Protocole IPv6

Marc SCHAEFER

20 janvier 2011

Résumé

Le but de ce cours est de présenter le protocole IPv6 (la version en test du nouveau protocole IP), qui devrait à terme remplacer le protocole IPv4. Les différences principales et améliorations ainsi que l'intégration et la transition seront présentés. Quelques informations sur d'autres protocoles également touchés seront également données.

Table des matières

1 Contexte général

– raréfaction des adresses et sous-réseaux disponibles en particulier pour les pays émergents et les nouveaux besoins
– augmentation de la taille des tables de routage dans les routeurs
– le NAT/PAT détruit la nature "end-to-end" d'Internet et pose des problèmes de performance
– le P2P va dans la direction d'une adresse par équipement, voire par application
– IPv4 a près de 30 ans: préparons-nous au changement, même si ce n'est pas pour demain!
– problème principalement stratégique

1.1 Notes

Pour contourner la limite actuelle des adresses IP (version 4) sur 4 octets, c'est finalement la proposition SIPP qui a été choisie plutôt que les projets TUBA (TCP/UDP with Bigger Address => OSI/CLNP) ou CATNIP (OSI/CLNP, Novell/IPX, IP). Les nouvelles adresses IP seront ainsi codées sur 16 octets, ce qui conduira à un nouvel entête IP simplifié sur 40 octets au lieu des 20 octets actuels!

2 IPv4: Rappels

– adressage 32 bits (environ 4 milliards d'adresses)
  – beaucoup de "gaspillage" (anciennes classe A, zones réservées, taille des tables de routage)
  – peu d'adresses véritablement utilisables
– entête complexe et signification des champs évolutive (peu claire parfois: p.ex. champ TOS)
– le passage du modèle des 5 classes (A à E) au subnetting/norme CIDR a augmenté la flexibilité mais encombre la mémoire des routeurs.
– la fragmentation dans les routeurs baisse la performance
– la transmission de données multimédia n'est pas réalisable hors de réseaux contrôlés (qualité de service)
– la sécurité et la cryptographie sont souvent relégués aux couches supérieures
– on peut tout résoudre, ou presque, en IPv4, mais c'est souvent complexe et lourd
  – qualité de service (DiffServ, RSVP)
  – mobilité IP
  – sécurité (IPsec, VPN, ...)
  – NAT/PAT

3 IPv6: le successeur

adressage
– étendu (128 bits)
– nouveau schéma (préfixes)
entête
– simplification (des champs, de la gestion des options, suppression du checksum, de la fragmentation par les routeurs, ...)
– but: gain de performance
nouveautés
– gestion optimisée des options par les routeurs (chaînage des entêtes)
– autoconfiguration des adresses et nouveau procédé ARP
– qualité de service (pour le multimédia), gestion de la congestion (ECN) et intégration facilitée à MPLS
– authentification et chiffrement (optionnels, mais implémentation obligatoire)
– mobilité IP
– ces améliorations ont été "back-portées" dans IPv4! (IPsec, QoS, NAT/Masquerading, ECN, don't fragment/PMTUDISC) comme options
– protection vie privée (chiffrement, adresses aléatoires, ...)

4 Intégration IPv4/v6

– les couches supérieures (4: TCP, UDP, ICMP; 7: HTTP, FTP, etc) sont pas ou peu touchées
  – tous les protocoles spécifiant des adresses dans leurs entêtes sont touchés (FTP PORT, PASV; SIP; etc)
  – quelques options particulières (p.ex. les jumbogramsnécessitent des changements à TCP et UDP
  – le protocole ICMPv6 a été largement étendu, reprenant des fonctions dévolues au protocole ARP p.ex.

SIT (Simple Internet Transition)

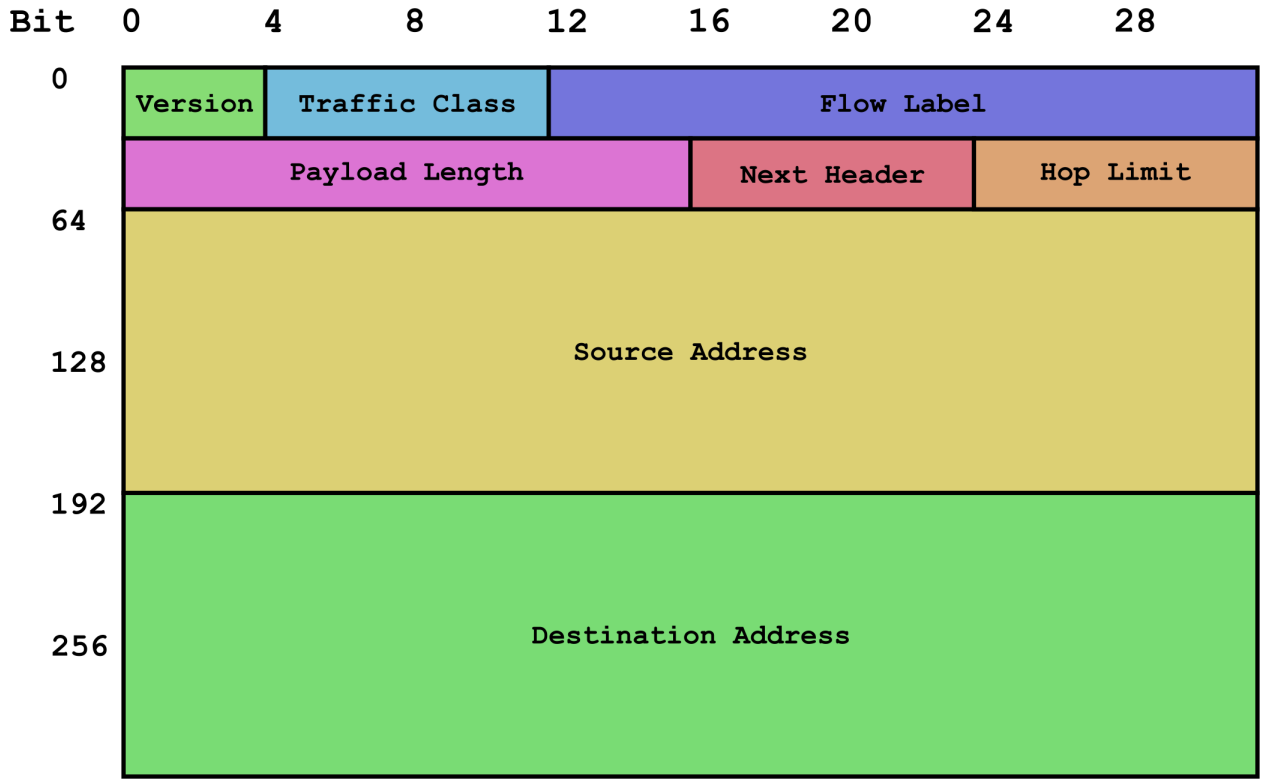
implémentations possibles
– dual-stack
– integrated dual-stack
– encapsulation / tunnelling
– passerelles couches supérieures

5 Entête IPv6

– simplifié et plus performant pour les routeurs
  – plus de checksum
  – options classées via chaînage, ne devant pas forcément être traitées par tous les routeurs
  – identification de classes de trafic
  – pas de fragmentation par les routeurs
– taille fixe: 40 octets
– chaînage d'entêtes (champ Next Header)

5.1 Notes

Entête IPv6



Description des champs
– Version: 6 (la 5 est réservée, la 4 est IPv4)
– Class et Flow Label: utilisé pour QoS
– Payload Length: 0 à 65535 (un mode étendu(jumbogram) existe via un entête d'options, voir RFC-2675), longueur des données (hors entête IPv6!)
– Next Header: type de l'entête qui suit immédiatement (ICMPv6, UDPv6, TCPv6 ou entêtes d'options et de contrôle)

– Hop Limit: décrémenté à chaque routeur, le datagramme est jeté lorsqu'il atteint zéro. Remplace le champ TTL. Utilisé également pour la sécurité: certains messages de contrôle ICMP au sein d'un sous-réseau (domaine de diffusion Ethernet) ne sont valides qu'avec Hop Limit = 255 (notamment ceux remplaçant le ARP).

Commentaires

– le checksum d'entête IP a disparu: on compte sur la couche 2 ainsi que sur la couche 4 (checksums TCP, et checksum UDP maintenant obligatoires).

6 Chaînage d'entête (Next Header)

IPv6 Header; Next Header = TCP	TCP Header	data

cas classique (le plus performant: le routeur peut simplement router le datagramme)

– le champ Next Header contient l'identificateur d'un protocole de couche 4 ou ICMP

ICMPv6	2
TCP	6
UDP	17

IPv6 Header; Next Header = Routing Routing Header; Next Header = Fragment Fragment Header; Next Header = TCP TCP Header data

cas étendu

– le champ Next Header contient l'identificateur d'un entête de contrôle

Hop-by-Hop	options à considérer par tous les équipements; jumbogram	0
Destination Options	à traiter par la première destination	60
Routing	source routing	43
Fragment	gestion de la fragmentation	44
IPsec Authentification Header (AH)	signature du datagramme	51
IPsec Encrypted Security Payload (ESP)	chiffrement du datagramme	50
Destination Options	à traiter par la dernière destination	60

Chaque entête contient potentiellement un entête suivant. L'ordre obligatoire des entêtes est spécifié ci-dessus, ce qui permet d'optimiser le routage (les routeurs s'arrêtent de traiter et transmettent simplement le datagramme).

7 Exemple de chaînage: Fragmentation

– en IPv4, si le MTU de l'interface de sortie est plus petit que la taille du datagramme IP, deux cas possibles:

1. le drapeau Don't fragment est mis: le datagramme est rejeté (erreur ICMP; voir MTUet RFC-1191: Path MTU Discovery / PMTUDISC)
2. sinon, le datagramme est fragmenté au routeur concerné par le problème du MTU, et défragmenté à l'arrivée (noeud final).

– en IPv6
  – les routeurs ne fragmentent jamais, ils se bornent à retourner le message ICMP Packet Too Big avec l'indication du MTU maximum.
  – c'est au système émetteur de reconnaître l'erreur et d'envoyer le datagramme de la bonne taille.
    – l'entête du datagramme est exactement le même, sauf que le champ Next Header de l'entête IP est mis à Fragment, et qu'un entête intercalaire Fragment est ajouté (avec son Next Header = TCP).
    – comme en IPv4 il y a les notions de "fragment offset", "message ID" et "last fragment", champs situés dans l'entête chaîné Fragment.
    – les routeurs transmettent tous les datagrammes en n'examinant pas les entêtes au-delà de l'entête IPv6 (gain de performance)
  – comme auparavant en IPv4, c'est le destinataire final qui défragmente (avec une minuterie, des tampons, etc).
  – comme auparavant, une perte d'un fragment signifie la réémission par la couche supérieure (p.ex. 4/TCP) de tous les fragments.

7.1 Notes

Si l'émetteur ne supporte pas non plus la fragmentation, IPv6 spécifie que le MTU minimal à supporter est de 1280 octets.

8 Adressage

16 bytes (128 bits)

Notation
- hexadécimale: **2001:0db8:85a3:0000:0000:8a2e:0370:7334**
- équivalente à: **2001:0db8:85a3:0:0:8a2e:0370:7334**
- forme simplifiée (élimination des zéros): **2001:0db8:85a3::8a2e:0370:7334**
  ou encore **2001:db8:85a3::8a2e:370:7334**
  - attention à l'ambiguité: 2001::FFD3::57ab (interdit!)
    - spécifier **2001:0:0:ffd3::57ab** ou **2001::ffd3:0:0:57ab**
- intégration IPv4 possible: **2001:0db8:85a3::192.168.42.24** (ou **2001:db8:85a3::c0a8:2a18** ou encore **2001:db8:85a3:0000:0000:c0a8:2a18**

Préfixes
- les adresses sont groupées par préfixes, noté comme une adresse IPv6, puis une spécification similaire au netmask d'IPv4.
- p.ex. **2001::/8** signifie que seuls les 8 premiers bits sont fixes, les 120 autres bits peuvent varier et former autant d'adresses de ce préfixe

9 Types d'adresses (variantes)

- unicast: désigne une interface réseau unique, dans la portée (scope) spécifiée
  - portée: global, link-local, unique local.
- anycast: désigne n'importe quelle interface réseau (UNE) parmi N (groupe); le choix est fait selon des critères de distance via le protocole de routage. On ne peut les distinguer dans leur forme des adresses unicast normales.
- multicast: toutes les interfaces du groupe considéré.
  - commence par le préfixe FF00::/8

Il n'y a pas à proprement parler de *broadcast* spécifique en IPv6!

10 Multicast en IPv6

Les adresses multicast ont également une portée (scope),

8	4	4	112 bits
11111111	0RPT	scop	group ID

- T=1: groupe multicast connu de l'IANA
- P et R: voir RFC-3306 et 3956
- usuellement, seuls les derniers 32 bits de l'adresse sont utilisés pour la spécification du groupe (comme pour les multicast IPv4)
  - groupes réservés: 1 pour tous les noeuds; 2 pour tous les routeurs
- les 4 derniers bits du deuxième octet de l'adresse indiquent la portée (scope)

0x1	interface-local (loopback)
0x2	link-local
0x3	admin-local
0x5	site-local
0x8	organisation-local
0xE	global

- par exemple **FF02::1:FF00:0/104**, avec les 24 derniers bits formés de l'adresse unicast (ou anycast) de l'interface
  - multicast, uniquement pour ce lien (domaine de diffusion Ethernet p.ex.)
  - sollicited-node multicast
  - sert aussi au successeur du protocole ARP: Neighbour Discovery Protocol (NDP)
  - avantage: toutes les interfaces ne sont pas "dérangées"

11 Problèmes des adresses "privées" d'IPv4

Cas classique:
– deux sites sont reliés par VPN
– on se rend compte que les deux sites utilisent en interne la même plage d'adresses privées IPv4 (p.ex. 192.168.1.0/24)
– en IPv4 on résoud cela par des règles complexes de firewall (NAT, PAT, masquerading)
Solution en IPv6
– les adresses privées sont remplacées par des adresses "uniques" locales, avec un préfixe les identifiant (scope, portée)
– l'allocation des plages peut se faire
  – via un registre centralisé pour tout Internet (état 2009: n'existe pas)
  – via des nombres aléatoires qui limitent les risques de collision!

12 Portée des adresses (scope) et préfixes globaux

portée	**préfixe (binaire)**	préfixe (hexadécimal)	**fraction**
non assigné	0000 0000	::0/8	1/256
réservé	0000 001		1/128
global unicast	001	2000::/3	1/8
link-local unicast	1111 1110 10	fe80::/10	1/1024
reservé (anc. site-local unicast)	1111 1110 11	fec0::/10	1/1024
local IPv6 address	1111 110	fc00::/7	
private admin	1111 1101	fd00::/8	
multicast	1111 1111	ff00::/8	1/256

(voir http://www.iana.org/assignments/ipv6-address-space)

13 Routage par préfixe

Une adresse de type "global unicast" se construit comme suit:

global routing prefix	subnet ID	interface ID

les deux premiers champs étant de longueur à définir; seul le global routing prefix est utilisé pour les protocoles de routages externes (inter-AS: EGP); le subnet ID est utilisé dans les protocoles de routages internes (intra-AS: RIP, OSPF).

L'interface ID quand à elle est définie de plusieurs manières (voir slide suivant).

14 Attribution des adresses

– comme en IPv4
  – attribution manuelle
– nouveau
  – chaque interface a automatiquement une adresse de portée (scope) link-local, unique pour l'interface considérée
    – préfixe: fe80::/10
    – 64 bits inférieurs: valeur EUI-64 (construite par l'adresse MAC): interface ID
  – attribution automatique depuis un préfixe communiqué par un routeur (router advertisement)
    – la nouvelle adresse est le préfixe + adresse MAC (RFC-2462 EUI-64; ou valeur aléatoire à des fins de protection de la vie privée, RFC-4941)
– similaire
  – DHCPv6
    – peut ne s'occuper que de la partie "non adresse" (serveurs DNS, indication serveur TFTP, etc)

14.1 Notes

Norme EUI-64: faire correspondre une adresse MAC (couche 2, p.ex. Ethernet à 6 x 8 = 48 bits) aux 64 bits de poids faibles d'une adresse IPv6

15 Adresses prédéfinies

unicast

0:0:0:0:0:0:0:0 ou ::0	adresse non définie (p.ex pour bind(2) à toutes les adresses, comme INADDR_ANY en IPv4)
0:0:0:0:0:0:0:1 ou ::1	adresse loopback

anycast

subnet prefix (N bits)	(128 - N) bits à zéro	sera traité par un des routeurs du subnet

multicast

ff00::/128 à ff0f::/128	réservées
ff01::1/128	all-nodes interface-local (loopback)
ff02::1/128	all-nodes link-local (broadcast Ethernet)
ff0X::2/128	all-routers: avec X 1, 2 ou 5 (interface-local, link-local ou site-local)
ff02::1:FFXX:XXXX	sollicited-node address: sert à la découverte d'une adresse MAC pour une adresse donnée, sans besoin de broadcast: seules les machines ayant des adresses ressemblant (24 bits inférieurs) à l'adresse cherchée vont recevoir le message

16 Résumé: les adresses d'un noeud

un noeud doit obligatoirement disposer de:
– une adresse de portée link-local pour chacune de ses interfaces
– des adresses unicast ou anycast configurées automatiquement ou manuellement
– l'adresse *loopback*
– s'abonner aux groupes multicast
  – all-nodes
  – sollicited-node (pour toutes les adresses de toutes les interfaces)
  – d'autres groupes desquels ce noeud fait partie

un routeur doit en plus s'abonner aux groupes
– anycast de type subnet router, pour toutes les interfaces pour lesquelles il agit comme routeur
– multicast all-routers

17 ICMPv6

Le rôle d'ICMP en IPv6 a été étendu

– attribution et validations d'adresses (Neighbour Discovery Protocol), remplaçant partiellement le DHCP

– remplacement de ARP

18 DHCPv6

séparation entre "attribution d'adresse proprement dite" (implémentable en DHCPv6 ou non) et "autres configurations" (DNS, NTP, TFTP, hostname, etc)

19 Autres différences

– NAT/PAT
  – le besoin de NAT/PAT est devenu inutile, sauf éventuellement dans des cas très particuliers
– fragmentation
  – déjà en IPv4, le Path MTU Discovery et le drapeau "don't fragment" rendait rare l'utilisation de la fragmentation

  - IPv6 la supprime. En cas de nécessité de taille de datagramme plus basse, un datagramme ICMPv6 Packet Too Big est envoyé, avec l'indication du MTU concerné
- options
  - le champ Next Header permet de chaîner des options dans un ordre défini
  - cette implémentation permet un traitement efficace par les routeurs
- Explicit Congestion Notification
  - permet à un routeur de marquer les datagrammes d'un flux en congestion avant de devoir jeter et donc permet d'informer l'émetteur via le récepteur.
  - option de TCP IPv4 (ECN); implémenté de base en IPv6
- DNS: champ AAAA

20 Transition de IPv4 à IPv6

SIT
- noeuds duaux compatibles IPv4 et IPv6
- réseau intermédiaire IPv4, IPv6, ou dual
- adresses IPv6 == préfixe de 96 bits à zéro, suivis des 32 bits de l'adresse IPv4 routable

conversion facilitée

21 Mobilité IP

Problème en IPv4
- routage triangulaire: le trafic émetteur vers mobile circule via le Home agent, mais le retour est direct, ce qui peut poser des problèmes de performance et d'anti-spoofing

Solution en IPv6
- les entêtes chaînés permettent de gérer le problème
- à l'avenir: RFC-3963

22 Qualité de service et intégration MPLS

(à compléter)

23 Freins à l'acceptation d'IPv6

- deux réseaux en parallèle à maintenir, peut-être des années durant!
- compétences à acquérir (plus de 40 RFCs!)
- coûts (formation, matériel)

24 Sécurité

- Microsoft Windows Vista et 7 implémentent par défaut http://en.wikipedia.org/wiki/Teredo_tunneling
- Impact
  - toute machine de ce type même derrière un firewall (ne filtrant pas l'UDP) est accessible directement par IPv6
- Voir aussi http://www.gont.com.ar/talks/lacnog2010/fgont-lacnog2010-ipv6-security.pdf

25 Références

- http://www.ipv6.org/
- IPv6 Essentials, Sivia HAGEN, ISBN 978-0-596-10058-2, O'Reilly, 2nd edition, 2006.
- LaboIPv6
- RFCs

  - RFC-2460 - Internet Protocol, Version 6 (IPv6) Specification
  - RFC-4291 - IP Version 6 Addressing Architecture

- utilitaire **ipv6calc** (calcul et validation d'adresses)
- http://www.shorewall.net/Linuxfest-2009.pdf
- http://www.blogg.ch/uploads/IPv6-deployment-for-the-IPv4-clueful-heise-ipv6-conference.pdf
- http://www.blogg.ch/uploads/Native-IPv6-via-xdsl-how-to-tweak-your-LNS.pdf