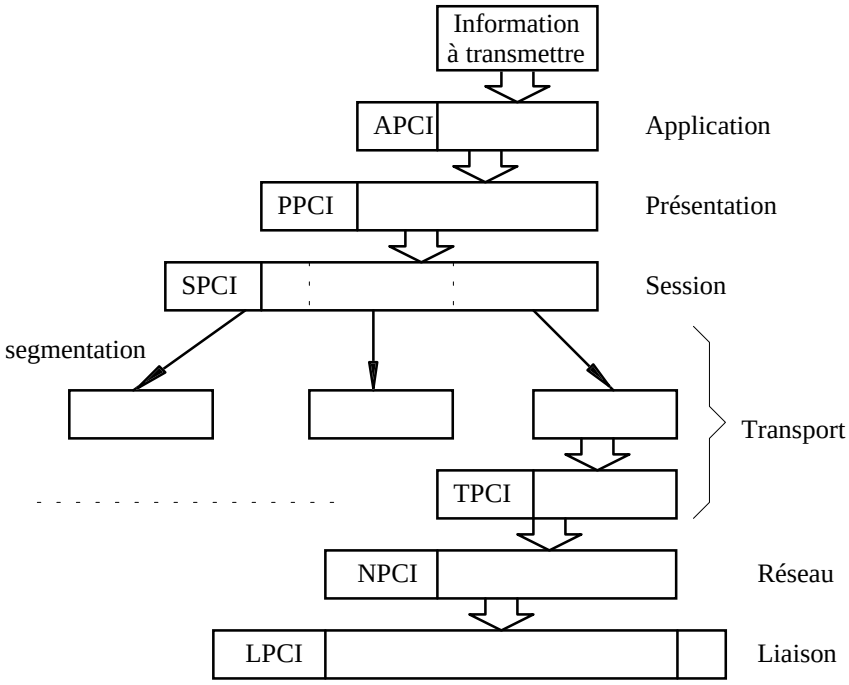
[1] Dans les protocoles actuels, la compression et le cryptage peuvent intervenir aussi à des niveaux plus bas dans la hiérarchie.

Architecture de principe d'un niveau (terminologie OSI) :



SAP : point d'accès au service (*Service Access Point*)
IDU : unité de données d'interface (*Interface Data Unit*)
ICI : informations de contrôle de l'interface (*Interface Control Information*)
PDU : données du protocole (*Protocol Data Unit*)
PCI : informations de contrôle du protocole (*Protocol Control Information*)
SDU : données du service (*Service Data Unit*)

Encapsulation successive des informations à transmettre :



Mode de transmission :

1° Avec connexion : établissement d'une connexion avant le transfert des données. Avantages : permet de s'assurer que le destinataire peut accepter les messages ; l'ordre des messages est respecté ("tuyau"). Désavantage : durée élevée d'établissement de la connexion. Mode intéressant uniquement pour le transfert de volumes importants de données (nombre élevé de messages ordonnés).

2° Sans connexion : les données sont envoyées sans qu'une connexion soit préalablement établie. L'ordre des messages n'est pas nécessairement respecté. Mode utilisable sur des réseaux à voie unique (l'ordre des messages est maintenu grâce à la structure du réseau) ou pour des messages individuels (l'ordre n'a aucune importance).

Qualité de service :

1° Service fiable : aucune perte de données grâce au contrôle des erreurs et à l'acquittement de chaque message (exemple : transfert de fichiers). Entraîne des délais supplémentaires.

2° Service non fiable : les erreurs ne sont pas détectées, il n'y a pas d'acquittement pour les messages (exemple : téléphone).

Le mode connecté et le service fiable ne sont en général pas utilisés dans toutes les couches ; si le support est très fiable, le contrôle des erreurs peut n'être effectué que pour le transfert de bout en bout (au niveau transport).

L'accès à un service utilise des primitives de service :

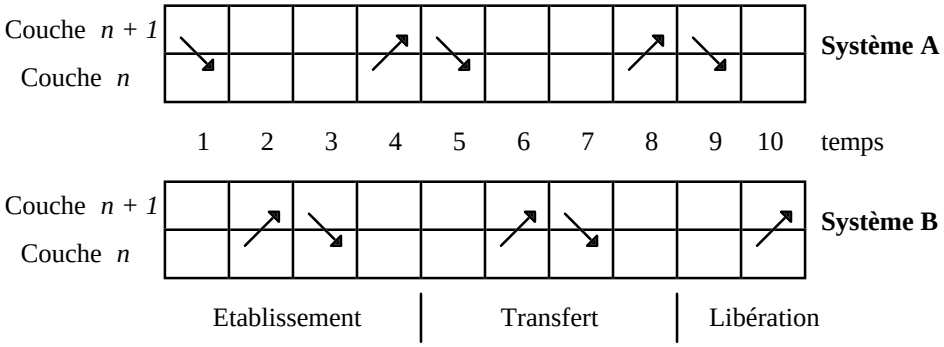
1° Requêtes (*request*) : une entité sollicite un service pour une activité.

2° Confirmation (*confirm*) : une entité est informée de sa demande de service.

3° Réponse (*response*) : une entité répond à un événement.

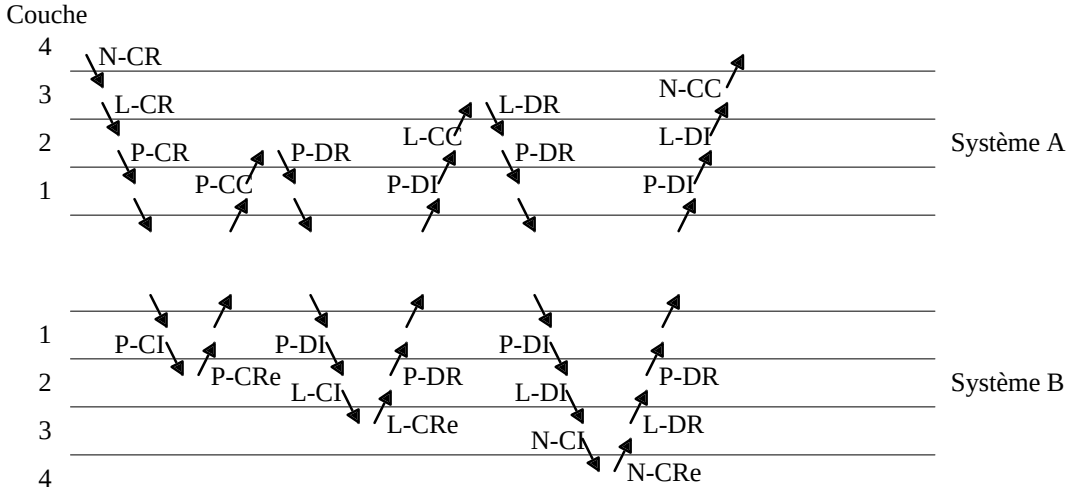
4° Indication (*indication*) : une entité est informée d'un événement.

Transfert de données *n+1* — *n+1* en mode connecté :



1°	CONNECT.request	6°	DATA.indication
2°	CONNECT.indication	7°	DATA.request
3°	CONNECT.response	8°	DATA.indication
4°	CONNECT.confirm	9°	DISCONNECT.request
5°	DATA.request	10°	DISCONNECT.indication

Etablissement d'une connexion entre deux couches transport (modèle OSI) :



CR : CONNECT.request CC : CONNECT.confirm

CI : CONNECT.indication DR : DATA.request
CRe : CONNECT.response DI : DATA.indication

Service confirmé : demande, indication, réponse et confirmation. Service non confirmé (ne pas confondre fiabilité et confirmation) : uniquement demande et indication. En général, un service DATA n'a pas besoin de confirmation, alors qu'un service CONNECT doit être toujours confirmé (les deux entités entre lesquelles s'établit la connexion doivent se mettre d'accord sur les paramètres de la connexion).

2. Réseaux locaux et systèmes d'exploitation réseau

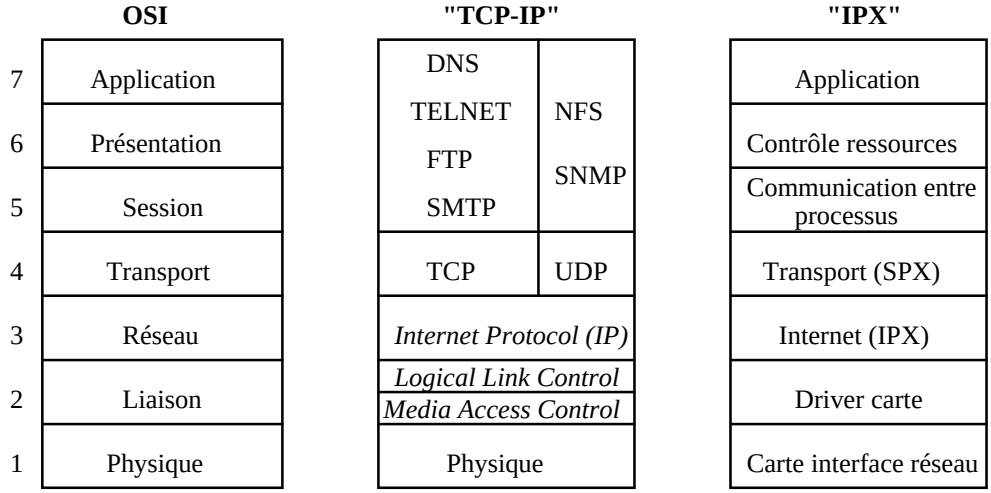
Différences LAN/WAN :

1° La structure des LAN est simple — en général, sur un LAN, la voie physique entre deux ordinateurs est unique. Le routage sur un LAN est donc simplifié. Des protocoles en mode non connecté peuvent être employés (exemple : *User Datagram Protocol*, UDP).

2° Le partage des ressources exige une réponse rapide et un débit élevé. Le mode connecté est souvent évité à cause des délais de connexion. Le contrôle des erreurs et le contrôle de flux sont parfois limités à une seule couche.

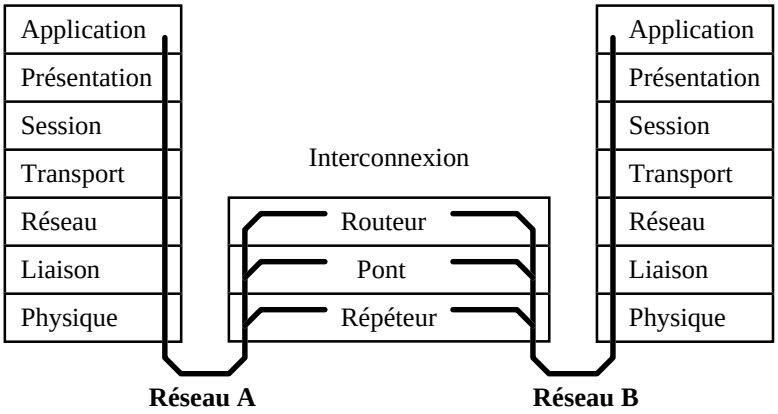
3° En général, un nombre élevé de machines sont connectées au même LAN qui est un réseau multipoint (canal de diffusion).

Organisation des modèles OSI, TCP/IP et IPX (Novell) :



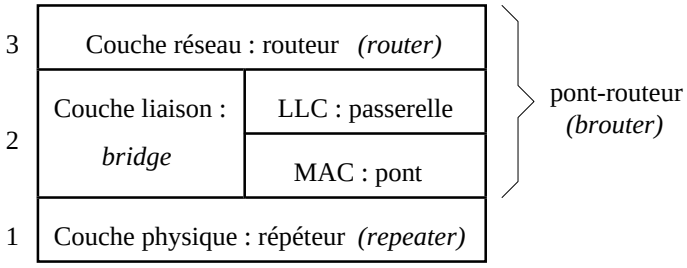
TCP : *Transport Control Protocol* (protocole transport mode connecté)
UDP : *User Datagram Protocol* (protocole transport mode non connecté)
DNS : *Domain Name Service* (service de noms de domaine)
TELNET : *TELecommunication NETwork* (terminal virtuel)
FTP : *File Transfer Protocol*
SMTP : *Simple Mail Transfer Protocol*
NFS : *Network File System*
SNMP : *Simple Network Management Protocol*

Interconnexion entre LAN :



Equipements d'interconnexion entre LAN :

1° Répéteur (*repeater*) : connexion au niveau du support physique (régénération du signal).
2° Pont (*bridge*) : interconnexion au niveau liaison (sous-couche MAC) entre 2 LAN de même type.
3° Passerelle : interconnexion au niveau liaison (sous-couche LLC) entre LAN respectant le même standard LLC.
3° Routeur (*router*) : routage et interconnexion entre réseaux utilisant le même protocole de couche 3 (réseau).
4° Pont-routeur (*brouter*) : par exemple, connexion entre LAN du même type plus connexion à un WAN.
5° *Gateway* : en plus du routage, peut effectuer des conversions de protocoles de niveau plus élevé. Employé souvent comme terme générique désignant un équipement d'interconnexion quelconque.



Systèmes d'exploitation réseau (*Network Operating Systems*, NOS) — mêmes fonctions que les OS classiques, mais sur un réseau (local) qui est en général rendu transparent à l'utilisateur :

1° Gestion et protection des ressources du réseau.
2° Partage de ressources matérielles et logicielles, partage de données.
3° Gestion et protection des utilisateurs.

Architecture globale du système informatique : une ou plusieurs machines-serveur (dédiées ou non), plusieurs machines-client. Certains NOS gèrent de façon exclusive toutes les machines (exemple : UNIX), d'autres travaillent en conjonction avec le OS local sur les machines-client (exemple : Novell NetWare).

Quelques types d'API (*Application Programming Interfaces*) :

1° Système d'exploitation/hardware drivers.
2° Système d'exploitation/réseau local.
3° Système d'exploitation/réseau distant.
4° GUI/système d'exploitation client.
5° Système d'exploitation client/système d'exploitation serveur.
6° Application/système d'exploitation.
7° Application/base de données.
8° Application/application.

Développement orienté objet utilisant les API : le développeur d'applications assemble des objets existants, les API restent présentes entre les objets mais sont utilisées uniquement par le développeur d'objets.

3. Architecture IEEE 802

IEEE 802.1 spécifie l'architecture générale (format adresses, techniques interconnexion réseaux, etc.).

IEEE 802.2 spécifie le niveau LLC, qui peut assurer les classes de services suivants :

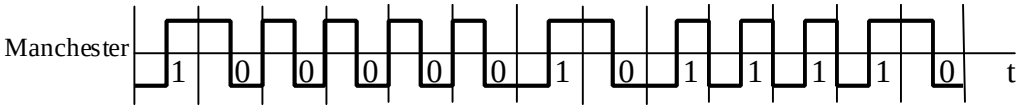
LLC1 — service simple en mode non connecté, sans acquittement (donc sans reprise sur erreur) et sans contrôle du séquencement et des duplications.
LLC2 — service en mode connecté, avec acquittement, reprise sur erreur, contrôle du séquencement et des duplications.
LLC3 — service en mode non connecté mais avec acquittement.

Caractéristique générale des protocoles : le contrôle est décentralisé ; quand un rôle de gestionnaire — à attributions limitées — existe (protocoles à jeton), il peut être rempli par une machine quelconque.

Les normes IEEE 802.3 et 802.5 sont actuellement utilisées uniquement en bande de base.

3.1. IEEE 802.3 et Ethernet

Codage en bande de base employé : Manchester. Des niveaux de tension symétriques sont employés afin de réduire la composante continue du signal transmis. Une transition est présente au milieu du temps-bit pour chaque bit transmis (codage biphase) afin d'améliorer la synchronisation (transition-horloge) ; si le bit est "1" la transition est ascendante, si le bit est "0" la transition est descendante (voir la figure suivante).



Le spectre de puissance d'un signal Manchester est :

$$\gamma(f) = b \cdot \left[\frac{2a}{\pi \cdot f}\right]^2 \cdot \sin^4\left(\frac{\pi \cdot f}{2b}\right)$$, b étant la fréquence de bit et a l'amplitude du signal.

La transmission a lieu à travers un milieu unique, câble ou milieu de transmission des ondes radio. Types de câblage utilisés : bus câblé en bus, bus câblé en étoile. Types de câble : coaxial (10 base 5 — ∅ 10 mm, 10 base 2 — ∅ 5 mm), paire torsadée (10 base T). Débit binaire général : 10 Mbps.

L'accès au bus correspond à un protocole CSMA/CD (*Carrier Sense Multiple Access/Collision Detection*) : chaque station qui désire émettre (*Multiple Access*) écoute le canal (le milieu commun) et essaie d'émettre uniquement quand le canal est libre (*Carrier Sense*) ; si plusieurs stations essaient d'émettre en même temps, une collision se produit sur le canal ; les stations écoutent le canal pendant l'émission, constatent la collision et arrêtent les émissions ; chaque station reprend l'émission après un intervalle aléatoire (entre des limites précises) afin de minimiser le risque d'une nouvelle collision entre les mêmes stations.

Paramètres caractéristiques IEEE 802.3 et Ethernet :

Time slot : temps nécessaire au signal pour parcourir deux fois la distance qui sépare les stations les plus éloignées. **La durée minimale d'une trame doit être supérieure au *time slot* pour que la détection des collisions soit possible** (dans tous les cas, une trame ne doit pas avoir une taille inférieure à 64 octets). Explication : après un retard de 1L la première trame arrive à l'autre bout, moment auquel la station à l'autre bout peut encore l'ignorer et émettre la sienne, plus un retard de 1L pour que la station ayant émis la première détecte la collision — car la station ayant émis la première doit être encore en émission — donc durée minimale trame > *time slot* (correspondant à 2L).

Taille minimale brouillage = 32 bits. A la détection de la collision, la station doit émettre une séquence de brouillage pour permettre aux autres stations de bien détecter la collision.

Durée minimale d'émission = 51,2 µs, correspond à la taille minimale de 64 octets ; si une collision est détectée, la séquence de brouillage doit couvrir ce qui reste de cette durée minimale.

Taille maximale trames = 1514 octets, permet d'éviter la monopolisation du canal.

Nombre d'essais avant abandon (la reprise peut être éventuellement demandée uniquement par les couches supérieures) = 16.

Intervalle de silence entre les messages = 9,6 µs.

Intervalle d'attente après détection d'une collision : $n \times time\ slot$, avec n tiré au sort dans l'intervalle $\left[0, 2^{\min(\text{nombre collisions successives}, 10)}\right]$.

Mécanismes CSMA/CA (*Collision Avoidance*) : les stations évitent les collisions en utilisant des délais différents, multiples du *time slot*, pour émettre (ces délais sont réattribués dynamiquement pour que la méthode soit équitable) ; ne sont pas retenus par les recommandations 802, mais 802.4 (bus à jeton) prévoit un mécanisme de *Collision Avoidance* basé sur la circulation d'un jeton qui accorde le droit d'émettre. Utilisations de CSMA/CA : réseaux locaux de taille réduite, réseaux locaux par ondes radio.

La trame physique 802.3 débute par un préambule de 7 octets 10101010, suivi par un SFD (*Start Frame Delimiter*) 10101011 — un seul octet. La trame Ethernet débute par un préambule sur 64 bits (8 octets).

Les types de trames niveau MAC (*Media Access Control*) :

Trame Ethernet :

Adresse destination	6 octets
Adresse source	6 octets
ID protocole	2 octets
Données	variable (la longueur de la trame sera un multiple de 8)
Contrôle	4 octets

Adresse destination, adresse source : adresses physiques, voir le format plus bas.

ID protocole : identifie le protocole de niveau supérieur (réseau) auquel les données sont destinées (ex. IP, ARP, AppleTalk).

Contrôle : un code polynomial est utilisé pour la détection des erreurs sur la trame en entier.

Trame IEEE 802.3 :

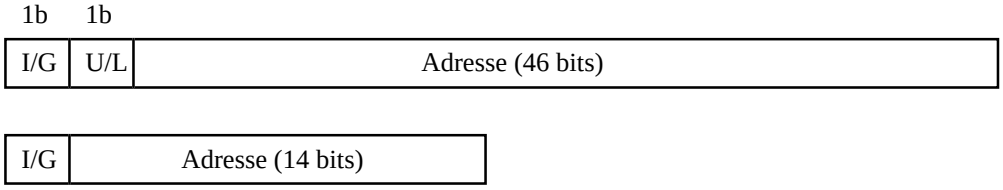
Adresse destination	2 ou 6 octets
Adresse source	2 ou 6 octets
Longueur données	2 octets
Données	variable (la longueur de la trame sera un multiple de 8)
Contrôle	4 octets

Adresse destination, adresse source : adresses physiques, voir le format plus bas.

Longueur données : la taille des données en octets. Les données d'une trame IEEE 802.3 sont toujours livrées à LLC ; c'est l'en-tête LLC qui permet d'identifier le protocole de niveau supérieur (réseau) destinataire.

Contrôle : un code polynomial est utilisé pour la détection des erreurs sur la trame entière.

Format des adresses physiques :



Bit I/G : 0 = adresse individuelle, 1 = adresse de groupe (*broadcast* à toutes les stations du réseau local — adresse FF FF FF FF FF FF, *multicast* à un groupe de stations).

Bit U/L : 0 = adresse universelle, 1 = adresse localement définie.

IEEE 802 définit des adresses sur 2 octets (utilisables pour des réseaux non connectés vers l'extérieur) et sur 6 octets, Ethernet utilise uniquement des adresses sur 6 octets.

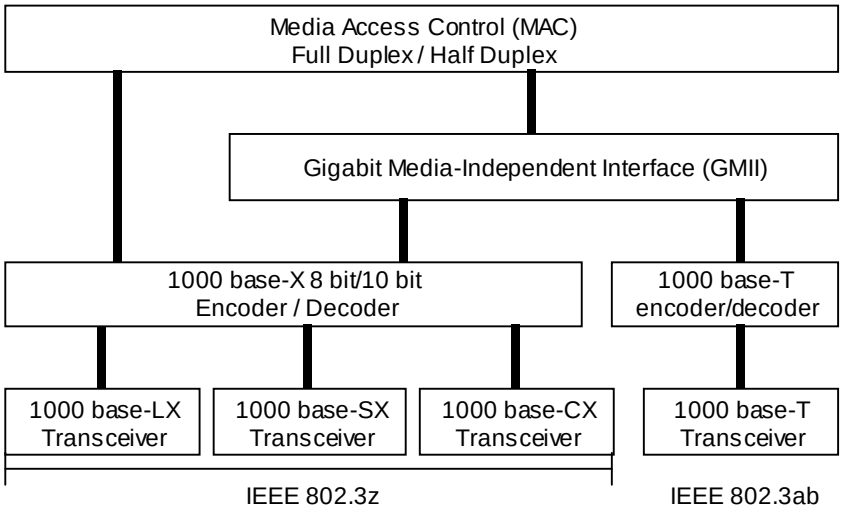
La tendance actuelle est d'utiliser la commutation (les concentrateurs sont remplacés par des commutateurs qui évitent la plupart des collisions). Pour l'interconnexion des commutateurs Ethernet, les propositions Fast Ethernet (100 Mbps) 100 base TX, 100 base T4 ou 100 base FX constituent une solution intermédiaire, avant le passage à Gigabit Ethernet.

3.2. Gigabit Ethernet

Gigabit Ethernet a été développé afin d'augmenter encore le débit des réseaux Ethernet, avec un minimum de modifications pour le protocole 802.3 et sans toucher aux couches supérieures. La structure de la trame est la même, ainsi que le format des adresses physiques. Plusieurs supports ont été définis : la fibre optique monomode, la fibre multimode (meilleur marché que la monomode), la paire torsadée blindée (*shielded*, pour de très courtes distances, correspondant aux connexions entre des équipements situés dans un même local technique) et la paire torsadée non blindée (100 Ω, *unshielded*, de catégorie 5, la plus courante dans le câblage actuel des immeubles de bureaux). Le tableau suivant reprend les principales caractéristiques des différents membres de la famille des technologies Gigabit Ethernet :

Nom	**Type du support**	**Distance maximale**	
		Duplex intégral	**Semi-duplex**
1000 Base-LX (wavelength 1300 nm)	62,5 µm multi mode fiber	440 m	320 m
	50 µm multi mode fiber	550 m	320 m
	10 µm single mode fiber	3000 m	320 m
1000 Base-SX (wavelength 850 nm)	62,5 µm multi mode fiber	260 m	260 m
	50 µm multi mode fiber	550 m	320 m
1000 Base-CX	Shielded twisted pair	25 m	25 m
1000 Base-T	Unshielded twisted pair (4 pairs), category 5	100 m	100 m

La structure des couches MAC (sous-couche de la couche liaison) et physique, pour les différents membres de la famille, est indiquée dans le schéma suivant :



Gigabit Ethernet a été défini dès le départ en deux versions, semi-duplex et duplex intégral. Dans la version semi-duplex le support est partagé et l'accès au support est géré par le même protocole CSMA/CD. Pour assurer la compatibilité avec les couches supérieures, la taille minimale des trames est maintenue à 64 octets par rapport à ces couches. Toutefois, afin de garder une longueur maximale raisonnablement élevée pour le support partagé, les trames de taille inférieure à 512 octets sont complétées par des octets vides (*padding*) pour atteindre les 512 octets. Après l'envoi de cette première trame qui permet à un équipement d'occuper le support partagé, ce même équipement a la possibilité d'envoyer d'autres trames, de taille inférieure cette fois-ci à 512 octets, pour une durée qui correspond à 8000 octets, sans être inquiété. Cette procédure de « réservation » temporaire du support par une trame de taille supérieure ou égale à 512 octets permet d'assurer un débit utile élevé même si les trames de faible taille (< 512 octets) dominent. La taille maximale des trames est la même que pour Ethernet (1514 octets).

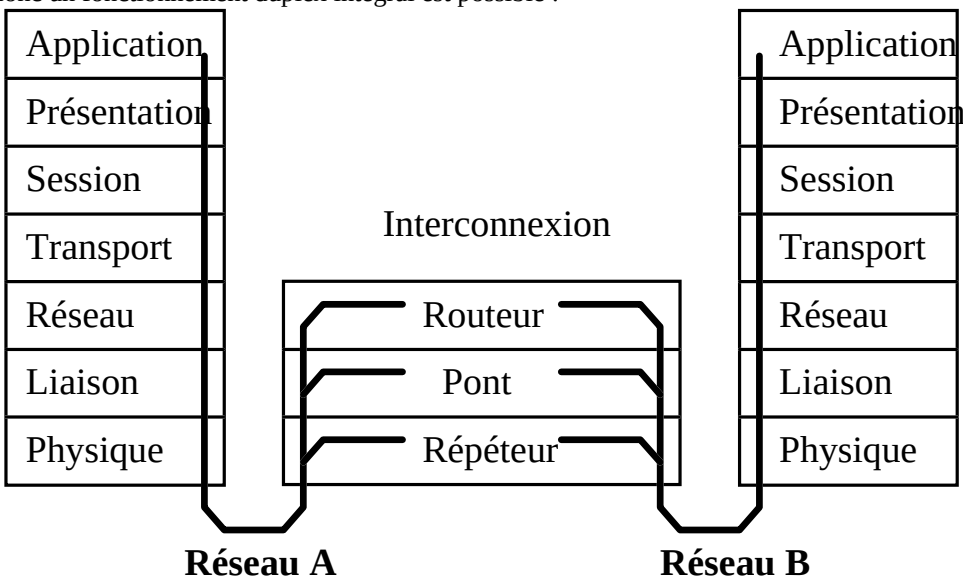
La version duplex de Gigabit Ethernet permet d'interconnecter des équipements à travers un support bidirectionnel (débit théorique 2x1 Gbit/s), utilisé par deux équipements à la fois, donc CSMA/CD devient inutile. Cela permet non seulement d'augmenter le débit utile, mais aussi de s'affranchir des contraintes de *time slot* et donc d'augmenter la taille maximale du réseau jusqu'aux limites imposées par la technologie de transmission (3000 m pour 1000 Base-LX sur fibre monomode). L'interconnexion à très haut débit d'équipements (commutateurs, par exemple) situés dans des immeubles distants sur un même campus devient alors possible sans passer par une technologie différente (ATM).

Pour les transmissions par fibre optique (LX, SX) ou paire torsadée blindée symétrique (*Balanced Shielded Twisted Pair*, 150 Ω, CX), un codage 8bits/10bits (2 bits de synchronisation pour 8 bits utiles) permet d'assurer la synchronisation des horloges d'émission et de réception. La présence de ce codage implique que pour atteindre un débit utile de 1000 Mbit/s, le support doit fonctionner à 1250 Mbit/s.

La norme 1000 Base-T a été proposée afin de permettre de tirer profit du câblage existant (4xUTP, 100 Ω, catégorie 5) dans la plupart des établissements. Plusieurs techniques sont combinées afin d'assurer un débit théorique de 1 Gbit/s à travers un support physique aussi peu performant :

1° Emploi des 4 paires torsadées du câble, avec pour chaque paire une vitesse de modulation de 125 Mbaud et 2 bit par période de modulation, donc 250 Mbit/s par paire.
2° Utilisation d'un codage PAM 5 (à 5 niveaux, 4 pour les 2 bits d'information par période de modulation, le cinquième utilisé par le code correcteur d'erreurs) complété par un codage *Forward Error Correction* (FEC).
3° Adaptation du spectre du signal émis aux caractéristiques du câble (*pulse shaping*) et égalisation du signal reçu par des filtres non linéaires.

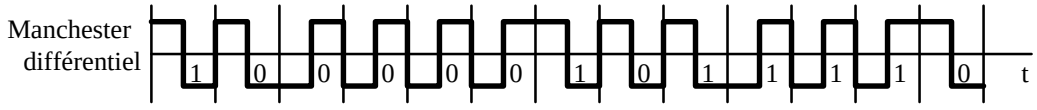
1000 Base-T est conçue pour fonctionner avec des concentrateurs Gigabit Ethernet, en semi-duplex, ou avec des commutateurs, en duplex intégral. Chaque paire torsadée est employée simultanément en émission et en réception, donc un fonctionnement duplex intégral est possible :



Pour garder le contact avec l'évolution de Gigabit Ethernet, consulter le site Internet de la Gigabit Ethernet Alliance, http://www.10gea.org/.

3.3. IEEE 802.5 et Token Ring

Codage en bande de base employé : Manchester différentiel. Les transitions au milieu du temps-bit codent la différence entre deux bits successifs : la transition est ascendante lorsque les deux bits sont identiques et descendante dans le cas contraire :



L'absence de transitions horloge est employée pour les marqueurs de début et de fin sur l'anneau à circulation de jeton — recommandation IEEE 802.5.

Un anneau est constitué de plusieurs lignes point à point reliées bout à bout par des dispositifs qui assurent la continuité de l'anneau lorsque la station connectée au répéteur est hors service ; la mise en service/hors service d'une station a toutefois des conséquences néfastes sur les signaux transmis et engendre la retransmission du message affecté. Câblage utilisé : l'anneau câblé + les câbles entre les stations et leur unité de raccordement à l'anneau. Type de câble : paire torsadée (blindée). Débit binaire employé : 4 Mbps.

Quelques techniques possibles pour la gestion **distribuée** des émissions sur l'anneau :

1° Multiplexage temporel avec affectation des tranches temporelles : le temps nécessaire pour un tour complet de l'anneau est divisé en tranches et une ou plusieurs tranches sont attribuées à chaque station (de façon statique ou dynamique). Chaque station retransmet sans délai tout ce qu'elle reçoit, en insérant éventuellement son propre message dans la (les) tranche(s) attribuée(s). Le débit pour une station dépend du nombre de tranches attribuées. Le délai de livraison est garanti mais l'efficacité est faible (tranches inutilisées).

2° Insertion de registre : l'anneau devient pratiquement une suite de liaisons point à point ; une station stocke chaque message qui lui arrive, le retire si le message lui est destiné, sinon le retransmet dès que la liaison de sortie est disponible (*store and forward*). L'efficacité est bonne mais les délais de garde et donc les délais de livraison sont élevés quand le réseau est chargé et diminuent fortement le débit réel.

3° Circulation de jeton : la station qui entre en possession du "jeton" (message spécifique) à le droit d'émettre pendant un certain temps et doit ensuite céder le jeton (émettre un message avec le jeton). Chaque station connectée retransmet toutes les trames qu'elle reçoit, en introduisant un **retard fixe**. Des techniques de gestion de priorités et de gestion du jeton sont implémentées. L'efficacité est raisonnable et **le délai de livraison est garanti** (contrairement à IEEE 802.3/Ethernet qui utilise l'accès CSMA/CD).

La recommandation IEEE 802.5 a retenu la troisième technique. Chaque station retarde la retransmission d'une trame de **1 bit**. En fonctionnement normal, c'est la station qui émet une trame qui est chargée de l'éliminer de l'anneau — chaque trame fait une rotation complète — et ensuite de libérer le jeton. La station qui émet une trame émet aussi des caractères de remplissage entre la fin de la trame et le début de l'émission du jeton qu'elle libère. Chaque station effectue un contrôle d'erreur et peut positionner un flag d'erreur de transmission qui invalide la trame. La station destinataire positionne des flags pour indiquer la façon dont elle a traité la trame. Différentes stations assurent différents services (identifiés par des adresses fonctionnelles, voir plus loin) sur l'anneau, le plus important étant le service de surveillance active ; toutes ou la plupart des stations se trouvent normalement en état de surveillance passive et peuvent assumer, de façon dynamique, la tâche de surveillance active. Pour permettre à tous les bits du jeton d'être sur l'anneau lorsque celui-ci est très court, la station de surveillance insère un tampon de 24 bits (3 octets, longueur du jeton) dans l'anneau.

Paramètres caractéristiques IEEE 802.5 :

Token Holding Time (THT) : durée de possession du jeton par une station, limitée à 10 ms (~ 5000 octets).

Ring latency = contenance de l'anneau. Chaque répéteur associé à une station introduit un retard d'un bit ; la longueur des câbles introduit aussi des retards (à 4 Mbps, 1 bit correspond à ~50 m). La somme des bits contenus dans l'anneau donne la *ring latency*, qui dépend donc du nombre de répéteurs (maximum 250) et de la longueur des câbles. La synchronisation entre deux stations successives se fait à partir du signal (codage Manchester différentiel) qui peut subir de légères distorsions, le résultat étant une variation de la durée des bits d'une machine à l'autre ; ces variations cumulées peuvent atteindre l'équivalent de 3 bits. La station de surveillance active fait varier la capacité du tampon inséré (*Latency buffer*) entre 24 et 30 bits pour maintenir la contenance de l'anneau à une valeur constante.

Le jeton :

Marqueur de début	1 octet
Contrôle d'accès	1 octet
Marqueur de fin	1 octet

Marqueur de début : l'octet VV0VV000, V étant des bits de violation de la convention utilisée par le codage Manchester différentiel — la transition en milieu de temps bit est absente.

Contrôle d'accès : octet **pppTMrrr**. **T** = jeton (**0**) ou trame (1) ; **M** = supervision (contrôle de la rotation des trames). **ppp** = niveau de priorité courante de la trame. **rrr** = niveau de réservation. Quand la station qui possède le jeton s'en sépare, **ppp**_jeton_émis **:= rrr**_trame_reçue.

Marqueur de fin : l'octet VV1VV111.

La trame 802.5 est le résultat de la transformation en trame du jeton reçu. Format trame :

Marqueur de début	1 octet
Contrôle d'accès	1 octet
Contrôle de la trame	1 octet
Adresse destination	2 ou 6 octets
Adresse source	2 ou 6 octets
Données	variable
Contrôle	4 octets
Marqueur de fin	1 octet
Statut de trame	1 octet

Marqueur de début : le même que pour le jeton.

Contrôle d'accès : octet **pppTMrrr**. **T** = jeton (0) ou trame (**1**) ; **M** = supervision — contrôle de la rotation des trames : positionné par la station qui assure le service de surveillance active ; une trame qui arrive à cette station avec le bit **M** déjà positionné indique une anomalie (deuxième rotation dans l'anneau) qui oblige la station de surveillance active à purger l'anneau et à régénérer un nouveau jeton. **ppp** = priorité courante de la trame ; lorsque le jeton arrive à une station qui désire émettre une trame, celle-ci compare le niveau de priorité de la trame à la valeur **ppp** ; si son niveau de priorité est supérieur, la station garde le jeton et émet la trame (plus précisément transforme le jeton en trame en positionnant le bit **T** et en ajoutant le reste de la trame). **rrr** = niveau de réservation ; une station qui désire émettre une trame peut positionner les bits **rrr** du jeton ou d'une trame qui passe — à condition que la nouvelle valeur soit supérieure à la valeur courante — pour faire connaître sa demande aux autres stations.

Contrôle de la trame — indique ce que transporte la trame (le champ "Données") : une PDU du niveau LLC ou des informations de contrôle de la couche MAC (informations de supervision de l'anneau).

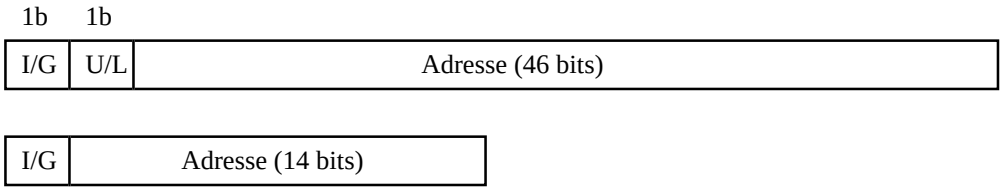
Adresse destination, adresse source : adresses physiques de stations, voir le format plus bas.

Contrôle : un code polynomial est utilisé pour la détection des erreurs sur la trame entière, à l'exception des marqueurs et de l'octet de statut de trame.

Marqueur de fin : l'octet VV1VV1**IE** ; **I** = bit trame intermédiaire, si = 1 la trame est intermédiaire dans une suite ; **E** = bit "erreur détectée", positionné par toute station qui détecte une erreur dans la trame (la trame est alors invalidée).

Statut de trame : octet **AC**xx**AC**xx, indique à l'émetteur (en fin de rotation) si le destinataire à reconnu la trame et quel sort lui à été accordé. Bit **A** (dupliqué) = 1 signifie adresse destinataire reconnue ; bit **C** (dupliqué) = 1 signifie trame copiée correctement par le destinataire. Avec le bit **E** du marqueur de fin, cet octet offre à la couche MAC la possibilité de vérifier la réception correcte d'une trame par le destinataire et de reprendre l'émission en cas d'insuccès.

Format des adresses :



Bit I/G : 0 = adresse individuelle, 1 = adresse de groupe (*broadcast* à toutes les stations du réseau local — adresse FF FF FF FF FF FF, *multicast* à un groupe de stations).

Bit U/L : 0 = adresse universelle, 1 = adresse localement définie.

IEEE 802 définit des adresses sur 2 octets (utilisables pour des réseaux non connectés vers l'extérieur) et sur 6 octets, Token Ring (IBM) utilise des adresses sur 6 octets.

Quand les deux premiers bits d'une adresse sur 6 octets sont positionnés à 1 nous avons une adresse **fonctionnelle** qui identifie un service ; ce service peut être assuré dans le réseau par différentes stations, qui possèdent différentes adresses physiques. Chaque bit dans les 4 octets de poids faible identifie un service, plusieurs services peuvent être assurés par une même station, donc à chaque service on associe un masque. Les principaux services :

Service	Masque (en hexadécimal)
Surveillance active	C000 0000 0001
Serveur de paramètres	C000 0000 0002
Enregistrement incidents	C000 0000 0008
Serveur de configuration	C000 0000 0010
Serveur NETBIOS	C000 0000 0080
Pont	C000 0000 0100
Serveurs personnalisés	de C000 0008 0000 à C000 4000 0000

4. Évolution de IP : IPv6

IP (*Internet Protocol*) : niveau réseau (3 OSI), mode sans connexion, non fiable et sans garantie de séquencement (*datagram*) ; permet la segmentation/le regroupement des messages.

TCP (*Transmission Control Protocol*) : niveau transport (4 OSI), mode avec connexion, fiable et avec contrôle de flux.

UDP (*User Datagram Protocol*) : niveau transport (4 OSI), mode sans connexion, non fiable et sans garantie de séquencement (*datagram*).

ARP (*Address Resolution Protocol*) : traduire une adresse IP en adresse physique.

RARP (*Reverse Address Resolution Protocol*) : traduire une adresse physique en adresse IP.

RDP (*Route Discovery Protocols*) : famille de protocoles permettant de diriger les messages et de mettre à jour les tables de routage.

4.1. Problèmes posés par IPv4

Quand IPv4 a été développé le nombre d'équipements connectés était relativement faible, les équipements mobiles étaient très rares et les délais de transmission n'avaient pas une grande importance car les données transmises n'étaient pas urgentes. Mais l'environnement d'utilisation a complètement changé et les caractéristiques de IPv4 posent actuellement des problèmes importants :

1° Epuisement des adresses, dû a l'explosion du nombre de sous-réseaux et d'équipements connectés.

2° Pour faire face à l'épuisement des adresses, l'utilisation des masques a remplacé les classes d'adressage, ce qui rend plus complexes les algorithmes et les tables de routage. Aucune technique de configuration automatique des espaces d'adressage n'a été définie.

3° Accommodation difficile des équipements mobiles, actuellement très nombreux : en effet, chaque mobile doit avoir une adresse IP complètement différente selon l'endroit où il se trouve, ce qui pose des problèmes très difficiles de gestion des adresses.

4° Indisponibilité de classes de service correspondant aux exigences imposées par des flots de données très divers : transfert de fichiers, sessions interactives, conversations téléphoniques, vidéoconférence, etc.

5° Certaines opérations effectuées dans les routeurs — recalcul du code de contrôle après la modification du champ durée de vie, fragmentation/réassemblage des messages — sont très coûteuses en temps de calcul et augmentent les délais d'acheminement.

4.2. IPv6

Paquet IPv6 (IP nouvelle génération, RFC 1752, offre commerciale disponible) :

Version	4 bits
Classe de trafic	4 bits
Etiquette de flot	3 octets
Longueur données	2 octets
En-tête suivant	1 octet
Nombre de sauts	1 octet
Adresse source	16 octets (128 bits)
Adresse destination	16 octets (128 bits)
SDU IPv6	variable

Version : la version du protocole IP qui a créé le paquet, 6 pour IPv6.

Classe de trafic : les valeurs de 0 à 7 sont employées pour les flots à paquets contrôlés (0 = trafic non caractérisé, 1 = trafic ne demandant pas de réponse, 2 = transfert de données intempestives, 3 et 5 sont réservés, 4 = transferts volumineux attendus, 6 = trafic interactif, 7 = trafic de contrôle Internet), les valeurs 8 à 15 pour les flots à paquets non contrôlés (par exemple conversations téléphoniques ; plus la valeur est élevée, moins l'utilisateur est disposé à accepter que les paquets soient jetés en cas de congestion).

Etiquette de flot : permet d'identifier un flot, qui est une séquence de paquets envoyés depuis une source particulière à une destination particulière, séquence pour laquelle la source désire un traitement particulier par les routeurs concernés.

Longueur données : longueur de ce qui suit l'en-tête, jusqu'à 64 Koct (codée sur 2 octets) ; une valeur supérieure (→ *Jumbogram*) peut être indiquée dans une option.

En-tête suivant : permet l'extension des en-têtes existants, en indiquant quelle entité de protocole doit être appelée afin de traiter l'en-tête suivant. Dans IPv6, les options sont indiquées dans des en-têtes supplémentaires, traités uniquement par le noeud identifié par l'adresse destination (à l'exception de l'option de routage par la source, dont l'en-tête est traité par chaque routeur intermédiaire). La longueur de chaque en-tête supplémentaire est un multiple de 8 octets.

Nombre de sauts (*hop count*), ancien champ durée de vie : chaque routeur réduit de 1 la valeur et jette le paquet si le résultat est nul.

Adresse source et adresse destination : adresses IPv6, voir les détails plus loin.

Quelques options proposées :

Authentification des utilisateurs et confidentialité des données.

Auto-configuration des adresses, permettant aux un stations connectées à un sous-réseau de se construire une adresse.

Routage par la source et marquage du chemin (modification des structures associées aux étiquettes de flot).

Fonctions de fragmentation.

Contrôle des erreurs.

4.2.1. Adressage

La définition et la gestion des adresses IPv6 doivent faciliter la tâche des utilisateurs/administrateurs des réseaux et aussi l'activité des routeurs. Cela est possible grâce notamment à la richesse de l'espace d'adressage, à la possibilité de définir de multiples niveaux hiérarchiques d'adresses, à la présence d'adresses de type *cluster* et aux mécanismes de configuration automatique des adresses.

Une adresse IPv6 est représentée comme une succession de 8 groupes de valeurs hexadécimales représentées sur 4 chiffres, par exemple 1080:222:AF45:FF:FE:143:4441:110. Une succession de 0 peut être

représentée par "::", comme dans FEDC::122:AD45:4555. Une adresse IPv4 encapsulée dans une adresse IPv6 sera x:x:x:x:x:x:d.d.d.d (x = valeur hexadécimale codant 16 bits, d = valeur décimale codant 8 bits), comme dans 0:0:0:0:0:0:192.22.128.1, représentée aussi ::192.22.128.1.

IPv6 possède 22 classes d'adresses, dont 17 sont réservées pour un usage futur. Les adresses IPv6 sont de trois types :

1° *Unicast* : adresse d'un correspondant unique, bien défini. Les adresses IPv6 de noeuds utilisant IPv4 ont pour préfixe 0000 0000.

2° *Cluster* : adresse d'un groupe de noeuds qui partagent un même préfixe d'adresse. Un paquet (datagramme) envoyé à une telle adresse sera livré au routeur le plus proche situé sur la frontière du domaine. Cela permet notamment de préciser de façon simple, en utilisant l'option de routage par la source, le ou les opérateur(s) télécom dont on veut utiliser les services. Dans une adresse *cluster* la partie de poids fort est le préfixe partagé par les adresses du *cluster* et la partie de poids faible est 0 (par conséquence, toutes les adresses IPv6 ayant une succession de 0 comme partie de poids faible sont réservées et ne doivent pas être employées comme adresses *unicast*).

3° *Multicast* : adresse de diffusion utilisée pour envoyer un datagramme à tous les membres d'un groupe multicast. Les adresses de *broadcast* sont remplacées dans IPv6 par des adresses *multicast*. Toutes les adresses multicast débutent par 1111 1111. Certaines adresses *multicast* prédéfinies permettent de simplifier le fonctionnement des protocoles (par exemple, FF0E::43 identifie tous les serveurs *Network Time Protocol* de l'Internet).

Contrairement aux adresses IPv4 qui sont totalement indépendantes des adresses de niveau inférieur (liaison/physique), une adresse *unicast* IPv6 est censée incorporer l'adresse de niveau inférieur. Par exemple, pour une station connectée via un réseau local IEEE 802, l'adresse MAC sur 48 bits (dont l'unicité dans le monde est garantie par les constructeurs des cartes réseau) forme les 48 bits de poids faible de l'adresse IPv6 (la source d'inspiration a été IPX). Ceci simplifie l'autoconfiguration des adresses et l'assignation dynamique d'une adresse à un mobile. Pour un réseau local qui n'est pas connecté à Internet, des adresses "lien local" peuvent être utilisées : une telle adresse est de type FEx0::<adresse MAC> (avec x = 1000 ou 1100) ; ces adresses peuvent être configurées par les stations connectées, en l'absence de toute intervention d'un utilisateur/administrateur ou d'un routeur.

Deux mécanismes existent pour former l'adresse Internet d'une station. Suivant le premier, la station forme son adresse en ajoutant un préfixe de réseau présent dans un message ICMPv6 *Router Advertisement* envoyé périodiquement par un routeur local à un suffixe qui est en général l'adresse MAC IEEE 802. La station peut aussi envoyer une demande ICMPv6 *Router Sollicitation* pour obtenir le préfixe sans plus attendre.

Suivant le deuxième mécanisme, l'adresse est assignée par le routeur ou un autre serveur d'adresses IP grâce au protocole DHCPv6 (*Dynamic Host Configuration Protocol*). L'administrateur du réseau local peut préciser non seulement l'ensemble de préfixes à utiliser, mais aussi les masques (pour IPv4), l'adresse du DNS, l'adresse du routeur par défaut, ainsi qu'un certain nombre de paramètres du protocole IP (taille des paquets, durée de vie conseillée, etc.). Ces différents paramètres peuvent être transmis à une station non seulement à la première connexion mais aussi ultérieurement, à la demande de la station. Les serveurs d'adresses gardent (*bindings*) la configuration IP de chaque station gérée : identification de la station, adresse IP, paramètres IP et durée de vie (*lease*) de l'association station−adresse. Un dialogue DHCP typique se déroule comme suit :

1° La station qui vient d'être connectée au réseau local envoie en *broadcast* un message *DHCP_Discover*.

2° Plusieurs serveurs d'adresse peuvent répondre par des messages *DHCP_Offer*. Chaque réponse contient l'adresse du serveur, l'adresse IP proposée à la station et plusieurs paramètres IP.

3° La station choisit le serveur DHCP qui lui convient et lui envoie un message *DHCP_Request*, en lui demandant éventuellement des paramètres IP supplémentaires.

4° Le serveur choisi sauvegarde le *binding* correspondant à la station et lui envoie les paramètres demandés dans un message *DHCP_ACK*.

5° Avant d'utiliser l'adresse IP obtenue, la station fait appel à ARP pour s'assurer de l'unicité de cette adresse.

6° Pour prolonger la durée de vie de l'adresse IP assignée, la station envoie un DHCP_Request avec la valeur de cette adresse. Si l'adresse n'est plus valide, le serveur répond par un *DHCP_NAK* et la station doit redémarrer une configuration d'adresse.

Le routage dans IPv6 est similaire, à quelques extensions près, à celui de IPv4. L'extension la plus importante concerne l'utilisation des adresses *cluster* dans l'option de routage par la source.

4.2.2. IPv6 et la mobilité

Un mobile est une station pour laquelle le point de rattachement à Internet peut changer souvent. Pour chaque mobile on définit un identificateur qui est similaire à une adresse (c'est en général l'adresse du mobile sur son réseau maison) et qui ne change pas dans le temps, ainsi qu'une adresse IP qui indique donc le point courant de

rattachement du mobile et qui peut donc évoluer. La correspondance entre l'identificateur et l'adresse courante est gardée dans des AMT (*Address Mapping Table*). Chaque entrée dans la table contient aussi un indicateur de la durée de maintien de la validité de cette entrée, ainsi qu'un numéro de version de l'adresse, numéro incrémenté à chaque changement d'adresse. Pour la création et la mise à jour d'entrées dans une AMT, une authentification (du mobile ou de la source du message) est effectuée au préalable sur les paquets reçus.

A son arrivée sur un réseau (ou avec une périodicité donnée), un mobile se fait attribuer une adresse IP adéquate ; si l'adresse MAC IEEE 802 est employée dans l'adresse IPv6, chaque mobile possède par défaut une adresse sur chaque réseau IEEE 802. Ensuite, il envoie un paquet — avec l'option de marquage de la route et le paramètre de contrôle de la mobilité — à son réseau maison. Chaque routeur qui intervient dans ce transfert met à jour son AMT (après authentification du mobile) à partir des informations présentes dans le paramètre de contrôle de la mobilité (identification, adresse, version de l'adresse, durée de maintien, etc.), si l'entrée est absente de sa table ou si le numéro de version de l'adresse est supérieur à celui présent déjà dans la table. Le réseau maison ne peut pas déterminer la nouvelle adresse d'un de ses mobiles si celui-ci ne se fait pas connaître en utilisant le mécanisme décrit.

Le protocole de niveau supérieur (comme TCP) emploie l'identifiant d'un mobile pour lui envoyer un paquet et c'est la couche IPv6 qui retrouve (éventuellement) l'adresse à jour du mobile dans son AMT.

4.2.3. IPv6 et la sécurité

Contrairement à IPv4 qui fait l'hypothèse que les mécanismes de sécurité sont présents à un niveau supérieur (niveau application pour l'environnement TCP/IP ou niveau présentation pour le modèle OSI), IPv6 inclut des mécanismes de sécurité à son niveau (réseau). Toutefois, ces mécanismes de sécurité ne protègent pas contre l'analyse du trafic : le contenu des messages peut être confidentiel mais certaines adresses, informations concernant les routes, le débit et les paramètres de la qualité de service restent accessibles dans les routeurs.

L'intérêt d'introduire la sécurité au niveau réseau est de fournir aux utilisateurs des niveaux de sécurité supplémentaires. Par exemple, une société implantée sur plusieurs sites peut utiliser la sécurité au niveau réseau pour la protection (automatique, non gérée par les utilisateurs) des communication inter-sites et la sécurité au niveau présentation ou application pour la protection (non automatique, gérée par les utilisateurs) des communications intra-site. Aussi, les mécanismes de sécurité au niveau réseau permettent de renforcer la protection des réseaux internes contre les intrusions.

Le premier mécanisme défini utilise un en-tête d'authentification et permet d'assurer l'authenticité de l'émetteur et l'intégrité (mais pas la confidentialité) du contenu des paquets IPv6. Seuls les champs qui sont modifiés dans chaque routeur (comme *Hop Count* ou *Next Address*) sont exclus de la vérification de l'intégrité. Ce mécanisme permet entre autres de s'assurer que les champs qui interviennent dans le filtrage d'accès des paquets (comme les adresses, le protocole de transport, le numéro de port, etc.) sont authentiques. La technique proposée pour l'authentification et contrôle de l'intégrité est *Message Digest 5* (MD5, développée par Rivest).

Le deuxième mécanisme utilise un en-tête de confidentialité (*Encapsulating Security Payload Header*, ESPH) qui permet en plus d'assurer la confidentialité du contenu à l'aide d'une technique de chiffrement. Le *datagram* IP en entier peut être encapsulé (*Tunnel mode*) et chiffré ou uniquement le segment de niveau transport (TCP ou UDP, *Transport mode*). La technique proposée est CBC-DES (*Cipher Block Chaining Data Encryption Standard*). La gestion des clés n'est pas incorporée dans IP mais est assurée par un protocole de niveau supérieur. Les deux mécanismes de sécurité sont des mécanismes de bout en bout (sauf si l'encapsulation d'un protocole de niveau réseau dans IPv6 est employée).

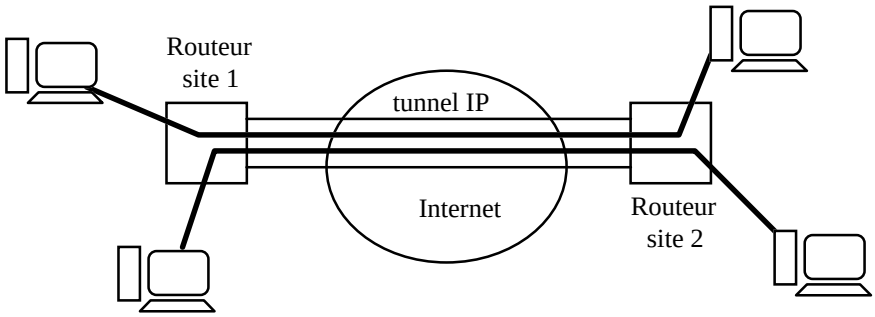
Chaque machine garde dans une table les informations de sécurité concernant ses possibles utilisateurs et interlocuteurs. A chaque interlocuteur correspond une ou plusieurs entrées (une entrée par utilisateur, usage, etc.) dans la table. Chaque entrée contient un numéro d'identification ou *Security Parameters Index* (SPI), l'adresse IP correspondante, l'identité des protocoles utilisés pour l'authentification et la confidentialité et les clés associées.

Considérons trois scénarios et les solutions correspondantes :

1° Une société désire sécuriser ses communications client↔serveur intra-site. L'utilisation de MD5 et de l'en-tête d'authentification permet de faire authentifier les clients par les serveurs et réciproquement. Ainsi, si les codes d'authentification sont différents (et confidentiels) pour toutes les machines, une machine ne peut pas se substituer à une autre en copiant son adresse IP. En même temps, grâce à la vérification de l'intégrité, une machine ne peut pas se substituer à une autre en copiant le bon en-tête et en remplaçant le contenu du message initial par des données fausses.

2° A l'intérieur d'une société, un administrateur est la seule personne habilitée à transférer des données très sensibles entre deux machines ; ces transferts doivent rester opaques. Un SPI distinct correspondra à l'utilisateur concerné et, par rapport aux autres utilisateurs des deux machines, une clé d'authentification différente sera employée. Aussi, la confidentialité sera utilisée (en-tête de confidentialité présent dans les paquets, contenu crypté en *Transport mode*) et donc une clé de cryptage sera présente.

3° Une société désire connecter ses différents sites distants via Internet, mais veut assurer une sécurité maximale au niveau réseau pour les transferts ainsi que pour les accès. Un tunnel IP sera créé : les routeurs "frontaliers" de chaque site utilisent l'authentification réciproque et la vérification de l'intégrité, ainsi que le cryptage et l'encapsulation de tous les paquets IP qui doivent circuler entre les sites (confidentialité *Tunnel mode*) dans des paquets IP inter-sites ; de cette façon, non seulement les données mais aussi les en-têtes IP (donc les informations sensibles comme les adresses IP, les paramètres de trafic, les débits, etc.) sont cachés vis à vis de l'extérieur. Sur chaque site, les machines emploient éventuellement une authentification et un cryptage poste à poste.



5. ATM

ATM (*Asynchronous Transfer Mode*) est une technique de transport asynchrone utilisant des paquets de 53 octets, développée pour servir de support au réseau numérique à intégration de services (RNIS) large bande (*Broadband* ISDN). Ceci signifie que ATM devrait permettre de transporter à la fois des informations isochrones (voix, sons, images animées) et asynchrones (données).

Structure trame (5 octets en-tête, 48 octets données) :

[*Generic Flow Control*	4 bits]
Virtual Path	12 [ou 8 bits]
Virtual Circuit	12 bits
Payload Type	3 bits
réservé	1 bit
Cell Loss Priority	1 bit
Header Error Control	1 octet
Données	48 octets

Generic Flow Control : utilisé uniquement dans l'interface utilisateur-réseau (absent au-delà), fournit des niveaux de priorité entre les différentes connexions de l'utilisateur qui partagent le même accès réseau.

Virtual Path : identifie une voie virtuelle qui peut contenir plusieurs circuits virtuels.

Virtual Circuit : identifie un circuit virtuel. Les numéros de voie et de circuit virtuel ont une signification locale au lien et sont donc modifiés dans chaque brasseur/commutateur ATM, selon des tables de routage dynamiques ; certaines valeurs sont réservées pour des usages spécifiques.

Payload Type : définit la nature de la charge utile (cellule utilisateur, cellule gestion réseau, cellule gestion congestion, etc.).

Cell Loss Priority : si le bit est 1, la cellule peut être jetée en priorité en cas de saturation d'un commutateur.

Header Error Control (HEC) : contrôle d'erreurs sur l'en-tête ; le polynôme employé permet de détecter certaines erreurs multiples et de corriger les erreurs simples. Les numéros de voie et de circuit virtuel étant modifiés dans chaque brasseur/commutateur, le HEC n'est pas calculé/vérifié par la couche ATM mais par la sous-couche de convergence (TC) de la couche physique !

Structure de la pile de protocoles :

<table>
<tr><td>Fonction des couches supérieures</td><td colspan="2">Couches supérieures</td></tr>
<tr><td>Convergence</td><td>CS</td><td rowspan="2">AAL</td></tr>
<tr><td>Segmentation et réassemblage</td><td>SAR</td></tr>
<tr><td>Contrôle de flux générique
Génération en-tête cellules
Commutation (traduction VPI/VCI)
Multiplexage/démultiplexage</td><td colspan="2">ATM</td></tr>
<tr><td>Génération/vérification contrôle erreur en-tête
Adaptation à la trame de transmission
Génération/récupération trame transmission</td><td>TC</td><td rowspan="2">Couche
physique</td></tr>
<tr><td>Synchronisation bit
Support physique</td><td>PM</td></tr>
</table>

Caractéristiques :

Le service offert par la couche ATM est en mode connecté (un circuit doit être établi avant que des données ne puissent être envoyées) mais sans acquittement, donc non fiable. Les cellules successives d'un même circuit virtuel arrivent toujours dans l'ordre d'émission (car elles prennent toutes le même chemin), mais certaines cellules peuvent être perdues (saturation d'un commutateur). C'est le niveau supérieur (AAL) qui, selon le type de service assuré, doit éventuellement s'occuper des retransmissions. Aussi, dans la couche ATM le contrôle d'erreurs porte uniquement sur l'en-tête, l'intégrité des données transportées doit être vérifiée par les niveaux supérieurs.

Chaque circuit virtuel est unidirectionnel.

Le routage est assuré par des tables dont le contenu est mis à jour de façon dynamique, en fonction des circuits virtuels ouverts. A l'ouverture d'un circuit virtuel, des ressources lui sont réservées dans chaque brasseur/commutateur traversé.

Des fonctions d'adaptation très différentes (couche AAL) doivent être utilisées pour différents services offerts : un transfert asynchrone de données n'impose pas les mêmes exigences qu'un transfert isochrone d'images, par exemple. Catégories de fonctions d'adaptation :

AAL"0" : la couche d'adaptation est absente, le contenu du champ données des cellules est transféré directement vers/depuis la couche supérieure.

AAL1 : trafic isochrone — la source produit à débit et rythme fixe des unités de données qui doivent être délivrées au même rythme à destination. Il faut pouvoir identifier la perte d'information sans chercher à y remédier. Les fonctions de la couche sont donc : segmentation et réassemblage des informations utilisateur, gestion de la variation du délai de propagation des cellules (gigue), gestion de la durée de constitution de la capacité utile des cellules, gestion des cellules perdues ou mal insérées, restauration de l'horloge source à la réception, récupération de la structure des données source en réception, recherche des erreurs dans les informations de contrôle de protocole AAL et traitement de ces erreurs, recherche des erreurs dans le champ informations utilisateur et traitement de ces erreurs sans retransmission (utilisation de codes correcteurs d'erreurs).

AAL2 : trafic de source isochrone — le rythme de la source est fixe mais le débit variable, le rythme doit être maintenu de façon **approximative** à destination. Il faut pouvoir identifier la perte d'information sans chercher à y remédier. Les fonctions assurées sont pratiquement les mêmes que celles de AAL1.

AAL3 : trafic asynchrone entre entités identifiées (en général applications transactionnelles). Le débit demandé peut être contraint dans les bornes négociées et les délais doivent être raisonnables. Les pertes de données (perte de cellules, erreurs de transmission) doivent être corrigées. Par rapport aux couches AAL précédentes, AAL3 peut assurer la retransmission des données corrompues à cause des erreurs de transmission ou de la perte de cellules ATM. AAL3 peut supporter le multiplexage.

AAL4 : trafic asynchrone en rafale (datagrammes sur réseau local). Le débit instantané demandé peut être très élevé, les délais sont peu importants. Les pertes de données doivent être corrigées. Les fonctions assurées sont les mêmes que celles de AAL3.

AAL5 : trafic asynchone à débit variable, les pertes de données doivent être corrigées. Par rapport à AAL3/4 les en-tête de protocole sont réduits, donc les transferts sont plus efficaces. AAL5 ne supporte pas le multiplexage.

Pour l'interface utilisateur↔réseau, deux débits sont disponibles : 155,520 Mbits/s sur deux câbles coaxiaux (un par sens, avis UIT-T G.703) ou deux fibres optiques monomode et 622,080 Mbits/s sur deux fibres optiques monomode (avis UIT-T G.957).

Avantages de ATM :

La granularité fine facilite le multiplexage des informations asynchrones avec des informations isochrones (une trame isochrone doit attendre au maximum une trame asynchrone avant d'accéder au réseau), ce qui n'est pas le cas pour des techniques comme *Frame relay* qui utilisent des trames de longueur variable, en général élevée.

La taille réduite et fixe des trames permet aussi un traitement rapide dans les commutateurs, et donc des délais de routage réduits.

Enfin, des cellules de taille réduite permettent un taux de remplissage plus élevé que des cellules de taille importante.

Problèmes posés par ATM :

La taille réduite des cellules fait baisser le rendement (taille données/taille cellule). La taille choisie est un compromis entre le rendement et le délai d'acheminement (avec la contrainte de pouvoir acheminer des signaux isochrones).

Le non respect par un utilisateur du débit négocié a une influence négative sur tous les utilisateurs qui emploient le même support. Une fonction de "police" complexe doit être implémentée.

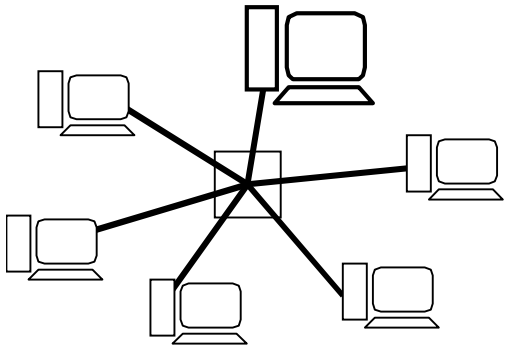
6. Évolutions dans les réseaux locaux

Les évolutions récentes dans le domaine des réseaux locaux ont pour but l'augmentation du taux de transfert utile, l'amélioration de la flexibilité, la simplification de l'administration et l'homogénéisation des supports. Le taux de transfert peut être augmenté par un changement technologique (passage de Ethernet à Fast Ethernet, Gigabit Ethernet ou ATM) ainsi que, dans certains cas, par une amélioration des techniques d'interconnexion entre les stations (segmentation d'un réseau Ethernet, remplacement des concentrateurs par des commutateurs). L'amélioration de la flexibilité et la simplification de l'administration (tout en restant compatible avec l'exigence d'augmentation du taux de transfert) impliquent en général la définition de réseaux locaux virtuels. L'homogénéisation des supports signifie souvent le passage à ATM (technologie développée au départ pour WAN ou MAN) sur le réseau local.

6.1. Augmentation du débit

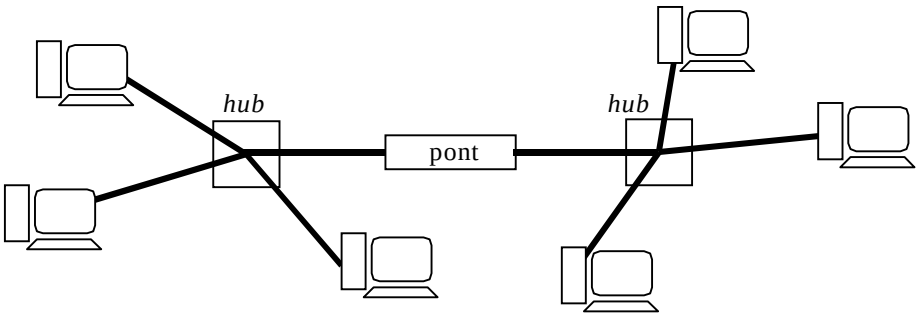
C'est le type de trafic constaté sur le réseau qui permet de faire le bon choix pour augmenter le débit. L'utilisation d'un analyseur de trafic est très utile, mais une simple analyse de l'utilisation du réseau peut donner de bons résultats. Considérons un réseau à collisions, qui fait appel à un concentrateur (*hub*) pour assurer l'interconnexion des machines. Le concentrateur est un équipement qui assure la connexion physique entre les différentes liaisons (donc de niveau 1). Voici, pour trois types d'utilisations du réseau, la solution à privilégier :

1° Les postes de travail communiquent presque exclusivement avec un serveur. Dans ce cas, l'augmentation du débit passe par une augmentation de la bande passante de l'accès au serveur, possible par une évolution vers une technologie à débit supérieur : Ethernet → Fast Ethernet, Fast Ethernet → Gigabit Ethernet. Le « domaine de collisions » reste donc le même. En plus du changement de concentrateur, cette évolution peut se faire sur toutes les liaisons avec le concentrateur, ou uniquement sur la liaison du serveur (concentrateur à multiplexage).



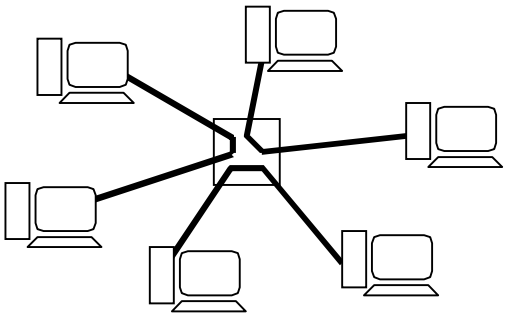
Interconnexion par concentrateur (*hub*)

2° Les postes et serveurs connectés au réseau peuvent être séparés en groupes, et les communications ont lieu surtout entre les membres d'un même groupe. Dans ce cas, pour augmenter le débit nous pouvons segmenter le réseau à l'aide d'un (ou plusieurs) ponts (*bridge*). Chaque groupe de postes correspond à un segment, séparé des autres segment par des ponts. Le pont permet de relier deux segments et est capable de filtrer les trames selon l'adresse physique du destinataire (équipement de niveau 2) : si le destinataire est sur l'autre segment, la trame peut traverser le pont ; si le destinataire est sur le même segment que l'émetteur, la trame ne peut pas traverser le pont. Le « domaine de collisions » initial est ainsi séparé en plusieurs sous-domaines (pour la majorité des trames), ce qui permet d'augmenter le débit utile sur chaque segment. Un pont laisse passer les *broadcasts*.



Segmentation du réseau à l'aide d'un pont (*bridge)*

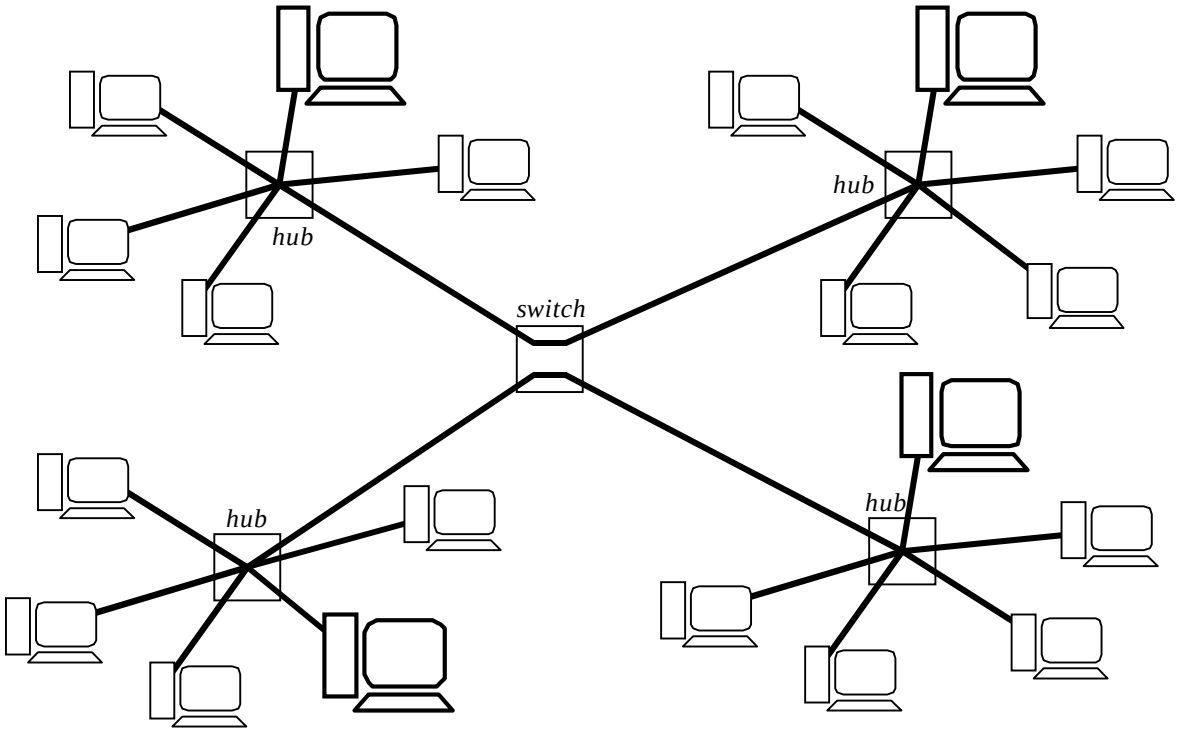
3° Le trafic est globalement homogène, tout le monde communique avec tout le monde et avec des intensités d'un même ordre de grandeur. Dans ce cas, afin d'augmenter le débit nous pouvons remplacer le concentrateur par un commutateur. Le commutateur permet de faire communiquer en même temps, au débit théorique permis par la technologie, plusieurs paires de machines. Des collisions peuvent encore se produire quand plusieurs machines essaient de communiquer en même temps avec la même, mais sont très rares. Le « domaine de collisions » est ainsi complètement éclaté, ce qui permet d'atteindre un débit global qui est un multiple du débit théorique de la technologie (plusieurs communications en même temps). Enfin, un commutateur transmet aussi les *broadcasts*.



Interconnexion par commutateur (*switch*)

Il faut insister sur le fait que si le trafic correspond au cas numéro 1 – tous les postes communiquent presque exclusivement avec le serveur – l'utilisation d'un commutateur à la place du concentrateur ne permet pas d'augmenter le débit de façon significative !

Bien évidemment, des solutions mixtes sont souvent mises en œuvre. Une des plus répandues est celle qui assure l'interconnexion entre plusieurs concentrateurs à travers un commutateur (interconnexion hiérarchique) :



Interconnexion hiérarchique

Si cette solution a le mérite de découper le « domaine de collisions » en plusieurs sous-domaines, le « domaine de *broadcast* » continue à couvrir le réseau entier, car le commutateur transmet les *broadcasts*. IP fait un usage très modéré du *broadcast* de niveau 2, ce qui n'est malheureusement pas le cas pour d'autres protocoles comme IPX (voir *service advertising*) ou Appletalk.

Nous nous sommes intéressés ici à l'augmentation du débit par une évolution technologique ou un découpage des « domaines de collisions », mais d'autres problèmes peuvent se poser, comme celui du filtrage des communications. Ces problèmes sont traités d'une façon plus globale dans le paragraphe suivant.

6.2. Réseaux locaux virtuels (VLAN)

Un autre problème posé par la solution mixte (hiérarchique) indiquée dans le paragraphe précédent est le fait qu'un commutateur (équipement de niveau 2) ne permet pas de filtrer les accès selon des informations contenues dans des en-têtes de niveau réseau, transport ou application (adresse IP, numéro de port TCP, etc.). Pour cette raison, la solution souvent préférée était de remplacer dans cette architecture le commutateur par un routeur. Le routeur est capable non seulement d'appliquer un filtrage évolué, mais aussi de découper le « domaine de *broadcast* » en sous-domaines. Malheureusement, l'utilisation d'un routeur pour interconnecter des sous-réseaux présente aussi des désavantages :

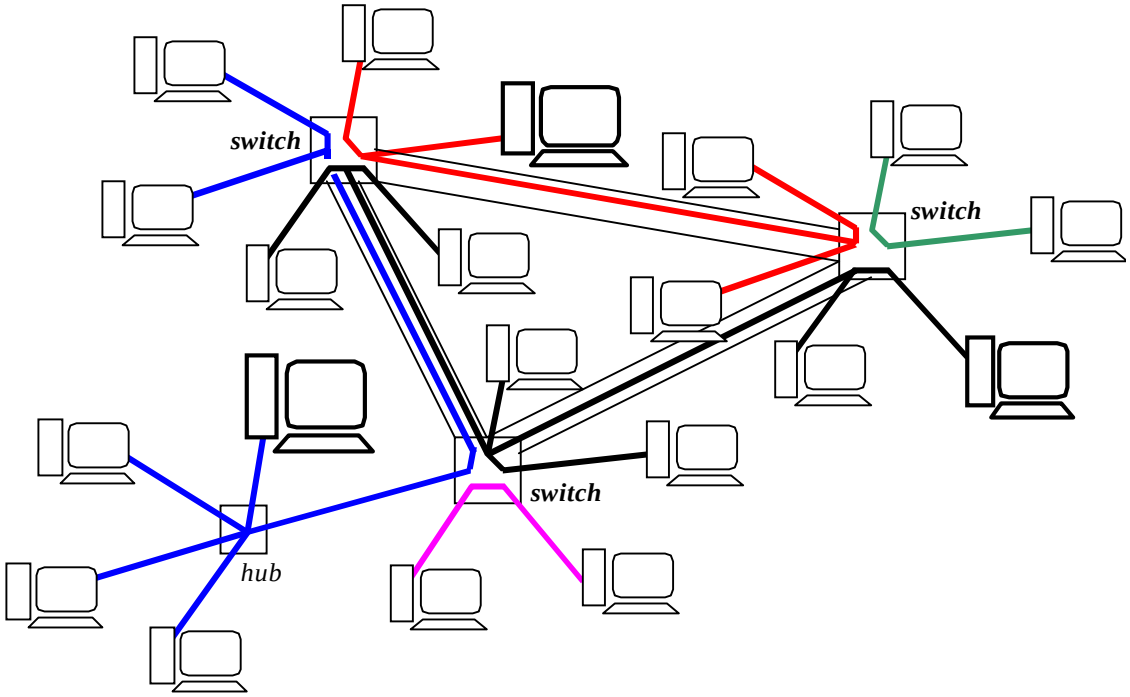
1° Une augmentation de la latence pour les communications entre les sous-réseaux : là où le commutateur se contentait d'aiguiller la trame, le routeur doit traiter l'en-tête du paquet contenu dans la trame.

2° Une gestion plus difficile des adresses IP, à cause du découpage en plusieurs domaines DHCP, chacun avec son serveur DHCP (à remarquer que certains routeurs peuvent assurer eux-mêmes la fonction de serveur DHCP).

3° Une flexibilité réduite : une machine qui se déplace entre deux sous-réseaux ne peut pas (en général) garder son adresse réseau (IP, IPX, etc.). Aussi, la connexion des machines à un concentrateur – et, par conséquence, le découpage en sous-réseaux – se fait en général sur un critère de proximité physique, qui ne correspond pas toujours au découpages organisationnels et donc aux besoins de filtrage d'accès.

Les problèmes sont donc complexes et les réseaux locaux virtuels (VLAN, *Virtual LAN*) ont été développés pour apporter des solutions (parfois partielles). Les VLAN permettent de déconnecter la structure logique des groupes de travail de la structure physique des réseaux supports. Si au départ les VLAN étaient basés sur des solutions propriétaires, le développement de normes IEEE (802.1Q pour le fonctionnement d'un VLAN de niveau 1 ou 2, 802.1p pour la gestion de priorités) permet de garantir aujourd'hui une certaine interopérabilité.

Un VLAN fait appel à des commutateurs de nouvelle génération, qui peuvent être tous du même type (par exemple Ethernet) ou non, et sont reliés entre eux par un réseau fédérateur à plus haut débit (Fast Ethernet, Gigabit Ethernet, FDDI, ATM). Les VLAN peuvent se limiter à un seul commutateur ou relier des machines distantes, connectées à des commutateurs différents. Aspect très important, les « domaines de *broadcast* » correspondent aux VLAN individuels. La figure suivante présente un exemple de configuration à trois commutateurs et 5 VLAN (l'appartenance à un même VLAN est représentée par des traits épais) :



Les concentrateurs de la solution précédente ont donc été remplacés par des commutateurs, reliés entre eux directement, sans passer par un nœud central supplémentaire. La connexion d'une machine à un commutateur se fait sur un critère de proximité physique, mais cela n'empêche pas cette machine de faire partie d'un même VLAN que des machines connectées à un autre commutateur. En plus des commutateurs qui permettent aux VLAN de fonctionner (*VLAN aware*), des équipements d'interconnexion qui ne connaissent pas l'existence des VLAN (*VLAN unaware*) peuvent continuer à être présents, à condition qu'ils soient dédiés à des VLAN individuels (le cas du *hub* dans le schéma précédent).

Quand un commutateur *VLAN aware* reçoit une trame en provenance d'une station, il ajoute une étiquette (*tag*) de format fixe à la trame. L'étiquette peut dépendre du numéro de port sur lequel la trame est arrivée, de l'adresse physique ou réseau de l'émetteur, ainsi que d'autres informations contenues dans les en-têtes à différents niveaux. Le commutateur est capable de prendre en compte des informations contenues dans des en-têtes de niveau supérieur à 2, et n'est donc plus vraiment un équipement de niveau 2 ; ceci n'est malheureusement pas sans conséquences sur la rapidité de commutation. L'étiquette permet d'indiquer l'appartenance de l'émetteur à un VLAN particulier et de donner éventuellement un niveau de priorité à la communication. La trame étiquetée est ensuite transmise au commutateur *VLAN aware* suivant, qui prend en compte les informations contenues dans l'étiquette pour lui appliquer le traitement approprié. Selon que le nœud suivant est *VLAN aware* ou *VLAN unaware*, l'étiquette reste associée à la trame ou est enlevée par le commutateur.

Tous les commutateurs qui participent à l'implémentation des VLAN sur un réseau support partagent des tables de filtrage (*filtering database*) qui indiquent l'appartenance des différentes machines aux VLAN et la localisation physique des machines. Ces tables sont en partie définies par l'administrateur du réseau (par exemple, appartenance à un VLAN), et en partie actualisées dynamiquement au cours du fonctionnement (par exemple, pour certains VLAN, en cas de déplacement d'une machine).

Avant de parler des différents types de VLAN, nous devons distinguer entre la **définition** des VLAN, liée à l'outil d'administration fourni, et le **fonctionnement** des VLAN, lié aux informations utilisées pour créer et traiter les étiquettes. Pour ce qui concerne le fonctionnement, nous pouvons distinguer entre :

1° Les VLAN de niveau 1 basés sur les numéros des ports. Dans ce cas, l'appartenance d'une machine à un VLAN dépend du numéro du port à travers lequel la machine est liée au commutateur. Cette technique est rigide, dans la mesure où chaque fois qu'une machine est déplacée, son appartenance à un VLAN doit être redéfinie. De plus, dans certaines configurations, il est difficile d'assurer une séparation stricte entre les VLAN : une machine peut éventuellement recevoir des trames qui ne sont pas destinées au VLAN auquel elle appartient.

2° Les VLAN de niveau 2 basés sur l'adresse MAC IEEE 802 (ou adresse « physique »). Dans ce cas, l'appartenance d'une machine à un VLAN dépend de l'adresse MAC (Ethernet, etc.) de la machine. Un très bon niveau de sécurité peut être assuré car l'adresse MAC est câblée dans la carte réseau de la machine, et ne peut donc pas être changée (pour changer de VLAN) par un utilisateur malveillant. De plus, si la machine est déplacée, les tables de filtrage peuvent être mises à jour de façon automatique, donc l'administration du VLAN est simplifiée. En revanche, définir au départ des VLAN à partir des adresses MAC est fastidieux, car les adresses MAC ne sont pas structurées et il faut donc les entrer une par une.

3° Les VLAN de niveau 2 basés sur l'identité du protocole de niveau supérieur (niveau 3), indiquée par l'en-tête IEEE 802.2. Cette technique peut être appliquée à condition d'avoir une hétérogénéité au niveau des protocoles de niveau 3. Elle présente un intérêt dans la mesure où elle permet de restreindre les « domaines de *broadcast* » aux machines dont les protocoles réseau font un usage fréquent du *broadcast* (IPX, Appletalk, etc.), diminuant ainsi l'impact négatif de celui-ci sur les autres machines du réseau.

4° Les VLAN de niveau 3 basés sur l'adresse de niveau 3 (IP, numéro réseau IPX, etc.). Toute l'adresse de niveau 3 ou une partie seulement (numéro de sous-réseau) peut être employée pour définir l'appartenance d'une machine à un VLAN. Même si la machine est déplacée, elle garde son adresse de niveau 3 et donc son appartenance à un VLAN. Dans la mesure où l'adresse de niveau 3 peut être modifiée par un utilisateur malveillant (pour changer de VLAN), cette technique peut poser des problèmes de sécurité. Aussi, l'obligation pour un commutateur de regarder l'adresse dans l'en-tête de niveau 3 augmente sa latence. En général, les commutateurs utilisés n'assurent aucune fonction de routage et font appel à l'adresse de niveau 3 uniquement pour déterminer le VLAN auquel appartient la machine. La définition d'un VLAN à partir des adresses de niveau 3 est simplifiée grâce à la structure logique hiérarchique de ces adresses (numéro de réseau, numéro de sous-réseau, … numéro de machine).

5° Les VLAN de niveau supérieur, basés sur différentes informations présentes dans les en-têtes successifs de niveau 3, 4, ou plus. La nécessité de rechercher des informations dans des en-têtes successifs, parfois de format variable, augmente de façon sensible la latence des commutateurs.

Si la définition de VLAN permet de séparer des groupes de machines du point de vue des accès physiques, la possibilité d' avoir des machines accessibles depuis (ou capables d' accéder à) plusieurs VLAN (serveurs, postes d' administration) est en général présente (bien que source de problèmes pour des VLAN de type 1).

Nous remarquerons que l'étiquetage peut être implicite (rien n'est ajouté à la trame) quand un seul commutateur intervient dans la communication, ou quand le fonctionnement du VLAN est de type 2, 3, 4 ou 5 (toute l'information est déjà présente dans la trame). Afin de diminuer la latence, il peut être intéressant d'utiliser des étiquettes explicites même pour ces VLAN : l'appartenance à un VLAN est choisie à l'entrée de la trame dans le premier commutateur (latence élevée), les autres commutateurs ne regardent plus que l'étiquette (latence minimale).

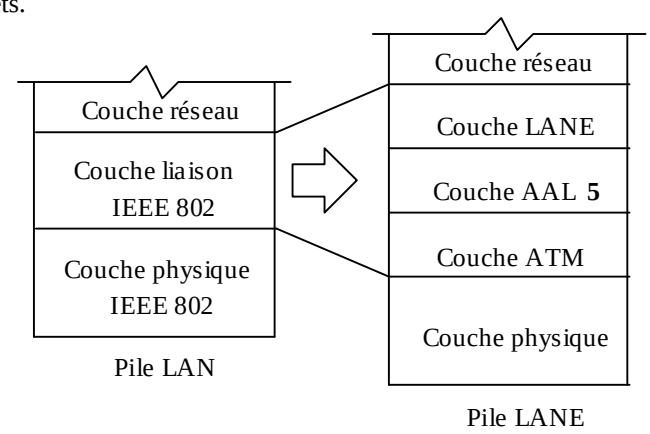
Le choix d'un type de fonctionnement et d'un mode de définition des VLAN doit être fait après une étude approfondie de l'utilisation du réseau. Une solution qui allie performance, flexibilité, sécurité et facilité d'administration est celle qui assure un fonctionnement de type 2 (basé sur l'adresse MAC), tout en permettant une définition des VLAN à partir des adresses réseau, plus faciles à gérer que les adresses MAC.

L'offre commerciale est arrivée à maturité et le choix est vaste ; les solutions non propriétaires sont à privilégier dans un milieu en évolution rapide, où l'interopérabilité est une contrainte forte.

6.3. Émulation LAN sur ATM (LANE)

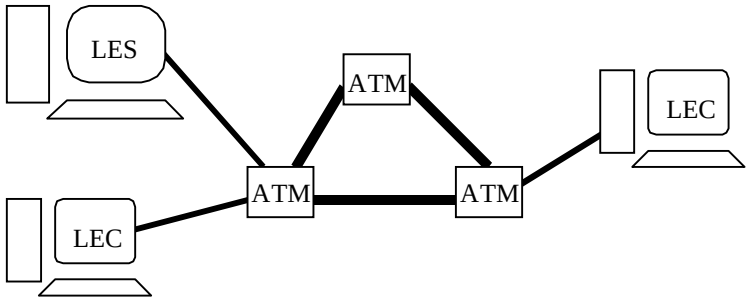
L'intérêt de l'émulation LAN est de permettre l'utilisation avec une nouvelle technologie de transport, ATM, des logiciels développés pour un LAN IEEE 802, et d'assurer ainsi une (éventuelle[2]) transition par étapes vers le tout ATM. Les composantes essentielles d'un LANE sont

Les LEC (*LAN Emulation Client*) sont des composants logiciels présents sur les machines connectées au réseau. Ces composants logiciels se situent dans la couche liaison du LAN émulé et permettent de garder l'interface avec les couches supérieures malgré le remplacement de la technologie support du LAN par l'ATM. Chaque LEC possède deux adresses, une adresse IEEE 802 MAC sur 48 bits et une adresse ATM sur 20 octets.



Le LES (*LAN Emulation Server*) est un composant logiciel (situé sur une machine connectée au réseau ou sur un commutateur ATM) qui garde les associations adresse IEEE 802 ↔ adresse ATM et qui répond donc à des requêtes LE-ARP (*LAN Emulation Address Resolution Protocol*) envoyées par les LEC.

Le BUS (*Broadcast and Unknown Server*) est un composant logiciel (situé sur une machine connectée au réseau ou sur un commutateur ATM) qui permet d'effectuer des envois de type *broadcast* sur le LANE. Pour cela, le BUS maintient des circuits virtuels (CV) bidirectionnels avec chaque LEC (pour le transfert des demandes de *broadcast*) et un CV unidirectionnel sous forme d'arbre en direction des LEC (pour l'envoi des *broadcast*).



Le LECS (*LAN Emulation Configuration Server*) est un composant logiciel (situé sur une machine connectée au réseau ou sur un commutateur ATM) qui garde les configurations des différents LANE présents sur le même réseau support ATM.

Quand la couche supérieure envoie une requête d'émission au LEC, accompagnée d'une adresse MAC de destination, le LEC doit trouver l'adresse ATM correspondante et ouvrir un circuit virtuel avec le destinataire. Si un circuit virtuel n'est pas déjà ouvert avec le destinataire et son adresse ATM n'est pas connue par le LEC, une requête LE-ARP est envoyée au LES du LANE.

A l'initialisation d'un LEC, celui-ci détermine l'adresse ATM du BUS grâce à une requête LE-ARP (avec l'adresse MAC de *broadcast*, FFFFFF) et met en place un CV bidirectionnel avec le BUS. Quand le LEC doit envoyer un message en *broadcast*, il l'envoie à travers le CV bidirectionnel au BUS qui se charge du *broadcast* (grâce au CV unidirectionnel qu'il maintient avec tous les LEC). De cette façon, il n'est pas nécessaire que chaque LEC qui désire envoyer un *broadcast* établisse un CV avec chaque autre LEC du réseau. Le désavantage

[2] Et de plus en plus improbable, en raison d'alternatives récentes, comme Gigabit Ethernet.

est qu'un LEC ne peut pas émettre un *broadcast* avant que toutes les cellules correspondant au *broadcast* d'un autre LEC aient été envoyées (le BUS emploie un CV unidirectionnel unique).

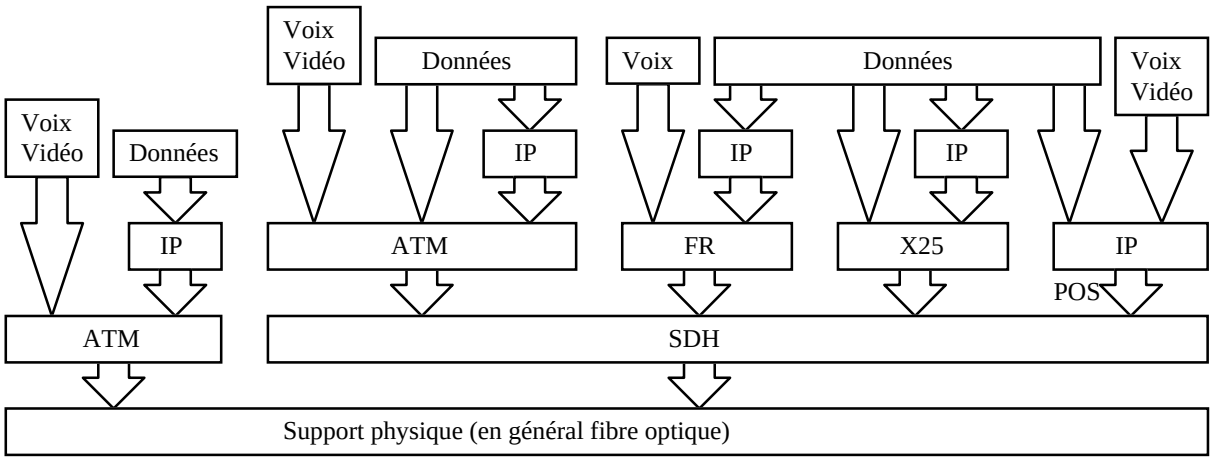
Le BUS peut aussi servir de relais pour l'envoi d'un message *unicast* à destination d'un LEC dont l'adresse ATM est pour le moment inconnue (le LES n'a pas encore répondu à la requête LE-ARP). Cette multiplicité des voies (voie directe LEC → LEC et voie indirecte LEC → BUS → LEC) peut contrarier le respect de l'ordre des messages, caractéristique générale des LAN IEEE 802. Une possibilité de *flush* demandé par le LEC émetteur a donc été prévue : dès qu'il obtient l'adresse ATM du LEC destinataire, l'émetteur envoie un *flush* au BUS qui le répercute en *broadcast* sur le réseau ; le LEC destinataire reconnaît son adresse, vide ses tampons de réception et répond à l'émetteur qui recommence l'émission, cette fois sur la voie directe.

7. IP et commutation

Nous avons étudié les techniques utilisées pour augmenter les performances et la flexibilité des réseaux locaux qui font appel à IP pour la couche réseau. Mais quelles sont les évolutions récentes des réseaux de transport longue distance des opérateurs ou des fournisseurs de services d'accès à Internet (ISP, *Internet Service Providers*) dans un monde des télécommunications où c'est Internet qui connaît la plus forte croissance, et où on assiste à une convergence « tout sur IP » ? Ces différentes évolutions sont en général liées à l'utilisation de la commutation pour assurer le trafic IP, et nous tenterons de donner d'autres éléments de réponse dans le reste de ce chapitre.

7.1. Empilement des protocoles et technologies

Avant d'étudier l'évolution des techniques utilisées pour transporter les paquets IP sur les réseaux longue distance, regardons l'empilement actuel des protocoles et technologies sur les réseaux des opérateurs. Le schéma suivant représente une partie des possibilités offertes ; les différents éléments d'adaptation n'ont pas été représentés.



Rappelons que pour ce schéma nous nous sommes volontairement limités au réseau de transport (WAN, *Wide Area Network*), car au niveau des boucles d'accès et même des réseaux locaux la diversité est encore plus importante.

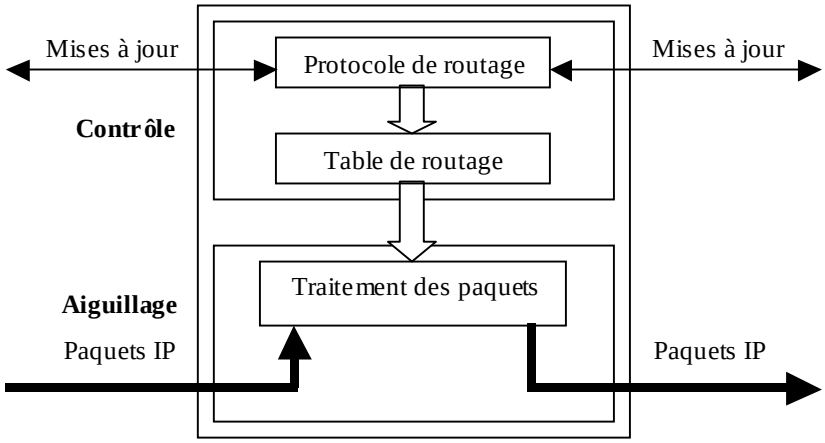
Il faut noter que l'introduction de couches intermédiaires produit une diminution du débit utile (données utiles transportées) pour un débit physique donné. Par exemple, avec ATM comme couche intermédiaire, cette diminution est d'au minimum 9,4% (en ignorant les couches d'adaptation à ATM).

La tendance actuelle des opérateurs français est d'utiliser ATM comme intermédiaire entre les différents trafics assurés (IP compris) et l'infrastructure SDH (*Synchronous Digital Hierarchy*, appellation internationale, ou SONET, *Synchronous Optical NETwork*, appellation américaine). Aux Etats-Unis, en revanche, profitant de l'évolution de IP au niveau des options de qualité de service (IPv6) et de l'introduction de protocoles de réservation (RSVP), certains opérateurs parmi lesquels Qwest, MCI ou Sprint ont déployé une solution IP natif sur SDH (*Packets Over SONET*, POS).

L'essor des techniques de multiplexage de longueurs d'ondes (HDWDM, *High Density Wavelength Division Multiplexing*) ouvre une nouvelle voie, dont l'exploration a commencé : l'utilisation sans intermédiaire (ATM, SDH) de IP sur la fibre optique. Parmi les techniques de commutation présentées dans ce qui suit, *Tag Switching* et MPLS sont les mieux adaptées à ce type de transmission.

7.2. Routage rapide

Remarquons, au préalable, que les fonctions d'un routeur peuvent être groupées en deux composantes bien distinctes et complémentaires : la composante de contrôle, qui implémente les algorithmes de création et gestion des tables de routage (OSPF, *Open Shortest Path First*, IS-IS, *Intermediate System-to-Intermediate System*, BGP-4, *Border Gateway Protocol*), et la composante d'aiguillage (*forwarding*), qui prend en charge les paquets entrants et les envoie sur le bon port de sortie, en respectant les tables de routage :



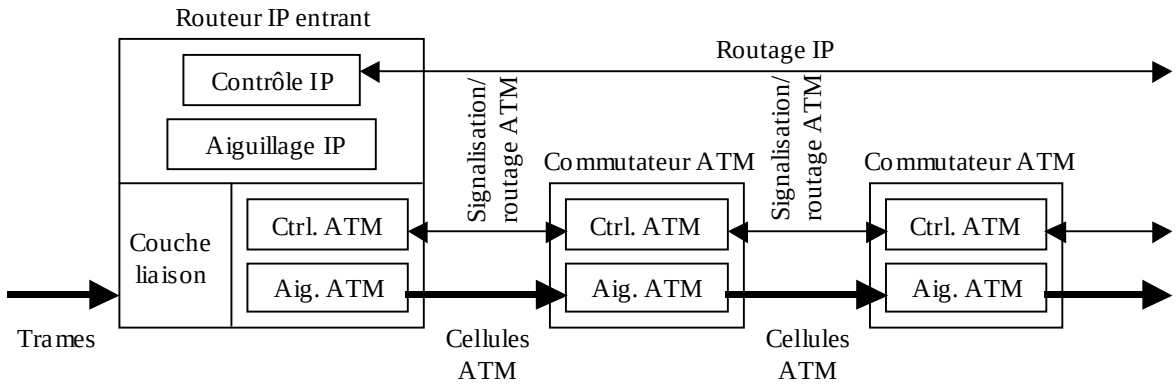
Représentation schématique d'un routeur

Une première évolution est la séparation entre ces deux composantes, et l'implémentation de la seconde (aiguillage) dans le matériel, grâce au développement de circuits ASIC (*Application Specific Integrated Circuits*) dédiés. On appelle parfois « commutation IP » cette approche, car elle permet d'accélérer le traitement des paquets IP (par rapport à une implémentation logicielle de l'aiguillage), mais ce nom n'est pas tout à fait approprié dans la mesure où la couche liaison continue à être présente comme une couche distincte.

7.3. IP sur ATM et ses difficultés

Le déploiement de ATM sur l'infrastructure des réseaux de télécommunications longue distance permet aux opérateurs et aux fournisseurs de services non seulement d'intégrer facilement des trafics de nature différente (voix, données, diffusion vidéo), mais aussi de gérer de façon plus transparente (donc plus efficace) les ressources disponibles et de garantir la qualité des transferts IP sur ATM.

La première solution employée dans le transport des paquets IP sur ATM (*IP-over-ATM*) maintient la séparation entre la couche réseau (IP) et la couche liaison (ATM). Cette solution fait appel aux protocoles habituels pour le routage IP et aux protocoles du Forum ATM pour la signalisation et le routage ATM :

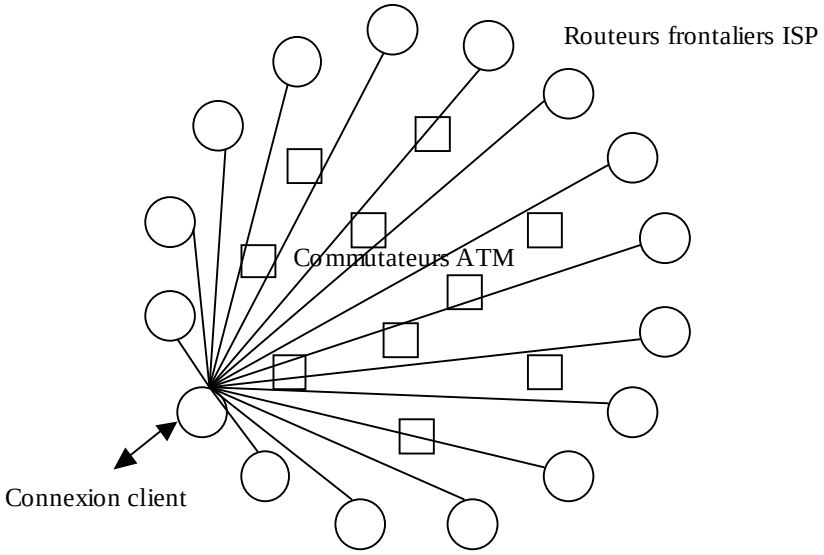


Cette solution permet une mise en œuvre rapide sur la base de protocoles normalisés et fait profiter les fournisseurs de services d'accès IP des avantages de ATM en termes de qualité de service (QoS) et de gestion du trafic (*Traffic Engineering*). Malheureusement, la solution présente aussi plusieurs défauts, et non des moindres :

1° L'encapsulation des paquets dans des cellules ATM (5 octets d'en-tête pour 48 octets de données) mène à une perte de bande passante d'au minimum 10% (en général plutôt proche de 20%).

2° L'administration est très lourde, dans la mesure où deux niveaux différents (ATM et IP), indépendants, doivent être gérés.

3° Au niveau IP, chaque routeur frontalier voit tous les autres routeurs frontaliers comme des voisins directs. Des circuits commutés permanents ATM sont établis entre chaque paire de routeurs pour assurer la circulation des paquets IP des utilisateurs et des informations de contrôle pour la gestion des tables de routage IP. Cela mène rapidement à une explosion du nombre de circuits virtuels à maintenir (leur nombre est proportionnel au carré du nombre de routeurs), or les étiquettes de flot ATM ne sont codées que sur 24 bits (12 pour VP et 12 pour VC) ; $\sqrt{2^{24}} = 2^{12} = 4096$.



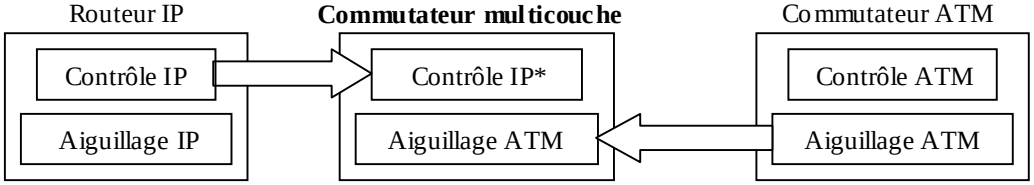
Représentation du réseau d'un ISP ayant adopté *IP-over-ATM*

Aussi, l'existence d'un voisinage direct aussi large pour un routeur pose des problèmes aux protocoles de gestion des tables de routage (IGP, *Interior Gateway Protocol*), qui n'ont pas été prévus pour une telle situation.

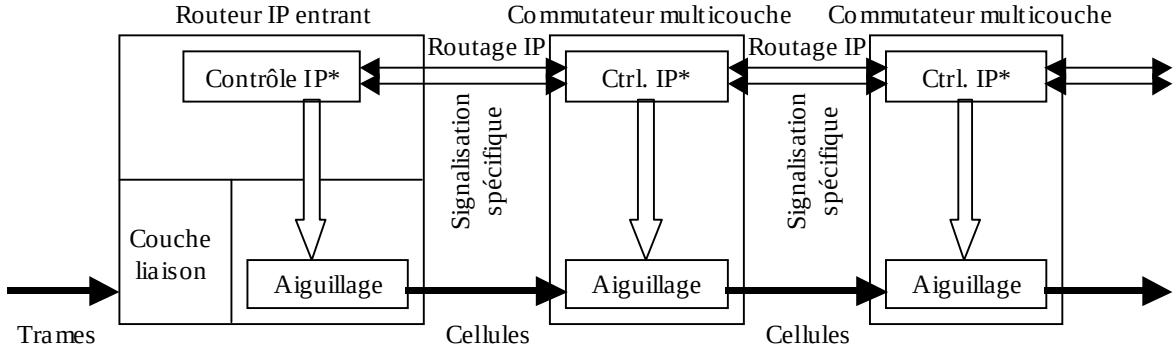
En conclusion, si la technique « IP sur ATM » présente certains avantages, elle pose notamment des problèmes d'échelle (*scalability*) qui restreignent son domaine d'application.

7.4. Évolutions – *MultiProtocol Label Switching*

Les différentes solutions propriétaires essayent d'adapter au monde IP les avantages de ATM dans le domaine de la qualité de service (QoS, *Quality of Service*) et de la gestion explicite du trafic. Le principe est d'intégrer dans un même équipement (appelé commutateur multicouche) la partie de contrôle IP et un aiguillage de type ATM (avec les adaptations qui s'imposent) :



Dans ces nouveaux équipements on retrouve donc les tables de routage et les algorithmes qui les gèrent (OSPF, IS-IS, BGP, etc.), mais aussi les tables d'association qui permettent à l'aiguillage de type ATM de fonctionner. Ces tables d'association sont gérées à partir des tables de routage, en utilisant des algorithmes de signalisation et de distribution d'étiquette spécifiques (autres que ceux prévus par le Forum ATM) :



MPLS (*MultiProtocol Label Switching*) doit permettre de normaliser la solution et d'assurer ainsi l'interopérabilité des équipements des différents constructeurs. Si MPLS s'inspire de ATM pour la partie commutation, le mécanisme d'assignation et distribution des étiquettes pour la définition des routes commutées est nouveau (différent de celui proposé pour ATM par le Forum ATM) et dirigé par les algorithmes classiques de l'environnement IP (OSPF, IS-IS, BGP). MPLS peut fonctionner avec pour LSR des commutateurs ATM (grâce à une mise à jour logicielle) ou des équipements non ATM (paquets sur SDH/SONET, ou directement sur la fibre optique). Si les trames de niveau 2 utilisées pour le transport des paquets IP contiennent dans l'en-tête un champ étiquette de longueur suffisante (comme c'est le cas pour ATM ou *Frame Relay*), MPLS l'emploie ; sinon, MPLS insère entre l'en-tête de niveau 2 et l'en-tête IP son propre en-tête, qui contient une étiquette, un champ de classe de service (CoS) et un champ durée de vie.

Regardons maintenant les étapes nécessaires pour transporter un paquet IP sur le réseau d'un ISP qui fait appel à MPLS. A son entrée sur le réseau, un paquet IP est associé par le routeur entrant (*ingress* LSR, *Label-Switching Router*) à une classe d'équivalence (FEC, *Forwarding Equivalence Class*) qui rassemble des trafics ayant comme destination le même sous-réseau et possédant les mêmes exigences de QoS. De nouveaux services peuvent être mis au point par une simple modification des éléments à prendre en compte pour l'assignation des paquets IP à un FEC. Chaque FEC est ensuite associé à un chemin commuté dans le réseau (LSP, *Label-Switched Path*), de façon automatique ou assistée (*traffic engineering*), selon ses exigences et la disponibilité des ressources dans le réseau. Dans chaque LSR qui suit, le LSP est défini par une succession de correspondances étiquette entrante $\rightarrow$ étiquette sortante (similaire à la définition d'un circuit virtuel ATM par des paires de VPI/VCI). Toutefois, dans un réseau complexe le LSP n'est pas de bout en bout, sinon les problèmes d'échelle de *IP-over-ATM* se manifestent ; là où le premier LSP s'arrête, le paquet IP subit une nouvelle opération de routage qui permet de l'associer à un nouvel LSP. Enfin, à l'arrivée sur le LSR sortant (*egress* LSR) le paquet IP est transmis selon la procédure de routage IP classique.

Les LSP sont définis à partir des informations de routage IP, des demandes transmises par des protocoles de réservation de ressources comme RSVP (*Ressource reSerVation Protocol*) ou en réponse à des interventions explicites pour la gestion du trafic.

Si la plupart des solutions propriétaires sont basées sur le même principe d'intégration du routage IP et de la commutation ATM, elles présentent aussi des différences :

- *IP Switching* (proposée par Ipsilon/Nokia) et *Cell Switching Router* (proposée par Toshiba) définissent un LSP pour un trafic après le transfert (par des opérations de routage classique) de plusieurs paquets IP appartenant au trafic. Si ce mécanisme de création de LSP dirigée par les données (*data-driven*) permet d'éviter la charge due à la gestion de LSP pour des trafics de très faible volume, il pose aussi des problèmes de latence (la création d'un LSP pour un trafic n'est pas immédiate) et d'échelle (le trafic de contrôle nécessaire est proportionnel au nombre de trafics de données).
- *IP Navigator* (proposée par Cascade/Ascend/Lucent), *Aggregate Route-based Ip Switching* (ARIS, proposée par IBM) et *Tag Switching* (proposée par Cisco), définissent les LSP à partir des informations de contrôle (*control-driven*), comme MPLS. *IP Navigator* est limitée à l'utilisation d'une encapsulation dans des cellules ATM. ARIS et *Tag* Switching sont les solutions les plus proches de MPLS.

Pour d'autres détails concernant l'utilisation de la commutation pour le transport des paquets IP voir par exemple http://infonet.aist-nara.ac.jp/member/nori-d/mlr/.

Bibliographie

******* ***01 Réseaux*****, supplément mensuel de l'hebdomadaire *01 Informatique*, collection 1998-2001.**

******* ***Décision Micro & Réseaux*****, supplément mensuel de l'hebdomadaire *Le Monde Informatique*, collection 1998-2001.**

Black, U. (1994) *TCP/IP and related protocols*, McGraw-Hill Series on Computer Communications, New York, 1994 (en bibliothèque, 628/5882).

En anglais. Relativement complète et souvent détaillée. Parle peu de la programmation mais présente les relations entre TCP/IP et d'autres protocoles. Les explications ne sont pas toujours claires.

Delahousse, A. (1997) *Câblage haut débit : voix, données, images*, Hermès, Paris, 1997 (en bibliothèque, 601/7579-0), 160 p.

Présentation rapide de supports de transmission et de leurs caractéristiques, présentation plus détaillée de normes de câblage et discussion de la conception ainsi que de la mise en oeuvre du câblage. Présentation très sommaire de quatre études de cas.

Feit, S. (1996) *TCP/IP: Architecture, Protocols and Implementation, with IPv6 and IP security*, McGraw-Hill, Inc., New York, 1996 (en bibliothèque, 628/6600).

En anglais. Complète et détaillée. Présente aussi la programmation (utilisation des sockets). Ne parle pas d'autres protocoles.

Händel, R., Huber, M. N., Schröder, S. (1995) *Comprendre ATM*, Addison-Wesley, Paris, 1995.

Très bonne présentation, avec des références aux normes existantes (la normalisation étant encore incomplète).

Huitema, C. (1995) *Le routage dans l'Internet*, Eyrolles, Paris, 1995 (en bibliothèque, 628/5584).

Présentation détaillée des algorithmes de routage (construction et mise à jour des tables de routage) utilisés en conjonction avec IP comme RIP, OSPF, EGP, BGP, CIDR.. Bibliographie très fournie.

Hunter, P. (1994) *Network Operating Systems: Making the Right Choices*, Addison-Wesley, Paris, 1994 (en bibliothèque, 628/5888).

Présente rapidement différents systèmes d'exploitation réseau (SER) et leur évolution (Novell Netware, LAN Manager, UNIX, OS/2, LANtastic, AppleTalk, etc.). Donne des critères de choix entre SER et présente quelques éléments concernant l'interopérabilité.

Millet, M. (1987) *Transmission et réseaux locaux : architecture IEEE 802*, Masson, Paris, 1987 (en bibliothèque, 628/3688).

Présentation des aspects physiques des réseaux locaux, des techniques d'accès les plus employées et des propositions IEEE 802. Assez peu de détails et d'explications sur IEEE 802...

Montagnier, J.-L. (1998) *Pratique des réseaux d'entreprise*, Eyrolles (en bibliothèque, 628/7355-0), 525 p.

Présentation des réseaux locaux ou étendus utilisés, en laissant de côté le modèle OSI. Détails sur Ethernet et Token-Ring, beaucoup moins sur FDDI ou ATM. Les protocoles de niveau réseau sont mentionnés, certains très rapidement (NetBIOS, AppleTalk, DECNet)... Bonne introduction aux réseaux télécom et au réseau téléphonique (classique et RNIS). Les études de cas détaillées à la fin de l'ouvrage sont ***très utiles*** *!*

Pujolle, G., Seret, D., Dromard, D., Horlait, E. (1989) *Réseaux et Télématique*, Tome 2, Eyrolles, Paris, 1989 (en bibliothèque, 628/5095-2).

Tour d'horizon relativement vaste des réseaux en général et des services disponibles (en 1986...). Survol rapide des réseaux locaux (presque exclusivement les aspects technologiques). Certains chapitres sont dépassés. Assez inégale.

Rolin, P. (1995) *Réseaux haut débit*, Hermès, Paris, 1995 (en bibliothèque, 628/5891).

Présentation à jour de différents protocoles et types de réseaux : Ethernet 100 Mbps, FDDI, Relais de trame, ATM (env. 200 pages), IPv6 (env. 50 pages). Description rapide de RIP et OSPF, algorithmes de construction de tables de routage employés avec IP. Présentation d'architectures d'interconnexion : relais de trame — ATM, LAN — relais de trame, X25 — relais de trame, LAN — ATM, émulation LAN.

Servin, C., Ghernaouti-Hélie, S. (1991) *Les hauts débits en télécoms*, InterEditions, Paris, 1998 (en bibliothèque, 628/8681.0).

Présentation rapide de Ethernet et Token Ring, des VLAN, de FDDI et DQDB, Frame Relay, et ATM.

Toutain, L. (1999) ***Réseaux locaux et Internet : des protocole*** *à l'interconnexion*, Hermès, Paris, 1999 (en bibliothèque, 628/8548.0).

Présentation des réseaux Ethernet et Token Ring, des protocoles IP et TCP, de protocoles de routage. Quelques détails sur les sockets. Présentation assez rapide des VLAN.