RFC 7421 : Analysis of the 64-bit Boundary in IPv6 Addressing

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 23 janvier 2015

Date de publication du RFC : Janvier 2015

https://www.bortzmeyer.org/7421.html

C'est vrai, ça, pourquoi 64 ? Les adresses du protocole IPv6 ont deux parties, une qui identifie le réseau et l'autre qui identifie la machine. Cette dernière est souvent considérée comme faisant forcément 64 bits. Est-ce vraiment une taille impérative ? Et, si oui, pourquoi ? (Si vous êtes pressé, les réponses du RFC sont « non, mais quand même oui » et « essentiellement en raison de la base installée ».)

Pendant la phase de conception d'IPv6, il y avait eu toute une discussion à l'IETF sur la taille idéale des adresses. La plupart des propositions initiales suggéraient 64 en tout, ce qui était déjà beaucoup pour les processeurs de l'époque (de même que, quand IPv4 avait été conçu, peu de processeurs pouvaient traiter 32 bits comme une donnée de base). Mais c'est finalement 128 bits qui a été choisi. 64 bits permettaient déjà d'adresser plus de machines qu'on ne pouvait en rêver. Mais 128 bits permettait plus de souplesse, et autorisait le développement de mécanismes d'adresse rigolos, par exemple d'utiliser des clés crytographiques comme adresses (RFC 7343 [1]).

Une fois ce choix fait, le RFC sur l'adressage IPv6, le RFC 4291, introduisait la notion d'**identificateur d'interface** ("*interface identifier*" ou IID). L'adresse comprend alors deux parties, n bits pour identifier le réseau et 128-n bits pour l'identificateur d'interface. Comme précisé par le RFC 7136, ces identificateurs d'interface n'ont pas de signification en dehors de la machine qui les a alloués et doivent donc être traités comme des valeurs opaques.

Le routage se fait sur la base du préfixe réseau, dont la longueur peut être quelconque, jusqu'à 128 bits. Tous les protocoles de routage gèrent des longueurs de préfixe variables, sans supposition a priori.

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc7343.txt

Bon, donc, le routage n'impose pas de préfixe de longueur 64, et se moque de la longueur du préfixe. Mais l'autoconfiguration d'adresse sans état (SLAAC), du RFC 4862, elle, considère la longueur 64 comme spéciale ? Non plus. Ce chiffre 64 n'apparait pas dans le RFC 4862 car la longueur de l'identificateur d'interface dépend du type de réseau sous-jacent et est donc spécifiée dans un autre RFC, spécifique à chaque technologie. Pour Ethernet, c'est le RFC 2464, et la longueur spécifiée est bien 64. Donc, pour Ethernet, si on veut de l'auto-configuration sans état, on ne peut pas avoir de réseau ayant un préfixe plus long que 64 bits. DHCP (RFC 8415) suit la même règle et n'a donc pas de chiffre « 64 » magique.

Si l'identifiant d'interface à 64 bits est aujourd'hui largement répandu, c'est en partie à cause de la part de marché d'Ethernet. Il était plutôt prévu de le mettre à 48 bits au début avant que les identifiants EUI-64 ne remplacent les EUI-48. Résultat, le RFC 4291 a sérieusement réduit la souplesse d'IPv6 en parlant d'identificateurs d'interface de 64 bits et en laissant entendre qu'ils étaient forcément dérivés d'une adresse MAC (le RFC 7136 a partiellement corrigé cette erreur mais trop tard, beaucoup de gens sont aujourd'hui persuadés que les 64 derniers bits d'une adresse IPv6 sont forcément une adresse Ethernet). Quant à l'autre choix, de fixer les identificateurs d'interface à 64 bits, il n'a pas empêché le RFC 6164 de recommander 1 seul bit pour cet identificateur, dans le cas de liaisons point-à-point.

Malheureusement, un certain nombre de logiciels et de matériels ont peut-être été bâtis en supposant un identificateur d'interface de 64 bits et, hélas, une autre longueur, bien que techniquement correcte, posera des problèmes si on s'approche de ces logiciels et matériels. Pourquoi diable le RFC 4291 a-t-il sanctifié cette valeur de 64 bits ? Et peut-on encore la changer ?

D'abord (section 2 du RFC), les avantages qu'il y a à avoir une frontière fixe entre le réseau et l'identificateur d'interface. Pour l'auto-configuration sans état (SLAAC pour "*StateLess Address AutoConfiguration*"), il faut que la longueur du préfixe, pour un type de réseau physique donnée, soit identique pour toutes les machines du réseau local, et connue à l'avance, donc fixe. Bien sûr, elle n'est pas forcément de 64 bits. Mais le poids d'Ethernet, où cette longueur est de 64 bits, et le désir de simplifier les choses (cf. RFC 5505) poussent à l'uniformisation.

Le RFC 4291 (section 2.5.1), on l'a vu, cite déjà cette limite fixe de 64 bits. Le préfixe du réseau local a donc forcément 64 bits de long, quelle que soit la taille du préfixe qui nous a été alloué (au moins un /56, si on suit le RFC 6177). Cela simplifie la conception des réseaux (préfixe de longueur identique sur tous les liens réseaux d'un campus, par exemple, un très gros avantage par rapport à IPv4, où le manque d'adresses oblige à calculer au plus juste chaque longueur de préfixe), la gestion des préfixes (pas besoin de faire un calcul à chaque allocation, un nouveau lien, paf, je lui donne un /64), la configuration des routeurs et la documentation du réseau de l'organisation sont plus simples.

Garantir une certaine taille pour l'identificateur d'interface permet de choisir le mécanisme qu'on veut pour leur allocation. Des préfixes longs risqueraient de réduire la taille de l'espace des identificateurs d'interface, au point de contraindre l'administrateur réseaux dans sa politique d'allocation.

Les sections suivantes du RFC examinent les arguments contre cette taille fixe de 64 bits mais je vous révèle la fin du RFC tout de suite : l'IETF estime que les avantages d'une longueur d'identificateur d'interface plus ou moins fixée à 64 bits l'emportent sur les inconvénients.

Bon, maintenant, pour la culture de mes lecteurs, quels auraient été les raisons de ne **pas** imposer une taille spéciale aux identificateurs d'interface ? D'abord, le /64 partout peut mener à du gaspillage. J'ai un réseau avec une demi-douzaine de serveurs statiquement configurés, ou bien utilisant DHCP (donc pas de SLAAC), même si chacun d'eux a plusieurs adresses IP, pourquoi ne pourrais-je pas utiliser un /96 ou similaire, pour économiser des adresses ? Bien sûr, l'espace d'adressage d'IPv6 est très grand mais

un utilisateur donné peut n'avoir reçu qu'un seul /64, et comment fait-il alors s'il veut mettre plusieurs réseaux physiques derrière ?

Le RFC 7368, portant sur les réseaux à la maison, écarte la solution des préfixes plus longs (cf. sa section 3.4.1) et préfère mettre la pression sur les FAI pour qu'ils distribuent plus d'un préfixe /64. Il n'y a en effet aucune raison de faire des économies : l'espace IPv6 actuellement utilisé (le `2000::/3`, qui ne représente que 13 % de l'espace d'adressage possible) permet 35 billions de /48 donc en donner un à chaque utilisateur humain ne pose aucun problème. Même avec un très mauvais « ratio HD », 0,89, on pourrait encore allouer un trillion de préfixes /48 (voir le RFC 4692, sur ces calculs et sur la notion de ratio HD).

Un autre argument a parfois été présenté pour justifier des préfixes plus longs que 64 bits : cela permet d'avoir davantage de réseaux locaux lorsqu'on a reçu une allocation trop petite. Ceci dit, avec un /48, on a 65 536 réseaux, ce qui est déjà énorme. Il a aussi été dit que des préfixes plus longs que 64 bits permettraient, sinon davantage de réseaux, en tout cas une meilleure agrégation des préfixes mais cet argument ne semble pas tenir non plus : même dans le pire cas (routage complètement à plat, zéro agrégation), router des milliers de préfixes ne pose aucun problème.

Une autre raison pour souhaiter des préfixes plus spécifiques que 64 bits vient des exigences de traçabilité. Moins de bits pour l'identifiant d'interface voudrait dire moins de possibilités qu'une machine prenne l'adresse qu'elle veut, ce qui faciliterait la tâche de l'administrateur réseaux. Comme le précédent, cet argument semble très faible : il existe d'autres méthodes pour surveiller ses machines (par exemple ndpmon <`http://ndpmon.sourceforge.net/`>).

Beaucoup plus sérieux est le risque d'attaque sur le protocole NDP par épuisement du cache (voir les supports d'exposé de Jeff Wheeler <`http://inconcepts.biz/~jsw/IPv6_NDP_Exhaustion.pdf`>). L'idée de l'attaquant est d'envoyer des paquets à des tas de machines non existantes. Le routeur menant au lien de ces machines va devoir faire une résolution d'adresse IP en adresse MAC, avec NDP et, la machine n'existant pas et ne répondant donc pas, le routeur va devoir mettre une demande en attente, dans une mémoire qui a une taille fixe et qui peut donc être vite remplie (le RFC 3756 détaille cette attaque). Avec un /120 (ou bien avec les préfixes typiques d'IPv4, pour qui cette attaque est également possible), l'attaquant ne pourra occuper que 256 entrées dans la mémoire. Avec un /64, il aura beau jeu pour la remplir complètement (même un /96 serait largement suffisant pour l'attaquant). C'est d'ailleurs un argument du RFC 6164 (section 5.2) pour justifier sa recommandation d'un /127.

Notre nouveau RFC estime que ce risque, quoique réel, n'est pas suffisant pour justifier des préfixes ultra-longs (genre /120) et qu'il vaut mieux déployer les recommandations du RFC 6583.

La section 3 parlait des arguments en faveur d'un préfixe plus long que 64 bits. Mais il y a aussi un autre débat, qui est celui sur les préfixes de longueur variable. Un argument en faveur de « tout le monde en /64 » était que cela fournissait une longueur de préfixe constante. Que se passe-t-il si elle ne l'est pas ? La section 4 du RFC se penche là-dessus. D'abord, par rapport aux normes. Malheureusement, la situation est confuse. Si les RFC 4291, RFC 6177, RFC 5453, RFC 6741 et RFC 7084 font tous référence à une longueur magique de 64 bits, ils ne sont pas forcément très clairs et précis sur le statut de cette référence : simple exemple, constatation de l'existant ou réelle normalisation ? Le RFC 5942 dit au contraire que les mises en œuvre d'IPv6 ne doivent pas supposer une longueur de préfixe fixe.

Les RFC décrivant le cas particulier d'une technologie de couche 2 donnée sont en général plus clairs, et beaucoup imposent un identificateur d'interface de 64 bits (et donc un préfixe de 64 bits) pour l'autoconfiguration sans état. C'est le cas des RFC 2464 (Ethernet, déjà cité), RFC 2467 (FDDI), RFC 4338 (Fibre Channel), RFC 5072 (PPP), etc.

D'autres RFC semblent (mais c'est souvent vague) supposer qu'un identifiant d'interface fait forcément 64 bits. C'est le cas du RFC 4862 pour les adresses locales au lien, du RFC 4429 pour la détection d'adresses dupliquées, du RFC 5969 sur 6rd, du RFC 6437 au sujet du "*flow label*", etc. Dans certains cas, la dépendance vis-à-vis de la longueur du préfixe est plus nette, comme une technique pour étendre un préfixe IPv6 reçu en 3G sur une interface WiFi (RFC 7278), dans les CGA du RFC 3972 ou dans les adresses protégeant la vie privée du RFC 8981.

Et si on est courageux, et qu'on essaie quand même de mettre des identificateurs d'interface dont la longueur est différente de 64 ? Que risque-t-on ? D'abord, certains routeurs peuvent avoir mal lu les spécifications et penser que des identificateurs d'interface entre 65 (RFC 7136) et 126 (RFC 6164) sont invalides, refusant de configurer ainsi une interface ou, pire, refusant ensuite de transmettre des paquets. (Je n'ai pas connaissance d'un étude systématique de routeurs à ce sujet, ni de récit détaillé d'un problème avec un routeur.) CGA, on l'a vu, ne marcherait pas du tout. On peut changer sa spécification mais attention, diminuer la taille de l'identificateur d'interface diminuerait la sécurité de CGA (même chose pour les techniques qui visent à protéger la vie privée par des identificateurs d'interface choisis aléatoirement, cf. section 4.5). Par contre, NAT64 (RFC 6146) marcherait sans doute, jusqu'à 96 bits de préfixe réseau (car il faut garder au moins 32 bits pour l'adresse IPv4 de destination). En revanche, NPT ("*Network Prefix Translation*", RFC 6296) est lié aux 64 bits. Même chose pour ILNP (RFC 6741).

Le mécanisme DAD ("*Duplicate Address Detection*", RFC 4862, section 5.4) pourrait avoir des problèmes si on réduisait trop la taille des identificateurs d'interface. Par exemple, avec un /120, et les adresses du RFC 7217, il n'y aurait que 256 identificateurs d'interface possible et donc des risques de collision élevés. Enfin, les adresses locales au lien, bien que prises dans un espace très large (`fe80::/10`) sont en fait quasiment toujours formées pour un préfixe de longueur 64 (et l'espace réel est donc `fe80::/64`) et il n'y a pas de moyen simple de le changer (ce n'est que rarement une option de configuration de la machine).

La section 4.3 contient les résultats d'essais systématiques d'essais de préfixes différents de /64, essais faits avec plusieurs systèmes d'exploitation. Avec des préfixes plus longs ou plus courts que 64 bits, et un routeur qui annonce, via les RA ("*Router Advertisement*"), un tel préfixe dans le champ "*Prefix Information Option*", tout marche bien sur tous les Unix et sur Windows (l'essai a également été fait sans le bit L dans l'option, bit qui indique que le préfixe pour être utilisé pour déterminer si le préfixe est sur le lien local et, dans ce cas, les machines ignorent à juste titre le préfixe). Par contre, l'option "*Route Information Option*" du RFC 4191 marche nettement moins bien sur les Unix, mais c'est le cas même avec une longueur de 64 bits.

D'autre part, les informations recueillies par les participants à l'IETF sur divers réseaux indiquent que les routeurs routent bien sur des préfixes de longueur quelconque (la recherche d'une route du préfixe le plus long est un algorithme de base d'IP, il est plus ancien qu'IPv6). Et, dans les expériences pratiques, DHCP ne semble pas avoir de problème non plus (un déploiement réel utilise des /120...)

Certains routeurs ont des problèmes de performance car ils traitent à part les préfixes plus longs que 64 bits. Cela fonctionne mais pas aussi vite.

Au moins un équipement réseau à une TCAM limitée à 144 bits ce qui fait que les ACL ne peuvent pas être définies pour un préfixe quelconque, puisqu'elles permettent également de mettre deux ports de 16 bits chacun. Avec un tel équipement, on ne peut pas utiliser d'ACL avec un préfixe plus long que 112 bits.

Bref, la situation des mises en œuvre d'IPv6 ne semble pas trop mauvaise. Contrairement à ce que l'on aurait pu craindre, peu ou pas de programmeurs ont, en lisant trop vite les spécifications, considéré

qu'un identificateur d'interface était toujours de 64 bits et, donc, peu de programmes ont cette taille « câblée en dur ». Évidemment, on ne peut pas être sûr tant qu'on n'a pas testé toutes les implémentations, une tâche impossible.

Il n'y a pas que le code IPv6 proprement dit : il y a aussi les outils auxiliaires comme les IPAM. Ont-ils également une taille de préfixe magique fixée dans leur code ? Apparemment, cela n'a pas été testé.

Il y a aussi les craintes liées aux humains : si on décidait de mettre des identificateurs d'interface d'une longueur différente de 64 bits, ne risque-t-on pas d'avoir à reprendre la formation de certains, qui ont mal compris leur cours IPv6 et croient que 64 bits a une signification spéciale ?

Un éventuel changement de la longueur des identificateurs d'interface a aussi des conséquences pour la sécurité. D'abord, la vie privée (section 4.5) : si on tire au sort un identificateur d'interface, pour le rendre difficile à deviner par un observateur indiscret, l'identificateur ne doit évidemment pas être trop petit, sans cela l'observateur pourrait essayer la force brute pour le deviner. Il est difficile de donner un chiffre précis mais notre RFC estime qu'un préfixe de plus de 80 bits, laissant moins de 28 bits pour l'identificateur d'interface, serait trop prévisible.

La longueur de l'identificateur d'interface a également des conséquences pour le balayage (RFC 7707) : il est beaucoup plus difficile de balayer systématiquement un réseau IPv6 qu'un réseau IPv4, en raison du nombre d'adresses possibles. Si on a un long préfixe, le balayage devient réaliste : un /120 IPv6 prendrait autant de temps à balayer qu'un /24 IPv4, c'est-à-dire très peu de temps. (Le problème est proche de celui des CGA, et de celui de la vie privée, déjà cité : beaucoup de techniques de sécurité dépendent de la taille de l'espace possible, en la réduisant, on facilite les attaques.)

En conclusion, si certains points d'IPv6 ne devraient absolument pas dépendre de la longueur du préfixe, qui est un simple paramètre (le routage, par exemple), d'autres sont bien plus liés au nombre magique de 64. Sans compter le risque qu'une partie de la base installée (et pas seulement logicielle, aussi les humains) ait attribué à ce nombre 64 encore plus d'importance que ce que les normes IPv6 prévoient. Le RFC décide donc d'entériner cet usage et de recommander qu'on ne s'éloigne pas de 64 sans de très bonnes raisons.