Le Protocole IP

EFORT
http://www.efort.com

TCP/IP désigne communément une architecture réseau, mais cet acronyme désigne en fait deux protocoles étroitement liés: un protocole de transport, TCP (Transmission Control Protocol) qu'on utilise "par-dessus" un protocole réseau, IP (Internet Protocol). Ce que l'on entend par "modèle TCP/IP" 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'implantation la plus courante. Par abus de langage, TCP/IP peut donc désigner deux choses : d'un côte un modèle architecture TCP/IP, et de l'autre, deux protocoles TCP et IP.
Le modèle TCP/IP, comme nous le verrons plus loin, 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 implantation ; la normalisation a suivi. Cet historique fait toute la particularité de ce modèle, ses avantages et ses inconvénients.
Le but de ce tutoriel est de présenter le protocole IP, l'adressage IP public et privé, le routage IP, l'obtention d'adresse IP et la translation entre adresse privée et adresse publique.

1 La place du protocole IP

IP correspond à la couche 3 du modèle OSI (Figure 1). IP peut fonctionner au-dessus de toute liaison de données: Ethernet, Fast Ethernet, Gigabit Ethernet, ATM, Frame Relay, etc.
La couche transport ou couche 4 peut être mise en oeuvre par TCP (Transmission Control Protocol), UDP (User Datagram Protocol) ou SCTP (Stream Control Transmission Protocol).
Un grand nombre d'applications se retrouvent à la couche 7: SMTP pour le transfert de mail, HTTP pour l'accès au WEB, FTP pour le transfert de fichier, RTP pour le transport de la voix sur IP, etc.

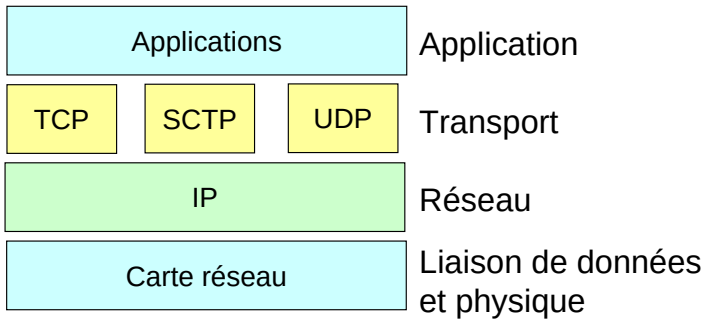

Figure 1 : La pile de protocoles TCP/IP

IP est le protocole qui cache le réseau physique sous-jacent en créant une vue de réseau virtuel.
Afin de permettre la communication entre systèmes ou stations dans le réseau, chaque station doit disposer d'une identité unique. Dans un réseau IP, il s'agit d'une adresse IP. Le protocole IP est donc responsable de l'adressage dans le réseau.
Une unité de données dans un réseau IP est appelée datagramme IP. C'est l'unité d'information de base transmise à travers des réseaux TCP/IP. Le protocole IP fournit les fonctions de routage afin d'acheminer ces datagrammes au destinataire correspondant. Le

protocole IP est non fiable (il offre un service best-effort) et fonctionne en mode non connecté.
"Non fiable" signifie que les datagrammes émis par le protocole IP peuvent être perdus, peuvent arriver en désordre, peuvent être dupliqués, ou peuvent arriver avec un contenu comportant des erreurs. Par ailleurs, il n'y a aucun mécanisme de contrôle de flux mis en oeuvre par IP. Le protocole IP considère que les protocoles des couches supérieures à savoir par exemple TCP, UDP ou SCTP corrigeront ses déficiences,.
La principale raison pour laquelle le mode non connecté a été préféré était de disposer d'un réseau qui puisse continuer à fonctionner même si une partie du réseau devient indisponible ou est détruite. Il s'agissait d'une exigence du DoD (United States Department of Defense).

2 Adressage IP (Classful)

Le plan d'adressage initial de l'Internet était divisé en trois classes (Figure 2):

- la classe A (commençant par la séquence binaire 0) réserve 7 bits pour les numéros de réseau et 24 bits pour les équipements;
- la classe B (commençant par la séquence binaire 10) réserve 14 bits pour les numéros de réseau et 16 bits pour les équipements;
- la classe C (commençant par la séquence binaire 110) réserve 21 bits pour les numéros de réseau et 8 bits pour les équipements.
- À cela s'ajoutent les classes D (commençant par la séquence binaire 1110), utilisées pour joindre un groupe d'équipements (ces adresses sont appelées multicast) et les classes E (commençant par la séquence binaire 1111) qui sont réservées à un usage futur.

L'allocation de la partie réseau était gérée par un organisme de l'Internet, ce qui garantissait l'unicité des adresses et permettait ainsi la construction d'un réseau mondial. Il est à noter que l'attribution des préfixes ne se faisait pas en tenant compte des contraintes géographiques ou suivant la topologie du réseau.
Il peut exister jusqu'à 127 réseaux de classe A contenant chacun jusqu'à 16 777 214 équipements.
Il peut exister jusqu'à 16 384 réseaux de classe B contenant chacun jusqu'à 65 534 équipements.
À l'opposé, il peut exister jusqu'à 2 097 152 réseaux de classe C, mais ceux-ci ne peuvent contenir que 254 équipements.
Le 1er octet de l'adresse est discriminant par rapport à la classe d'adresse associée : 0-127 : Class A ; 128-191 : Class B; 192-223 : Class C; 224-239 : Class D; 240-255 : Class E

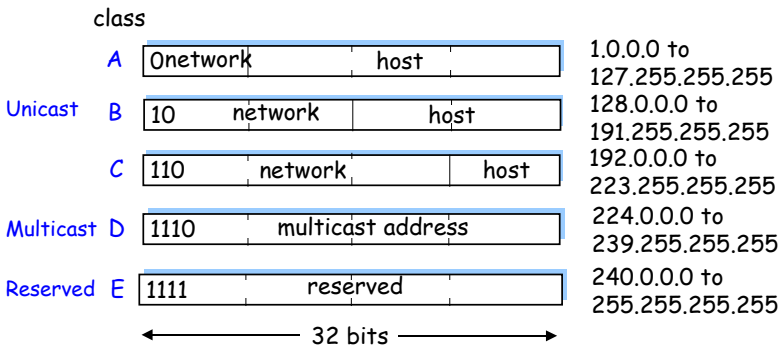

Figure 2 : Plan d'adressage IP

3 Subnetting

Le subnetting est le mécanisme qui permet à une entreprise de partager un numéro de réseau unique entre plusieurs réseaux internes.
L'organisation décompose la partie host en deux parties : parties subnet (sous-réseau) et subnet host (host de sous-réseau).
Un masque de sous-réseau (subnet mask) est utilisé afin d'indiquer quelle partie variable de l'adresse est considérée comme "network ID+SubnetworkID" par le site local.
Les sous-réseaux (Subnets) sont invisibles en dehors du site
Dans le réseau Internet public, les tables de routage des routeurs stockent uniquement les correspondances (network number, next hop).
Dans l'entreprise, les tables de routage des routeurs stockent les correspondances (subnet number, subnet mask, next hop)

Un masque de sous-réseau (souvent désigné par ses homonymes anglais : subnet mask, netmask et address mask) est un masque indiquant le nombre de bits d'une adresse IPv4 utilisés pour identifier le sous-réseau, et le nombre de bits caractérisant les hôtes (ce qui indique aussi le nombre d'hôtes possibles dans ce sous-réseau).
Les masques de sous-réseau utilisent la même représentation que celles des adresses IPv4. En IPv4, une adresse IP est codée sur 4 octets, soit 32 bits (représentés en notation décimale à point). Un masque de sous-réseau possède lui aussi 4 octets.
Dans l'exemple ci-dessous le masque de sous-réseau est 255.255.255.128 ou 11111111 11111111 11111111 10000000 en notation binaire. Il faut noter que par convention, l'adresse de réseau (network address) est incluse dans le masque de sous-réseau. En résumé, dans le masque de sous-réseau, les bits à 1 identifient la partie réseau et sous-réseau et les bits à 0 identifient la parte hôte.
Le numéro de sous-réseau d'un host est obtenu en effectuant une opération logique "AND" entre le masque et l'adresse IP du host (adresse IP de destination du datagramme reçu par le routeur).

Considérons un réseau de classe C, 192.44.77.0: Si le réseau contient deux sous-réseaux, le masque de sous-réseau est : 11111111 11111111 11111111 10000000 égal à 255.255.255.128. Chaque sous-réseau est capable de supporter 126 hosts (il faut soustraire 2 de 128, les valeurs 0 et 127 correspondant respectivement à l'adresse du sous-réseau et à l'adresse broadcast dans le sous-réseau). .
Les adresses des sous-réseaux sont 192.44.77.0 et 192.44.77.128.

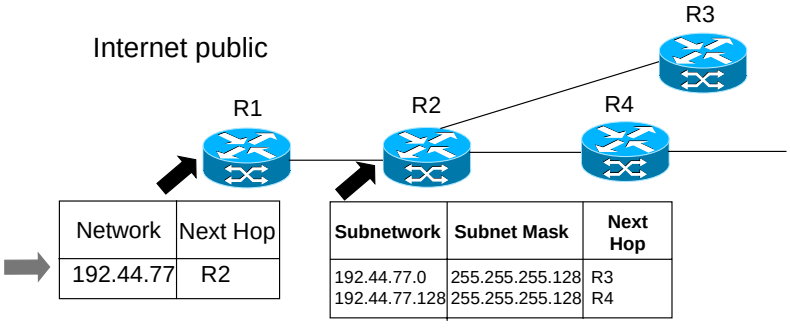

Figure 3 : Exemple de subnetting

Un datagramme IP a pour destination 192.44.77.23 (Figure 3). Le réseau Internet public route le datagramme au routeur R2 puisqu'une des entrées de la table de routage du routeur R1 indique que le saut suivant est R2 pour le réseau 192.44.77.0.

Le routeur R2 reçoit le datagramme. Puisque des sous-réseaux sont présents, une opération logique AND est réalisée entre l'adresse IP de destination 192.44.77.23 du datagramme et le masque spécifié par la table de routage, à savoir 255.255.255.128. Le résultat est un sous-réseau, 192.44.77.0. R2 route alors le datagramme au sous-réseau 192.44.77.0 via R3 comme indiqué par la table de routage.

Pour un réseau de class C, si l'on considère réserver 1 bit de la partie host pour configurer des sous-réseaux, alors 2 sous-réseaux sont possibles. Le masque de sous-réseau est 255.255.255.128 (le dernier octet étant 10000000). Chaque sous-réseau dispose de 126 (128-2) adresses (Figure 4).
Avec 2 bits réservés dans la partie host pour les sous-réseaux, il est possible de configurer 4 sous-réseaux. Le masque est alors 255.255.255.192 (le dernier octet étant 11000000). Chaque sous-réseau peut disposer de 62 (64-2) adresses.
Avec 3 bits de sous-réseaux dans la partie host, 8 sous-réseaux sont alors possibles. Le masque est 255.255.255.224 (le dernier octet étant 11100000). Chaque sous-réseau peut disposer de 30 (32-2) adresses.
Le maximum de bits pouvant être réservés dans la partie host pour les sous-réseau est 6. Il y a aura alors 64 sous-réseaux chacun disposant de 2 (4-2) adresses. Le masque de sous réseau associé est 255.255.255.252 (le dernier octet étant 11111100).

Bits Host	Bits Subnet	Nombre de Hosts	Plage de valeurs sur un Subnet	Masque subnet
7	1	126	x.y.z.1-x.y.z.127	255.255.255.128
6	2	62	x.y.z.1-x.y.z.62	255.255.255.192
5	3	30	x.y.z.1-x.y.z.30	255.255.255.224
4	4	14	x.y.z.1-x.y.z.14	255.255.255.240
3	5	6	x.y.z.1-x.y.z.6	255.255.255.248
2	6	2	x.y.z.1-x.y.z.2	255.255.255.252

Figure 4 : Sous-réseaux possibles dans un réseau de Classe C

4 Adressage CIDR : Classless Inter Domain Routing

La structure des adresses IP basée sur la classe (classful) rend l'allocation de ces adresses difficiles.
Adresses Classe A : Il existe trop peu de réseaux de Classe A
Adresses Classe C : un réseau de Classe C n'offre que 254 adresses (rappelons que bien qu'il existe un octet pour le host, ce qui représente 256 valeurs, il faut toujours ôter 2 valeurs, 0 et 255); ceci est insuffisant pour la majorité des organisations. Plusieurs réseaux de classe C par organisation rend les tables de routage trop grandes.
Adresses Classe B : un réseau de Classe B correspond au type de réseau idéal mais n'est plus disponible.
Pour un site avec 500 hôtes, l'usage de l'espace d'adressage de la Classe B est: $500/2^{16} < 1\%$
La solution est d'allouer des adresses de réseau par bloc d'adresses IP contiguës sous la forme <IP address, mask>
Le masque identifie la taille du bloc, qui doit être une puissance de 2.
exemple de la figure 5: allocation de 4 blocs d'adresses de classe C 192.4.16.0-192.4.19.255 : <192.4.16.0, 255.255.252.0>, ou 192.4.16/22.

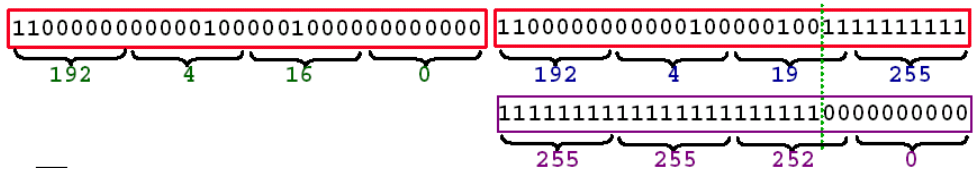

Figure 5 : Exemple CIDR

CIDR (Classless Inter Domain Routing) étend le concept de sous-réseau. à la partie réseau et non plus qu'à la partie host.
La partie réseau de l'adresse a donc une longueur arbitraire. Le format d'adresse est a.b.c.d/x, où x est le nombre de bits de la partie réseau de l'adresse.

Avec le principe d'adresse Classful, une entreprise qui disposerait d'un réseau IP constitué de 4 réseaux de Classe C serait déclaré dans les routeurs de l'Internet public via 4 entrées.
Avec le principe d'adresse CIDR, une seule entrée est nécessaire comme le montre la figure 6.

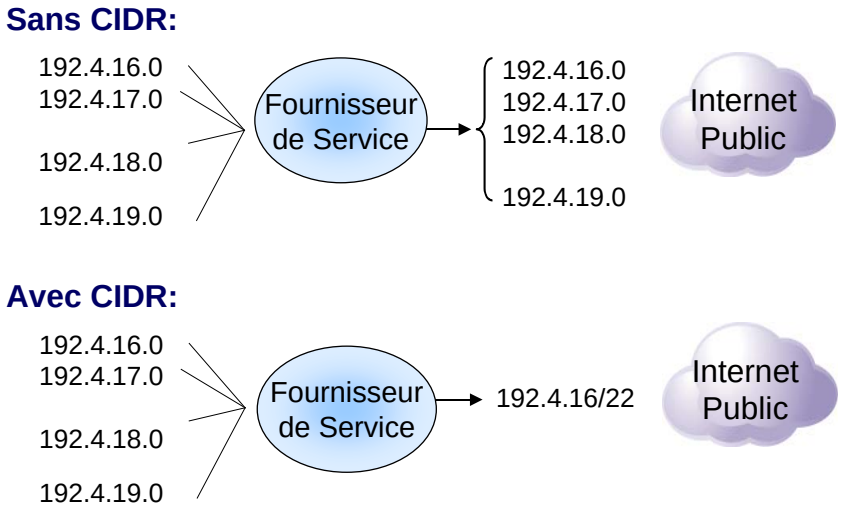

Figure 6 : Adressage Classful versus adressage Classless (CIDR)

5 Adressage privé

La RFC 1918 définit un espace d'adressage privé permettant à toute organisation d'attribuer des adresses IP aux machines de son réseau interne sans risque d'entrer en conflit avec une adresse IP publique allouée par les organismes officiels. Ces adresses dites non-routables correspondent aux plages d'adresses suivantes :
Classe A : plage de 10.0.0.0 à 10.255.255.255 ;
Classe B : plage de 172.16.0.0 à 172.31.255.255 ;
Classe C : plage de 192.168.0.0 à 192.168.255.55 ;
Toutes les machines d'un réseau interne, connectées à Internet par l'intermédiaire d'un routeur et ne possédant pas d'adresse IP publique doivent utiliser une adresse contenue dans l'une de ces plages. Pour les petits réseaux domestiques, la plage d'adresses de 192.168.0.1 à 192.168.0.255 est généralement utilisée.

6 Format du datagramme IP

Le datagramme IP est formé de deux parties (Figure 8):

- l'en-tête du datagramme, d'une taille au minimum de 20 octets. C'est là que sont inscrites toutes les informations du protocole (adresse, segmentation, options, etc.).
- les données utiles ou champ de données, d'une taille maximum de 65536 octets - les octets d'en-tête et d'options. Elle véhicule les données provenant de la couche supérieure (e.g., un segment TCP ou UDP).

L'entête IP, comme tout le paquet IP est structurée en mots de 4 octets. Il contient les champs suivants :

- VER (4 bits)- La version du protocole IP (actuellement 0100 (4) mais peut aussi être la version 6).

- HLEN (4 bits)- Longueur de l'en-tête en mots de 4 octets; donc si la longueur est **6**, alors 6*4 = 24 octets représente la longueur de l'en-tête. La taille de l'en-tête est 20 octets ou **5** mots de 4 octets si il n'y a pas d'options.

- Type of Service (8 bits) - Contient l'information de qualité de service qui peut induire un traitement particulier pour le datagramme. Ce champs contient les bits suivants (Figure 7):

- Les 3 bits de PRECEDENCE qui définissent la priorité du datagramme.
- Les Bit D, T, R qui indiquent le type d'acheminement désiré pour le datagramme, permettant à un routeur de choisir entre plusieurs routes (si elles existent).
- Le bit D si positionné à 1 signifie que le délai doit être court pour le datagramme,
- Le bit T si positionné à 1 signifie que le débit doit être élevé,
- Le bit R si positionné à 1 signifie que la fiabilité doit être importante.

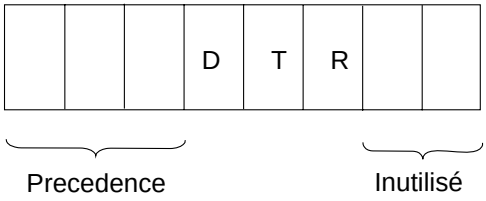

Figure 7 : Le champ ToS (Type of Service)

Lorsqu'un routeur a ses files d'attente saturées, il doit supprimer des datagrammes. Il peut alors considérer par exemple traiter avec priorité les datagrammes dont le bit fiabilité est positionné à 1 par rapport à ceux dont le bit de fiabilité est à 0.

- Total Length (16 bits) - Nombre d'octets du datagramme IP en incluant l'en-tête. La taille maximum est 65,535 octets.

- Identification (16 bits) - L'identification est l'identité unique assignée à un datagramme afin d'aider au réassemblage de datagrammes fragmentés.

- Flags (3 bits) - Le bit 0 est toujours positionné à 0 et correspond à un bit réservé. Le bit 1 (DF, Don't Fragment) indique si un datagramme peut être fragmenté (0) ou non (1). Le bit 2 (MF, More Fragment) indique au récepteur si le fragment est le dernier du datagramme initial (0) ou s'il y a encore des fragments à venir (1).

- Frag Offset - En unités de 8 octets (64 bits), le champ frag offset spécifie une valeur pour chaque fragment dans le processus de réassemblage.

- TTL (Time to Live, durée de vie) - Limite la durée maximum de maintien du datagramme dans le réseau. Le champ TTL est positionné par l'hôte émetteur et est décrémenté à chaque routeur qui traite le datagramme. Un datagramme qui n'a pas encore atteint sa destination lorsque le TTL a atteint 0, est supprimé. La valeur par défaut pour TTL est 64.

- Protocol - Utilisé afin d'indiquer quel est le protocole de la couche supérieure destinataire du datagramme (i.e., TCP, UDP, OSPF, ou autre protocole). Si la valeur de ce champ est 6, la couche supérieure est alors TCP. La valeur 17 représente UDP.

- Header Checksum - Ce champ de 16 bits (2 octets) contient un checksum qui est calculé sur l'en-tête du datagramme uniquement et non pas sur les données utiles qui proviennent de la couche supérieure. Le checksum doit être recalculé et mis à jour à chaque saut du datagramme car la valeur TTL est modifiée par chaque routeur sur la route.

- Source and destination IP addresses - Correspondent aux adresses IP de la source et de la destination du datagramme. L'adresse IP de destination permet d'acheminer le datagramme à la destination. L'adresse IP source permet à la destination de retourner une réponses à la source. Par exemple, si la source émet une requête HTTP afin d'obtenir une page WEB, alors la requête HTTP GET est d'abord encapsulée dans un segment TCP puis dans un datagramme IP dont l'adresse source est celle du client WEB et l'adresse de destination, celle du serveur WEB. Lorsque le serveur WEB retourne la réponse HTTP 200 OK contenant la page WEB, le datagramme encapsulant cette réponse à pour adresse IP de destination le client WEB et pour adresse IP source celle du serveur WEB.

- IP Options - Le datagramme IP peut contenir des informations optionnelles. Cependant, il est rare de trouver des options dans un en-tête IP. Parmi les options prévues par le protocole, certaines concernent le routage :

- l'option router alert permet d'indiquer aux routeurs qu'ils doivent examiner le contenu du datagramme. Cette option est utilisée par le protocole de gestion des groupes multicast (IGMP : Internet group message protocol) et le protocole de signalisation et de réservation de ressource (RSVP : reservation protocol) ;
- l'option routage par la source (source routing). Cette option permet de préciser un chemin dans le réseau, c'est-à-dire l'adresse IP des routeurs par lesquels le datagramme devra passer.

- Padding - Un bourrage est souvent rajouté afin de s'assurer que la longueur du datagramme sera multiple de 4 octets.

A la suite de l'en-tête décrit précédemment, sont présentes les données utiles provenant de la couche supérieure, e.g., TCP, UDP, SCTP, etc.

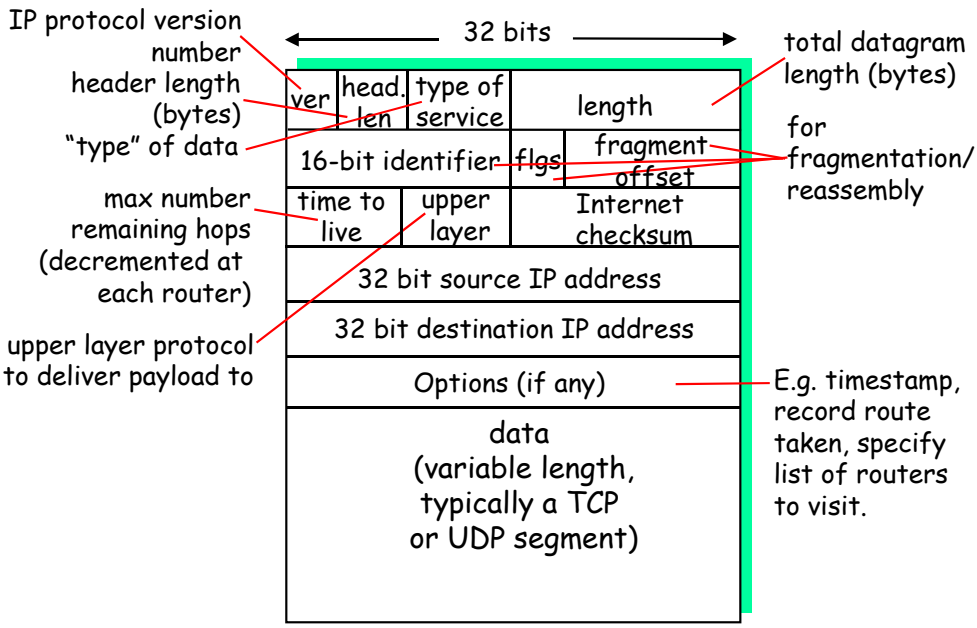

Figure 8 : Format du datagramme IP

7 Le protocole DHCP

DHCP signifie Dynamic Host Configuration Protocol. Il s'agit d'un protocole qui permet à un hôte qui se connecte sur un réseau local d'obtenir dynamiquement et automatiquement ses paramètres de configuration IP. Le but principal étant la simplification de l'administration d'un réseau IP.
DHCP supporte trois modes d'allocation d'adresse IP :

- En mode "automatic allocation", DHCP affecte une adresse IP permanente à un client quand celui-ci contacte pour la première fois le serveur DHCP.
- En mode "dynamic allocation", DHCP affecte une adresse IP à un client pur une durée limitée (ou jusqu'à ce que le client relache l'adresse).
- En mode "manual allocation", l'adresse IP est assignée au client par l'administrateur réseau, et DHCP est juste utilisé pour fournir l'adresse affectée au client.

Un réseau utilise un ou plusieurs de ces modes, en fonction des politiques décidées par l'administrateur de réseau.

8 Network Address and Port Translation (NAPT)

Le NAT (Network Address Translation) dynamique permet de partager une adresse IP routable (ou un nombre réduit d'adresses IP routables) entre plusieurs machines en adressage privé. Ainsi, toutes les machines du réseau interne possèdent virtuellement, vu de l'extérieur, la même adresse IP. C'est la raison pour laquelle le terme de "mascarade IP" (en anglais IP masquerading) est parfois utilisé pour désigner le mécanisme de translation d'adresse dynamique.
Afin de pouvoir " multiplexer " (partager) les différentes adresses IP sur une ou plusieurs adresses IP routables le NAT dynamique utilise le mécanisme de translation d'adresse et de port (NAPT - Network Address and Port Translation), c'est-à-dire l'affectation de port source/adresse IP source différents à chaque requête de telle manière à pouvoir maintenir une correspondance entre les requêtes provenant du réseau interne et les réponses des machines sur Internet, toutes adressées à l'adresse IP du routeur.

Considérons la configuration à la figure 9 qui utilise NAPT. Un client dispose d'une BOX à laquelle l'opérateur a affecté une adresse IP publique à savoir 138.76.29.7. Le client possède 3 notebooks, tous connectés au réseau local et disposant chacun d'une adresse IP privée, respectivement 10.0.0.1, 10.0.0.2 et 10.0.0.3. Lorsque le client souhaite envoyer une requête HTTP à un serveur WEB depuis le notebook 1, il encapsule la requête dans un segment TCP. L'en-tête du segment rappelle les numéros de port source (port du notebook 1) à savoir 3345 pour cette session HTTP et le port destination (port du serveur WEB) 80. Le segment TCP est ensuite encapsulé dans un datagramme IP dont l'en-tête inclut les adresses IP source (10.0.0.1) et destination (128.119.40.186). L'adresse IP source est privée, alors que l'adresse IP destination est publique (c'est celle du serveur WEB obtenue par interrogation du DNS). Lorsque le datagramme est reçu par la BOX, sa fonction NAPT remplace le port source par un autre port 5001 et l'adresse IP source privée par l'adresse de la BOX, à savoir 138.76.29.7, quant à elle publique. La correspondance est mémorisée dans la table de traduction NAPT de la BOX (138.76.29.7, 5001; 10.0.0.1, 3345). Le datagramme est routé via l'Internet public au serveur WEB. Le serveur exécute la requête HTTP GET et retourne une réponse HTTP 200 OK. Cette réponse est encapsulée dans un segment TCP dont l'en-tête inclut les port source et destination, à savoir 80 et 5001. Par ailleurs, le segment est encapsulé dans un datagramme IP dont les adresses source et destination sont respectivement 128.119.40.186 et 138.76.29.7. Ce datagramme est routé via l'Internet publique à la BOX qui est le destinataire. Le mécanisme NAPT est essentiel car le réseau Internet public ne sait router des datagrammes que si leur adresse IP de destination est publique. La BOX grâce à sa fonction NAPT remplace le couple (138.76.29.7, 5001) du datagramme à l'aide de sa table NAPT par le couple (10.0.0.1, 3345). Le datagramme est ensuite routé à sa destination, le notebook 1.

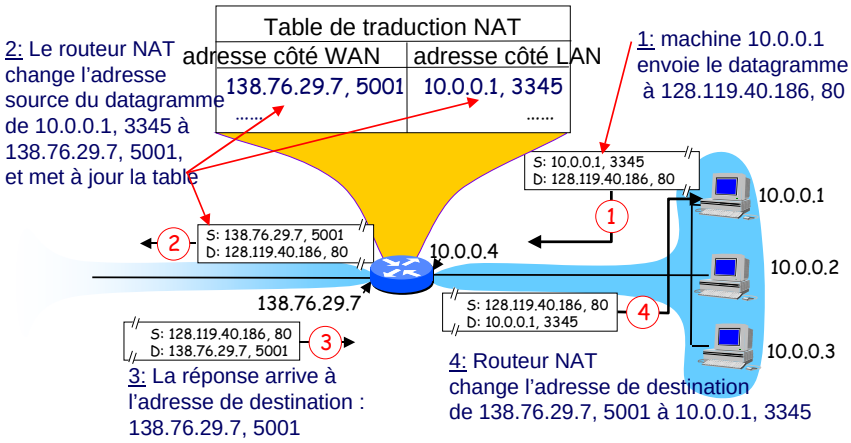

Figure 9 : Exemple de translation d'adresse et de port

La translation d'adresse ne permet de relayer que des requêtes provenant du réseau interne vers le réseau externe. Cela signifie qu'il est impossible en tant que tel pour une machine externe d'envoyer un datagramme vers une machine du réseau interne. En d'autres termes, les machines du réseau interne ne peuvent pas fonctionner en tant que serveur vis-à-vis de l'extérieur.
Pour cette raison, il existe une extension du NAT appelée “ redirection de port ” (en anglais Port Forwarding ou Port mapping) consistant à configurer la passerelle pour transmettre à une machine spécifique du réseau interne, tous les paquets réçus sur un port particulier. Ainsi, si l'on souhaite pouvoir accéder de l'extérieur à un serveur web (port 80) fonctionnant sur la machine 10.0.0.5, il sera nécessaire de définir une règle de redirection de port sur la passerelle, redirigeant tous les paquets TCP reçus sur son port 80 vers la machine 10.0.0.5.

9 Routage IP

Afin de relayer un datagramme IPv4, un routeur réalise les fonctions suivantes :

- Vérifie le champ Checksum en calculant son propre checksum et en comparant son résultat avec celui trouvé dans l'en-tête du datagramme IPv4.
- Vérifie la valeur du champ Version.
- Décrémente la valeur du champ TTL.
- Vérifie la présence d'options dans l'en-tête du datagramme IPv4. Si des options sont présentes, elles sont traitées.
- Utilise la valeur du champ Adresse Destination et les informations de la table de routage locale afin de déterminer une interface pour relayer le datagramme à l'adresse IPv4 du prochain saut.
- Fragmente le datagramme si nécessaire
- Recalcule le nouveau checksum sur l'en-tête du datagramme IPv4 et met à jour le champ correspondant.
- Relaye le datagramme en utilisant l'interface de relai appropriée.

Parallèlement au traitement des datagrammes reçus, le routeur doit mettre à jour les routes disponibles dans sa table de routage. Pour ce faire, il doit échanger avec les autres routeurs du réseau des informations de routage.

Sur Internet, un Système Autonome (Autonomous System ou AS) est un ensemble de réseaux IP sous le contrôle d'une seule et même entité, typiquement un fournisseur d'accès à Internet ou une plus grande organisation qui possède des connexions redondantes avec le reste du réseau Internet.
La notion de système autonome est donc administrative et non technique. Ce qui caractérise le système autonome, c'est l'unité de décision en son sein. Certains protocoles de routage sont utilisés pour déterminer les routes dans l 'AS. D 'autres protocoles de routage sont utilisés afin d 'interconnecter un ensemble d'ASs.
Au sein d'un système autonome, le protocole de routage est qualifié d'interne (par exemple, Open Short Path First (OSPF) and Routing Information Protocol (RIP)).
Entre deux systèmes autonomes, le routage est externe (par exemple Border Gateway Protocol (BGP)).
Il existe différentes raisons qui justifient une multiplicité de protocoles de routage :

- Le routage à l'intérieur d'un réseau et le routage entre réseaux ont des contraintes en terme de sécurité, de scalabilité et de stabilité différentes. Différentes protocoles de routages ont été développés pour répondre à ces besoins.
- Des réseaux de différentes tailles utilisent des protocoles de routage différents. Des réseaux de petite taille et de taille moyenne utilisent généralement des protocoles de routage simples mais qui ne sont pas scalables. Pour des réseaux de grande taille, des protocoles de routage complexes sont nécessaires pour répondre au besoin de scalabilité.

9.1 Protocole de routage à vecteur de distance

Les protocoles de routage à vecteur de distance sont des exemples de protocoles de routage dynamique dont font partie les protocoles RIP et BGP. Les algorithmes associés permettent à chaque routeur dans le réseau de construire et maintenir automatiquement une table de routage IP locale. Le principe sous-jacent au routage à vecteur de distance est simple. Chaque routeur met à jour la distance ou le coût de lui-même à toute destination connue. La route représentée par le coût le moins élevé devient la route préférée pour atteindre la destination.

L ’information est maintenue dans une table de routage à vecteur de distance. La table est périodiquement émise aux routeurs voisins. Chaque routeur traite les tables reçues pour déterminer les meilleures routes à travers le réseau.
Le principal avantage des algorithmes à vecteur de distance est sa simplicité. Ils sont très utilisés dans des petits réseaux avec une redondance limitée.
Les principaux point faibles sont le temps excessif pour la convergence, le manque de scalabilité et le grand volume d ’information de routage échangé.

9.2 Protocole de routage à état de liaison

Les protocoles de routage à vecteur de distance sont des exemples de protocoles de routage dynamique dont fait partie le protocole OSPF.
Chaque routeur construit un message contenant la liste de ses voisins immédiats ainsi que le coût associé à la liaison. Ce message est appelé LSP (Link State Packet).
Ce paquet est transmis à tous les autres routeurs du réseau (LSP broadcast). Cette transmission a lieu lorsque le routeur découvre qu ’il a un nouveau voisin, lorsque le coût d ’un liaison vers un voisin a changé ou lorsque l'état d'un lien vers un voisin a changé.
Chaque routeur met à jour sa base de données, ce qui lui donne une vision globale du réseau et il peut en déduire ses tables de routage en appliquant l ’algorithme à état de liaison (link state).

Critère	Vecteur de distance	Etat de liaison
Algorithme de base	Algorithme de Bellman-Ford	Algorithme de Djikstra
Scalabilité	Difficile avec de grands réseaux	Les tables de routage convergent rapidement même avec de grands réseaux
Complexité	L ’implantation est simple et aisée	L ’implantation est complexe et demande de la CPU pour le calcul des routes.
Trafic généré	Tendance à produire un gros trafic de routage	Mises à jour sont minimisées.
Fonctionnement	Le coût le plus petit est déterminée par le nombre de sauts le plus petit	Le coût le plus petit est calculé à partir de l ’exécution d ’un algorithme sur la base de données de la topologie.

Figure 9 : Comparaison entre protocoles à vecteur de distance et à état de liaison