

Les réseaux

Cette section s'attache à décrire et expliquer les réseaux, leur structure et leur fonctionnement général. Il n'est pas question ici de traiter des protocoles réseaux ; ces derniers disposent en effet d'une section qui leur est propre.

- introduction à la notion de réseau
- aspects sociaux et économiques des réseaux
- topologie des réseaux locaux
- classification des réseaux
- principaux concepts des réseaux
- le logiciel de réseau
- le modèle OSI
- fonctionnement et critique du modèle OSI
- le modèle TCP/IP

► Introduction

Introduction

Tout comme le 19^{ème} siècle fut le siècle de la machine à vapeur, le 20^{ème} fut le siècle de la collecte, du traitement et de la distribution de l'information. Ce dernier siècle vit donc l'apparition et le déploiement d'un réseau téléphonique à l'échelle planétaire, de l'invention de la radio et de la télévision, des satellites de télécommunication et l'explosion de l'informatique.

Une des caractéristiques de ces technologies est qu'elles ont eu tendance, petit à petit, à converger. On est progressivement passé des systèmes centralisés, avec généralement un unique ordinateur central que plusieurs informaticiens "alimentaient" de leurs programmes, à une interconnexion globale des équipements, aux réseaux informatiques et systèmes répartis qui permettent de répartir les puissances de calcul et de stockage.

Quelques dates

Début des années 60 :

systèmes de réservation de places d'avion
répartition des opérations bancaires

1969 : naissance du réseau ARPA aux Etats-Unis qui relie les principaux centres de recherche américains. C'est la première architecture formalisée en couche d'abstraction.

1974 : IBM invente SNA (system Network Architecture) qui permet la communication entre un système central et des périphériques distants

1975 : écriture de la norme Ethernet par Intel et Xerox

1980 : début de l'Internet

1982 : IBM commercialise la norme Token Ring

1985 : l'IEEE publie les normes de la famille 802 (Ethernet et Token Ring notamment)

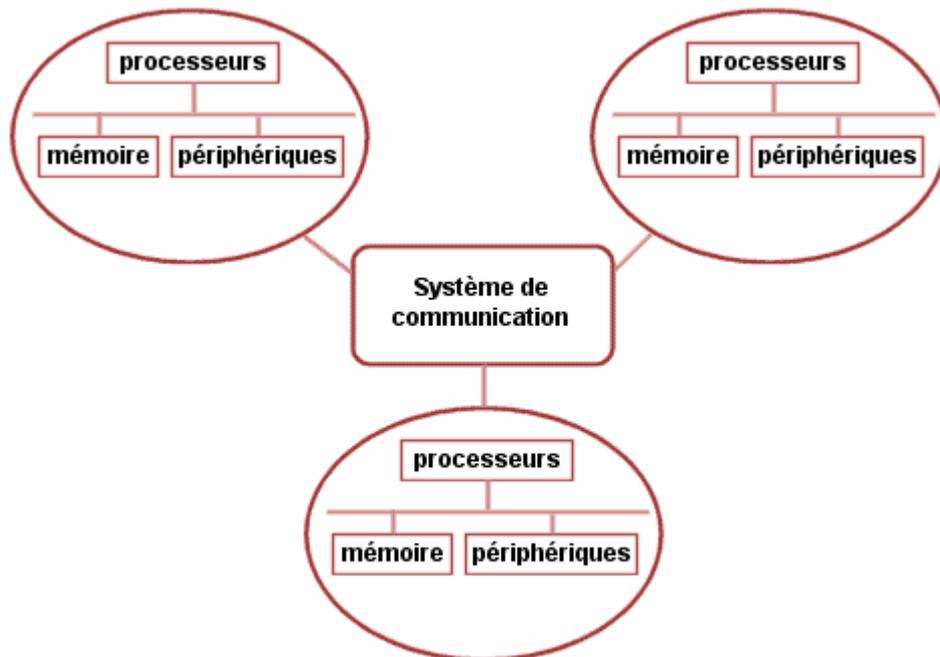
1989-1991 : naissance du Web, l'application majeur de l'Internet

1990 : apparition du RNIS (Réseau Numérique à Intégration de Services)

A partir du début des années 80, les réseaux du monde entier se sont progressivement interconnectés pour former **Internet**. Internet signifie "interconnexion de réseaux" ; c'est un réseau mondial fournissant toute une série de services, série qui ne se limite pas qu'au Web, comme on aurait tendance à le croire de nos jours.

Définitions

Un **réseau d'ordinateurs** est un ensemble d'ordinateurs **indépendants** appelés **noeuds**, sites ou **ordinateurs hôtes**, interconnectés et pouvant échanger des informations par des **lignes de communication** :

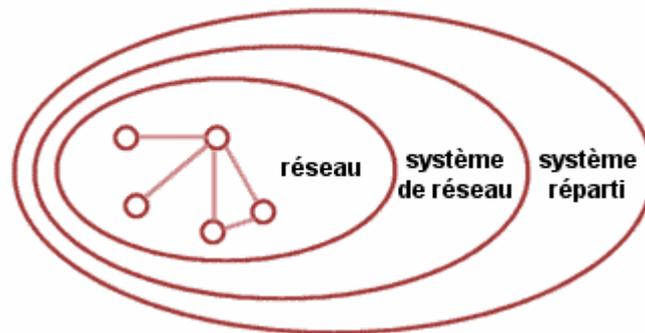


L'autonomie des ordinateurs exclut toute relation du type maître/esclave ; est exclu, donc, de l'appellation "réseau" tout système composé d'une unité de contrôle et de plusieurs unités esclaves (ordinateur ayant des imprimantes et des terminaux distants).

On appelle **système de réseau** ou **logiciel de réseau** un ensemble de modules logiciels qui gèrent les communications entre les sites. Ce logiciel est implanté sur chaque site du réseau.

On distingue les réseaux d'ordinateurs des systèmes répartis ou distribués : un **système réparti ou distribué** est un ensemble d'ordinateurs **indépendants** interconnectés possédant un **logiciel** tel que la répartition en ordinateurs autonomes est **transparente et invisible** à l'utilisateur. Concrètement, lorsque l'on veut exécuter un programme, le système d'exploitation se débrouille pour sélectionner le meilleur processeur, trouver et transporter les fichiers d'entrée vers ce processeur et mettre les résultats au bon endroit.

Dans le cas d'un réseau, tout est fait *explicitement* (on se connecte d'abord explicitement à une machine donnée, on soumet explicitement une commande à exécuter sur cette machine donnée, on doit explicitement déplacer les fichiers...), alors que dans le cas d'un système réparti, tout est réalisé automatiquement par le système d'exploitation. La différence est donc une différence au niveau du logiciel système, et non pas au niveau du matériel (type d'ordinateur, type de connexion...). On peut résumer ceci par le schéma suivant :



Objectifs et utilisation des réseaux

Les réseaux ont été et sont toujours développés pour un certain nombre de raisons. Il y en a en fait 4 principales.

Partage des ressources

Les réseaux permettent de rendre accessible un certain nombre de ressources (logiciels, bases de données, imprimantes...) indépendamment de la localisation géographique des utilisateurs. Le partage des données commerciales d'une entreprise en est une illustration : chaque employé d'une multinationale peut accéder aux derniers comptes de résultat de l'entreprise.

Augmentation de la fiabilité et des performances

Les réseaux permettent par exemple de dupliquer en plusieurs endroits les fichiers vitaux d'un projet, d'une entreprise ; en cas de problème, la copie de sauvegarde est immédiatement disponible. L'augmentation des performances vient également du fait qu'il est relativement facile d'augmenter les performances d'un système en réseau en ajoutant tout simplement un ou deux autres ordinateurs supplémentaires. Ce dernier point associé à un constat économique (voir objectif suivant) rend presque obsolètes les grosses installations.

Réduction des coûts

En effet, les ordinateurs individuels coûtent bien moins cher que les gros systèmes centralisés (1000 fois moins environ), et ce pour une baisse des performances d'à peine un facteur 10. Les réseaux permettent également de réagir plus vite à certains événements (un appel d'offres par exemple), et donc faire gagner (ou économiser) de l'argent.

Accès à l'information

Avec les réseaux et en particulier Internet, il est très facile de s'informer sur toute sorte de sujets très rapidement. Ce dernier objectif joue en fait un rôle capital dans l'utilisation que les gens ont des réseaux. C'est peut-être même l'utilisation principale aujourd'hui.

Autres utilisations

Au delà de ces quatre points, il existe quelques autres objectifs aux réseaux, mais ces objectifs sont apparus récemment avec la démocratisation des réseaux et l'émergence d'Internet notamment, et ne correspondent pas véritablement à un besoin des professionnels. Les réseaux vont servir par exemple de support pour des jeux interactifs et autres divertissements, ainsi que de médium de communication. Ces dernières raisons ont des conséquences sociales relativement importantes car elles influencent énormément le comportement des gens.

► Aspects sociaux et économiques

Aspects sociaux

Une réponse à des besoins vitaux

Les conséquences de l'utilisation des réseaux sont loin d'être négligeables. Ces conséquences sont directement les utilisations que les gens en ont, utilisations prévues ou inattendues. Le point important est que les réseaux sont devenus le support de deux besoins vitaux de l'homme : le besoin de rassemblement et de communication ainsi que le besoin de se divertir.

La réponse au besoin de rassemblement et de communication va de soi, car c'est tout simplement dans ce but que les réseaux ont été inventés. Initialement, les réseaux fournissaient le support de communication pour les entreprises et les plus fortunés seulement, puis progressivement, le téléphone puis l'Internet sont devenus accessibles à tous. Cette facilité de communication a quelques conséquences intéressantes : certains individus ont maintenant tendance à s'enfermer chez eux car leur besoin de contact humain est entièrement tari par les contacts virtuels qu'ils entretiennent sur Internet. Heureusement, ces cas sont rares. Autre exemple typique : le téléphone portable qui a complètement libéré l'homme. Chaque instant de notre vie peut être partagé avec n'importe qui, quelque soit l'endroit (même le cinéma ou le théâtre, ce qui est fort regrettable : certains ne savent pas se modérer et voir où sont les limites de la liberté). De manière plus générale, la virtualisation totale des rapports humains et l'affranchissement des limites géographiques ont des conséquences (parfois lourdes) que l'on a encore du mal à mesurer. L'aspect "réponse au besoin de se divertir" n'est apparue que depuis peu ; à vrai dire au début des années 90 avec la naissance du Web. Le Web n'a fait qu'apporter un côté ludique à l'Internet, et c'est peut-être finalement ce côté divertissant et novateur qui a fait ce que le Web est devenu aujourd'hui. Les réseaux ont fait naître une nouvelle catégorie de jeux, les jeux en réseaux, et ces derniers apportent une dimension interactive sans précédents, à l'échelle de la planète.

L'homme face à ses pairs

Le point faible des réseaux vient hélas de la réponse qu'ils apportent à ces deux besoins : la perversion de certains individus est en train de tout remettre en cause...

L'utilisation illicite des réseaux et notamment Internet pour convoier et transmettre des informations prohibées ou que la morale réprouve (la pornographie ou le nazisme, voire plus grave, la pédophilie) a fait changer d'avis les gens qui ne voyaient que du bien dans ces technologies. Le piratage de logiciel ou d'oeuvres (musique, films) est devenu relativement commun et des entreprises qui avaient bâti des empires dessus commencent à se poser des questions. De la même façon, on est en droit de s'interroger sur l'équilibre précaire de certains marchés et économies, telles que ceux de la musique et des arts picturaux.

A travers ces actes de barbarie, certains cherchent en fait à atteindre une reconnaissance sociale. A défaut de devenir une pop star, certains préfèrent utiliser leurs compétences en informatique pour pirater des sites ou les lignes de téléphone, voire en programmant des virus informatiques destructeurs. Cela peut aussi être par simple jeu intellectuel ou pour la compétition. Les raisons profondes de tels actes sont très variées et font ressortir un problème anthropologique fondamental : quelle est ma position dans la société, par rapport aux autres. Dans la plupart des cas, même si ce n'était pas malveillant à la base, on peut ressentir un profond besoin d'expression vis à vis de l'extérieur : j'ai besoin de parler, dire que j'existe, alors me voilà. C'est souvent une sorte de petit jeu sournois : il est très facile de faire des choses de manière anonyme sur Internet, mais finalement, on ne veut pas être si anonyme que ça, donc on s'invente un nom de code et on pirate des sites. Vouloir devenir célèbre n'est pas gênant en soit ; ce qui est répréhensible, c'est la façon de le devenir : en nuisant à autrui.

Deux des facteurs tendant à favoriser ces actes sont l'augmentation de la bande passante (il devient de plus en plus facile de transférer de grosses quantités de données) et le nombre croissant d'utilisateurs. Ce dernier facteur est particulièrement intéressant puisqu'il signifie que l'Internet rassemble de plus en plus de gens susceptibles de réaliser des actes de piratage. Avec le taux de croissance que connaît l'Internet aujourd'hui, on comprend pourquoi les grosses compagnies misent énormément d'argent dans des systèmes de sécurité de plus en plus sophistiqués.

La société en question

Les questions que nous sommes en droit de nous poser aujourd'hui sont d'un intérêt capital, puisqu'il s'agit par certains abords de remettre en cause tout une partie des fondements de la société actuelle : l'accès à l'information est un droit, alors qui doit-on punir pour la mauvaise utilisation des ressources ? D'autant plus qu'il est relativement facile d'usurper une identité (si ce n'est pas devenir complètement anonyme) sur Internet. Puisqu'il est difficile de savoir qui est responsable d'actes licencieux et qu'un crime ne peut rester impuni, doit-on sanctionner

l'hébergeur du site qui a été trop "laxiste" en autorisant ce site, ou plus simplement le fournisseur d'accès Internet qui a tout bonnement favorisé l'accès au site ?

La question des libertés individuelles est également au centre des débats. L'exemple typique est le travailleur qui a accès à l'Internet depuis son lieu de travail : quels sites peut-il visiter ? Les sites d'information, certes, mais au delà ? Des logiciels de filtrage apparaissent mais sont loin d'être efficaces. Ce dernier marché est, selon les spécialistes, sur le point d'exploser...

Une des toutes dernières conséquences sur notre société concerne le travail. Certes, les réseaux favorisent et améliorent les conditions de travail, mais de nouveaux types de travail commencent à apparaître (le télétravail par exemple, ou le travail à domicile) : les réseaux apportent en fait un aspect liberté que le travail n'avait pas auparavant. C'est très bien, et il faut souhaiter que cela continue, mais une fois de plus, la responsabilité de l'individu face à cette liberté retrouvée est très grande, et il ne s'agirait pas que les télétravailleurs parce qu'ils travaillent depuis chez eux ne soient plus efficaces...

Aspects économiques

Initialement, les réseaux devaient apporter une solution simple et **économique** aux besoins en communication des entreprises. Seulement, les réseaux ne font plus économiser de l'argent et du temps, ils en font gagner.

Il existe deux principaux marchés dans le monde des réseaux : le marché de la fourniture du support des réseaux (c'est la charge des opérateurs) et le marché de la fourniture de services sur ces réseaux.

Le marché du support est maintenant peut-être fermé ; seuls les plus gros opérateurs sont en mesure d'apporter leur pierre à l'édifice ou d'avoir une position de leader sur ce créneau. On raisonne en terme de milliards de dollars ou d'euros. C'est un marché relativement concurrentiel où l'on fait la chasse aux octets. Un des gros enjeux est la facturation : les opérateurs voudraient facturer à la quantité d'informations, mais avec les technologies actuelles, c'est relativement délicat. Ce marché est également le marché sur lequel les innovations techniques sont les plus importantes mais aussi les plus longues à mettre en place et cela ennuie fortement les opérateurs qui, à défaut de pouvoir proposer la meilleure qualité avant les autres, doivent trouver des parades tarifaires.

Le second marché important des réseaux est **le marché des services**. Il s'agit tout simplement d'apporter le ou les services les plus rentables et utiles sur les réseaux. Comme pour beaucoup de choses maintenant, la création de besoin dans ce domaine est très importante. Ces services vont des services pour les professionnels aux services pour les particuliers. Les services pour les professionnels passent principalement par ce qu'on appelle les **systèmes d'information**. Il est question ici de trouver le moyen le plus rationnel d'organiser la circulation et l'utilisation de l'information au sein d'une entreprise. La solution passe par l'utilisation des réseaux et surtout le service que l'on peut développer avec. Une autre catégorie de services pour professionnels, mais cette fois plus spécifique aux opérateurs, est la facturation : il est en effet difficile de facturer au volume car on ne peut pas facilement deviner à qui appartient tel ou tel paquet d'information. La dernière catégorie de services que l'on peut fournir est principalement à destination des particuliers. C'est en fait tout le reste : achats en ligne, mais aussi consultation de ses mails sur son téléphone portable, factures téléphoniques détaillées, localisation par balise GPS... Beaucoup de ces services existent déjà depuis longtemps (en particulier pour ce qui est du téléphone). Cependant, depuis quelques années, les services de proximité via le web sont devenus très importants et font l'objet d'une concurrence acharnée. De plus en plus d'utilisateurs se laissent séduire par les achats groupés sur Internet ou bien d'autres choses ; c'est clairement un marché prometteur.

En bref, les réseaux constituent un formidable marché économique qui est loin de s'essouffler, tant au niveau des fournisseurs de support qu'au niveau des services fournis sur ces réseaux. C'est hélas aussi un marché relativement fragile comme ont pu le démontrer courant 2000 la mini crise boursière ainsi que les pertes économiques engendrées par les attaques de déni de service (DoS) sur les gros sites marchands de l'Internet. Ce marché franchira une étape importante lorsque les pays en voie de développement (en particulier l'Afrique et l'Amérique du sud) commenceront à rentrer dans le jeu des télécoms : les entreprises n'auront plus la possibilité de s'étendre à volonté.

L'avenir des réseaux

Une chose est claire : les réseaux et toutes les technologies adjacentes ont de beaux jours devant eux ! Augmentation de la bande passante, de plus en plus d'utilisateurs (i.e. de clients)... autant d'éléments motivant les entreprises dans la réalisation de solutions techniques innovantes. Le but

n'est plus de proposer un moyen de connecter les gens, il est de fournir la meilleure connexion et le meilleur service possibles au moindre coût.

La concurrence sur le marché des réseaux est telle qu'il est à prévoir que seules les grosses entreprises ayant suffisamment d'argent pour investir dans les technologies innovantes pourront sortir leur épingle du jeu. Ceci dit, le monde des réseaux est relativement ouvert et beaucoup de choses sont encore à découvrir. Les entreprises de petite ou moyenne taille vont se recentrer sur des activités de services et utiliser ce que les grosses compagnies vont fournir. Le marché du service sur les réseaux est certainement le plus prometteur car étant celui qui est le plus à même d'évoluer rapidement.

Ce secteur est gouverné par deux facteurs principaux : le nombre croissant d'utilisateurs et l'innovation technologique qui tend à connecter en réseau tout ce qui peut l'être (cela va de l'ordinateur de bureau au téléphone portable, en passant par le réfrigérateur et la voiture). La miniaturisation et l'émergence des appareils portatifs (téléphones portables et autres PDA) constituent également d'autres facteurs d'orientation dans le développement des réseaux : les marchés de l'informatique embarquée et du sans fil s'ouvrent progressivement, et le but est finalement de fournir ce service de proximité dont tout le monde rêve.

Les réseaux vont s'étendre pour finalement s'interconnecter en une structure hypercomplexe et devenir de plus en plus gros et rapide (le haut débit n'en est qu'à son début). Parallèlement, on se dirige vers le monde du sans fil, où tout le monde pourra acheter depuis son téléphone portable ou sa voiture le dernier CD de Madonna qu'on vient tout juste d'entendre à la radio.

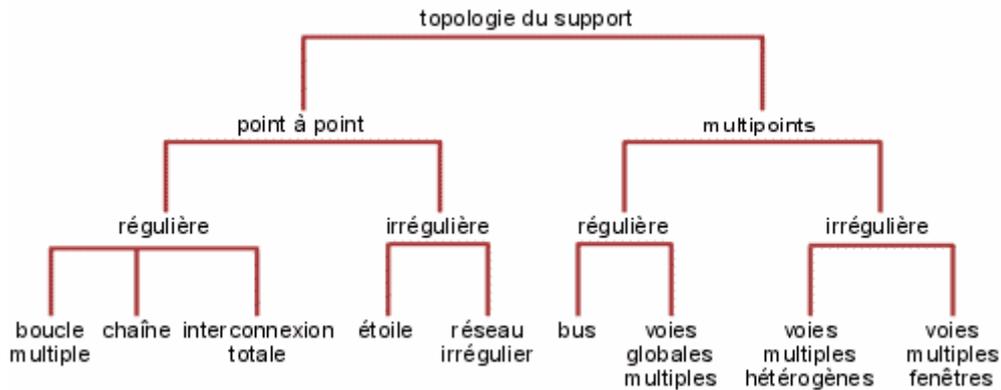
► Topologie des réseaux locaux

Classification des topologies

Les réseaux locaux sont généralement regroupés en 2 familles (voir classification des réseaux) :

- **réseau point à point**,
- **réseau de diffusion** (broadcast).

De cette première classification, on peut tirer une hiérarchie des topologies pour les réseaux locaux :



Les interconnexions sont de deux types :

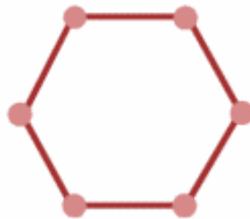
- **régulières** : relation de connexion constante entre différents modules voisins,
- **irrégulières** : irrégularité de connexion, due à un besoin de spatialisation de la topologie vis-à-vis de l'application. Le système n'est pas flexible.

Finalement, la topologie choisie pour un réseau dépend principalement :

de l'efficacité du réseau (débit, rapidité),
du degré de synchronisation entre les tâches réparties,
de la fiabilité et de la sûreté de transmission,
du coût de l'installation.

Les différentes topologies

Réseau en anneau ou en boucle



Boucle simple

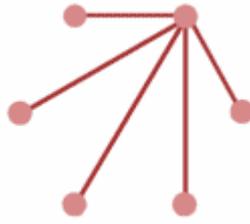
Ce type de réseau est peu efficace et peu fiable mais relativement économique. Dès que 2 lignes sont coupées, le réseau ne marche plus.



Boucle multiple

Ce type de topologie est utilisé pour des raisons de sécurité. Ce schéma est utilisé pour les réseaux locaux FDDI à fibre optique.

Réseau en étoile



Réseau en étoile

C'est un type de réseau relativement efficace et économique. La plupart des petits réseaux locaux fonctionnent (principalement pour des raisons de coût) sur ce principe, en utilisant un hub central reliant tous les ordinateurs entre eux. Les réseaux téléphoniques privés (PABX) tournent également avec cette topologie. La faiblesse en est le noeud central qui ne doit jamais tomber en panne.

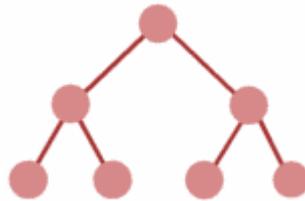
Réseau complet ou à maillage régulier



Réseau complet

C'est la topologie la plus fiable, mais la plus coûteuse aussi. Chaque élément est relié à tous les autres. N'est jamais utilisé en pratique.

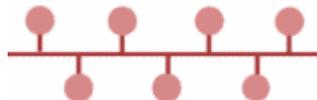
Structure hiérarchique



Structure hiérarchique

Parfois utilisée pour copier une structure hiérarchique donnée. En pratique, pas très efficace car les noeuds intermédiaires peuvent être des goulets d'étranglement.

Réseau en bus



Réseau en bus

La grande caractéristique de cette topologie est que c'est une structure passive : si un hôte tombe en panne, le réseau n'est pas affecté. Le signal circule sur l'ensemble du bus et disparaît une fois aux extrémités (c'est le rôle joué par les bouchons sur les réseaux Ethernet coaxiaux).

Réseau quelconque ou à maillage irrégulier



Réseau quelconque

C'est (hélas) la topologie la plus fréquemment rencontrée. La communication se fait rarement en direct, il faut passer par des intermédiaires. Ce n'est pas forcément la configuration la plus efficace et la plus fiable (les intermédiaires jouent une fois de plus un rôle très important). Ce type de structure peut se rencontrer lors de l'interconnexion de sous-réseaux locaux.

► Classification des réseaux

Technique de transmission

Une première façon de classer les réseaux peut se faire en considérant la technique de transmission utilisée :

- la diffusion (broadcast),
- le point à point.

Les **réseaux à diffusion** se caractérisent de la façon suivante. Toutes les machines se partagent un seul et unique canal de communication (sa forme physique peut être quelconque, c'est une description logique). Résultat : lorsqu'une machine émet un message sur ce réseau, toutes les autres machines, sans exception, le reçoivent également. Les machines sont alors différenciées en utilisant ce qu'on appelle une **adresse**. Une variante de ce type de réseaux est le réseau à **diffusion restreinte** ou **multipoint** (multicast). Le paquet n'est alors envoyé qu'à un certain nombre de machines, lesquelles machines doivent être généralement abonnées à un groupe (de multicast) pour recevoir le message.

Les **réseaux point à point** sont cette fois caractérisés par un canal de communication ne reliant que 2 machines, c'est-à-dire que pour arriver à sa destination, un message doit transiter par plusieurs machines intermédiaires.

Ainsi, de manière générale, on peut considérer que les réseaux de dimensions limitées (réseaux locaux notamment, voir ci-dessous) utilisent plutôt la diffusion, alors que les réseaux de grande taille sont plutôt du type point à point.

Classement par dimension

Classification générale

On classe plutôt les réseaux et les systèmes de multi-traitement en fonction de la distance d'interconnexion des processeurs. On établit ainsi la classification suivante :

distance	catégorie
< 1 cm	machines massivement parallèles ou à flots de données
< 1 m	machines multiprocesseurs
< 200 m	réseau local d'entreprise (RLE ou LAN)
< 1 km	réseau de campus
< 10 km	réseau métropolitain (MAN)
< 1.000 km	réseau grande distance (WAN)
> 1.000 km	interconnexion de réseaux grande distance (Internet)

Cette classification fait aussi apparaître une classification technique. En effet, les solutions techniques mises en oeuvre à chacun de ces niveaux sont relativement bien distinctes les unes des autres. Par exemple, sur réseau local, les connexions seront plutôt réalisées par câble coaxial ou câble à paires torsadées, alors que les connexions sur WAN se feront plutôt par fibre optique.

Réseau local d'entreprise (RLE ou LAN)

Aussi appelé **LAN** pour Local Area Network.

Les grandes caractéristiques des RLE sont :

faible étendue géographique (moins de 200 m),

débit élevé (généralement entre 10 et 100 Mbps mais ça peut parfois aller au delà),

intégration globale : tout est géré par une organisation unique.

Ce type de réseau est utilisé pour relier entre eux les ordinateurs d'une entreprise. Ces réseaux ont par ailleurs la particularité d'être placés sous une **autorité privée**.

Du fait de la faible dimension de ce type de réseau, les délais de transmission sont courts, avec peu d'erreurs, ce qui a l'avantage d'en simplifier l'administration.

Réseau métropolitain (MAN)

Aussi appelé **MAN** pour Metropolitan Area Network.

Métropolitain car ils ont généralement une étendue suffisante pour couvrir l'ensemble d'une ville.

Caractéristiques principales :

étendue géographique moyenne et limitée (étendue inférieure à 10 km de diamètre environ),

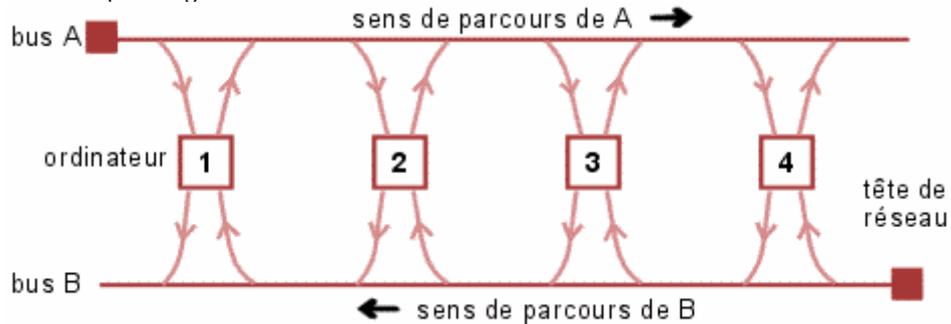
débit élevé (entre 10 et 100 Mbps),

capacité d'interconnexion de réseaux locaux.

Ces réseaux peuvent cette fois être placés sous une **autorité publique**.

Ce type de réseau peut être utilisé pour transmettre voix et données.

Ces réseaux sont généralement construits autour de 1 ou 2 câbles (bus) de transmission sans élément de routage. Tous les ordinateurs sont en effet directement reliés aux bus, ce qui a l'avantage d'en simplifier grandement la structure :



Architecture DQDB

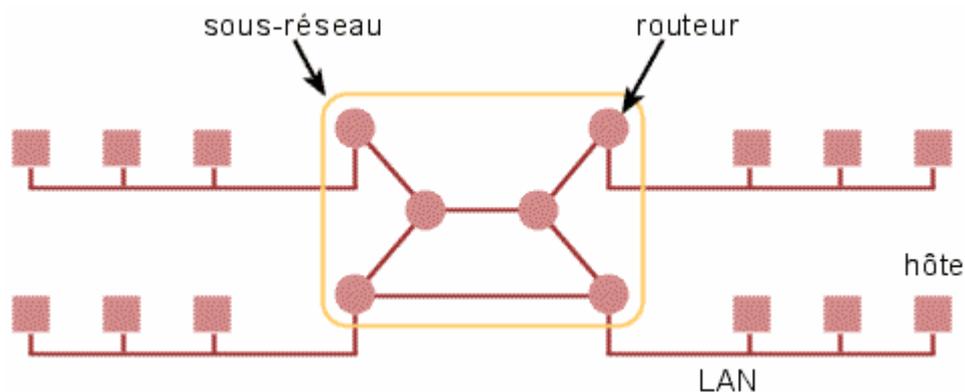
Ils utilisent alors souvent la norme DQDB (Distributed Queue Dual Bus) pour les transmissions.

Réseau grande distance (WAN)

Aussi appelé **WAN** pour Wide Area Network.

Ce type de réseau couvre une (très) large étendue géographique, de l'ordre parfois du millier de kilomètres. Les infrastructures utilisées sont en fait celles déjà existantes, à savoir celles des télécommunications. C'est du coup l'un des points de convergence de l'informatique et des télécoms. Actuellement, ces infrastructures sont renouvelées pour parfaire cette convergence et améliorer la qualité du service fourni par de tels réseaux, mais cela coûte très cher (voir plus bas, les problèmes de performance).

D'un point de vue plus technique, ces réseaux grande distance sont construits autour de noeuds de commutation (appelés commutateurs ou routeurs) reliés entre eux par des voies. Ces noeuds forment ce qu'on appelle des **sous-réseaux de communication** ou plus simplement des **sous-réseaux**. Le rôle profond de ces sous-réseaux est de relier les **hôtes** (des ordinateurs) entre eux. Ces hôtes se trouvent généralement reliés sur un LAN, lequel est relié via un routeur à un sous-réseau :



WAN : relation sous-réseau/hôte

Très souvent, les routeurs sont reliés entre eux par d'autres routeurs. Dans ce cas, un paquet reçu par un routeur n'est renvoyé au routeur suivant que lorsque le paquet a été entièrement reçu et que le premier routeur est en mesure de joindre le deuxième. Ce type de fonctionnement caractérise ce qu'on appelle un sous-réseau **point à point**, à **commutation de paquets** ou encore **en mode différé**. Mise à part les réseaux satellite, tous les WAN sont de ce type.

De par les grandes distances qu'ils couvrent, ces réseaux sont sujets à certaines contraintes et problèmes de fiabilité (le bruit augmente avec la longueur des lignes) et de performance (le délai de transmission augmente proportionnellement avec la longueur), ce qui entraîne une

augmentation des coûts non négligeable lorsqu'il s'agit d'en augmenter la qualité. Heureusement, l'apparition des fibres optiques polymères très performantes et bon marché permet de réduire significativement les coûts d'installation.

Exemples de WAN : Arpanet (le réseau à l'origine d'Internet), Internet, Transpac...

Autre classification

L'amélioration des technologies et donc des débits suggère une classification des réseaux non pas par leur dimension, mais par leur débit (classification plus "logique"). Ainsi, on trouve trois grandes familles :

- les réseaux à faible et moyen débits (débits < 200 kbps),
- les réseaux à haut débit (200 kbps < débit < 20 Mbps),
- les réseaux à très haut débit (débit > 20 Mbps).

Les réseaux locaux ont de plus en plus tendance à ressembler à des réseaux très haut débit, surtout depuis l'avènement du Fast Ethernet devenu très bon marché et surtout très fiable.

► Concepts d'un réseau informatique

Problématique de base

Le problème de base est la communication et l'échange de données entre utilisateurs. Notons au passage qu'un utilisateur peut être une personne physique, un programme, un terminal...

L'émetteur du message impose un certain nombre de contraintes quant à la transmission des informations. Il souhaite en particulier les choses suivantes :

que les informations parviennent à un correspondant parfaitement identifié,
que les informations transmises arrivent intégralement et sans erreur,
que les informations arrivent dans l'ordre d'émission.

En sus de ces contraintes viennent s'ajouter un certain nombre d'autres problèmes, notamment liés aux moyens dont nous disposons pour transmettre ces informations. En effet, la transmission va s'effectuer sur base des infrastructures de télécommunication existantes, et ces infrastructures présentent quelques défauts : il n'existe généralement pas de ligne directe entre 2 utilisateurs, et ces lignes, par des limitations physiques, ne permettent pas de garantir une qualité de transmission absolue (les lignes vont altérer les signaux).

Dernier point important : il faut passer de l'utilisateur et des données (logiques) qu'il veut transmettre à une représentation physique utilisable sur les précédentes infrastructures.

Clairement, on cherche à établir des "ponts" entre les différents éléments constitutifs de la chaîne de transmission. Selon toute vraisemblance, les données à transmettre seront transformées pour passer les différents ponts, mais à la fin, tout devra être correctement reconstitué selon les souhaits de l'émetteur.

La transformation de l'information

Cette transformation va s'effectuer progressivement en trois principales étapes :

transformation par les interlocuteurs : les 2 interlocuteurs doivent en fait se mettre d'accord sur les procédures d'échange, et ils doivent trouver une signification commune aux informations échangées. Un certain nombre de modalités sont à mettre en place, en particulier, ils vérifient en permanence que tout se passe bien.

conversion parallèle-série : c'est une transformation "raisonnée". En effet, il est beaucoup plus économique de transmettre des données en série plutôt qu'en parallèle. Les interlocuteurs doivent donc pouvoir faire la conversion.

transformation numérique-analogique : ceci vient tout simplement du fait que les signaux numériques ne peuvent pas être transmis sans déformation sur des lignes physiques. Les signaux numériques sont donc transformés en signaux analogiques. plus transportables sur des lignes classiques et sur grande distance.

Le mode d'acheminement

Mode connecté et non connecté

Il existe deux modes (ou services) de transmission : le **mode connecté** et le **mode non connecté**. Pour expliquer la différence, prenons deux exemples : le téléphone et le système postal. Le téléphone est typique du mode connecté. Pour dialoguer avec quelqu'un, on décroche son téléphone, on compose un numéro et on parle. Lorsque la conversation est terminée, et seulement à ce moment, on raccroche le combiné. Concrètement, le mode connecté se caractérise par l'ensemble suivant de processus : l'utilisateur doit d'abord établir une connexion, puis à la fin de la transmission il coupe la connexion.

A l'inverse, le courrier est typique du mode non connecté. Chaque message porte l'adresse de destination et est transporté indépendamment de tous les autres. En toute rigueur, le premier message, si la poste fonctionne bien, doit arriver avant le second et ainsi de suite. Seulement, du fait de l'indépendance des transmissions, et en particulier des intermédiaires et des routes suivies, il n'est pas impossible que le second message arrive avant le premier. Ce genre de cas est purement impossible en mode connecté. Notons au passage que le service non connecté est parfois appelé **service datagramme**, en référence au vieux système des télégraphes qui n'utilise pas de système d'acquiescement.

Chaque service peut se caractériser par ce qu'on va appeler une **qualité de service**. La qualité de service, c'est la capacité à garantir qu'un message complet sera transmis sans erreurs (pas d'inversion ni d'erreur de contenu). Le mode connecté se caractérise donc par une qualité de service élevée, alors que le mode non connecté se caractérise par une qualité de service médiocre. Pour garantir une haute qualité de service, on utilise notamment ce qu'on appelle des

acquiescements : on signale à l'émetteur par un nouveau message que son envoi est bien arrivé. Généralement, l'émetteur attend l'acquiescement avant de renvoyer un second message. Ce système d'acquiescement est très efficace mais hélas coûteux en charge réseau et en temps. C'est

pourquoi l'émetteur peut envoyer plusieurs paquets à la suite sans attendre les acquittements, histoire d'améliorer les performances.

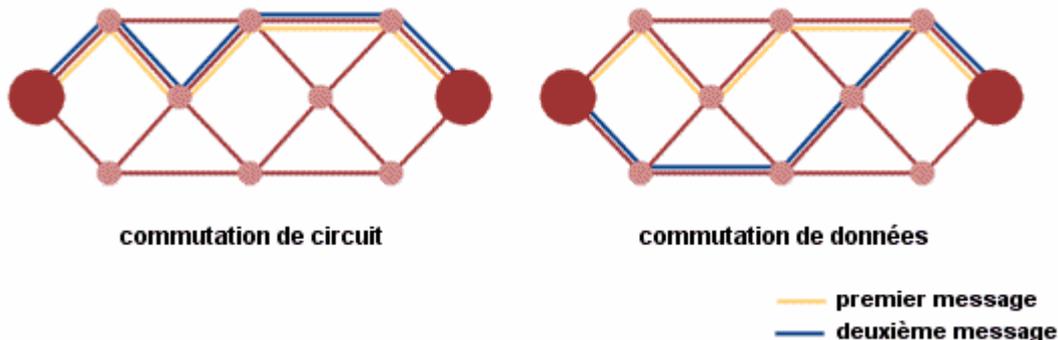
On aurait tendance à demander systématiquement la qualité de service maximale, mais il faut bien se rendre compte que ceci peut imposer des coûts de structures élevés et parfois non justifiés. Il faut donc faire un choix, et ce choix dépend à la fois des infrastructures de transmissions et du réel besoin de l'utilisateur. Par exemple, un transfert de fichier s'effectuera préférentiellement en mode connecté, car on demande une fiabilité importante. A l'inverse, la transmission de messages électroniques va se faire en mode non connecté car il n'y a pas de réel impératif dans l'ordre d'arrivée (on peut toujours trier à la main après) et la probabilité que le message soit perdu est (très) faible. En fait, l'acquiescement peut être utilisé en mode datagramme comme il peut ne pas être utilisé en mode connecté. On peut ainsi définir 4 grandes classes de qualité de service.

Conséquences sur les réseaux

Le fait de proposer 2 modes de transmission permet de distinguer 2 classes de réseaux, en fonction justement du type de connexion choisie :

réseaux à commutation de circuit : ils sont typiques du mode connecté. Un chemin est choisi au moment de l'établissement de la connexion et il est conservé sur toute la durée de la transmission. Les messages empreintent tous le même chemin.

réseaux à commutation de données : ils sont typiques du mode datagramme. Chaque voix peut être utilisée par plusieurs communications indépendantes. Deux messages peuvent empreinter des chemins différents.



Fonctions d'un système de réseau

Les fonctions d'un système de réseau sont des fonctions qui permettent :

- **la connexion** physique entre 2 utilisateurs directement ou indirectement,
- **l'émission** par un utilisateur d'un message que le système se charge de transmettre à un ou plusieurs destinataires,
- **la réception** par un utilisateur des messages dont il est le destinataire,
- **l'acheminement** des messages sans erreur, sans perte, sans déséquencement, sans duplication et en temps utile,
- **l'optimisation** des lignes par le partage du support physique,
- **le contrôle de flux** et le stockage des messages en transit avant leur utilisation,
- **le choix** entre différentes méthodes de dialogue,
- **la gestion** et le contrôle de l'utilisation des fonctions réseaux.

La diversité des besoins en communication entraîne des différences dans les fonctionnalités requises au niveau du système de réseau. Par exemple, les besoins d'une application télématique (le minitel) sont bien différents de ceux pour le transfert de fichiers bancaires. Les différences de besoins entre les applications se situent au niveau des vitesses de transfert, de la tolérance aux fautes et de l'intégrité des données.

On découpe ainsi le logiciel de réseau en un ensemble d'activités de base et d'activités secondaires. Pour un réseau donné ou une application particulière, on ne choisira d'implémenter qu'une partie des activités.

Ainsi, les activités de base sont :

- **la sûreté** (correction des erreurs, maintien de l'ordre des messages),
- **le contrôle de flux**,
- **l'adressage** (pour désigner les interlocuteurs),
- **connexion et déconnexion**,
- **l'acheminement**.

Les activités secondaires sont par exemple :

- prise en compte de l'hétérogénéité du réseau,
- la synchronisation de la transmission,
- multiplexage-démultiplexage,
- reprise des erreurs,

- contrôle de congestion,
- traitement des incidents,
- ...

► Logiciel de réseau

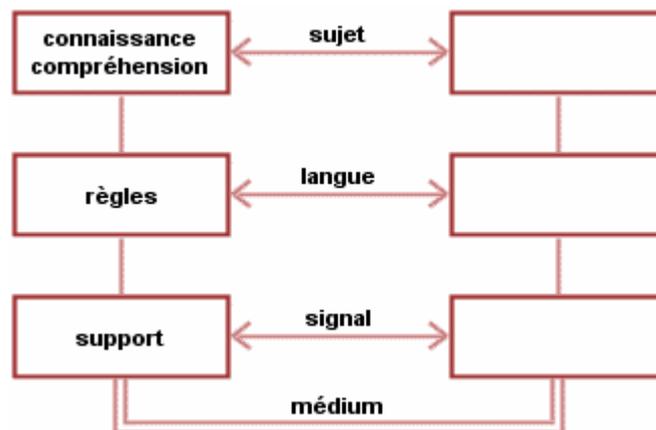
Notion de protocole

Au chapitre précédent, nous avons vu que le système de communication prenait à sa charge le transfert de l'information et toute la gestion nécessaire à sa remise correcte et compréhensible à l'utilisateur. Dans ce contexte, il est nécessaire d'établir un certain nombre de règles de manière à ce que tout se passe bien. De la même façon, nous avons vu qu'il était nécessaire que les interlocuteurs se mettent d'accord sur la signification des messages transmis. De manière plus générale, l'ensemble de ces processus de transmission sont gérés par des mécanismes appelés **protocoles**.

Exemple simple de protocole : une conversation courante entre deux amis. En fait, un certain nombre de règles implicitement établies régule la conversation. Ces règles sont par exemple : ne pas parler en même temps, parler la même langue et parler du même sujet. Ces règles constituent un protocole.

Structure en couches et hiérarchies de protocoles

La structure en couche apparaît déjà dans l'exemple précédent : les règles de communication peuvent se superposer comme le montre le schéma suivant :



Les communications sur réseau fonctionnent exactement de la même façon. Afin de réduire la complexité de conception du logiciel de réseau et de réaliser l'indépendance logiciel/matériel, on effectue un découpage fonctionnel de l'ensemble du processus communicant en **couches ou niveaux d'abstraction**. Une couche correspond à un ensemble de fonctions ou de processus cohérents entre eux et assurant une fonction précise globale. Deux couches adjacentes sont indépendantes dans leurs fonctions mais elles s'interconnectent par ce qu'on appelle une **interface** : c'est au niveau des interfaces que les couches s'échangent des données. Pour donner un exemple concret : une machine à laver ! Considérons la machine à laver comme une couche (sa fonction est alors "laver le linge"). Elle est reliée à deux autres couches : le réseau électrique et le réseau d'eau. Les interfaces sont alors la prise de courant et le tuyau d'alimentation en eau. Les protocoles correspondent à une implémentation donnée d'une couche. Ce modèle est finalement fondé sur le principe de "diviser pour régner". L'ensemble de ces couches et protocoles est appelé **architecture du réseau**. L'ensemble des protocoles utilisés par un système ayant un protocole par couche est appelé **pile de protocoles** (exemple de pile : la pile TCP/IP).

Pour fonctionner, chaque couche s'appuie sur la précédente. Chaque couche doit fournir un certain nombre de services aux couches précédentes, sans que ces dernières aient à connaître les détails de l'implémentation de ces services. Ceci garantit une certaine modularité du système dans son ensemble et surtout l'indépendance des implémentations. Dans un tel système, la couche *N* d'une machine gère la conversation avec la couche *N* d'une autre machine en utilisant un ou des protocoles et les couches sous-jacentes.

Entre chaque paire de couches adjacentes se trouve une **interface**. C'est cette interface qui définit les services que la couche inférieure fournit à la couche supérieure. Au moment de la conception d'une architecture réseau et toujours dans le respect de l'indépendance des implémentations, il est très important de bien définir ces interfaces. Les couches doivent donc réaliser un ensemble de fonctions bien définies. Nous verrons plus bas que ces interfaces sont construites autour de **primitives**.

Principes de conception des couches

Du fait qu'une machine peut faire fonctionner plusieurs processus (ou programmes) accédant à des ressources réseau, et que par ailleurs une machine peut avoir plusieurs interlocuteurs (au moins un

par application tournante), il est indispensable que chaque couche possède un mécanisme d'identification des émetteurs (les processus) et des récepteurs (les interlocuteurs distants). Il convient également de s'interroger sur les règles de transfert des données. En effet, une communication peut s'effectuer de différentes façons :

- **communication simple** : les données ne voyagent que dans un seul sens,
- **communication semi-duplex** : les données peuvent voyager dans les 2 sens mais pas en même temps,
- **communication duplex** : les données peuvent voyager dans les 2 sens en même temps.

Ensuite intervient un problème de qualité de transmission : une couche doit pouvoir détecter et corriger les erreurs dues aux circuits physiques de communication. Il existe beaucoup de codes de détection et correction, mais les 2 extrémités doivent se mettre d'accord sur le code à utiliser. Par ailleurs, les extrémités doivent disposer d'un système permettant de dire à l'émetteur que le message a été correctement reçu. Deuxième problème de qualité de transmission : la conservation de l'ordre d'arrivée des messages. L'émetteur doit donner les moyens au récepteur de pouvoir remettre les paquets dans l'ordre, en numérotant les paquets par exemple, mais cela ne dit pas comment effectivement reconstruire le message.

Encore un problème important : certains processus ne sont pas en mesure de gérer des messages de longueur quelconque. Il faut donc fournir des mécanismes de désassemblage, de transmission et de réassemblage des messages. Le cas inverse peut également se produire : certains processus, pour des raisons de performance entre autre, vont vouloir transmettre des messages plus petits que ce qu'il pourrait se faire. Il faut alors être capable de réassembler ces petits messages en des messages plus gros.

Voici les principaux problèmes qui interviennent au moment de la conception de couches ; on peut toutefois en citer encore quelques uns tout aussi importants : problèmes d'engorgement du récepteur (lorsque l'émetteur est trop rapide par exemple), problèmes de routage rationnel de messages suivant certaines règles, problème du multiplexage (notamment sur la couche physique) lorsque le nombre de canaux de communication de la couche sous-jacentes sont limités...

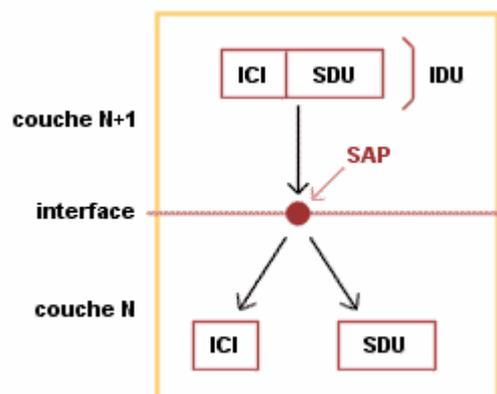
Les services

Interfaces et services

Le but de chaque couche est de fournir un certain nombre de services à la couche supérieure. On dit alors que la couche sous-jacente est **fournisseur de service** à la couche supérieure, **utilisateur de service**. Les éléments actifs d'une couche sont appelés **entités**. Ces entités peuvent être des entités logicielles (un processus par exemple) ou matérielles (une puce électronique par exemple). Les entités de la même couche N de 2 machines différentes sont appelés **entités paires**.

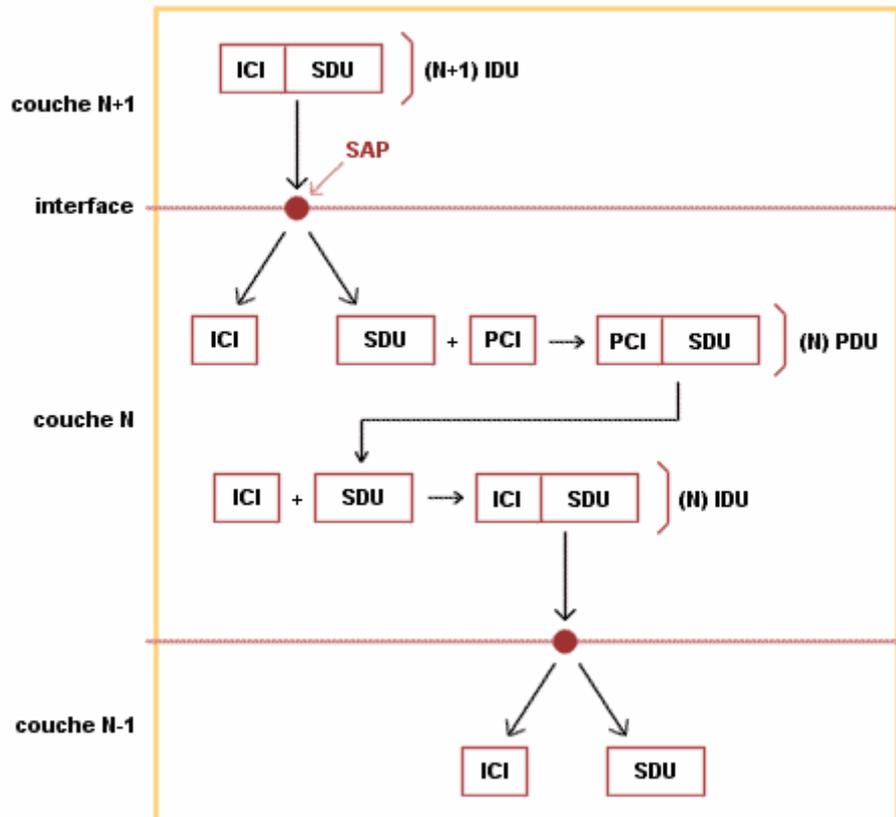
Les services d'une couche N sont accessibles par ce qu'on appelle des **points d'accès aux services**, ou **SAP** (Service Access Point). Chaque SAP est identifié par une adresse unique. Typiquement, les SAP du réseau téléphonique sont les prises de téléphone et les adresses sont les numéros de téléphones.

Pour que 2 couches adjacentes puissent communiquer, un certain nombre de règles doivent être mises en place à propos de l'interface. L'entité de la couche $N+1$ va donner à l'entité de la couche N une **unité de données d'interface** ou **IDU** (Interface Data Unit) à travers le SAP. L>IDU est en fait constitué de 2 choses : une **unité de données de service**, ou **SDU** (Service Data Unit) et certaines **informations de contrôle**, ou **ICI** (Interface Control Information) :



Le SDU constitue l'information que 2 entités paires échangent, mais c'est également ce que la couche $N+1$ du récepteur va transmettre à la couche N . L'information de contrôle est là pour

assister la couche inférieure dans son travail. Elle va par exemple contenir le nombre d'octets contenues dans le SDU (cela peut servir dans la fonction de contrôle de l'intégrité de l'information). Pour transmettre une SDU, il se peut qu'une entité de la couche N ait besoin de la fragmenter en plusieurs morceaux. Au moment de l'échange avec son entité paire, chaque morceau reçoit des **informations de contrôle de protocole**, ou **PCI** (Protocol Control Information) dans un **en-tête**, et le tout est envoyé séparément comme **unité de données de protocole**, ou **PDU** (Protocol Data Unit). Les en-têtes sont utilisés par les entités paires pour transporter leur protocole pair. Ce PDU devient alors le SDU de la couche N qui sera transmis à la couche $N-1$ via le SAP. Le fonctionnement général est donc le suivant :



Au bout du compte, on se retrouve avec un emboîtement des messages les uns dans les autres. Ce mécanisme d'emboîtement est souvent qualifié de mécanisme d'**encapsulation**.

Les primitives de service

Un service est défini formellement par un ensemble de **primitives** ou opérations qu'un utilisateur ou d'autres entités peuvent utiliser pour accéder au service. C'est ce qui concrétise l'interface. On classe habituellement les primitives de service suivant 4 classes :

primitive	signification
request (requête)	une entité sollicite un service (on demande une connexion auprès d'un ordinateur distant)
indication	une entité est informée d'un événement (le récepteur reçoit une demande de connexion)
response (réponse)	une entité répond à un événement (le récepteur envoie l'autorisation de la connexion)
confirm (confirmation)	une entité accuse réception de la réponse à sa demande (l'émetteur reçoit confirmation de la connexion)

La plupart des primitives possèdent des paramètres. Par exemple, les paramètres d'un **CONNECT.request** (requête d'établissement d'une connexion) vont être la machine à laquelle on veut se connecter, le type de service désiré (FTP...) et la taille maxi des paquets échangés.

Un **service confirmé** est un service pour lequel il y a une demande (request), une indication (indication), une réponse (response) et une confirmation (confirm). Un **service non confirmé** est un service pour lequel il y a juste une demande (request) et une indication (indication). Typiquement, le service d'établissement de connexion est un service confirmé car l'entité paire doit être d'accord pour établir la connexion. En revanche, le transfert de données peut être non confirmé, suivant que l'on veuille un acquittement ou non. D'un point de vue implémentation, les primitives correspondent aux fonctions que l'on peut utiliser dans un programme pour accéder au service en question.

Relations des services aux protocoles

Un service est un ensemble de primitives qu'une couche fournit à la couche supérieure. Le service définit les opérations qu'une couche peut effectuer, mais cela ne renseigne pas sur la façon dont sont réalisées ces opérations. L'élément le plus caractéristique du service est l'interface entre deux couches adjacentes.

En revanche, un protocole est un ensemble de règles qui s'appliquent à la signification et au format des messages échangés entre entités paires. Les entités utilisent les protocoles pour **implémenter** les spécifications de service. Un service peut donc toujours être le même avec 2 protocoles différents.

Protocoles et services sont donc différents, mais ils sont étroitement liés. Il ne faut donc pas les confondre. Le service est une notion plutôt abstraite, alors que le protocole correspond véritablement à ce qui se passe physiquement. Cette distinction répond en fait à des exigences de programmation et d'implémentation modernes. Cela revient à faire la différence entre algorithme et implémentation de l'algorithme.

Services en mode connexion et mode déconnexion

Ces services sont tout simplement ceux abordés dans le chapitre précédent. On va donc juste rappeler ici quels en sont les grands principes.

Le **service en mode connecté** nécessite d'abord d'établir une connexion entre les 2 interlocuteurs. Le récepteur s'attend alors à recevoir des données de la part de l'émetteur. A la fin de la transmission, la connexion est coupée. Un exemple typique est le téléphone : pour l'utiliser, il faut d'abord décrocher le téléphone et composer un numéro. La personne appelée décroche alors son combiné et la connexion est établie. Les 2 personnes échangent jusqu'à ce qu'elles raccrochent. Le **service en mode non connecté** se caractérise par une indépendance des messages transmis. Le récepteur reçoit un message alors qu'il n'a pas forcément été prévenu auparavant par l'émetteur. Les messages peuvent alors suivre des chemins différents, d'où une possible inversion de l'ordre d'arrivée des messages. L'exemple typique de ce service est la poste : quelqu'un écrit une lettre et l'envoie sans prévenir son destinataire. Cette lettre peut alors arriver après la seconde lettre que lui a écrite la même personne. Les deux chemins suivis par les 2 lettres peuvent être différents.

► Le modèle OSI

Introduction

Les constructeurs informatiques ont proposé des architectures réseaux propres à leurs équipements. Par exemple, IBM a proposé SNA, DEC a proposé DNA... Ces architectures ont toutes le même défaut : du fait de leur caractère propriétaire, il n'est pas facile des les interconnecter, à moins d'un accord entre constructeurs. Aussi, pour éviter la multiplication des solutions d'interconnexion d'architectures hétérogènes, l'ISO (International Standards Organisation), organisme dépendant de l'ONU et composé de 140 organismes nationaux de normalisation, a développé un modèle de référence appelé **modèle OSI** (Open Systems Interconnection). Ce modèle décrit les concepts utilisés et la démarche suivie pour normaliser l'interconnexion de **systèmes ouverts** (un réseau est composé de systèmes ouverts lorsque la modification, l'adjonction ou la suppression d'un de ces systèmes ne modifie pas le comportement global du réseau).

Au moment de la conception de ce modèle, la prise en compte de l'hétérogénéité des équipements était fondamentale. En effet, ce modèle devait permettre l'interconnexion avec des systèmes hétérogènes pour des raisons historiques et économiques. Il ne devait en outre pas favoriser un fournisseur particulier. Enfin, il devait permettre de s'adapter à l'évolution des flux d'informations à traiter sans remettre en cause les investissements antérieurs. Cette prise en compte de l'hétérogénéité nécessite donc l'adoption de règles communes de communication et de coopération entre les équipements, c'est à dire que ce modèle devait logiquement mener à une normalisation internationale des protocoles.

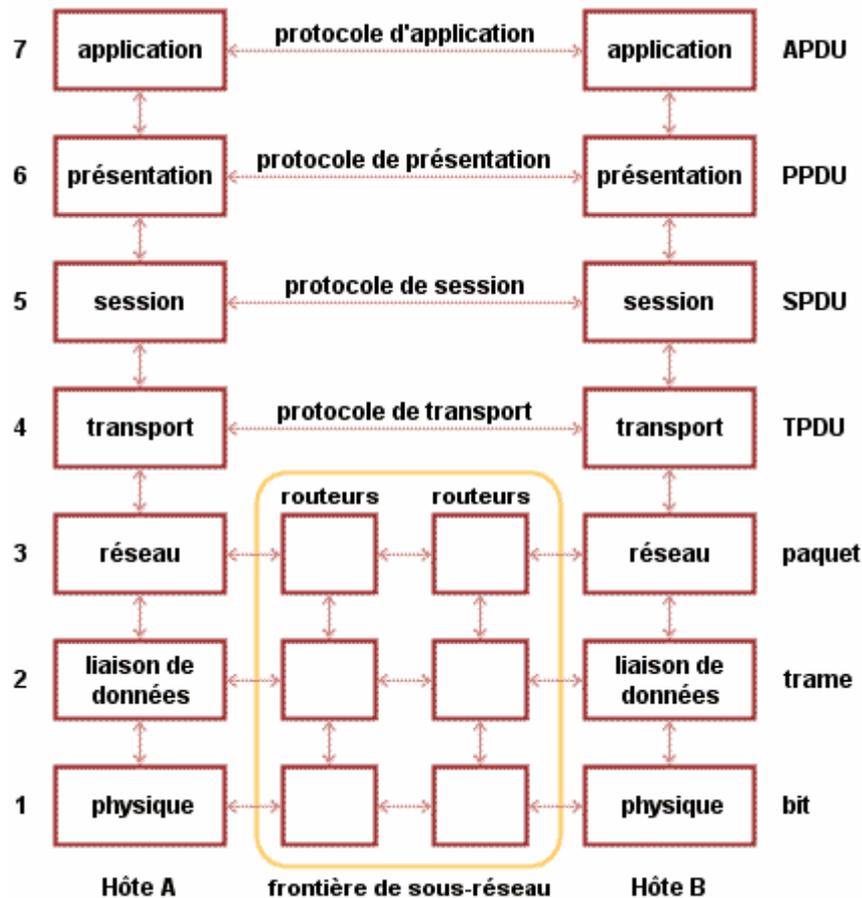
Le modèle OSI n'est pas une véritable architecture de réseau, car il ne précise pas réellement les services et les protocoles à utiliser pour chaque couche. Il décrit plutôt ce que doivent faire les couches. Néanmoins, l'ISO a écrit ses propres normes pour chaque couche, et ceci de manière indépendante au modèle, i.e. comme le fait tout constructeur.

Les premiers travaux portant sur le modèle OSI datent de 1977. Ils ont été basés sur l'expérience acquise en matière de grands réseaux et de réseaux privés plus petits ; le modèle devait en effet être valable pour tous les types de réseaux. En 1978, l'ISO propose ce modèle sous la norme ISO IS7498. En 1984, 12 constructeurs européens, rejoints en 1985 par les grands constructeurs américains, adoptent le standard.

Les différentes couches du modèle

Les 7 couches

Le modèle OSI comporte 7 couches :



Les principes qui ont conduit à ces 7 couches sont les suivants :

- une couche doit être créée lorsqu'un nouveau niveau d'abstraction est nécessaire,
- chaque couche a des fonctions bien définies,
- les fonctions de chaque couche doivent être choisies dans l'objectif de la normalisation internationale des protocoles,
- les frontières entre couches doivent être choisies de manière à minimiser le flux d'information aux interfaces,
- le nombre de couches doit être tel qu'il n'y ait pas cohabitation de fonctions très différentes au sein d'une même couche et que l'architecture ne soit pas trop difficile à maîtriser.

Les couches basses (1, 2, 3 et 4) sont nécessaires à l'acheminement des informations entre les extrémités concernées et dépendent du support physique. Les couches hautes (5, 6 et 7) sont responsables du traitement de l'information relative à la gestion des échanges entre systèmes informatiques. Par ailleurs, les couches 1 à 3 interviennent entre machines voisines, et non entre les machines d'extrémité qui peuvent être séparées par plusieurs routeurs. Les couches 4 à 7 sont au contraire des couches qui n'interviennent qu'entre hôtes distants.

La couche physique

La couche physique s'occupe de la transmission des bits de façon brute sur un canal de communication. Cette couche doit garantir la parfaite transmission des données (un bit 1 envoyé doit bien être reçu comme bit valant 1). Concrètement, cette couche doit normaliser les caractéristiques électriques (un bit 1 doit être représenté par une tension de 5 V, par exemple), les caractéristiques mécaniques (forme des connecteurs, de la topologie...), les caractéristiques fonctionnelles des circuits de données et les procédures d'établissement, de maintien et de libération du circuit de données.

L'unité d'information typique de cette couche est le **bit**, représenté par une certaine différence de potentiel.

La couche liaison de données

Son rôle est un rôle de "liant" : elle va transformer la couche physique en une liaison a priori exempte d'erreurs de transmission pour la couche réseau. Elle fractionne les données d'entrée de l'émetteur en **trames**, transmet ces trames en séquence et gère les trames d'acquittement renvoyées par le récepteur. Rappelons que pour la couche physique, les données n'ont aucune signification particulière. La couche liaison de données doit donc être capable de reconnaître les frontières des trames. Cela peut poser quelques problèmes, puisque les séquences de bits utilisées pour cette reconnaissance peuvent apparaître dans les données.

La couche liaison de données doit être capable de renvoyer une trame lorsqu'il y a eu un problème sur la ligne de transmission. De manière générale, un rôle important de cette couche est la détection et la correction d'erreurs intervenues sur la couche physique. Cette couche intègre également une fonction de contrôle de flux pour éviter l'engorgement du récepteur.

L'unité d'information de la couche liaison de données est la **trame** qui est composées de quelques centaines à quelques milliers d'octets maximum.

La couche réseau

C'est la couche qui permet de gérer le sous-réseau, i.e. le routage des paquets sur ce sous-réseau et l'interconnexion des différents sous-réseaux entre eux. Au moment de sa conception, il faut bien déterminer le mécanisme de routage et de calcul des tables de routage (tables statiques ou dynamiques...).

La couche réseau contrôle également l'engorgement du sous-réseau. On peut également y intégrer des fonctions de comptabilité pour la facturation au volume, mais cela peut être délicat.

L'unité d'information de la couche réseau est le **paquet**.

Couche transport

Cette couche est responsable du bon acheminement des messages complets au destinataire. Le rôle principal de la couche transport est de prendre les messages de la couche session, de les découper s'il le faut en unités plus petites et de les passer à la couche réseau, tout en s'assurant que les morceaux arrivent correctement de l'autre côté. Cette couche effectue donc aussi le réassemblage du message à la réception des morceaux.

Cette couche est également responsable de l'optimisation des ressources du réseau : en toute rigueur, la couche transport crée une connexion réseau par connexion de transport requise par la couche session, mais cette couche est capable de créer plusieurs connexions réseau par processus de la couche session pour répartir les données, par exemple pour améliorer le débit. A l'inverse, cette couche est capable d'utiliser une seule connexion réseau pour transporter plusieurs messages à la fois grâce au multiplexage. Dans tous les cas, tout ceci doit être transparent pour la couche session.

Cette couche est également responsable du type de service à fournir à la couche session, et finalement aux utilisateurs du réseau : service en mode connecté ou non, avec ou sans garantie d'ordre de délivrance, diffusion du message à plusieurs destinataires à la fois... Cette couche est donc également responsable de l'établissement et du relâchement des connexions sur le réseau. Un des tous derniers rôles à évoquer est le contrôle de flux.

C'est l'une des couches les plus importantes, car c'est elle qui fournit le service de base à l'utilisateur, et c'est par ailleurs elle qui gère l'ensemble du processus de connexion, avec toutes les contraintes qui y sont liées.

L'unité d'information de la couche réseau est le **message**.

La couche session

Cette couche organise et synchronise les échanges entre tâches distantes. Elle réalise le lien entre les adresses logiques et les adresses physiques des tâches réparties. Elle établit également une liaison entre deux programmes d'application devant coopérer et commande leur dialogue (qui doit parler, qui parle...). Dans ce dernier cas, ce service d'organisation s'appelle la **gestion du jeton**.

La couche session permet aussi d'insérer des points de reprise dans le flot de données de manière à pouvoir reprendre le dialogue après une panne.

La couche présentation

Cette couche s'intéresse à la syntaxe et à la sémantique des données transmises : c'est elle qui traite l'information de manière à la rendre compatible entre tâches communicantes. Elle va assurer l'indépendance entre l'utilisateur et le transport de l'information.

Typiquement, cette couche peut convertir les données, les reformater, les crypter et les compresser.

La couche application

Cette couche est le point de contact entre l'utilisateur et le réseau. C'est donc elle qui va apporter à l'utilisateur les services de base offerts par le réseau, comme par exemple le transfert de fichier, la messagerie...

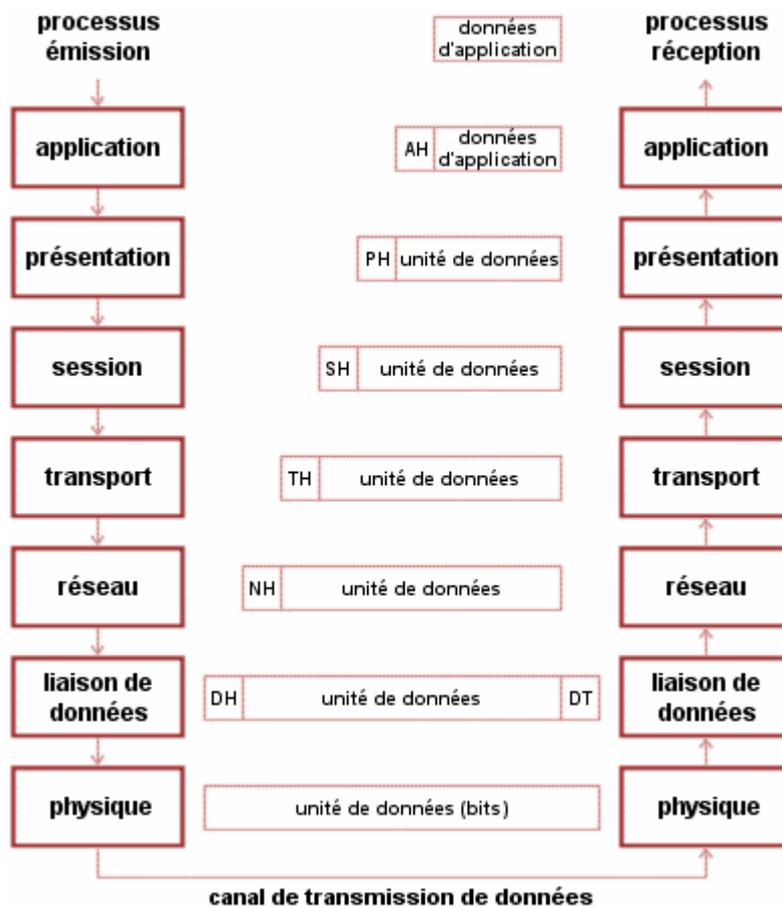
► Le modèle OSI, suite

Transmission de données au travers du modèle OSI

Le processus émetteur remet les données à envoyer au processus récepteur à la couche application qui leur ajoute un en-tête application AH (éventuellement nul). Le résultat est alors transmis à la couche présentation.

La couche présentation transforme alors ce message et lui ajoute (ou non) un nouvel en-tête (éventuellement nul). La couche présentation ne connaît et ne doit pas connaître l'existence éventuelle de AH ; pour la couche présentation, AH fait en fait partie des données utilisateur. Une fois le traitement terminé, la couche présentation envoie le nouveau "message" à la couche session et le même processus recommence.

Les données atteignent alors la couche physique qui va effectivement transmettre les données au destinataire. A la réception, le message va remonter les couches et les en-têtes sont progressivement retirés jusqu'à atteindre le processus récepteur :



Le concept important est le suivant : il faut considérer que chaque couche est programmée comme si elle était vraiment horizontale, c'est à dire qu'elle dialoguait directement avec sa couche paire réceptrice. Au moment de dialoguer avec sa couche paire, chaque couche rajoute un en-tête et l'envoie (virtuellement, grâce à la couche sous-jacente) à sa couche paire.

Critique du modèle OSI

La chose la plus frappante à propos du modèle OSI est que c'est peut-être la structure réseau la plus étudiée et la plus unanimement reconnue et pourtant ce n'est pas le modèle qui a su s'imposer. Les spécialistes qui ont analysé cet échec en ont déterminé 4 raisons principales.

Ce n'était pas le bon moment

David Clark du MIT a émis la théorie suivant quant à l'art et la manière de publier une norme au bon moment. Pour lui, dans le cycle de vie d'une norme, il y a 2 pics principaux d'activité : la recherche effectuée dans le domaine couvert par la norme, et les investissements des industriels pour l'implémentation et la mise en place de la norme. Ces deux pics sont séparés par un creux

d'activité qui apparaît être en fait le moment idéal pour la publication de la norme : il n'est ni trop tôt par rapport à la recherche et on peut donc assurer une certaine maîtrise, et il n'est ni trop tard pour les investissements et les industriels sont prêts à utiliser des capitaux pour l'implémenter. Le modèle OSI était idéalement placé par rapport à la recherche, mais hélas, le modèle TCP/IP était déjà en phase d'investissement prononcé (lorsque le modèle OSI est sorti, les universités américaines utilisaient déjà largement TCP/IP avec un certain succès) et les industriels n'ont pas ressenti le besoin d'investir dessus.

Ce n'était pas la bonne technologie

Le modèle OSI est peut-être trop complet et trop complexe. La distance entre l'utilisation concrète (l'implémentation) et le modèle est parfois importante. En effet, peu de programmes peuvent utiliser ou utilisent mal l'ensemble des 7 couches du modèle : les couches session et présentation sont fort peu utilisées et à l'inverse les couches liaison de données et réseau sont très souvent découpées en sous-couches tant elles sont complexes.

OSI est en fait trop complexe pour pouvoir être proprement et efficacement implémenté. Le comité rédacteur de la norme a même du laisser de côté certains points techniques, comme la sécurité et le codage, tant il était délicat de conserver un rôle bien déterminé à chaque couche ainsi complétée. Ce modèle est également redondant (le contrôle de flux et le contrôle d'erreur apparaissent pratiquement dans chaque couche). Au niveau de l'implémentation, TCP/IP est beaucoup plus optimisé et efficace.

La plus grosse critique que l'on peut faire au modèle est qu'il n'est pas du tout adapté aux applications de télécommunication sur ordinateur ! Certains choix effectués sont en désaccord avec la façon dont les ordinateurs et les logiciels communiquent. La norme a en fait fait le choix d'un "système d'interruptions" pour signaler les événements, et sur des langages de programmation de haut niveau, cela est peu réalisable.

Ce n'était pas la bonne implémentation

Cela tient tout simplement du fait que le modèle est relativement complexe, et que du coup les premières implémentations furent relativement lourdes et lentes. A l'inverse, la première implémentation de TCP/IP dans l'Unix de l'université de Berkeley (BSD) était gratuite et relativement efficace. Historiquement, les gens ont donc eu une tendance naturelle à utiliser TCP/IP.

Ce n'était pas la bonne politique

Le modèle OSI a en fait souffert de sa trop forte normalisation. Les efforts d'implémentation du modèle étaient surtout "bureaucratiques" et les gens ont peut-être vu ça d'un mauvais œil. A l'inverse, TCP/IP est venu d'Unix et a été tout de suite utilisé, qui plus est par des centres de recherches et les universités, c'est-à-dire les premiers à avoir utilisé les réseaux de manière poussée. Le manque de normalisation de TCP/IP a été contre-balançé par une implémentation rapide et efficace, et une utilisation dans un milieu propice à sa propagation.

L'avenir d'OSI

Au niveau de son utilisation et implémentation, et ce malgré une mise à jour du modèle en 1994, OSI a clairement perdu la guerre face à TCP/IP. Seuls quelques grands constructeurs dominant conservent le modèle mais il est amené à disparaître d'autant plus vite qu'Internet (et donc TCP/IP) explose.

Le modèle OSI restera cependant encore longtemps dans les mémoires pour plusieurs raisons. C'est d'abord l'un des premiers grands efforts en matière de normalisation du monde des réseaux. Les constructeurs ont maintenant tendance à faire avec TCP/IP, mais aussi le WAP, l'UMTS etc. ce qu'il devait faire avec OSI, à savoir proposer des normalisations dès le départ. OSI marquera aussi les mémoires pour une autre raison : même si c'est TCP/IP qui est concrètement utilisé, les gens ont tendance et utilisent OSI comme le modèle réseau de référence actuel. En fait, TCP/IP et OSI ont des structures très proches, et c'est surtout l'effort de normalisation d'OSI qui a imposé cette "confusion" générale entre les 2 modèles. On a communément tendance à considérer TCP/IP comme l'implémentation réelle de OSI.

► Le modèle TCP/IP

Introduction

TCP/IP désigne communément une architecture réseau, mais cet acronyme désigne en fait 2 protocoles étroitement liés : un protocole de transport, **TCP** (Transmission Control Protocol) qu'on utilise "par-dessus" un protocole réseau, **IP** (Internet Protocol). Ce qu'on entend par "modèle TCP/IP", c'est en fait une architecture réseau en 4 couches dans laquelle les protocoles TCP et IP jouent un rôle prédominant, car ils en constituent l'implémentation la plus courante. Par abus de langage, TCP/IP peut donc désigner deux choses : le modèle TCP/IP et la suite de deux protocoles TCP et IP.

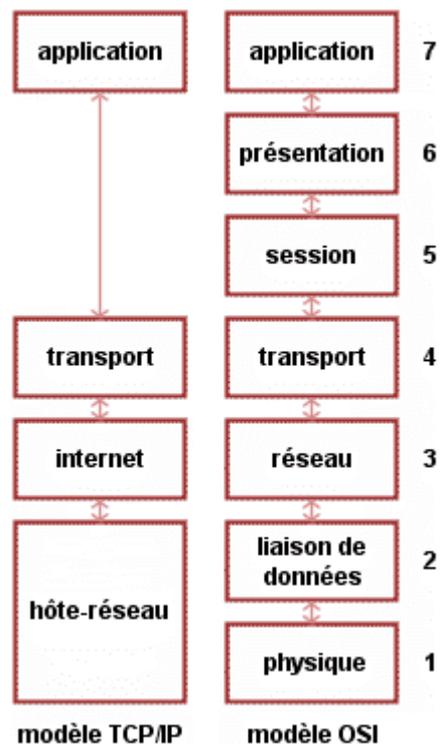
Le modèle TCP/IP, comme nous le verrons plus bas, s'est progressivement imposé comme modèle de référence en lieu et place du modèle OSI. Cela tient tout simplement à son histoire. En effet, contrairement au modèle OSI, le modèle TCP/IP est né d'une implémentation ; la normalisation est venue ensuite. Cet historique fait toute la particularité de ce modèle, ses avantages et ses inconvénients.

L'origine de TCP/IP remonte au réseau ARPANET. ARPANET est un réseau de télécommunication conçu par l'ARPA (Advanced Research Projects Agency), l'agence de recherche du ministère américain de la défense (le DOD : Department of Defense). Outre la possibilité de connecter des réseaux hétérogènes, ce réseau devait résister à une éventuelle guerre nucléaire, contrairement au réseau téléphonique habituellement utilisé pour les télécommunications mais considéré trop vulnérable. Il a alors été convenu qu'ARPANET utiliserait la technologie de commutation par paquet (mode datagramme), une technologie émergente promettante. C'est donc dans cet objectif et ce choix technique que les protocoles TCP et IP furent inventés en 1974. L'ARPA signa alors plusieurs contrats avec les constructeurs (BBN principalement) et l'université de Berkeley qui développait un Unix pour imposer ce standard, ce qui fut fait.

Description du modèle

Un modèle en 4 couches

Le modèle TCP/IP peut en effet être décrit comme une architecture réseau à 4 couches :



Le modèle OSI a été mis à côté pour faciliter la comparaison entre les deux modèles.

La couche hôte-réseau

Cette couche est assez "étrange". En effet, elle semble "regrouper" les couches physique et liaison de données du modèle OSI. En fait, cette couche n'a pas vraiment été spécifiée ; la seule contrainte de cette couche, c'est de permettre un hôte d'envoyer des paquets IP sur le réseau. L'implémentation de cette couche est laissée libre. De manière plus concrète, cette implémentation est typique de la technologie utilisée sur le réseau local. Par exemple, beaucoup de réseaux locaux utilisent Ethernet ; Ethernet est une implémentation de la couche hôte-réseau.

La couche internet

Cette couche est la clé de voûte de l'architecture. Cette couche réalise l'interconnexion des réseaux (hétérogènes) distants sans connexion. Son rôle est de permettre l'injection de paquets dans n'importe quel réseau et l'acheminement des ces paquets indépendamment les uns des autres jusqu'à destination. Comme aucune connexion n'est établie au préalable, les paquets peuvent arriver dans le désordre ; le contrôle de l'ordre de remise est éventuellement la tâche des couches supérieures.

Du fait du rôle imminent de cette couche dans l'acheminement des paquets, le point critique de cette couche est le **routing**. C'est en ce sens que l'on peut se permettre de comparer cette couche avec la couche réseau du modèle OSI.

La couche internet possède une implémentation officielle : le **protocole IP** (Internet Protocol). Remarquons que le nom de la couche ("internet") est écrit avec un i minuscule, pour la simple et bonne raison que le mot internet est pris ici au sens large (littéralement, "interconnexion de réseaux"), même si l'Internet (avec un grand I) utilise cette couche.

La couche transport

Son rôle est le même que celui de la couche transport du modèle OSI : permettre à des entités paires de soutenir une conversation.

Officiellement, cette couche n'a que deux implémentations : le **protocole TCP** (Transmission Control Protocol) et le **protocole UDP** (User Datagram Protocol). TCP est un protocole fiable, orienté connexion, qui permet l'acheminement *sans erreur* de paquets issus d'une machine à un internet à une autre machine du même internet. Son rôle est de fragmenter le message à transmettre de manière à pouvoir le faire passer sur la couche internet. A l'inverse, sur la machine destination, TCP replace dans l'ordre les fragments transmis sur la couche internet pour reconstruire le message initial. TCP s'occupe également du contrôle de flux de la connexion.

UDP est en revanche un protocole plus simple que TCP : il est non fiable et sans connexion. Son utilisation présuppose que l'on n'a pas besoin ni du contrôle de flux, ni de la conservation de l'ordre de remise des paquets. Par exemple, on l'utilise lorsque la couche application se charge de la remise en ordre des messages. On se souvient que dans le modèle OSI, plusieurs couches ont à charge la vérification de l'ordre de remise des messages. C'est là un avantage du modèle TCP/IP sur le modèle OSI, mais nous y reviendrons plus tard. Une autre utilisation d'UDP : la transmission de la voix. En effet, l'inversion de 2 phonèmes ne gêne en rien la compréhension du message final. De manière plus générale, UDP intervient lorsque le temps de remise des paquets est prédominant.

La couche application

Contrairement au modèle OSI, c'est la couche immédiatement supérieure à la couche transport, tout simplement parce que les couches présentation et session sont apparues inutiles. On s'est en effet aperçu avec l'usage que les logiciels réseau n'utilisent que très rarement ces 2 couches, et finalement, le modèle OSI dépouillé de ces 2 couches ressemble fortement au modèle TCP/IP.

Cette couche contient tous les protocoles de haut niveau, comme par exemple Telnet, TFTP (trivial File Transfer Protocol), SMTP (Simple Mail Transfer Protocol), HTTP (HyperText Transfer Protocol). Le point important pour cette couche est le choix du protocole de transport à utiliser. Par exemple, TFTP (surtout utilisé sur réseaux locaux) utilisera UDP, car on part du principe que les liaisons physiques sont suffisamment fiables et les temps de transmission suffisamment courts pour qu'il n'y ait pas d'inversion de paquets à l'arrivée. Ce choix rend TFTP plus rapide que le protocole FTP qui utilise TCP. A l'inverse, SMTP utilise TCP, car pour la remise du courrier électronique, on veut que tous les messages parviennent intégralement et sans erreurs.

Comparaison avec le modèle OSI et critique

Comparaison avec le modèle OSI

Tout d'abord, les points communs. Les modèles OSI et TCP/IP sont tous les deux fondés sur le concept de pile de protocoles indépendants. Ensuite, les fonctionnalités des couches sont globalement les mêmes.

Au niveau des différences, on peut remarquer la chose suivante : le modèle OSI faisait clairement la différence entre 3 concepts principaux, alors que ce n'est plus tout à fait le cas pour le modèle

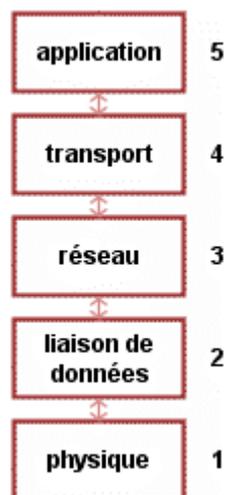
TCP/IP. Ces 3 concepts sont les concepts de *services*, *interfaces* et *protocoles*. En effet, TCP/IP fait peu la distinction entre ces concepts, et ce malgré les efforts des concepteurs pour se rapprocher de l'OSI. Cela est dû au fait que pour le modèle TCP/IP, ce sont les protocoles qui sont d'abord apparus. Le modèle ne fait finalement que donner une justification théorique aux protocoles, sans les rendre véritablement indépendants les uns des autres.

Enfin, la dernière grande différence est liée au mode de connexion. Certes, les modes orienté connexion et sans connexion sont disponibles dans les deux modèles mais pas à la même couche : pour le modèle OSI, ils ne sont disponibles qu'au niveau de la couche réseau (au niveau de la couche transport, seul le mode orienté connexion n'est disponible), alors qu'ils ne sont disponibles qu'au niveau de la couche transport pour le modèle TCP/IP (la couche internet n'offre que le mode sans connexion). Le modèle TCP/IP a donc cet avantage par rapport au modèle OSI : les applications (qui utilisent directement la couche transport) ont véritablement le choix entre les deux modes de connexion.

Critique

Une des premières critiques que l'on peut émettre tient au fait que le modèle TCP/IP ne fait pas vraiment la distinction entre les spécifications et l'implémentation : IP est un protocole qui fait partie intégrante des spécifications du modèle.

Une autre critique peut être émise à l'encontre de la couche hôte-réseau. En effet, ce n'est pas à proprement parler une couche d'abstraction dans la mesure où sa spécification est trop floue. Les constructeurs sont donc obligés de proposer leurs solutions pour "combler" ce manque. Finalement, on s'aperçoit que les couches physique et liaison de données sont tout aussi importantes que la couche transport. Partant de là, on est en droit de proposer un modèle hybride à 5 couches, rassemblant les points forts des modèles OSI et TCP/IP :



Modèle hybride de référence

C'est finalement ce modèle qui sert véritablement de référence dans le monde de l'Internet. On a ainsi gardé la plupart des couches de l'OSI (toutes, sauf les couches session et présentation) car correctement spécifiées. En revanche, ses protocoles n'ont pas eu de succès et on a du coup gardé ceux de TCP/IP.