2 Le protocole IP

1. Présentation de TCP/IP

TCP/IP est un sigle très connu dans le domaine des réseaux. Au sens strict, TCP/IP est un ensemble de deux protocoles :
- IP (*Internet Protocol*) qui est un protocole de niveau 3 (Réseau)
- TCP (*Transmission Control Protocol*) qui est un protocole de niveau 4 (Transport)

En pratique, les réseaux TCP/IP ne sont pas limités à ces deux protocoles mais recouvrent en fait toute une famille de protocoles (on parle de pile de protocoles).
L'adoption quasi universelle de TCP/IP en fait son principal intérêt. Les technologies TCP/IP couvrent en effet aussi bien les réseaux intranets (LAN) que l'Internet.

TCP/IP est le fruit des recherches qui ont été menées par le DARPA (*Defense Advanced Research Projects Agency*) dès la fin des années 60. En 1969, dans le cadre du projet ARPAnet (*Advanced Research Projects Agency Network*) une première expérimentation permit de relier quatre sites entre eux.

En 1983, les protocoles TCP/IP devinrent des standards militaires. Peu à peu, le réseau ARPAnet fut remplacé par l'Internet. Celui-ci dépassa le domaine exclusif des universités et passa très vite dans le domaine commercial.

TCP/IP fut intégré très tôt en standard sous Unix (dans le noyau BSD, *Berkeley Software Distribution*), ce qui contribua inévitablement à sa popularité.

Depuis 1990, on assiste à une véritable explosion d'IP en Europe (dans le monde non académique). Trois facteurs principaux expliquent la montée en puissance de TCP/IP :

- **l'interopérabilité** : un protocole commun sur des produits provenant de différents constructeurs.
- **l'intérêt commercial** sur l'Internet : l'Internet est basé sur les protocoles et services TCP/IP.
- l'augmentation du nombre **d'outils de gestion de réseau** : le plus important protocole de gestion de réseau est actuellement SNMP (*Simple Network Management Protocol*).

1.1. Architecture de TCP/IP

Comme les protocoles TCP/IP ont été historiquement créés à la demande du ministère de la Défense des Etats-Unis, on les désigne souvent sous le nom de modèle DoD.

La plupart des descriptions de TCP/IP définissent une architecture de protocoles comportant quatre niveaux fonctionnels (du bas vers le haut) :

- Couche Accès réseau : comporte les routines permettant d'accéder aux réseaux physiques

- Couche Internet : définit le **datagramme** et prend en charge le **routage** des données
- Couche Transport Hôte à Hôte (TCP / UDP) : assure les services de **transmission de données de bout en bout**
- Couche Application : comporte les **applications** et processus utilisant le réseau

1.1.1. Couche Accès réseau

La couche la plus basse représente la connexion physique avec les câbles, les circuits d'interfaces électriques (transceivers), les cartes coupleurs, les protocoles d'accès au réseau.
La couche Accès réseau est utilisée par la couche Internet. La couche Accès réseau TCP/IP intègre généralement les fonctions des deux couches inférieures du modèle de référence OSI (Liaison de Données et Physique).

Les utilisateurs ignorent souvent cette couche, la conception de TCP/IP cachant les fonctions des couches inférieures. Etant donné que les protocoles de cette couche font partie intégrante d'Unix, ils apparaissent souvent sous la forme d'une combinaison de pilotes de périphériques (« drivers ») et de programmes associés.

1.1.2. Couche Internet

La couche Internet doit fournir une **adresse logique** pour l'interface physique. L'implémentation du modèle DoD de la couche Internet est IP (*Internet Protocol*). Cette couche fournit un mappage entre l'adresse logique et l'adresse physique fournie par la couche Accès réseau grâce aux protocoles ARP (*Address Resolution Protocol* – RFC 826) et RARP (*Reverse Address Resolution Protocol*).

Les incidents, les diagnostics et les conditions particulières associées au protocole IP relèvent du protocole ICMP (*Internet Control Message Protocol*), qui opère aussi au niveau de la couche Internet.

La couche Internet est aussi responsable du **routage** des paquets de données, les **datagrammes**, entre les hôtes. Cette couche est utilisée par les couches plus hautes du modèle DoD.

1.1.3. Couche Transport hôte à hôte

La couche Transport hôte à hôte ou Transport en abrégé, définit les connexions entre deux hôtes sur le réseau. Le modèle DoD comprend deux protocoles hôte à hôte :

- TCP (*Transmission Control Protocol*) : protocole responsable du service de transmission fiable de données avec détection et correction d'erreurs. TCP permet aussi les connexions simultanées. Plusieurs connexions TCP peuvent être établies sur un hôte, et les données sont envoyées simultanément. TCP permet des connexions full-duplex.

- UDP (*User Datagram Protocol*) : protocole peu fiable, utilisé par des applications qui n'exigent pas la fiabilité de TCP.

1.1.4. Couche Application

La couche Application constitue le sommet de l'architecture TCP/IP. Elle permet aux applications d'utiliser les protocoles de la couche hôte à hôte (TCP et UDP) pour transmettre leurs données. Parmi les protocoles d'applications les plus répandus (orientés utilisateurs), on trouve :

- TELNET, le protocole de terminal de réseau (*Terminal Emulation*), permettant l'ouverture d'une session à distance sur un réseau.
- FTP, le protocole de transfert de fichiers (*File Transfert Protocol*)
- SMTP, le protocole de transfert de courrier électronique (*Simple Mail Transfer Protocol*)

Mais aussi (orientés administrateurs) :

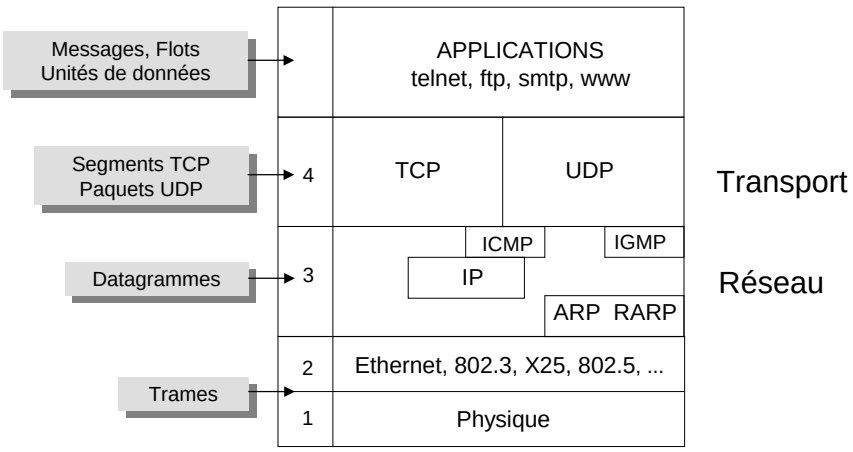
- DNS (*Domain Name Service*) - également appelé *Name service* qui permet d'établir la correspondance entre les adresses IP et les noms attribués aux hôtes du réseau.
- RIP (*Routing Information Protocol*), permettant de gérer la fonction de routage
- NFS (*Network File System*), permettant de partager des fichiers entre différentes machines-hôtes du réseau

1.2. Comparaison des modèles OSI et DoD

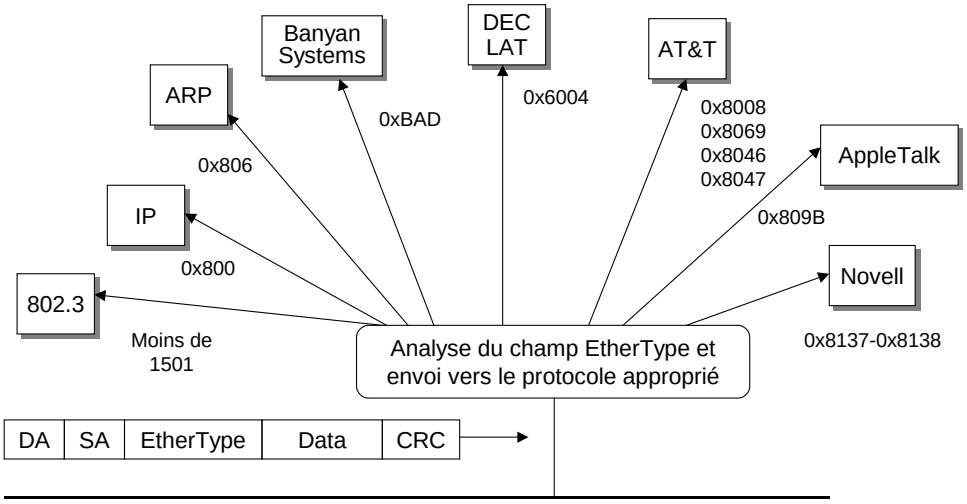
Comme dans le modèle OSI, les données sont transmises de haut en bas dans la pile lors de leur envoi sur le réseau, et de bas en haut dans la pile lors de leur réception à partir du réseau. Chaque couche de la pile ajoute des informations de contrôle, un en-tête, de manière à garantir une transmission des données correcte.

Chaque couche possède ses propres structures de données indépendantes. La terminologie utilisée pour décrire les données au niveau de chaque couche diffère dans les deux modèles :

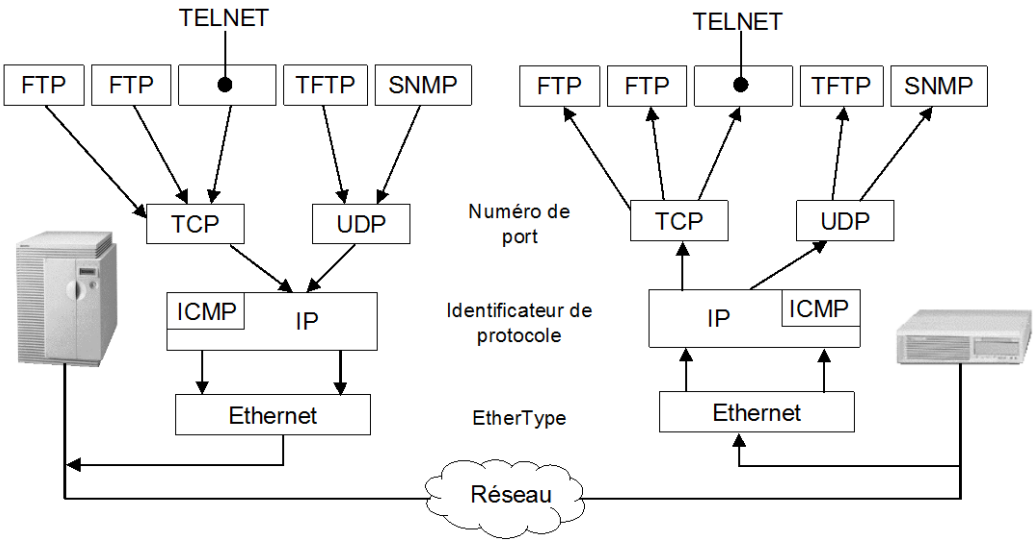
- Dans le modèle OSI, l'expression PDU (*Protocol Data Unit*) est employée pour décrire les données d'une couche.
- Dans le modèle DoD, les termes *message* et *flot* (*stream*) sont utilisés au niveau de la couche application ; les termes *segment* et *paquet*, au niveau de la couche hôte à hôte ; le terme *datagramme*, au niveau de la couche Internet ; et le terme *trame*, au niveau de la couche accès réseau.



Le multiplexage/démultiplexage permet à plusieurs protocoles des couches hautes d'utiliser un protocole commun d'une couche basse.
Un réseau Ethernet supportant IP (et ICMP), peut potentiellement supporter d'autres protocoles comme IPX. C'est le champ EtherType qui permet de savoir à quel protocole supérieur (réseau) la trame Ethernet est destinée. Il autorise, au niveau de la couche liaison de données, le multiplexage de plusieurs protocoles réseau au niveau de la source et le démultiplexage au niveau de la destination.



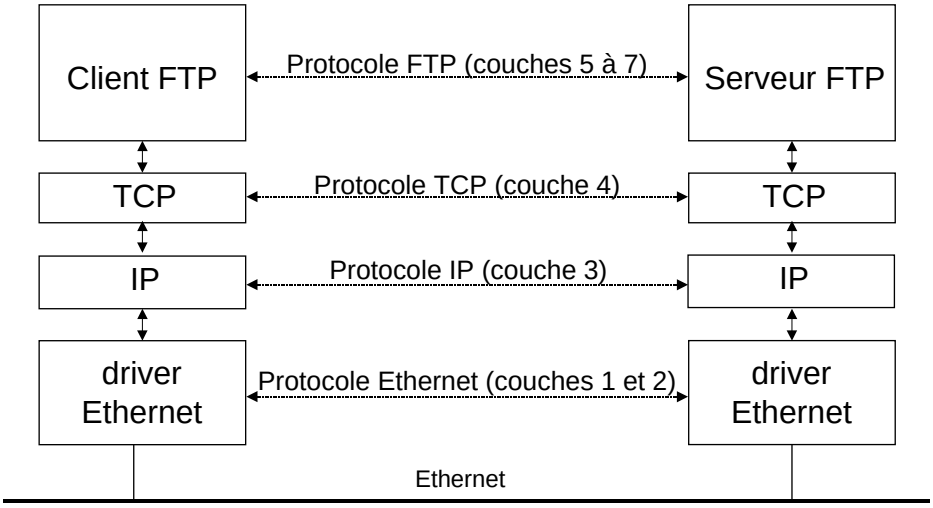
Quand la couche IP reçoit un paquet Ethernet, elle doit distinguer les paquets pour TCP ou UDP grâce à un champ d'identification de protocole de 8 bits situé dans le paquet.



2. Les fonctions d'IP

Le protocole Internet, RFC 791, fournit le service de transmission de paquets de base sur lequel les réseaux TCP/IP sont construits. Tous les protocoles, figurant dans les couches au-dessus et en-dessous d'IP, utilisent le protocole Internet pour transmettre des données. Toutes les données de TCP/IP traversent IP, entrant et sortant, indépendamment de leur destination finale.

La couche IP s'appuie sur le matériel réseau sous-jacent pour sa transmission. IP peut fonctionner au-dessus de plusieurs technologies différentes : Ethernet, Token-Ring, FDDI, ATM, …



Ses fonctions incluent :

- La définition du **datagramme**
- La définition du plan d'**adressage** Internet
- La circulation de données entre la couche Accès réseau et la couche Transport machine-hôte à machine-hôte
- L'**acheminement** (routage) des datagrammes vers les ordinateurs à distance
- La **fragmentation** et le **réassemblage** des datagrammes.

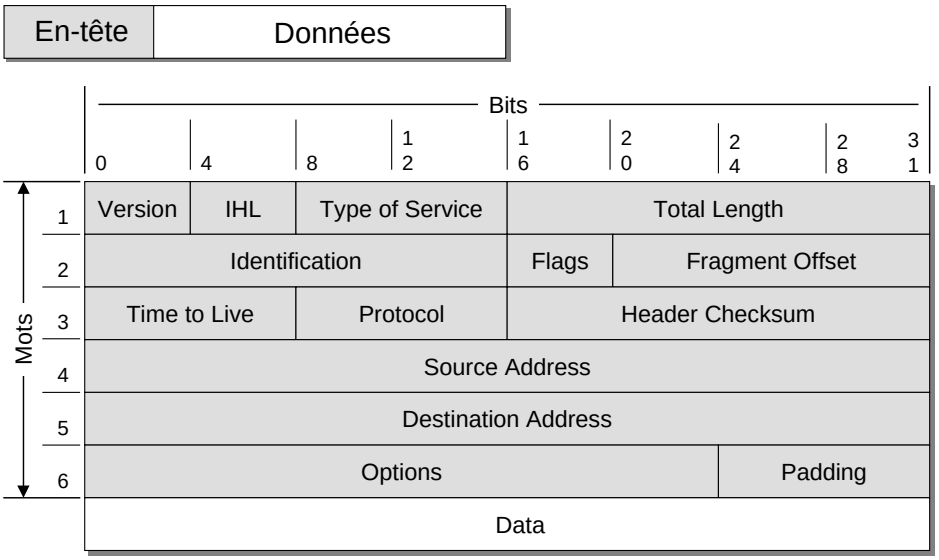
IP est avant tout un **protocole orienté non connexion**, et chaque datagramme est routé indépendamment des autres. Le protocole Internet délègue aux protocoles des autres couches le soin d'établir la connexion s'ils requièrent un service orienté connexion. Le réseau IP emploie pour la livraison la méthode de *Best Effort*, c'est-à-dire la meilleure livraison possible du datagramme, sans que cette livraison soit garantie pour autant.

IP dépend également des protocoles des autres couches pour garantir la détection éventuelle d'erreurs et leur correction. Le protocole contient uniquement un total de contrôle d'en-tête (« header checksum ») et est dépourvu de tout code de vérification des données (code de détection d'erreurs et de correction). Les protocoles des autres couches de l'architecture TCP/IP permettent d'effectuer cette vérification lorsqu'elle s'avère nécessaire.

Les protocoles de couches supérieures peuvent espérer une certaine qualité de service (QoS). La couche supérieure passe à la couche IP les paramètres QoS en même temps que les données ; la couche IP peut alors tenter de faire correspondre ces paramètres et les services fournis par le matériel réseau sous-jacent, capable ou non de fournir les services requis.

2.1. Format du datagramme

Le datagramme IP contient un en-tête IP suivi des données IP provenant des protocoles des couches supérieures.



- Par défaut, la longueur de l'en-tête est de 5 mots de 32 bits (soit 20 octets) ; le sixième mot est facultatif. Puisque la longueur de l'en-tête est variable, elle inclut un champ appelé *Internet Header Lenght* (IHL - longueur de l'en-tête Internet) en mots. L'en-tête contient toutes les informations nécessaires à la transmission du paquet.

- Le champ *Version* fait quatre bits de long et indique le format de l'en-tête IP : le numéro de version actuel est 4 (IPv4) ; la version suivante est la version 6 (IPv6) et permet des adresses IP à 128 bits. Ce champ est utilisé par l'émetteur, le récepteur et tout routeur intermédiaire pour déterminer le format de l'en-tête IP.

- Le champ *Type of Service* (TOS) informe les réseaux de la qualité de service désirée, spécifiant ainsi la préséance, les délais, le débit et la fiabilité. La plupart des implémentations de TCP/IP et des protocoles de routage ignorent ce champ ; il est pourtant probable qu'il sera amené à jouer un rôle plus important dans l'avenir.

- Le champ *Total length* (longueur totale) contient la longueur de l'en-tête et des données IP, en octets. L'Internet ne limite pas les datagrammes à une taille précise mais suggère que les réseaux et les passerelles puissent supporter ceux de 576 octets (512 octets de données + 64 octets d'en-tête et de données propres au protocole) sans les fragmenter. La taille des datagrammes IP sur la plupart des réseaux et des hôtes dépasse rarement 16 Ko.

- La *durée de vie* (*Time To Live*) se mesure en secondes et représente la durée maximale de vie d'un datagramme sur le réseau. Cette valeur est décrémentée à chaque routeur. Lorsque le champ TTL tombe à 0, le temporisateur TTL expire et le datagramme IP est

écarté par le routeur (et pas par l'hôte de destination). Le champ TTL a une double fonction :
- limiter la durée de vie des segments TCP
- éliminer les boucles de routage Internet

Lorsque le temporisateur TTL expire, un message ICMP en avertit la source. Une valeur par défaut initiale de 32 ou 64 est courante.

- IP transmet le datagramme en utilisant *l'adresse de destination* contenue dans le cinquième mot de l'en-tête. L'adresse de destination est une adresse IP standard de 32 bits permettant d'identifier le **réseau de destination** et la **machine-hôte** connectée à ce réseau.
  - si l'adresse de destination correspond à l'adresse d'une machine-hôte connectée au réseau local, le paquet est transmis directement vers la destination.
  - sinon, le paquet est envoyé vers une **passerelle** afin d'être transmis à la machine-hôte.

  L'acheminement ou **routage** correspond à la sélection de la passerelle à utiliser pour la transmission des données. IP détermine le routage approprié pour chaque paquet.

2.2. Acheminement des datagrammes

Les passerelles Internet sont généralement (et plus exactement) référencées comme étant des **routeurs IP** puisqu'elles utilisent le protocole Internet pour acheminer les paquets entre les réseaux. Dans le jargon traditionnel de TCP/IP, il n'existe que deux types de machines-périphériques de réseau :
- les passerelles
- les machines-hôtes.

Les passerelles transmettent les paquets entre réseaux, tandis que les machines-hôtes non. Toutefois, si une machine-hôte est connectée à plusieurs réseaux (appelée une *machine-hôte multiconnectée*), elle peut transmettre des paquets entre les réseaux.
Lorsque une machine-hôte multiconnectée transmet des paquets, son fonctionnement est alors identique à celui d'une passerelle et est utilisée comme une passerelle.

2.3. Fragmentation des datagrammes

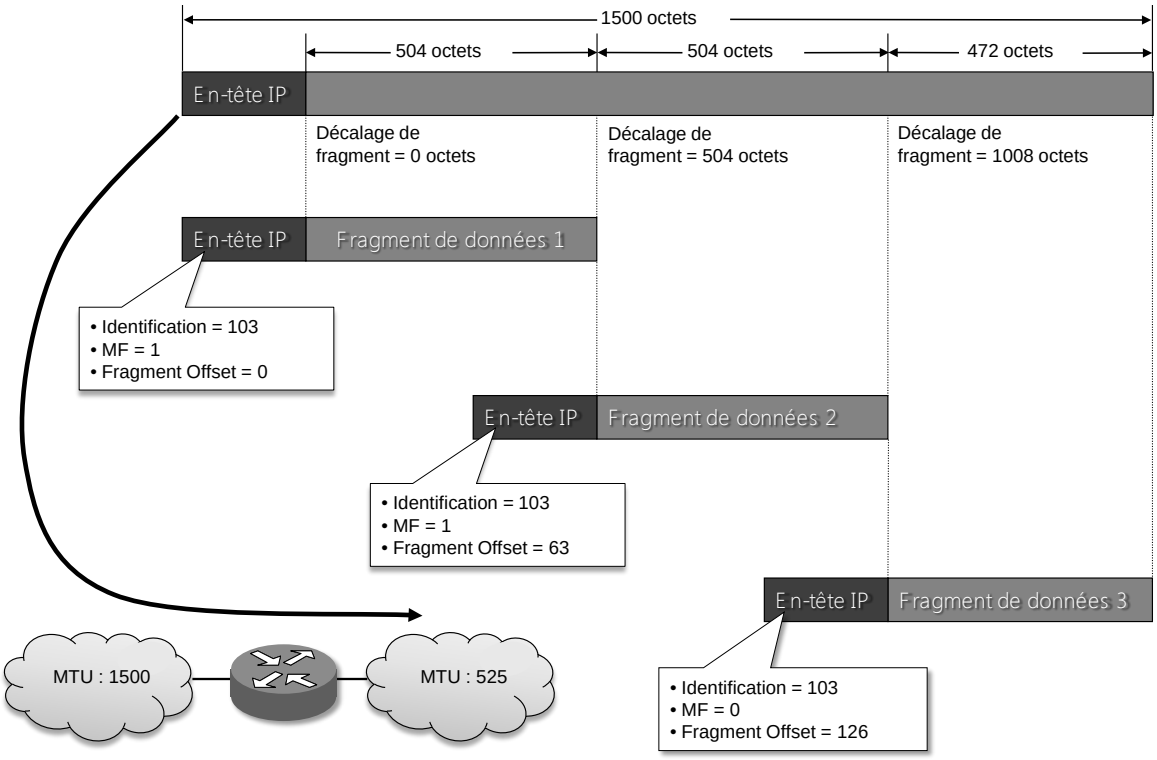
Lorsqu'une passerelle interconnecte des réseaux physiques différents, il est parfois nécessaire de diviser le datagramme en éléments de plus petite taille pour passer d'un réseau à l'autre.

Chaque type de réseau se caractérise par une *unité de transfert maximale* (MTU *: Maximum Transmission Unit*), correspondant au plus grand paquet que celui-ci puisse transférer. Si la longueur du datagramme provenant d'un réseau est supérieure à la MTU de l'autre réseau, il est alors nécessaire de diviser le datagramme en *fragments* de plus petite taille afin de permettre la transmission des données. Cette procédure s'appelle la *fragmentation*.

Type de réseau	MTU (en octets)
Ethernet	1 500
IEEE 802.3	1 492
Token-Ring	4 440 à 17 940
FDDI	4 352
IEEE 802.4	8 166

Le format de chaque fragment est identique à celui de tout datagramme normal. Le deuxième mot de l'en-tête contient les informations permettant de reconnaître chaque fragment de datagramme et fournit les informations relatives à la procédure de réassemblage des différents fragments en un datagramme original.

Le champ *Identification* indique à quel datagramme le fragment appartient, et le champ *Fragment Offset* (décalage de la fragmentation, en valeur multiple de 8 octets) précise à quelle partie du datagramme correspond ce fragment. Le champ *Flags* (Drapeaux) possède un élément binaire « *More Fragments bit* » qui indique à IP s'il a assemblé tous les éléments du datagramme (MF=0).



Le réassemblage est effectué par le **module IP de destination**, jamais par les routeurs intermédiaires.

2.4. Transmission de datagrammes à la couche Transport

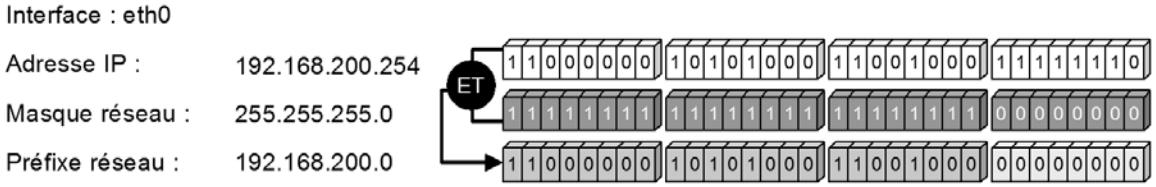
Le champ *Protocol* indique quel protocole de couche supérieure recevra les données IP. Ce champ est utilisé pour le multiplexage/démultiplexage des données vers des protocoles de couche supérieure. Chaque protocole possède un numéro de protocole unique qui permet à IP de le reconnaître.
Par exemple, la valeur du champ Protocol est 6 pour TCP, 17 pour UDP et 1 pour ICMP. Sous Unix, ces valeurs sont stockées dans un fichier spécial `/etc/protocols`.

3. Adressage

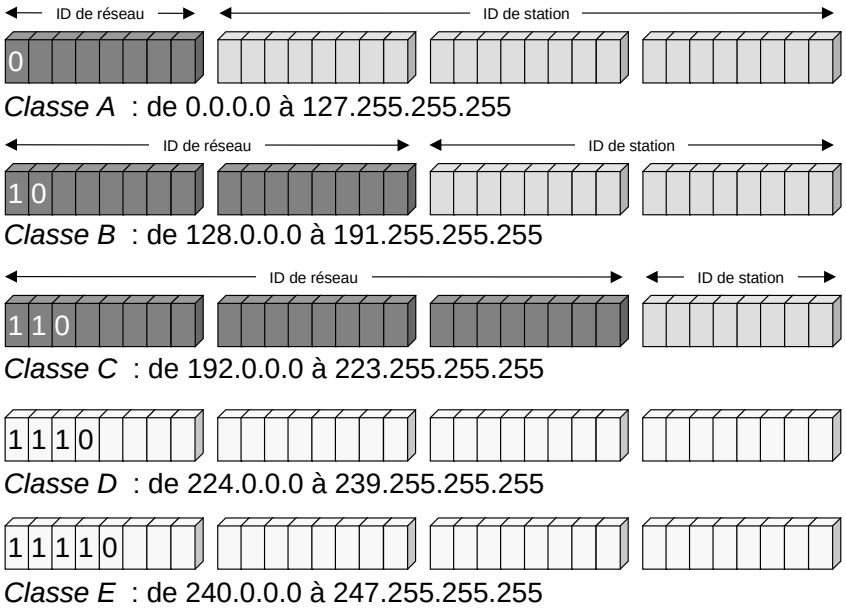
3.1. L'adresse IP

TCP/IP offre une vision **logique** du réseau, rendant ainsi ce dernier indépendant de la technologie matérielle sous-jacente (Ethernet, Token-Ring, FDDI …). Les nœuds du réseau TCP/IP sont identifiés et accédés grâce à une **adresse logique**, l'adresse IP. Comme les adresses IP ne dépendent pas des adresses physiques, on peut modifier le matériel sous-jacent sans modifier l'adresse logique.

L'adresse IP (ou IPv4) est définie sur 32 bits ; elle comporte deux parties : la **partie réseau** (*netid*, ou ID de réseau) et la **partie hôte** (*hostid* ou ID d'hôte). Un **masque** (*netmask*) est associé à cette adresse. Il permet au logiciel IP de déterminer le préfixe de réseau d'une adresse en calculant un ET logique avec le masque.



Historiquement le réseau Internet était découpé en **classes d'adresses**. Le nombre de bits d'adresse utilisé pour identifier le réseau et le nombre utilisé pour reconnaître l'hôte variaient en fonction de la classe.
Les trois principales classes d'adresses étaient : la classe A, la classe B et la classe C. La classe D était réservée à la multidiffusion (*multicasting*), la classe E était réservée à un usage ultérieur. Les classes devaient permettre de répondre aux besoins d'entreprises de différentes tailles.
Un niveau de division complémentaire peut être inséré avec le **sous-adressage** (*subnetting*), permettant de segmenter une plage d'adresse en sous-réseaux IP.



- Si le premier bit d'une adresse est positionné sur 0, il s'agit alors d'une adresse d'un réseau de classe A. Le premier bit d'une adresse de classe A permet d'identifier la classe de l'adresse. Les 7 bits suivants permettent de déterminer le réseau et les 24 derniers bits de reconnaître l'hôte. Il existe moins de 128 numéros de réseau de classe A ; chaque réseau de classe A peut être constitué de millions d'hôtes.

- Si les 2 premiers bits de l'adresse sont 1 0, il s'agit alors d'une adresse de réseau de classe B. Les 2 premiers bits permettent de déterminer la classe, les 14 bits suivant de déterminer le réseau. Les 16 derniers bits permettent de reconnaître l'hôte. Il existe des milliers de numéros de réseau de classe B et chaque réseau de classe B peut être constitué de milliers d'hôte.

- Si les 3 premiers bits de l'adresse sont 1 1 0, il s'agit alors d'une adresse de réseau de classe C. Les 3 premiers bits identifient la classe ; les 21 suivants correspondent à l'adresse du réseau, et les 8 derniers bits permettent d'identifier l'hôte. Il existe des millions de numéros de réseau de classe C. Chaque réseau de classe C comporte moins de 254 hôtes.

- Si les 3 premiers bits sont 1 1 1, il s'agit d'une adresse spéciale réservée. Ces adresses ne correspondent pas à des réseaux spécifiques.

La notion de classe d'adresses a été rendue obsolète pour l'adressage des nœuds du réseau Internet car elle induisait une restriction notable des adresses IP affectables par l'utilisation de masques spécifiques. Les documents RFC 1518 et RFC 1519 publiés en 1993 spécifient une nouvelle norme : l'adressage CIDR (*Classless Internet Domain Routing* ou « routage de domaine Internet sans classe » cf. ci-après). Ce nouvel adressage précise qu'il est possible d'utiliser un masque quelconque appliqué à une adresse quelconque. Il organise par ailleurs le regroupement géographique des adresses IP pour diminuer la taille des tables de routage des principaux routeurs du réseau Internet.

Les adresses IP sont représentées sous la forme de quatre nombres correspondant à la valeur décimale des quatre octets qui composent l'adresse, séparés par des points (**notation décimale pointée**).
Le masque peut être spécifié soit en notation décimale pointée, soit sous forme condensé c'est-à-dire en indiquant simplement le nombre de bits à 1 qu'il contient :
Exemple :
192.168.200.254/255.255.255.0 (notation « classique »)
ou 192.168.200.254/24 (notation CIDR).

3.2. Adresses IP spéciales

- Une adresse IP dont l'hostid vaut 0 indique le réseau lui-même.

- Une adresse de **diffusion dirigée** (*directed broadcast*) est une adresse dans laquelle, tous les bits de l'hostid sont à 1 (soit 255 en décimal) : XXX.XXX.XXX.255 pour un réseau de classe C, XXX.XXX.255.255 pour un réseau de classe B. Une adresse de diffusion dirigée peut apparaître dans l'adresse IP de destination d'un datagramme IP, mais jamais comme source.

- Une adresse de **diffusion limitée** (*limited broadcast*) représentée par l'adresse 255.255.255.255, permet d'atteindre tous les nœuds du réseau. La diffusion limitée peut être utilisée dans les réseaux locaux, pour lesquels une diffusion ne franchit jamais la frontière du routeur. Elle ne peut jamais apparaître comme adresse IP source, mais seulement comme adresse de destination.

- Une adresse IP ne comportant que des 0 (0.0.0.0), est généralement utilisée lorsqu'un nœud IP essaye de déterminer sa propre adresse IP, comme le fait le protocole BOOTP qui permet aux nœuds d'un réseau de se voir affecter une adresse IP.
  On utilise aussi l'adresse 0.0.0.0 dans les tables de routages pour indiquer l'entrée de réseau de l'adresse IP du routeur par défaut (souvent appelé « passerelle par défaut »).
  L'adresse IP 0.0.0.0 ne peut être utilisée que comme adresse IP source, jamais comme adresse IP de destination.

- Adresse de bouclage : tout paquet envoyé par une application TCP/IP vers une adresse de type 127.X.X.X, où X est un nombre entre 0 et 255, a pour conséquence le renvoi de ce paquet à l'application sans que le paquet n'atteigne le support du réseau. Le paquet est copié du buffer de transmission au buffer de réception sur la machine elle-même.
  L'adresse de bouclage logiciel permet de vérifier rapidement que TCP/IP est bien configuré (on peut faire un « ping » sur sa propre machine). Elle est également utile lorsqu'un client et un serveur TCP/IP s'exécutent sur une même machine.
  Bien que toute adresse de format 127.X.X.X indique une adresse de bouclage, les adresses de bouclage usuelles sont 127.0.0.1 et 127.1 dans la plupart des implémentations de TCP/IP.

3.3. Unicast, broadcast, multicast

Lorsqu'un datagramme IP est envoyé à une adresse IP individuelle, on dit qu'il s'agit d'un datagramme IP *unicast*. L'envoi de ce datagramme est appelé *unicasting* ; on utilise l'unicasting lorsque deux nœuds IP communiquent ensemble.
Lorsqu'un datagramme IP est envoyé à tous les nœuds d'un réseau spécifique, il s'agit d'une diffusion (*broadcasting*).

Lorsqu'un datagramme IP est envoyé à un groupe de destinataires, on parle de *multicasting* (multidiffusion). Dans ce cas, on utilise une adresse de classe D comme adresse de destination. Les systèmes qui appartiennent au même groupe de multidiffusion (qui ont donc une même adresse de classe D), doivent aussi se voir affecter une adresse IP (de classe A, B ou C).
Le groupe de multidiffusion peut recevoir un datagramme IP de deux manières :
- des datagrammes IP envoyés directement aux adresses IP individuelles (classes A, B ou C)
- des datagrammes IP envoyés à l'adresse de multidiffusion (classe D)

Pour que la multidiffusion fonctionne bien, un hôte doit avoir la capacité de se joindre à un groupe de multidiffusion et de le quitter. Le logiciel de couche IP doit être en mesure de reconnaître les adresses de multidiffusion des datagrammes IP entrants ou sortants (ce qui n'est pas le cas des vieilles implémentations d'IP).
N'importe quel hôte sur un inter-réseau IP peut se joindre à un groupe de multidiffusion. Il n'est pas nécessaire que les hôtes du groupe appartiennent à un même réseau local ; ils peuvent être séparés par des routeurs.

3.4. Affectation des adresses IP

Si le réseau local doit être connecté à d'autres réseaux tels que l'Internet, il faut obtenir un netid distinct qui n'est utilisé par personne d'autre. Une fois ce numéro obtenu, il incombe à l'administrateur réseau d'affecter les numéros d'hôte à partir de son numéro de réseau. C'est l'*Internet Address Network Authority* (IANA) qui définit les procédures et constitue l'autorité suprême sur les numéros affectés. L'IANA a délégué la zone européenne à un organisme : le RIPE NCC (*Réseaux IP Européens - Network Coordination Centre*). Cet organisme distribue les adresses IP aux fournisseurs d'accès à Internet.

Pour réduire le besoin en nouvelles adresses IP, la RFC 1918 concerne l'allocation d'adresses pour les réseaux privés. Ce sont des réseaux qui ne sont pas connectés à d'autres réseaux, ou dont les hôtes et les services ont une interaction limitée avec l'Internet.

- classe A : 10.0.0.0 à 10.255.255.255
- classe B : 172.16.0.0 à 172.31.255.255
- classe C : 192.168.0.0 à 192.168.255.255

Pour des raisons de sécurité, de nombreuses entreprises utilisent des « passerelles » logicielles (tels des *firewalls* ou garde-barrière) pour connecter leur réseau local à l'Internet. Le réseau interne n'a généralement pas un accès direct à l'Internet et seuls un ou plusieurs hôtes sont visibles depuis l'Internet.

Il est également possible d'utiliser des mécanismes de translation d'adresses (NAT, *Network Address Translation*) qui peuvent faire correspondre un ensemble d'adresses IP privées et un ensemble d'adresses IP distinctes (ce qui permet aussi de faire des économies au niveau des classes de réseaux).

Si on veut utiliser l'espace d'adressage privé, il s'agit de bien déterminer quels sont les hôtes qui n'ont pas besoin de bénéficier d'une connectivité de couche de Réseau, et ceux qui ont besoin d'adresses globalement distinctes.

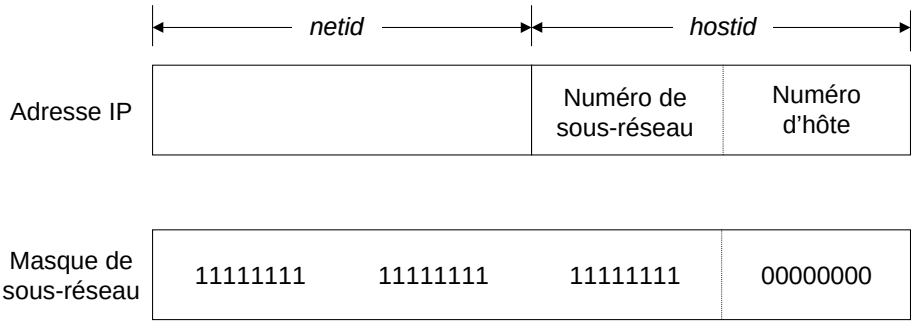
3.5. Adressage de sous-réseau

Avec IP, il est possible de diviser une adresse de réseau en plusieurs adresses de sous-réseaux (cf RFC 950). La création de sous-réseaux permet de résoudre des problèmes organisationnels ou de topologie.

L'application d'un masque de bit, appelé *masque du sous-réseau* (*subnet mask*), à l'adresse IP permet de définir un sous-réseau. Le masque de sous-réseau divise le champ hostid en un numéro de sous-réseau et un numéro d'hôte. Le masque est un nombre à 32 bits dont les valeurs sont définies selon les règles suivantes :

- si un bit déterminé est activé dans le masque (est à 1), son pendant dans l'adresse est interprété comme étant un bit de réseau.

- si un bit du masque est désactivé (est à 0), ce bit appartient alors à la partie hôte de l'adresse.



Le sous-réseau n'est reconnu que **localement**. Dans le reste d'Internet, l'adresse est toujours interprétée comme étant une adresse IP standard.

D'après la RFC 950, les sous-réseaux dont les bits sont tous à 0 ou tous à 1 ne devraient pas être utilisés pour éviter les erreurs d'interprétation sur les adresses réservées.

3.6. Sur-réseaux

Le concept de sous-réseau a été conçu en 1985 pour optimiser l'utilisation de l'espace d'adressage IP. Les adresses de classes A et B venaient en effet à manquer, mais les adresses de classe C disponibles étaient encore en nombre appréciable.

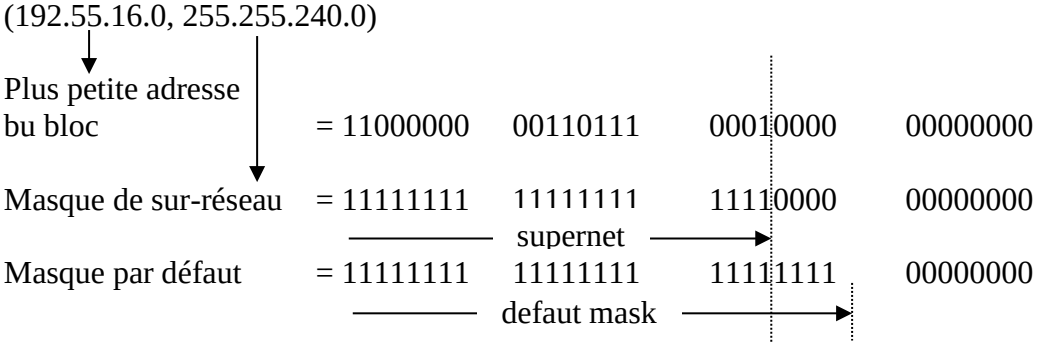
La mise en sur-réseau, ou adressage en sur-réseau, consiste à affecter un bloc d'adresse de classe C plutôt qu'une adresse de classe B unique, afin de créer une classe d'adresses virtuelle située à mi-chemin entre un réseau de classe C et un réseau de classe B.
L'adressage en sur-réseau est conçu notamment pour les fournisseurs d'accès Internet (FAI).

Comme un bloc de 256 adresses de classe C est nécessaire pour prendre en charge l'équivalent d'une seule adresse de classe B, une table de routage doit contenir 256 adresses de réseau pour chacun des réseaux de classe C. Ce qui pose des problèmes de taille.

La technique du CIDR permet de résumer un bloc d'adresses de classe C en une seule entrée de table de routage. Cette entrée de table est constituée ainsi :

(plus basse adresse du bloc, masque de sur-réseau)

La plus basse adresse du bloc est l'adresse de départ de ce bloc, et le masque de sur-réseau spécifie le nombre d'adresses de classe C dans le bloc.



L'intervalle d'adresses de classe C est délimité par les adresses suivantes :
11000000 00110111 00010000 00000000 = 192.55.16.0
11000000 00110111 00011111 11111111 = 192.55.31.255

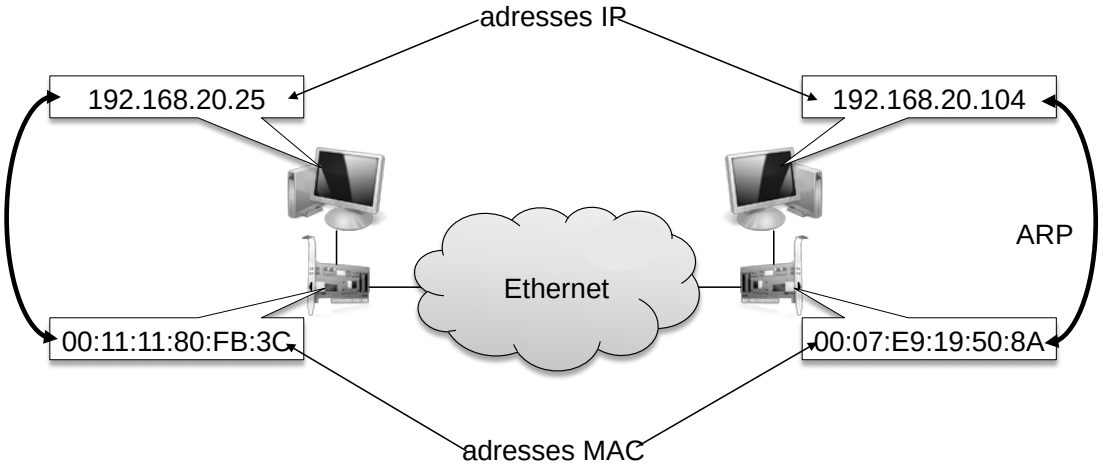
Combien d'adresses de classe C contient ce bloc ? Il suffit de regarder le nombre de bits pouvant varier dans la partie réseau de l'adresse de classe C (le masque par défaut indique la partie netid et hostid). Ce sont les 4 bits du masque de sur-réseau :
11111111 11111111 1111**0000** 00000000
soit $2^4 = 16$ adresses de classes C.
Les routes d'un bloc CIDR peuvent être résumées en une unique entrée de table de routage appelée un **agrégat**.

Il existe une autre notation utilisable pour les blocs CIDR :

plus basse adresse du bloc/nombre de bits de préfixe commun
(192.55.16.0, 255.255.240.0)
192.55.16.0/20
Le masque de sous-réseau par défaut d'une adresse de classe C est 255.255.255.0, soit pour les hôtes d'une seule adresse de classe C, un préfixe commun de 24 bits. Lorsque ce nombre de bits est inférieur à 24, c'est l'adressage en sur-réseau qui est utilisé (adressage en sous-réseau lorsqu'il est supérieur à 24).

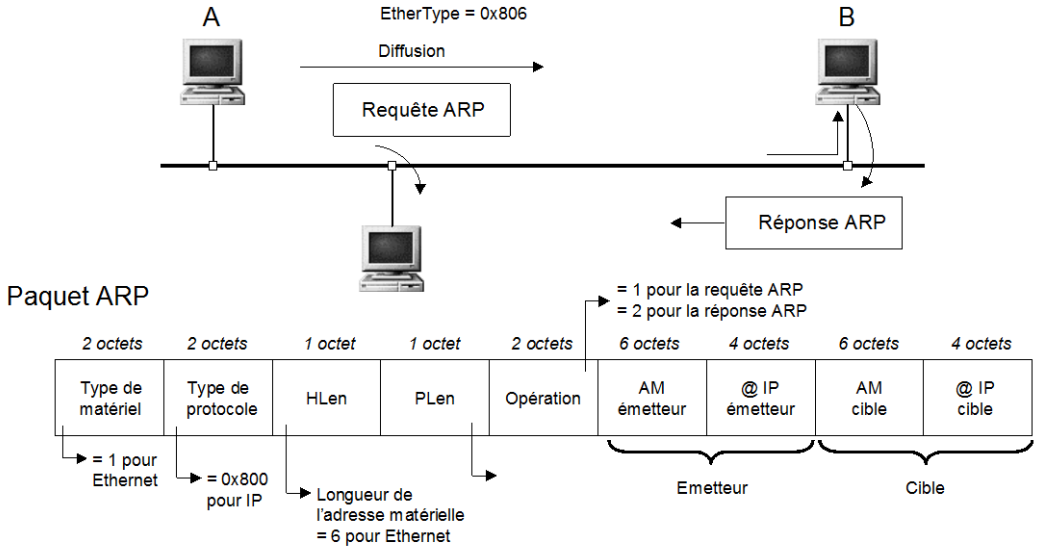
3.7. Protocoles de résolution d'adresse

En TCP/IP les interfaces du réseau sont modélisées par un unique identificateur à 32 bits, l'adresse IP. Or la transmission des datagrammes IP sur le réseau physique nécessite que ces datagrammes soient encapsulés dans des trames de couche de liaison de données (couche 2), telles que Ethernet ou Token-Ring, qui elles contiennent des adresses « physiques » (adresses MAC).



Un mécanisme souple, implémenté sous forme d'un protocole distinct et appelé ARP (*Address Resolution Protocol*) permet de déterminer dynamiquement l'adresse MAC à partir de l'adresse IP d'un hôte.

Pour déterminer l'adresse matérielle de l'hôte B avant de lui envoyer son message, la station A envoie sur le réseau une trame MAC de diffusion, appelée trame ARP de requête. Celle-ci contient les adresses IP et MAC de l'hôte A émetteur, ainsi que l'adresse IP de la destination B. La trame inclut un champ destiné à contenir l'adresse MAC de B. Tous les nœuds du réseau physique reçoivent la trame ARP. Seul l'hôte dont l'adresse IP correspond à l'adresse requise dans la trame de requête ARP répond en encodant sa propre adresse matérielle dans une trame de réponse ARP. L'hôte A initialise alors sa table cache ARP (conservée en mémoire) en utilisant la réponse fournie. Les entrées dans cette table expirent après une temporisation donnée qui peut être configurée dans certaines implémentations de TCP/IP (généralement 15 mn). Le cache ARP est consulté par un hôte juste avant l'envoi d'une requête ARP ; si la réponse se trouve dans le cache, la requête n'est pas effectuée.



La requête ARP est envoyée avec une adresse matérielle de diffusion. Le protocole ARP part du principe que le réseau physique sous-jacent prend en charge la diffusion.
Le paquet ARP est encapsulé dans la trame de couche de liaison de données du réseau. La valeur 0x806 d'EtherType est réservée pour les trames ARP.

RARP (*Reverse ARP*) est un mécanisme utilisé par les stations sans disques (terminaux X) pour obtenir leur adresse IP auprès d'un serveur distant. Le nœud qui veut connaître sa propre adresse envoie en diffusion une requête RARP. En Ethernet, la valeur 0x8035 d'EtherType est réservée pour les trames RARP. S'il existe plusieurs serveurs RARP, chacun d'eux tente de traiter la requête RARP. Généralement, le client RARP accepte la première réponse reçue et ignore silencieusement les suivantes.

Le serveur RARP tient à jour une table des adresses IP des nœuds du segment ; cette table est indexée par un identificateur distinct pour chaque machine, son adresse matérielle.

Dans la plupart des implémentations, RARP n'est pas automatiquement fourni par le module ARP ou IP, mais est exécuté sous forme de processus séparé (un démon) sur la machine qui tiendra lieu de serveur RARP.

4. Le protocole ICMP : Internet Control Message Protocol

Une partie de IP correspond au protocole ICMP (*Internet Control Message Protocol*) défini dans le RFC 792. Ce protocole est une partie intégrante de la couche Internet.

Toute implémentation de IP peut générer un message ICMP, c'est en particulier le cas des hôtes et des routeurs. Les messages ICMP sont encapsulés dans IP (le message ICMP est inclus dans la partie données du datagramme IP ; le champ *Protocol* vaut alors 1). Il n'y a jamais de réponse à un message ICMP (sauf *echo request*) pour ne pas engendrer d'autres messages en cascade.

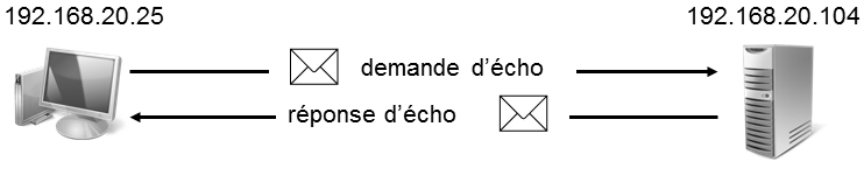
ICMP envoie des messages qui réalisent les fonctions suivantes de contrôle, de détection des erreurs, et de transmission d'informations pour TCP/IP. Il existe 18 types de messages ICMP.

- *Contôle de flux* : lorsque les datagrammes arrivent trop rapidement afin d'être traités, la machine-hôte de destination ou une passerelle intermédiaire renvoie un message de congestion de la source ICMP à l'émetteur. Ce message indique à la source de suspendre temporairement l'envoi de datagrammes.

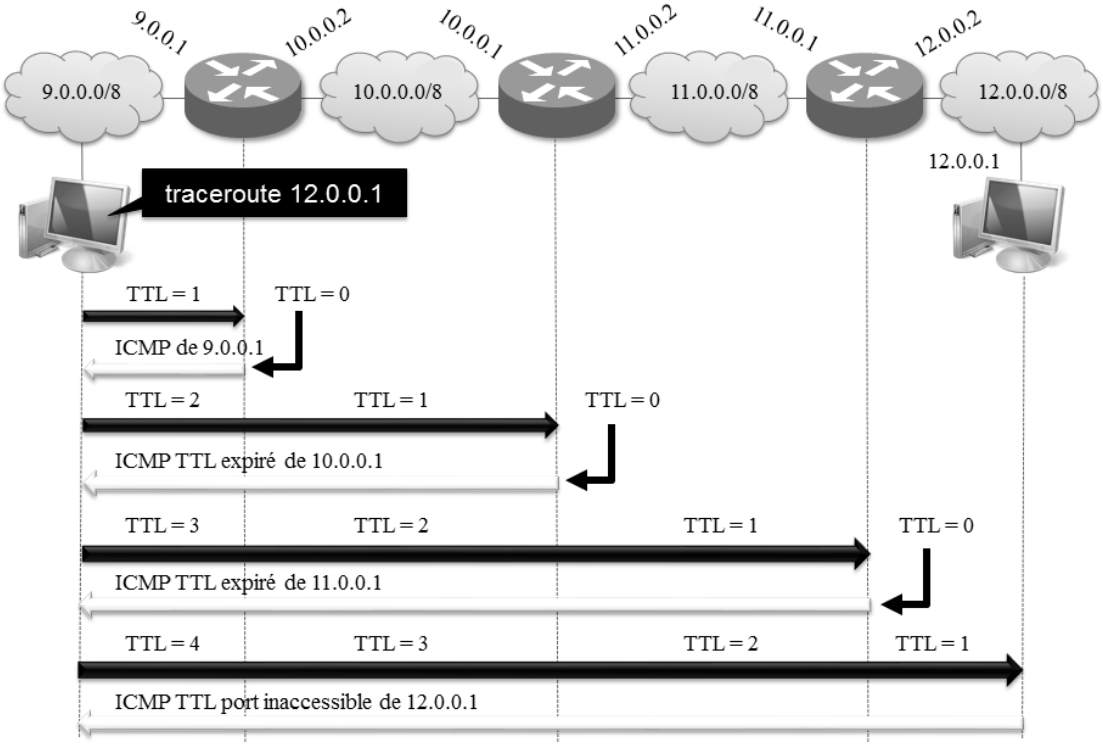
- *Détection de destinations inaccessibles* : lorsqu'une destination s'avère inaccessible, le système qui détecte le problème envoie un message Destination inaccessible à la source du datagramme. Si la destination inaccessible est un réseau ou une machine-hôte, le message est alors envoyé via une passerelle intermédiaire. En revanche, si la destination est un port inaccessible, la machine-hôte de destination envoie le message.

- *Redirection des voies* : une passerelle envoie le message de redirection ICMP afin d'indiquer à une machine-hôte d'utiliser une autre passerelle, probablement parce que l'autre passerelle constitue un meilleur choix. Ce message peut être uniquement utilisé si la machine-hôte source est connectée au même réseau que les deux passerelles.

- *Vérification des machines-hôtes à distance* : une machine-hôte peut envoyer le message d'écho ICMP pour constater si le protocole Internet du système à distance est opérationnel. Lorsqu'un système reçoit un message de demande d'écho (*echo request*, type 8), il renvoie un message de réponse d'écho (*echo reply*, type 0) à la machine-hôte source. La commande `ping` d'Unix utilise ce message ; elle mesure aussi le temps moyen d'accès (latence) à l'hôte distant.



  La commande `traceroute` utilise également des messages ICMP ; elle permet de connaître la route exacte empruntée par les datagrammes. `traceroute` envoie 3 paquets UDP avec un TTL égal à 1 puis recommence en augmentant le TTL de 1 à chaque envoi. A chaque fois que le TTL arrive à 0, le routeur renvoie un message ICMP d'erreur.



5. IPv6

IPv6 est le successeur de la version actuelle d'IP (IPv4). Il maintient les meilleures fonctions d'IPv4, en écarte ou minimise les mauvaises, et en ajoute de nouvelles quand elles sont nécessaires. Les principales améliorations de cette version sont :

- des adresses codées sur 16 octets (128 bits) et notées sous la forme de 8 groupes de 4 chiffres hexadécimaux séparés avec le symbole deux-points, offrant un espace d'adressage quasi illimité (environ $3{,}4\times10^{38}$ adresses)
  exemple : `8000:0000:0000:0000:0123:4567:89AB:CDEF`

- la simplification de l'en-tête des datagrammes avec 7 champs (contre 14 pour IPv4) permettant aux routeurs de traiter les datagrammes plus rapidement et améliorant globalement leur débit.

- plus de souplesse accordée aux options : les champs obligatoires de l'ancienne version sont maintenant devenus optionnels. De plus, la façon dont les options sont représentées est différente ; elle permet aux routeurs d'ignorer plus simplement les options qui ne leur sont pas destinées. Cette fonction accélère le temps de traitement des datagrammes.

- une plus grande sécurité (authentification et confidentialité)

- une plus grande attention que par le passé accordée aux types de services.