TCP/IP

Les interconnexions de réseaux

Compilation par Pierre-Alain Muller

Avant propos

Au siècle dernier, le chemin de fer a révolutionné le monde (avènement de la société industrielle). Aujourd'hui, les réseaux amorcent une nouvelle révolution (avènement de la société de l'information). Les réseaux gomment les distances géographiques et créent de nouvelles communautés qui interagissent fréquemment.

Petite histoire des réseaux

avant 1960 : comment acheminer de façon fiable des bits le long d'un support de communication => théorie de l'information (traitement de signal)

vers 1965 : comment acheminer de façon fiable et efficace des paquets => commutation de paquets, réseaux locaux, analyse statistique du temps de réponse.

de 1975 à nos jours : comment fournir des services de communication aux travers d'une interconnexion de réseaux => techniques d'interconnexion, modèles de protocoles en couches, services de transport en mode connecté, et le modèle client-serveur.

Dans ce qui suit, nous étudierons l'architecture des réseaux interconnectés

1. Introduction et présentation générale

Les agences gouvernementales américaines (NSF, DoE, DoD, HHS, NASA, DARPA, ...) ont subventionné la réalisation d'une interconnexion qui couvre tout le territoire américain. La DARPA (Defense Advanced Research Projects Agency) définit la suite de protocoles TCP/IP.

1.1 Les services d'Internet

Internet définit un ensemble de normes, aussi appelées protocoles, c'est-à-dire de règles indépendantes des constructeurs pour échanger des informations. Les avantages de TCP/IP sont multiples :

Indépendance par rapport aux technologies des réseaux

Les protocoles définissent une unité de transmission appelée « datagram » et spécifie comment transmettre les datagrammes sur un type de réseau donné.

Connectivité universelle

Une interconnexion TCP/IP permet à toute paire de machine qui y est reliée de communiquer. Une adresse (unique au sein de l'Internet) est affectée à chaque machine. Chaque datagramme contient les adresses source et destination. Les noeuds de commutation intermédiaires utilisent ces adresses pour prendre leurs décisions de routage.

Accusés de réception de bout en bout

Même lorsque les machines source et destination ne sont pas reliées au même réseau physique.

Protocoles d'application normalisés

TCP/IP inclut de nombreux protocoles d'application pour la messagerie, le transfert de fichiers, la connexion à distance, ...

1.1.1 Services d'interconnexion de niveau application

TCP/IP se présente comme un ensemble de programmes d'application qui utilisent le réseau pour effectuer des tâches utiles. Les utilisateurs de ces programmes n'ont pas besoin de connaître TCP/IP, ils n'utilisent que les programmes d'application (courrier électronique, transfert de fichiers, connexion à distance)

1.1.2 Services d'interconnexion de niveau réseau

Il existe deux grands types de services.

Service de remise de paquets en mode non connecté

Ce service, à la base de tous les autres services, assure le routage de petits messages d'une machine à une autre, en utilisant les adresses qui sont contenues dans les messages. Chaque paquet est routé indépendamment, il n'y a pas de garantie de fiabilité ou de séquentialité. Ce service est extrêmement efficace, car il est en correspondance directe avec le matériel sous-jacent.

Service de remise en mode connecté

Ce service assure le transport fiable des messages, il simule une connexion physique de bout en bout.

1.2 Les appels à commentaires Internet (RFC)

La documentation relative aux protocoles, normes et stratégies, peut être obtenue auprès du NIC (Network Information Center), sous la forme de RFC (Request for comments), par courrier électronique ou transfert de fichiers.

1.3 Evolution

Chaque jour de nouveaux groupes rejoignent l'Internet. Physiciens, chimistes, scientifiques de l'espace manipulent et échangent des informations en quantité beaucoup plus importante que les informaticiens. Demain, en fait aujourd'hui, les particuliers se connectent également à Internet.

	Nombre de réseaux	de machines	d'administrateurs
1980	10^1	10^2	10^0
1990	10^3	10^5	10^1
1995	10^5	10^{10}	10^2

Cet accroissement important a demandé de passer d'une gestion centralisée à une gestion hiérarchique et décentralisée du réseau Internet.

2. Présentation des technologies sous-jacentes

L'Internet ne constitue pas un nouveau type de réseau physique. La technologie matérielle joue un rôle mineur dans l'ensemble, mais il important de faire la distinction entre

mécanismes de bas niveau, assurés par le matériel, et services de haut niveau, fournis par le logiciel.

2.1 Deux approches des réseaux de communication

La commutation de circuits

Repose sur l'établissement d'une connexion dédiée entre deux points (téléphone). Cette technique présente l'intérêt d'offrir une capacité garantie, mais détriment d'un coût fixe, indépendant du trafic (en fait cela peut être perçu comme un avantage, car on sait à l'avance ce que l'on paye !)

La commutation de paquets

Le trafic émis sur le réseau est tronçonné en petits morceaux, appelés paquets, qui sont multiplexés sur les liens à haut débits qui relient les machines. Chaque paquets contient des informations qui identifient le destinataire. La commutation de paquets permet des communications simultanées entre machines. Ce type de réseaux est devenu extrêmement populaire.

2.2 Types de réseaux

WAN	Wide Area Network	9,6 Kbits/s	à	45 Mbits/s
MAN	Metropolitan Network	56 Kbits/s	à	100 Mbits/s
LAN	Local Area Network	4 Mbits/s	à	2 Gbits/s

2.3 La technologie Ethernet

La technologie Ethernet utilise la commutation de paquets ; elle a été inventée par Xerox au PARC, au début des années 70, puis normalisée en 1978 par Xerox, Intel et Digital Equipment.

Un réseau Ethernet est constitué d'un câble coaxial d'environ 1,27 cm de diamètre et d'une longueur de 500 m au plus. Chaque extrémité est terminée par une résistance d'impédance identique à l'impédance caractéristique du câble (52 Ohms). Le câble lui-même, appelé « ether » est entièrement passif.

Les réseaux Ethernet peuvent être étendus au moyen de répéteurs. Deux répéteurs au plus peuvent être placés entre deux machines quelconques, ce qui limite la longueur d'un réseau Ethernet à environ 1500 mètres. Les connexions sont réalisées au moyen de robinets « taps », appelés également vampires lorsqu'ils mordent le câble. Chaque connexion comporte deux composants électroniques majeurs : un « transceiver » et un coupleur.

2.3.1 Propriétés d'un réseau Ethernet

Un réseau Ethernet est un bus à contrôle distribué, d'une capacité nominale de 10 Mbit/s, basé sur la notion de remise pour le mieux (« best effort delivery »). Les stations sont en compétition les unes avec les autres et l'accès au support n'est pas garanti, des données peuvent donc être perdues. Les informations sont diffusées (« broadcast »), les transceivers ne filtrent pas les données, ce sont les coupleurs qui éliminent les paquets non désirés. Il n'y a pas d'autorité centrale responsable de l'allocation du support, la méthode d'accès est dite CSMA/CD (« Carrier Sense Multiple Access with Collision Detect »).

CSMA : plusieurs machines peuvent accéder simultanément au câble, chaque machine détermine si le câble est disponible en y détectant une onde porteuse. En l'absence de transmission, l'interface de la machine commence à transmettre pour une durée limitée (afin de ne pas monopoliser le support)

CD : Le signal se propage dans le câble à environ 80 % de la vitesse de la lumière, et n'atteint pas simultanément les deux extrémités du câble. Deux TRX peuvent donc décider en même temps que le support est libre, ce qui engendre un mélange des signaux électriques. Ce genre d'incident est appelé « collision ». Les collisions sont détectées par chaque transceiver (par comparaison entre le signal émis et le signal sur le câble). En cas de collision, le coupleur interrompt la transmission pendant un temps aléatoire, avant de réémettre. En cas de nouvelle collision, la plage de délai est doublée, quadruplée, ...

2.3.2 Variantes de l'Ethernet

La technologie Ethernet a été amélioré de multiples manières :

Paires torsadées : câble de type téléphone, bon marché, sans blindage 10baseT et 100baseT.

Ethernet fin : câble coax moins cher, connexion BNC, transceiver et coupleur intégrés sur la même carte.

Ethernet bande large : multiplexage de plusieurs Ethernet sur un même câble (porteuses différentes)

2.3.3 L'adressage Ethernet

L'interface filtre les paquets qui ne sont pas destinés à la machine. Chaque machine (en fait chaque coupleur) connecté à un réseau Ethernet possède un identificateur unique (un entier sur 48 bits) qui constitue son adresse Ethernet.

Les adresses Ethernet sont partagées en 3 types :

- ♣ adresse d'une interface de réseau
- ♦ adresse de diffusion sur un réseau
- ♥ adresse de diffusion de groupe

Par convention, l'adresse de diffusion (tout à 1) est réservée à l'émission vers toutes les machines du réseau.

2.3.4 Structure des trames Ethernet

Le réseau Ethernet est une connexion entre machines au niveau liaison, les données échangées sont structurées en paquets appelés trames (« frames »). Les trames ont une longueur variable entre 64 octets et 1518 octets.

Préambule	Destination	Source	Type	Données	CRC
64 bits	48 bits	48 bits	16 bits	368-1200 bits	32 bits

Le préambule contient une alternance de 64 bits à 1 et 0, terminée par 2 bits à 1. Le type de trame permet l'utilisation simultanée de plusieurs protocoles, le CRC (Cyclic Redundancy Check) contrôle la validité des informations.

2.3.5 Les ponts et leur importance

Un pont est un calculateur rapide, doté de deux interfaces réseau.

Les ponts sont supérieurs aux répéteurs car ils ne reproduisent pas les bruits et les erreurs, ni les trames mal-formées. De plus, les ponts respectent les règles CSMA/CD de sorte que les collisions et le délai de propagation d'un segment demeure indépendant de l'autre segment. Grâce aux ponts, un nombre quasi illimité de réseaux Ethernet peuvent être reliés. Les ponts masquent les détails des interconnexions, le tout se comporte comme un réseau unique.

La plupart des ponts prennent des décisions intelligentes relatives aux trames à acheminer. Ces ponts sont appelés « ponts filtrants » ou « ponts auto-adaptatifs » (« learning bridges »).

3. Les adresses Internet

Ce chapitre traite de l'adressage qui permet à TCP/IP de masquer les détails des réseaux physiques afin de faire apparaître les interconnexions de réseaux comme un réseau unique et uniforme.

3.1 Identificateurs universels

Un système de communication fournit un service de communication universel s'il permet à toute machine de communiquer avec toute autre machine. Il faut donc une méthode générale pour identifier les machines.

3.2 Les trois premières classes d'adresses IP

Une adresse « Internet », aussi appelée adresse IP, est un entier sur 32 bits, soigneusement choisi pour assurer l'efficacité du routage. En particulier, une adresse IP inclut l'identification du réseau et d'une machine particulière. Conceptuellement, chaque adresse IP est une paire (nom de réseau, nom de machine). Les adresses IP sont regroupées en 5 classes (A, B, C, D, E).

```
Classe    1 2 3 4 5 6 7 8         16        24        32
A         0 <--réseau--><-------------machine--------->
B         1 0 <------réseau--------><-------machine---->
C         1 1 0 <---------réseau------------><-machine>
D         1 1 1 0 <---adresse de diffusion de groupe--->
E         1 1 1 1 0 <------réservée pour le futur------>
```

Les deux bits de poids fort suffisent à différencier les 3 classes primaires d'adresses.

Les adresses de classe A sont utilisées pour quelques réseaux de plus de 2^{16} machines. Il n'est aujourd'hui pratiquement plus possible d'obtenir des numéros de classe A ou B. Les organisations demandent donc 1 ou plusieurs numéros de classe C (nous en avons 5 à l'essaim).

3.3 Adresses de réseau et adresses de diffusion

En plus de permettre un routage efficace, le codage des informations réseau dans les adresses permet également de référencer les réseaux indépendamment des machines. L'adresse réseau est obtenu en plaçant tous les bits machine à 0.

Un autre avantage significatif du plan d'adressage IP est d'inclure une adresse de diffusion (« broadcast address ») qui référence toutes les machines du réseau. Conformément à la norme, tout identificateur de machine ne comportant que des 1 est réservé à la diffusion. Avec la technologie Ethernet, la diffusion est aussi efficace que la transmission normale.

3.4 Faiblesse de l'adressage Internet

Le fait de coder l'information réseau dans l'adresse IP présente le désavantage d'obliger à changer l'adresse IP d'une machine lorsqu'elle est déplacée d'un réseau à un autre, ce qui pose un réel problème pour l'informatique nomade.

3.5 La notation décimale pointée

Les adresses IP sont généralement représentées sous forme de quatre entiers décimaux, séparés par un point décimal.

```
1000 0000 0000 1010 0000 0010 0001 1110
128.10.2.30
```

3.6 La boucle locale (« loopback »)

Il s'agit d'une adresse particulière qui n'apparaît jamais sur un réseau (127.0.0.0) et qui est destinée à permettre les communications inter-processus sur la machine locale. Lorsque n'importe quel programme d'application utilise cette adresse, le logiciel de l'interface réseau renvoie les données, sans les émettre sur le réseau.

4. Association des adresses IP et des adresses physiques (ARP)

Dans le plan d'adressage IP, chaque machine est décrite par une adresse IP sur 32 bits. Il nous faut maintenant étudier comment une machine ou une passerelle met une adresse IP en

correspondance avec la bonne adresse physique (qui dépend de la technologie de réseau employée).

4.1 Résolution par association dynamique

Le protocole ARP (« Address Resolution Protocol ») fournit un mécanisme efficace et simple pour résoudre les adresses dynamiquement. Lorsqu'une machine A veut résoudre une adresse IP I_B, elle diffuse (donc vers tout le monde) un paquet spécial qui demande, à la machine dont l'adresse IP est I_B, de répondre en indiquant sont adresse physique P_B. Toutes les machines, B incluse, reçoivent la requête mais seule la machine B reconnaît son adresse IP et renvoie donc un message contenant son adresse physique.

Il n'est pas réaliste de diffuser chaque fois qu'une machine A veut atteindre une machine B, car c'est beaucoup trop coûteux (toutes les machines doivent traiter le paquet de diffusion). Pour réduire les coûts de communication, les machines qui utilisent ARP gèrent un tampon dans lequel elles enregistrent les associations entre adresses physiques et adresses IP les plus récentes.

Plusieurs améliorations sont possibles :

l'association entre adresse IP et Physique de l'émetteur est incluse dans chaque demande ARP diffusée, et tous les récepteurs extraient l'information

une machine qui démarre diffuse son adresse physique.

4.2 Encapsulation et identification de ARP

Les messages ARP sont véhiculés dans le champs de données d'une trame :

```
En tête de trame Zone de données de trame
<--------------><--------Message ARP------->
```

Le champs de type vaut : 16 #0806#

5. Détermination d'une adresse IP au démarrage

L'adresse IP des machines est souvent enregistrée dans la mémoire permanente et est lue durant le démarrage de la machine. Les machines dépourvues de mémoire secondaire (TX, Diskless) doivent obtenir leur adresse avant que le système d'exploitation ne commence à s'exécuter. La procédure peut sembler paradoxale : une machine communique avec un serveur distant pour obtenir une adresse nécessaire pour communiquer. En fait, la machine cliente recours temporairement à son adresse physique, pour communiquer sur un réseau physique donné avec une autre machine qui contient les associations entre adresses physiques et adresses IP. Le serveur n'est pas connu du client, de nouveau la diffusion est employée est un ou plusieurs serveurs répondent.

5.1 Le protocole de résolution d'adresse inverse (RARP)

Le protocole RARP est l'inverse du protocole ARP dont il est dérivé. Un message RARP permet à une machine de demander l'adresse IP d'une machine tierce aussi facilement que sa propre adresse. Un message RARP est acheminé d'une machine à une autre, encapsulé dans le champs de données d'une trame Ethernet. Le type de trame contient la valeur 16#8035#.

Lorsqu'une machine diffuse une demande RARP, toutes les machines du réseau reçoivent la requête, mais seules des machines particulières, dénommées serveurs RARP, répondent à la

demande. L'adresse IP, une fois obtenue, est enregistrée en mémoire centrale et le protocole de résolution d'adresse inverse n'est plus utilisé jusqu'à la prochaine mise sous tension.

6. Remise de datagrammes en mode non connecté

6.1 Système de remise en mode sans connexion

Le système de remise de paquets est le service fondamental des interconnexions de réseaux. Il s'agit d'un service de remise pour le mieux (« best effort delivery »), non fiable et sans connexions.

Le service est dit non fiable car la remise des paquets n'est pas garantie. Des paquets peuvent être perdus, dupliqués, retardés, altérés ou remis dans le désordre. Le service ne détecte pas ces conditions et n'en informe ni l'émetteur, ni le destinataire.

Le service est dit sans connexion car chaque paquet est traité indépendamment de tous les autres. Les paquets d'une suite, émis depuis une machine vers une autre, peuvent emprunter des chemins différents, certains peuvent être perdus tandis que d'autres sont correctement remis.

Enfin, le service assure une remise dite pour le mieux parce que le logiciel d'interconnexion de réseaux ne détruit pas de paquets de sa propre initiative. La remise est non fiable uniquement lorsque toutes les ressources sont saturées, ou lorsque le réseau tombe en panne.

Le protocole qui définit le système de remise non fiable en mode sans connexion est appelé « Internet Protocol » -> IP.

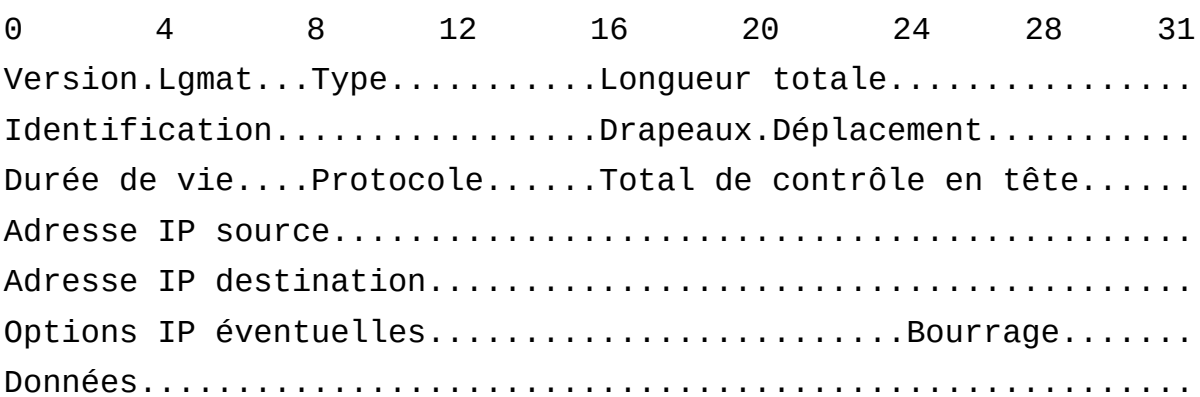
IP donne trois définitions importantes :

- ♣ l'unité de donnée transférée (la structure de la donnée)
- ♦ la fonction de routage
- ♥ des règles sur le concept de remise non-fiable

6.2 Datagrammes IP

L'analogie entre un réseau physique et une interconnexion TCP/IP est forte. Sur un réseau physique, l'unité de donnée transférée est la trame. Pour IP, l'unité de donnée transférée est appelée datagramme (« datagram »). Un datagramme contient les adresses source et destination, ainsi qu'un champs de type, qui identifie la nature des données contenues dans le datagramme.

6.2.1 Structure des datagrammes



```
Données.....................................................
Données.....................................................
```

Tous les champs de l'entête sont de taille fixe, sauf les options et le bourrage (« padding »).

Description des champs :

Version : numéro de version du protocole IP utilisé pour créer le datagramme.

Lgmat : longueur de l'entête en mots de 32 bits.

Lgr Totale : longueur du datagramme (en octets), entête et données incluses. La taille max du datagramme est de 2^{16} =65536 octets.

Type : type de service, indique comment le datagramme doit être géré. il se décompose en 5 parties.

```
0    1    2    3    4    5    6    7
priorité..D....T....R....Inutilisés..
```

La priorité du datagramme qui varie de 0 à 7 (0 = valeur normale, 7 = supervision du réseau) permet d'assurer la priorité des informations de contrôle par rapport aux données. Ainsi, les algorithmes de contrôle de congestion ne sont pas affectées par les phénomènes qu'ils tentent de maîtriser.

Les bits D, T, R indiquent le type d'acheminement désiré pour le datagramme :

D = délai court, T = débit élevé, R = grande fiabilité

Ces informations sont destinées aux algorithmes de routage, ce ne sont pas des exigences.

6.2.2 Taille des datagrammes, MTU réseau et fragmentation

Dans le cas idéal, l'intégralité du datagramme tient dans le champs de données d'une trame physique, ce qui rend la transmission efficace (Ethernet utilise la valeur 16#0800# pour indiquer que des données sont encapsulées dans un datagramme IP).

Chaque réseau physique définit une taille limite qui lui est propre (Ethernet 1500 octets, Pro-Net 2044, X25 128, ...). Au lieu de choisir des tailles de datagrammes compatibles avec les contraintes des réseaux physiques, les concepteurs de TCP/IP ont choisi une taille initiale pratique et ont mis au point une technique permettant de découper les grands datagrammes en morceaux plus petits, si besoin est. Ce processus est dénommé fragmentation.

La taille de chaque fragment est choisie de façon à permettre la transmission de chaque sous-fragment dans une seule trame du réseau physique sous-jacent. De plus, comme IP représente le déplacement sous forme de multiples de 8 octets, la taille des fragments doit être un multiple de 8.

Chaque fragment contient une entête de datagramme qui reproduit la majeure partie de l'entête initiale du datagramme (sauf un but dans le champ drapeaux qui indique qu'il s'agit d'un fragment.

6.2.3 Réassemblage des fragments

Dans une interconnexion TCP/IP, les fragments sont acheminés comme autant de datagrammes indépendants, jusqu'à leur destination finale. Le réassemblage au niveau de la

destination finale permet le routage indépendant des fragments, les passerelles intermédiaires ne sont pas obligées de stocker ou réassembler les fragments.

L'entête contient des informations pour le contrôle de la fragmentation :

Identification : un entier unique qui identifie le datagramme, et qui est recopié dans chaque fragment.

Déplacement fragment : déplacement (en multiple de 8 octets) des données transportées dans le fragment courant par rapport au datagramme initial. Les fragments n'arrivent pas nécessairement dans le bon ordre et il n'y a aucune communication entre la passerelle qui a réalisé la fragmentation et la machine qui effectue le réassemblage.

Les drapeaux : le premier bit (dit de non fragmentation) indique si le datagramme peut être fragmenté. Le deuxième bit indique si le fragment contient des données issues du début, du milieu ou de la fin du datagramme.

Durée de vie : indique en seconde la durée maximale de transit du datagramme dans l'interconnexion de réseaux. La durée de vie est définie par la machine qui émet le datagramme dans l'interconnexion de réseaux. Les machines et les passerelles qui traitent le datagramme doivent décrémenter la durée de vie du datagramme au fur et à mesure que le temps s'écoule. Le datagramme est détruit lorsque sa durée de vie expire.

Protocole : analogue au champs type de trame du protocole Ethernet.

Total de contrôle : assure l'intégrité des données contenues dans l'entête. Les données ne sont pas contrôlées à ce niveau de protocole.

Adresses IP : source et destination.

6.2.4 Les options des datagrammes IP

Les options sont incluses essentiellement à des fins de test ou de mise au point. Il est par exemple possible d'enregistrer la route suivie par un datagramme dans l'interconnexion de réseaux (suite des adresses IP des passerelles traversées).

7. Routage des datagrammes IP

Dans un système de commutation de paquets, le routage fait référence au processus de choix d'un chemin pour transmettre les paquets, et le terme de routeur désigne toute machine effectuant un tel choix.

7.1 Types de routage

Sur un réseau physique donné, une machine peut envoyer directement une trame physique à une autre machine reliée à ce même réseau.

Routage indirect : l'émetteur doit identifier la passerelle vers laquelle envoyer un datagramme, puis la passerelle doit acheminer le datagramme vers le réseau correspondant à la destination. Les algorithmes de routage usuels utilisent, sur chaque machine, une table de routage Internet. Cette table contient des adresses de réseaux (valables pour toutes les machines de ces réseaux) accessibles par l'intermédiaire des passerelles directement reliées au réseau physique.

Routage par défaut : lorsque la destination n'apparaît pas dans la table de routage, le datagramme est envoyé à une passerelle particulière, désignée par la route par défaut.

7.2 Routage avec les adresses IP

Le routage IP n'altère pas le datagramme initial. Les adresses IP, calculées par l'algorithme de routage, sont connues sous le nom d'adresse de prochain saut (« next hop address ») et ne sont pas conservées après transmission du datagramme sur le réseau physique.

7.3 Gestion des datagrammes entrants

Un datagramme reçu par une machine est remis, par le logiciel de l'interface réseau, au logiciel IP pour être traité. Si l'adresse IP du datagramme coïncide avec l'adresse IP de la machine, le logiciel IP accepte le datagramme et le transmet au protocole de plus haut niveau pour complément de traitement. Sinon, le datagramme est détruit (les machines ne corrigent pas les erreurs de routage).

8. Protocole de datagramme utilisateur

Jusqu'à présent, nous avons considéré la machine destinataire comme une entité globale. En fait, les machines exécutent des processus qui sont eux les destinataires réels de datagrammes. Toutefois, indiquer un processus particulier, sur une machine spécifique, n'est pas d'une grande souplesse. Pour cette raison, les concepteurs de TCP/IP ont introduit une abstraction intermédiaire, qualifiable de destination abstraite, les « ports de protocole ».

Chaque port est identifié par un entier positif, et le système d'exploitation local fournit une interface que les processus utilisent pour spécifier un port. La plupart des OS assurent un accès synchrone aux ports. Les données reçues sont conservées un certain temps dans une file d'attente, jusqu'à ce qu'un processus vienne les extraire.

Pour communiquer avec un port distant, l'émetteur a besoin de connaître à la fois l'adresse IP de la machine destination et le numéro de port associé au protocole sur cette machine. Chaque message émis contient à la fois le numéro de port destination sur la machine destination, et le numéro de port source sur la machine source à qui les réponses seront adressées.

8.1 Le protocole UDP

Dans la suite de protocole TCP/IP, UDP (« User Datagram Protocol ») assure le mécanisme de base qui permet aux programmes d'application d'envoyer des datagrammes à d'autres programmes d'application. Le protocole UDP fournit des ports de protocoles qui permettent de distinguer différents programmes d'application.

En plus des données émises, les messages UDP contiennent les numéros de port destination et source.

UDP assure un service de remise non fiable, en mode sans connexion, qui utilise IP pour acheminer les messages entre machines. Un programme qui utilise UDP doit gérer les problèmes de fiabilité, pertes de messages, duplications, retards, dé-séquencements et pertes de connectivité.

Piège : le logiciel fonctionne bien sur un réseau local simple mais pas sur une grande interconnexion.

8.2 Structure des messages UDP

Un datagramme utilisateur comporte deux parties : un entête UDP et une zone de données.

```
0         8         16                  31
Port UDP source     Port UDP destination
Longueur msg UDP    Total de contrôle
Données .................................
.........
```

Le calcul du total de contrôle est facultatif => pour gagner en vitesse.

8.3 Numéros de ports réservés et non réservés

Certains ports sont réservés, mais la plus grande partie est laissée à la disposition des sites ou des programmes d'application locaux. Les numéros réservés sont affectés en croissant à partir des valeurs les plus faibles. Les valeurs les plus grandes peuvent être allouées dynamiquement.

Ex : les TX utilisent `tftp` (port 69)pour charger leur logiciel de serveur X.

9. Transfert fiable en mode connecté (TCP)

TCP a pour objet de permettre le transfert d'information de façon fiable quel que soit le débit, le volume ou la distance, et ceci indépendamment des réseaux formant l'interconnexion.

9.1 Propriétés des services de remise fiable

Orientation connexion. Lors de transfert de gros volume de données, les flots de bits sont décomposés en octets (8 bits). Un service de remise fiable transmet à la machine destination autant de bits que l'émetteur en a envoyé depuis la machine source.

Circuits virtuels. Les programmes d'application voient les connexions comme des circuits matériels dédiés. La fiabilité est une illusion assurée par le service de remise fiable.

Transferts tamponnés. Les applications utilisent toute taille de données qui leur parait adaptée. Le logiciel de communication est libre de composer ou de décomposer le flot en paquets indépendants du choix de l'application (à des fins d'efficacité). Les applications peuvent forcer l'émission des données (même si le tampon n'est pas plein) au moyen d'un mécanisme dénommé « push ». La commande pousser agit de bout en bout, elle force la transmission et la remise. Par contre, la taille des enregistrements reste du ressort des couches plus basses.

Connexions non structurées. Les applications qui utilisent TCP doivent définir et comprendre la structure des données applicatives échangées.

Connexions bidirectionnelles simultanées (« full duplex »). Les programmes d'applications voient deux flots indépendants, de sens contraire. Le protocole sous-jacent peut envoyer des informations de contrôle relatives à un sens de transfert dans l'entête TCP des segments émis en sens inverse (« piggy backing »).

9.2 Assurer la fiabilité

La technique fondamentale est nommée acquittement positif avec retransmission (« positive acknowledgment with retransmission »). Le récepteur dialogue avec l'émetteur et lui envoie des accusés de réception à chaque arrivée de données. La source conserve une copie des paquets émis et attend un accusé de réception avant d'émettre le paquet suivant. L'émetteur arme une temporisation lorsqu'il émet un paquet et retransmet le paquet si la temporisation expire avant l'arrivée de l'accusé de réception.

Le schéma se complique lorsque la couche sous-jacente duplique un paquet (données ou accusé de réception). En général, les protocoles fiables détectent les paquets dupliqués en affectant à chaque paquet un numéro de séquence et en exigeant que le récepteur se souvienne des numéros reçus. Inversement, pour éviter les ambiguïtés dues aux accusés de réception retardés ou dupliqués, les protocoles d'acquittement positif et de retransmissions mentionnent les numéros de séquence dans les accusés de réception. Le récepteur peut ainsi associer correctement les accusés de réception aux paquets qui leur correspondent.

9.3 Le concept de fenêtre glissante (« sliding window »)

Le protocole d'acquittement positif gaspille une quantité significative de bande passante, car la transmission d'un nouveau paquet est retardée, tant que l'accusé de réception correspondant n'a pas été reçu par l'émetteur du paquet précédent.

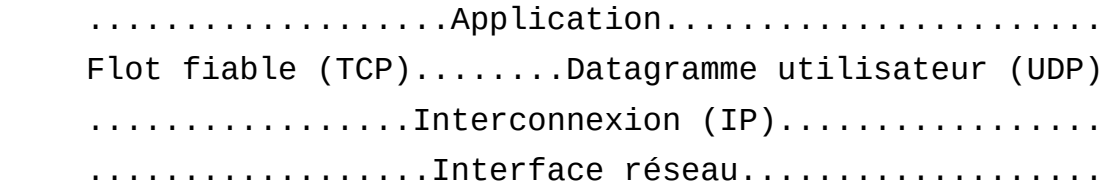
La technique de la fenêtre glissante consiste à envoyer un petit nombre de paquets sans attendre l'accusé de réception d'un paquet pour envoyer le suivant. Ce système, basé sur le fait qu'un réseau fonctionne plus souvent qu'il ne fonctionne pas, permet d'obtenir un débit sensiblement plus élevé qu'un simple protocole d'accusé positif.

9.4 Le protocole de contrôle de transmission (TCP)

TCP est un protocole, pas une application. Le document relatif au protocole définit uniquement les fonctions du protocole, pas la façon précise d'accéder à ces fonctions. TCP ne suppose que peu de choses sur le système de communication sous-jacent. Il est utilisable avec un grand nombre de systèmes de remise de paquets, parmi lesquels le service de remise de datagramme IP.

9.5 Ports, connexions et extrémités de connexions

Comme UDP, TCP se trouve au-dessus de IP dans le modèle en couche. TCP (comme UDP) permet à plusieurs applications de communiquer en même temps, en démultiplexant les informations reçues vers les différents programmes utilisateurs, à partir des numéros de port pour identifier les destinations finales sur une machine donnée.



TCP utilise les connexions, et non le port, comme concept principal. Les connexions sont alors identifiées par leurs deux extrémités. Une extrémité de connexion est définie par une paire de nombres entiers `(machine, port)`. Une connexion est définie par une paire d'extrémités `((machine i, port i), (machine j, port j))`.

Le concept de connexion permet de partager une même extrémité entre plusieurs connexions. Ainsi, il est possible de concevoir des programmes serveurs, qui assurent un service simultané de plusieurs connexions, sans avoir besoin de numéros de port locaux spécifiques pour chaque connexion.

9.6 Taille de fenêtre variable et contrôle de flux

Avec TCP la taille de la fenêtre varie dans le temps. Chaque accusé de réception indique le nombre d'octets correctement reçus et contient une indication de taille de fenêtre (« window advertisement ») qui indique le nombre d'octets supplémentaires que le récepteur est prêt à accepter. Lorsque les indications de taille de fenêtre sont inférieures à la valeur précédente, la fenêtre se réduit au fur et à mesure qu'elle se déplace.

Le mécanisme de taille de fenêtre variable permet d'effectuer du contrôle de flux, ainsi que du transfert fiable. Dans les cas extrêmes, le récepteur donne une indication de taille de fenêtre nulle pour arrêter la transmission des données.

9.7 Structure des segments

9.8 Temporisations et retransmissions

9.9 Echantillonage précis du temps de boucle

9.10 Algorithme de Karn et augmentation des temporisations

9.11 Etablissement d'une connexion TCP

TCP utilise un processus en trois temps :

Emetteur	Réseau	Récepteur
---->	SYN séq x	----->
<----	SYN séq y, ACK x+1	<-----
---->	ACK y+1	----->

Le processus en trois temps permet d'établir une synchronisation correcte entre les deux extrémités de la connexion, car TCP ignore les demandes de connexions supplémentaires reçues après l'établissement de la connexion (du fait de la retransmission possible de la demande initiale). Chaque extrémité est prête à transférer des données et les deux parties se sont accordées sur les numéros de séquence de départ.

9.12 Libération d'une connexion TCP

TCP utilise un processus en trois temps modifié pour libérer une connexion :

Emetteur	Réseau	Récepteur
---->	FIN séq x	----->
<----	ACK x+1	<-----
		Messages vers les applications
		pour avertir de la fin
		<----- réponses des applications
<----	FIN y, ACK x+1	<-----
----->	ACK y+1	----->

9.13 Numéros de ports TCP réservés

Comme UDP, TCP combine l'association statique et dynamique. Moins de 256 ports sont réservés, le reste est laissé à la disposition des applications. (FTP 21 et 20, Telnet 23).

9.14 Performance de TCP

Malgré sa complexité, TCP est un protocole très efficace. Le logiciel TCP qui fonctionne sur l'Internet peut assurer un débit soutenu de 8 Mbit/s entre deux stations reliées par un réseau Ethernet à 10 Mbit/s. Cray Research Inc. a obtenu des débits de plus de 600 Mbit/s, soit une valeur du même ordre de grandeur qu'un bus.

10. Le système de noms de domaine

Les noms de domaines permettent d'attribuer des noms symboliques significatifs à un grand nombre de machines, à travers une correspondance entre noms et adresses IP.

10.1 Les noms hiérarchiques

Le mécanisme de nommage est décentralisé à des organismes responsables d'un sous-ensemble de l'espace de nommage. D'un point de vue syntaxique, la subdivision de l'espace des noms introduit une partition supplémentaire dans le nom. Ainsi :

```
lsi.essaim.univ-mulhouse.fr
```

10.2 Noms de domaines officiels et officieux

L'administration d'Internet à choisi de diviser les domaines supérieurs selon la liste suivante :

```
Nom de domaine   Correspondance
com              entreprises commerciales
edu              enseignement
gov              gouvernement (US)
mil              militaires (US)
```

net	sites réseau d'importance majeure
org	organisation
int	organisation internationales
code du pays	chaque pays

Notons qu'il n'est pas possible de distinguer les noms de sous-domaines des noms d'objets individuels, d'après le seul système des noms.

11. La connexion à distance (Telnet, rlogin)

Telnet permet d'établir une connexion TCP avec le serveur Telnet d'une autre machine. Les caractères frappés sur le clavier de la machine locale sont transmis à la machine distante, qui les reçoit comme s'ils avaient été frappés sur son propre clavier. Telnet achemine également les réponses du système distant vers l'écran du terminal de l'utilisateur.

11.1 Les services de Telnet

Telnet propose trois services de base :

un terminal virtuel réseau : avec une interface normalisée

un mécanisme de négociation d'options entre client et serveur

un traitement symétrique des deux extrémités de la connexion

Un point d'accès au système d'exploitation est appelé « pseudo-terminal ».

Pour prendre en compte l'hétérogénéité des équipements du réseau, telnet spécifie les séquences de commandes et la façon dont les données sont émises sur le réseau.

```
Représentation client
     Client
          Représentation virtuelle
               Serveur
                    Représentation serveur
```

Telnet utilise un signal « hors bande » lors de l'émission des fonctions de contrôle (à partir du mécanisme TCP de données urgentes).

11.2 Les options de Telnet

Telnet permet de négocier les options de manière symétrique entre un client et un serveur, ceci permet de reconfigurer une connexion au mieux. Comme tous les logiciels Telnet comprennent un protocole de terminal virtuel de base, clients et serveurs peuvent coopérer même si l'un d'entre eux comprend une option que l'autre ne comprend pas.

Dans la terminologie Telnet, les termes de la négociation se présentent comme suit :

Will X	Do X
	Don't X

Ainsi, il est possible de faire inter-opérer des versions récentes et sophistiquées avec des versions plus anciennes et rudimentaires.

11.3 rlogin

Rlogin est un système de connexion à distance spécifique des machines UNIX, par opposition à Telnet, plus général (indépendant des types de machine et des OS). Rlogin intègre la notion de machine privilégiée, ce qui permet à un utilisateur de se connecter sur une autre machine sans devoir entrer son mot de passe.

12. Transfert de fichiers et accès aux fichiers

12.1 FTP, le principal protocole de transfert de fichiers

Le transfert de fichiers entre machines hétérogènes connectées par TCP est moins évident qu'il n'y parait. En effet, les détails des droits d'accès et du système de nommage rendent ce protocole très complexe. De plus, FTP offre de nombreuses options qui dépassent le seul transfert de fichier.

Accès interactif. En plus de l'interface programmable, FTP offre une interface interactive qui permet à l'utilisateur d'interagir facilement avec des serveurs distants.

Spécification de représentation. FTP permet aux clients de préciser le type et la représentation des données transférées (texte, binaire, ASCII ou EBCDIC).

Vérification des utilisateurs. FTP exige des clients qu'ils envoient un nom d'utilisateur et un mot de passe au serveur, avant d'être autorisé à transférer des fichiers.

FTP crée deux connexions entre la machine cliente et la machine serveur, la première pour les données (port 20) la deuxième pour le contrôle (port 21). Les serveurs FTP communiquent sur la liaison de contrôle par l'intermédiaire de NVT (Network Virtual Terminal) déjà défini par Telnet, mais sans négociation d'option.

Afin de simplifier l'accès aux fichiers de nombreux sites adoptent la convention du « FTP anonyme » :

```
login > anonymous
passwd > guest        ou alors aussi        pam@univ-mulhouse.fr
```

12.2 TFTP Trivial File Transfer Protocol

La famille TCP/IP contient un deuxième protocole de transfert de fichier peu coûteux, qui assure les services de base. Ce protocole est destiné aux applications ne nécessitant pas d'interaction complexes entre client et serveur (amorçage des terminaux X par exemple). Il limite le fonctionnement au seul transfert de fichiers et n'effectue pas de contrôle d'accès.

Notons que TFTP ne nécessite pas un transfert fiable, il s'exécute donc au-dessus de UDP. L'émetteur envoie les fichiers sous forme de blocs de taille fixe (512 octets) et attend d'avoir un accusé de réception pour chaque bloc, avant d'envoyer le bloc suivant. Le récepteur acquitte les données dès réception.