

RFC 5963 : IPv6 Deployment in Internet Exchange Points (IXPs)

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 31 août 2010

Date de publication du RFC : Août 2010

<https://www.bortzmeyer.org/5963.html>

Le fonctionnement de l'Internet aujourd'hui repose largement sur des points d'échange où les différents opérateurs se connectent pour échanger du trafic IP. Le point d'échange typique fournit un réseau Ethernet où chacun connecte son routeur, et alloue des adresses IP pour configurer ces dits routeurs, qui vont ensuite établir des liens BGP entre eux. La principale conclusion de ce nouveau RFC est que la très grande majorité des points d'échange fournissant un service de couche 2, le fait d'utiliser IPv6 au lieu d'IPv4 ne change pas grand'chose à la gestion du point d'échange.

La section 1 résume l'état actuel du monde des points d'échange. Presque toujours, le service rendu est une connexion de niveau 2, quasiment uniquement en Ethernet. Le principal service de niveau 3 rendu par les gérants du point d'échange est l'attribution d'adresses IP aux routeurs des opérateurs. Curieusement, ce service n'est pas mentionné dès la section 1, qui cite à la place des fonctions moins vitales comme le serveur de routes ou bien les statistiques <<http://www.ams-ix.net/statistics/>> (globales ou bien par protocole <<http://www.ams-ix.net/sflow-stats/ipv6/>>).

La section 2 du RFC passe ensuite aux questions liées à l'interface entre la couche de liaison et la couche réseau. IPv6 sur Ethernet doit se faire selon le RFC 2464¹. Le commutateur Ethernet lui-même, travaillant en couche 2, n'a rien de spécial à faire. On peut se poser la question de séparer les trafics v4 et v6. Cela peut se mettre en œuvre avec deux ports physiques sur le commutateur ou bien avec deux VLAN séparés. Mais cette séparation n'est pas indispensable. (Le faire avec des ports séparés consomme des ressources matérielles sur le routeur et le faire avec des VLAN impose aux routeurs de gérer 802.1Q.) Elle complique la configuration mais peut simplifier certaines fonctions comme les statistiques.

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc2464.txt>

La section 3, plutôt descriptive que normative, décrit le mécanisme d'adressage à un point d'échange. Chaque RIR a sa politique <<http://www.nro.net/documents/comp-pol.html#3-4-2>> pour l'allocation d'adresses IPv6 aux points d'échange. Ce sont typiquement des préfixes de longueur /48 qui sont alloués. Ces adresses ont besoin d'être résolubles en nom par le DNS et de pouvoir être trouvées dans la base des RIR via whois. Donc, un point d'échange n'utilise pas les ULA du RFC 4193. (Voyez par exemple les adresses IP allouées sur FranceIX <<http://france-ix.fr/wp-content/uploads/2009/11/Connected-partners-2010-08-16.pdf>>.)

Par rapport à un réseau local IPv6 typique, il faut aussi noter que l'autoconfiguration des adresses par l'envoi de RA ("*Router Advertisement*") n'est typiquement pas utilisée. La configuration des routeurs est faite manuellement puisque, de toute façon, la configuration de BGP dépend de l'adresse. Puisqu'on n'utilise pas l'autoconfiguration, que mettre dans les 64 bits les plus à droite ? En IPv4, les routeurs à un point d'échange sont en général numérotés séquentiellement mais l'espace d'adressage bien plus grand d'IPv6 permet des plans d'adressage plus informatifs. Il existe plusieurs mécanismes acceptables :

- Encoder le numéro d'AS en décimal dans les 64 bits. Ainsi, si le point d'échange utilise le préfixe `2001:db8:f00f::/64` et qu'un opérateur connecté a le numéro d'AS 64496, son adresse IP sera `2001:db8:f00f::6:4496:1` (le 1 tout à fait à droite vient de la réservation des 16 derniers bits pour mettre plusieurs routeurs par opérateur connecté, si nécessaire).
- Certains préfèrent l'hexadécimal, moins lisible mais plus compact, donc ici `2001:db8:f00f::fbf0:1`.
- Une autre méthode est de dériver l'adresse IPv6 de la v4. donc un routeur qui aurait `192.0.2.123` en v4 recevra `2001:db8:f00f::123` en v6 (ce n'est qu'un exemple, le RFC en cite un autre, qui permet des points d'échange de plus de 256 routeurs, contrairement à mon choix de ne garder que le dernier octet).
- etc (d'autres méthodes sont possibles).

Ces adresses IP du point d'échange doivent-elles être routées globalement ? Le débat a toujours fait rage pour IPv4 et n'est pas différent ici. Les adresses routées globalement facilitent la configuration et le débogage mais peuvent rendre le point d'échange plus vulnérable à certaines attaques. Si on ne route pas globalement ces adresses, les participants au point d'échange peuvent toujours le faire eux-même dans leur réseau, en prenant soin d'utiliser des méthodes comme la communauté `no-export` de BGP, pour éviter que les annonces des routes vers le point d'échange ne se propagent trop loin. Enfin, le point d'échange a aussi des services publics (pages Web, serveur DNS, serveurs NTP, etc) et ceux-ci doivent évidemment être installés sur des adresses routables, que ce soit dans le même préfixe que celles des routeurs des participants ou bien dans un préfixe différent.

Une des particularités d'un point d'échange est que les routeurs qui y sont présents appartiennent à des organisations différentes, souvent concurrentes. Le réseau Ethernet partagé n'est donc pas forcément peuplé que par des gentils paquets, on y trouve un peu de tout, des annonces OSPF aux serveurs DHCP illicites... La section 4 mentionne ce problème et ses conséquences pour la sécurité et note que le gérant du point d'échange peut, par exemple, limiter l'utilisation de la diffusion (qui sont transmis à tous) aux paquets de découverte des voisins (RFC 4861. En tout cas, bloquer les paquets "*Router Advertisement*" (qui ne devraient jamais apparaître sur le réseau du point d'échange) est conseillé.

Tiens, puisqu'on a parlé du DNS, la section 5 lui est consacrée. Elle recommande que les adresses IP du point d'échange aient un enregistrement « inverse » (enregistrement PTR) dans le DNS, pour faire de plus jolis traceroute et, de manière générale, pour faciliter le débogage.

Un autre service très populaire sur les points d'échange est le serveur de routes, discuté en section 6. Il sert parfois à échanger réellement des routes entre pairs, et parfois simplement de "*looking glass*". Voir par exemple le "*looking glass du DE-CIX*" <http://www.de-cix.net/content/network/looking_glass.html>. Notre RFC recommande qu'il soit accessible en IPv6 mais surtout qu'il gère les extensions à BGP des RFC 2545 et RFC 4760, qui permettent de gérer plusieurs familles d'adresses IP, donc de gérer des routes IPv6. Une autre recommandation est que les routes IPv6 soient échangées sur des sessions IPv6 (ce qui n'est pas obligatoire avec BGP), pour améliorer la cohérence de l'information de

routage (si un pair BGP reçoit des informations sur IPv6, il peut être sûr que le pair qui lui annonce des routes IPv6 est lui-même effectivement joignable par ce protocole).

Enfin, la section 7 traite les cas « divers ». Par exemple, un des points peu spectaculaire, mais souvent critique, d'une transition vers IPv6 est l'adaptation des systèmes d'avitaillement (la base de données qui stocke les adresses des participants au point d'échange...) : ces systèmes doivent eux aussi être migrés de manière à pouvoir gérer des adresses IPv6.

La section 8 couvre le cas des politiques d'accès au point d'échange. Les règles d'utilisation doivent bien préciser si le contrat concerne seulement IPv4, seulement IPv6 ou bien les deux. Je me souviens, il y a quelques années (les choses ont peut-être changé depuis) que le contrat avec le Sfinx ne couvrait qu'IPv4 (alors que l'organisme qui gère le Sfinx se vantait de son rôle pionnier en IPv6) et qu'il fallait une paperasserie longue et compliquée pour pouvoir faire de l'IPv6. Notez que notre RFC ne formule pas de recommandation précise sur la politique idéale. Pour moi, le contrat devrait couvrir **IP**, quelle que soit sa version, car il n'existe aucune justification opérationnelle pour traiter IPv6 comme un « plus », imposant un contrat nouveau.

Pour un bon survol des points d'échange IPv6, voir <<https://prefix.pch.net/applications/ixpdir/summary/ipv6/>>.