NAT à l'échelle d'une université

Didier BENZA, benza@univ-tln.fr
CRI de l'université de Toulon

L'article présente la technologie Network Address Translation et les raisons qui peuvent conduire un établissement universitaire à la mettre en œuvre sur un routeur de site. Il expose les tâches d'administration quotidiennes, les problèmes rencontrés et l'impact de NAT sur la sécurité et sur le raccordement au Mbone. Il évoque enfin quelques pistes probables d'évolution de cette technologie.

Historique à l'université

Etat des lieux en 1997

En septembre 1997, le réseau de l'université de Toulon s'étendait sur 4 sites. Cinq cents machines étaient raccordées à ce réseau qui utilisait 3 plages d'adresses de classe C sur le site principal.

Pour créer des séparations fonctionnelles de type enseignants, étudiants, administration, nous avions organisé les 3 classes C en sous-réseaux. Nous approchions de la limite de plusieurs des sous-réseaux définis et nous prévoyions de faire une nouvelle demande de classe C.

L'introduction d'une adresse de classe C allait nous obliger à jongler avec l'adressage du site et ceci pour une durée indéterminée. Pour nous aider dans le choix de la marche à suivre, nous avons réalisé une étude statistique de l'utilisation des adresses du campus et nous avons essayé d'évaluer le nombre de raccordements à venir.

Nous avons donc écrit un script qui testait cycliquement la présence des machines sur le réseau (à l'aide de simples *ping*). Ce script a démontré qu'en moyenne seules 40 % des machines déclarées étaient en fonctionnement, donc 60 % des adresses IP étaient inutilisées. Par ailleurs, le taux de raccordement au réseau, observé sur les derniers mois était d'environ 13 nouvelles machines par semaine. Le nombre de salles qui n'étaient pas encore raccordées au réseau était important, nous savions donc que nous aurions besoin d'une nouvelle adresse de classe C au bout de 4 mois environ.

Les Choix possibles

Nous avons déterminé que les 3 choix suivants s'offraient à nous :

- *Demander plusieurs adresses de classe C à RENATER.* Cette solution permettait de résoudre temporairement le problème de la pénurie d'adresses. Mais les nouvelles adresses de classe C devaient alors être affectées selon des analyses de flux pour un routage efficace. Ceci impliquait la renumérotation d'une bonne partie des machines du site. Nous ne savions pas pendant combien de temps les adresses demandées suffiraient avant que nous ayons de nouveau besoin d'adresses.
- *Utiliser DHCP*[1]. DHCP permettait d'économiser des adresses en n'utilisant à un moment donné qu'une partie des adresses allouées à l'université. Mais L'économie d'adresses ne serait pas suffisante pour nous éviter de faire une demande de classe C. De plus, il faudrait déployer plusieurs serveurs DHCP, ce qui engendrerait une consommation importante de ressources humaines. Enfin, nous serions contraints de faire un nouveau paramétrage de toutes les machines du site.
- *Utiliser NAT*[2]. NAT proposait une solution durable, puisque nous pouvions utiliser des adresses de classe A ou B pour renuméroter notre site. La pléthore d'adresses permettrait de réaliser le design de notre « réseau idéal ». Un plan d'adressage dont nous étions certains qu'il ne serait pas remis en cause par une pénurie d'adresses IP. La mise en place de NAT ne demandait aucun autre investissement matériel que l'achat de mémoire sur le routeur. Notre routeur de site était un routeur Cisco, il suffisait juste d'acheter une nouvelle version de l'IOS[3]. L'effort de renumérotation n'était pas plus important que le déploiement de DHCP et le plan d'adressage résultant de la mise en place durerait bien plus longtemps que celui obtenu en demandant de nouvelles classes C à RENATER.

Nous avons donc choisi NAT.

Mise en place

Le choix de NAT étant fait, nous avons entrepris les actions permettant de le mettre en œuvre :

- Concertation avec les correspondants techniques des UFR et services pour définir un nouveau plan d'adressage et établissement de la liste des machines qui continueraient d'être accessibles depuis l'Internet. 10 machines ont été recensées.
- Mise en place de messages d'information sur le serveur WEB du campus et écriture de scripts CGI informant la machine se connectant dessus de ses nouveaux paramètres réseaux, tels qu'ils devraient être configurés après la mise en place de NAT : nouvelle adresse IP, masque de sous-réseau, passerelle par défaut, serveur DNS primaire et secondaire, etc.
- Ecriture de la nouvelle configuration du routeur plusieurs semaines avant sa mise en place. Le jour J, chargement de la configuration sur le routeur. 90 % des configurations des machines ont été modifiées directement par leur utilisateur, le reste étant traité par des informaticiens des UFR ou du CRI.
- Restitution de 3 adresses de classe C devenues inutiles à RENATER.

Etat actuel

Deux ans après la mise en place de NAT sur le routeur de site la situation est la suivante :

- 1 500 machines sont connectées au réseau, l'équivalent de 5 classes C sans sous-réseaux ou de 8 classes C avec des subnets. Le plan d'adressage mis en œuvre il y a 2 ans n'a pas été modifié depuis.
- 20 machines ont une traduction statique. Les ajouts sont essentiellement des machines accédant à des bases de données payantes dont le contrôle d'accès se fait sur l'adresse IP, moyen d'authentification plutôt bas de gamme très en vogue actuellement.

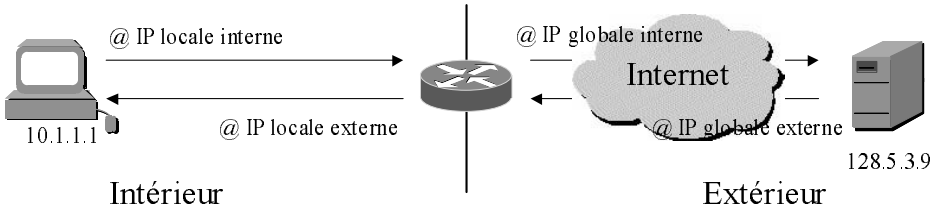
■ Description du fonctionnement

Terminologie

Routeur NAT : Routeur sur lequel NAT a été activé. *Site NAT :* Site ayant mis en place un routeur NAT au point de connexion de son réseau avec l'Internet

La terminologie suivante est celle présentée par le groupe de travail sur NAT mis en place conjointement par le CRU_4 et l'$UREC_5$ sur le mécanisme NAT en février 1998_6. Cette terminologie est directement issue de celle employée par Cisco. Dans les termes suivants, les mots *interne* et *externe* désignent l'origine d'une adresse par rapport au routeur NAT. Les mots *local* et *global* désignent le côté du routeur où cette adresse évolue.

Espace d'adressage privé L'espace des adresses utilisées en interne par un site NAT. Ces adresses sont généralement celles définies par le $RFC1918_8$, mais pas obligatoirement		**Espace d'adressage public** L'espace des adresses gérées par l'$IANA_7$. Ces adresses sont globalement uniques, elles sont routées sur l'Internet	
Adresse locale interne (ALI)	**Adresse locale externe (ALE)**	**Adresse globale interne (AGI)**	**Adresse globale externe (AGE)**
@ IP d'une machine sur un site NAT	@ IP d'une machine externe vue de l'intérieur (overlapping)	@IP d'une machine interne vue de l'extérieur	@IP d'une machine externe



Mise en œuvre sur les routeurs Cisco

Fonctionnement général

Un routeur NAT maintient une table d'équivalence entre l'adresse IP locale interne (ALI) d'une machine et son adresse IP globale interne (AGI). Si une *traduction statique* a été définie pour une ALI, une AGI unique lui est associée. Cela est utile pour les machines devant recevoir des connexions de l'extérieur. Si aucune traduction statique n'a été définie pour l'ALI source d'un paquet qui se présente sur une interface interne, une AGI est prise dans un *pool*, c'est une *traduction dynamique*.

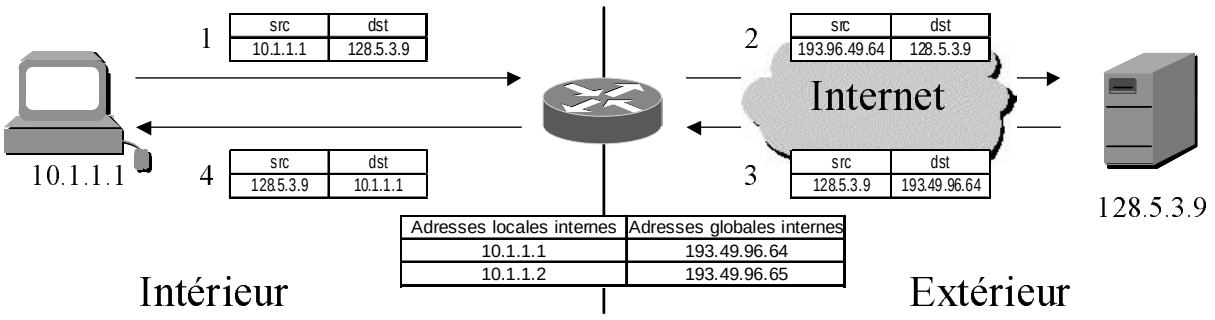


Table maintenue par le routeur

Admettons que la machine avec l'ALI 10.1.1.1 se connecte sur le serveur WEB ayant l'adresse IP 128.5.3.9. Le routeur réécrit le paquet en changeant l'adresse source. Pour cela, il prend la première adresse libre dans le pool des AGI, disons l'adresse 193.49.96.64. Il ajoute à la table des traductions dynamiques la paire (10.1.1.1, 193.49.96.64). Le paquet sortant du site a maintenant pour adresse source l'AGI 193.49.96.64, il atteint le serveur WEB qui répond à la requête. Le routeur voyant arriver un paquet à destination de la machine 193.49.96.64 lit la table de correspondance, trouve l'ALI associée, il réécrit le paquet et envoie le paquet modifié à la machine 10.1.1.1.

Traduction dynamique étendue ou Port Address Translation (PAT)

Pour éviter de tomber en panne d'adresses dans le pool dans le cas d'un nombre de connexions trop important, on peut valider PAT. Dans ce mode, ce n'est plus seulement l'ALI qui est traduite, mais aussi le port source de la connexion. Une fois PAT mis en œuvre et en excluant les ports réservés, un routeur Cisco peut associer environ 4000 ALI à une même AGI prise dans le pool.

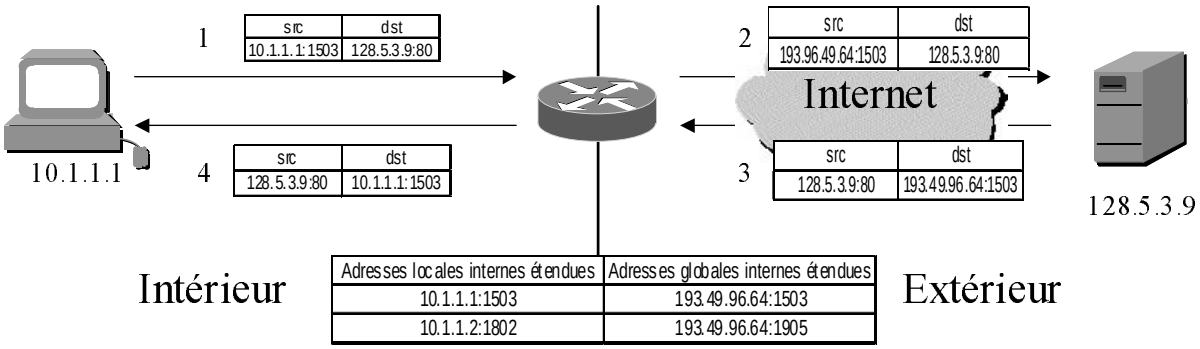
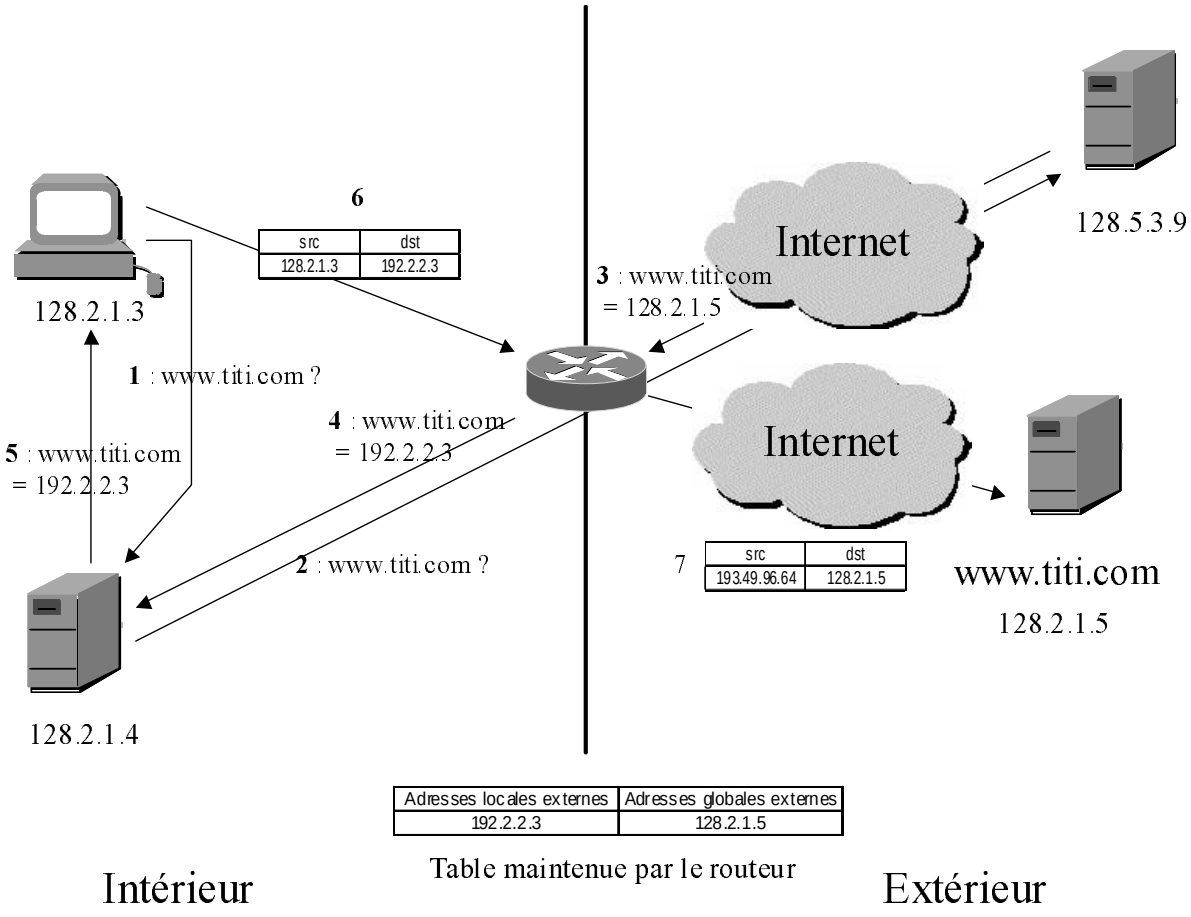


Table maintenue par le routeur

Pour reprendre notre exemple, si la machine ayant l'ALI 10.1.1.1 établit une connexion HTTP (80) vers le serveur WEB 128.5.3.9 à partir du port 1503, le routeur va associer la paire (10.1.1.1, 1503) à la paire (193.49.96.64, 1503). Lorsque le serveur renvoie un paquet en retour, le routeur réécrit ce paquet en remplaçant l'adresse destination et le port destination (193.49.96.64, 1503) en (10.1.1.1, 1503).

Fonctionnement particulier : overlapping

Si un site utilise des adresses publiques attribuées par l'IANA à un autre domaine, le schéma que nous venons de voir se complique quelque peu.



Prenons un autre exemple dans lequel nous éliminerons la traduction des adresses internes pour nous intéresser exclusivement aux adresses externes. Un site NAT a adressé les machines de son site avec des adresses IP publiques attribuées par l'IANA au domaine titi.com. L'utilisateur de la machine ayant l'ALI 128.2.1.3 désire se connecter sur le serveur WEB du domaine titi.com. Il tape l'url http://www.titi.com/. Son navigateur fait une demande de résolution DNS au serveur DNS interne du site. Celui-ci retransmet la demande à un serveur externe qui lui renvoie l'AGE 128.2.1.5. Lorsque cette réponse arrive au routeur, celui-ci modifie la partie données de la réponse DNS, il remplace l'adresse 128.2.1.5 par la première adresse libre du pool des ALE, disons l'adresse 192.2.2.3. Il envoie la réponse DNS modifiée au serveur DNS interne ayant fait la requête. Le serveur renvoie la réponse DNS à la machine cliente. Celle-ci initie la connexion HTTP vers le serveur www.titi.com en se connectant à l'adresse 192.2.2.3. Lorsque le routeur voit arriver des paquets de cette connexion, il change l'ALE 192.2.2.3 en l'AGE 128.2.1.5.

Administration au quotidien

L'administration d'un routeur NAT n'est pas très différente de celle d'un autre routeur. Il n'est pas fréquent de devoir toucher aux traductions statiques et encore moins aux traductions dynamiques. Lorsqu'une traduction statique doit être ajoutée, il faut accomplir les tâches suivantes :

- Entrer la commande `ip nat source static` adéquate.
- Exclure l'ALI de l'access-list de sélection des traductions dynamiques.

Lorsqu'une traduction doit être détruite, il suffit d'accomplir les tâches inverses en annulant la traduction et en sélectionnant de nouveau la machine pour les traductions dynamiques.

Pour modifier les règles des traductions dynamiques ou recharger la configuration du routeur, il est nécessaire d'entrer une commande `clear ip nat translations *` car sinon le routeur n'accepte pas les modifications. Cette commande a pour effet de supprimer toutes les traductions en cours. Pour éviter que de nouvelles traductions ne soient immédiatement créées, il faut aussi enlever les commandes `ip nat inside` sur les interfaces internes. Evidemment ces commandes coupent les connexions en cours. Pour éviter ce désagrément aux utilisateurs, il est utile de prendre l'habitude de modifier la configuration par morceaux et de ne plus faire un `configure network` pour changer une simple access-list.

Vous pourrez trouver sur le WEB une version de cet article détaillant les commandes d'un routeur NAT_{13}.

Impact sur les services Internet

Pour être compatible avec NAT, un protocole ne doit pas transporter d'adresses IP dans la partie données de ses paquets.

Protocoles supportés par construction	Protocoles supportés par un traitement particulier	Protocoles non supportés
HTTP, TFTP, TELNET, ARCHIE, FINGER, NTP, NFS, RLOGIN, RSH, RCP	ICMP, FTP, NetBios, RealAudio, CuSeeMe, StreamWorks, les requêtes A et PTR du DNS, H 323, Netmeeting, VDOLive, Vxtreme, Multicast IP	Mise à jour des tables de routage, transferts de zones DNS, BOOTP, talk & ntalk, SNMP, Netshow

Le DNS est profondément modifié par la mise en place de NAT. Les ALI doivent être gérées par un serveur DNS interne, les AGI et adresses de DMZ sont gérées par un serveur externe. Les routeurs NAT Cisco en traduisant les requêtes A et PTR du DNS permettent à certains sites de ne mettre en œuvre qu'un seul serveur. En effet, si le routeur voit arriver sur son interface interne une réponse à une requête DNS contenant une ALI pour laquelle il a une traduction en cours, il remplace dans la réponse l'ALI par l'AGI associée. A contrario, s'il reçoit sur une interface externe une requête PTR pour une AGI pour laquelle il a une traduction en cours, il la remplace par l'ALI associée.

Dans notre communauté, nous utilisons des serveurs DNS secondaires à l'extérieur de nos sites pour éviter des ruptures de services. Les serveurs primaires et secondaires communiquent par des transferts de zones[10]. Or NAT ne traite pas ce type de transfert, ce qui nous interdit de n'utiliser qu'un seul serveur. Nous devons donc mettre en œuvre 2 serveurs DNS, l'un interne, l'autre avec une adresse IP publique, généralement dans une DMZ. Les machines à l'intérieur du site NAT n'utilisent que le serveur DNS interne, elles n'ont pas à utiliser ni à connaître l'existence du serveur externe. Les machines externes n'ont pas à connaître l'existence et à utiliser le serveur DNS interne, elles n'utilisent que le serveur DNS externe.

Impact sur la sécurité

NAT offre un avantage indéniable à l'administrateur d'un site : il cache l'ensemble de ses machines pour ne laisser apparaître que les serveurs hébergeant des services Internet.

Mais NAT pose des problèmes réels de sécurité. Il est impossible à l'administrateur d'un site NAT de remonter à l'origine d'une attaque venant de son site. En effet, sauf traduction statique, ce qui par définition est rare, l'adresse IP que l'administrateur extérieur aura détectée comme posant un problème ne permettra pas de remonter jusqu'à la machine à l'origine du problème. La seule commande permettant de tracer les traductions effectuées par le routeur est la commande `debug ip nat` qui produit beaucoup trop de résultats pour pouvoir être utilisée et logguée. Cela écroulerait le routeur.

Il faut donc interdire en sortie d'un site NAT ce qui n'est pas strictement nécessaire en accordant une attention particulière aux protocoles dangereux comme les *rcommands* telles *rsh, rlogin*, etc… Les messages d'insultes que peuvent envoyer les utilisateurs d'un site par mail ou par formulaire sur le WEB peuvent aussi poser des problèmes, accentués par la mise en place de NAT. Pour pouvoir toujours remonter à la machine à l'origine de tels messages, il est nécessaire d'interdire les connexions SMTP sortantes sauf à partir des MX officiels. Des versions de *sendmail* écrivant dans les logs et dans les en-têtes des messages l'adresse IP réelle de la machine émettrice doivent être installés sur ces serveurs. Il faut installer un serveur cache HTTP qui loggue les accès au WEB pour pouvoir remonter à une machine postant des insultes sur des formulaires. Sur l'interface externe, il faut filtrer les ALI en plus des règles de mascarade IP classiques si votre adressage interne est conforme au RFC1918. Dans le cas contraire, ce problème est réglé par la mise en place de l'overlapping

En cassant la connectivité de bout de bout, NAT empêche la mise en place de protocoles de type IPSec.

Connexion au Mbone

Les versions actuelles de l'IOS, à partir de la version 12.0(5.5)T offrent la compatibilité entre NAT et le multicast IP[11]. Avec ces versions, si le routeur détecte une machine qui s'abonne avec une ALI à un groupe multicast, il lui attribue dynamiquement une AGI par une traduction simple. C'est-à-dire qu'il utilise une adresse du pool pour chaque machine abonnée à un groupe du Mbone[9] sur le site.

Le traitement du trafic multicast par NAT supporte ce qui suit : Adresses sources des paquets de données, les paquets de contrôle de PIM (charge utile de PIM), parmi eux Auto-rp, PIMv2 BSR. Mstat/mrinfo/requêtes mtrace /réponses mtrace et les annonces SDR (charge utile de l'application SDR).

Les adresses dans la charge utile de RTP/RTCP et des autres applications ne sont pas traduites.

NAT, IPv6, conclusion et perspectives

En ces périodes de baisse des prix des commutateurs de niveau 3 et de pénurie d'adresses IPv4, nous assistons à deux tendances contradictoires. D'un côté nous pouvons introduire du routage de plus en plus bas dans nos réseaux sans pertes de performances grâce à l'évolution technologique mais d'un autre côté, nous ne pouvons pas découper nos réseaux en sous-réseaux pour réaliser du routage sans perdre des parties importantes de notre espace d'adressage.

En éliminant les limites liées à la dépendance par rapport aux adresses IPv4 distribuées au compte-gouttes par un ISP, NAT offre à l'administrateur réseau la possibilité de créer un design correspondant aux besoins de son organisation.

Pour un nouveau site, il est probablement préférable d'adopter immédiatement un schéma IPv6 et de mettre en œuvre un boîtier de type NAT-PT_{12} (encore à l'état de draft) pour un résultat et un effort de mise en œuvre quasiment identique à NAT v4. De telles solutions devraient être disponibles bientôt au catalogue des grands constructeurs. NAT-PT permet la communication des mondes IPv6 et IPv4 en réalisant la traduction des adresses v6 & v4 et en ajoutant une traduction de protocole. Par exemple, les nouvelles commandes EPRT et EPSV de FTP-IPv6 sont traduites en des commandes correspondantes PORT et PASV et réciproquement. De la même façon, les requêtes DNS A IPv4 sont traduites en leurs équivalents IPv6, AAAA ou A6 et réciproquement.

Pour un site ancien dont le réseau est devenu, au fil du temps, un véritable casse-tête et dans lequel chaque introduction d'une nouvelle plage d'adresse oblige à la renumérotation d'une partie de ses machines; il est certain que la mise en œuvre de NAT permet d'augmenter les performances en laissant l'administrateur introduire de la commutation de niveau 3 sans craindre la pénurie d'adresses IP.

NAT sera probablement mis en œuvre sur de nombreux sites et pendant longtemps encore. Soit sous sa forme actuelle, soit sous la forme d'un routeur implémentant NAT-PT pendant la période transitoire entre IPv4 et IPv6. Au début de cette période, ce mécanisme permettra à des sites ayant adopté IPv6 pour leur adressage interne de conserver la connectivité avec le monde IPv4. Plus tard, il permettra à des sites restés IPv4 de bavarder avec un monde IPv6.

Donc loin d'être concurrents, ces protocoles sont complémentaires l'un de l'autre, NAT permettant une transition en douceur vers IPv6.

Références

1. RC2131 : Dynamic Host Configuration Protocol
   ftp://ftp.imag.fr/pub/archive/IETF/rfc/rfc2131.txt
   Ralph Droms, Mars 1997
2. RFC1631 : The IP Network Address Translator (NAT)
   ftp://ftp.imag.fr/pub/archive/IETF/rfc/rfc1631.txt
   Kjeld Borch Egevang, Paul Francis, Mai 1994
3. Cisco IOS Software
   http://www.cisco.com/public/sw-center/sw-ios.shtml
4. Comité Réseau des Universités
   http://www.cru.fr/
5. Unité Réseau du CNRS
   http://www.urec.fr/
6. NAT ou Network Address Translation, compte-rendu du groupe de travail
   http://www.urec.fr/nat/
   Claudine Chassagne, Août 1998
7. Internet Assigned Number Authorithy
   http://www.iana.org/
8. RFC1918 : Address Allocation for Private Internets
   ftp://ftp.imag.fr/pub/archive/IETF/rfc/rfc1918.txt
   Yakov Rekhter, Robert G Moskowitz, Daniel Karrenberg, Geert Jan de Groot, Eliot Lear, Février 1996
9. La diffusion multipoint, le Mbone
   http://www.cru.fr/multicast/
   C. Claveleira, Septembre 1999
10. RFC2182 : Selection and Operation of Secondary DNS Servers
    ftp://ftp.imag.fr/pub/archive/IETF/rfc/rfc2182.txt
    Robert Elz, Randy Bush, Scott Bradner, Michael A. Patton, Juillet 1997
11. Brief Overview on Multicast NAT (12.0T)
    ftp://ftpeng.cisco.com/ipmulticast/Multicast-NAT.txt
    Cisco Systems, Septembre 1998
12. Network Address Translation - Protocol Translation (NAT-PT)
    ftp://ftp.imag.fr/pub/archive/IETF/internet-drafts/draft-ietf-ngtrans-natpt-06.txt
    George Tsirtsis, Pyda Srisuresh, Juin 1999
13. NAT à l'échelle d'une université - version détaillée
    http://www.univ-tln.fr/~benza/jres99/
    Didier Benza, Novembre 1999