

# RFC 6227 : Design Goals for Scalable Internet Routing

Stéphane Bortzmeyer  
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 7 mai 2011

Date de publication du RFC : Mai 2011

<https://www.bortzmeyer.org/6227.html>

---

Le problème fondamental du routage dans l'Internet est connu depuis longtemps : le système actuel a du mal à fournir certains services ("*multi-homing*", mobilité, ingénierie de trafic, etc) d'une manière qui passe à l'échelle, c'est-à-dire qui puisse grandir au rythme de la croissance de l'Internet. Un groupe de travail IRTF, le "*Routing Research Group*" <<http://www.irtf.org/rrg>> travaille donc sur ce sujet et ce RFC 6227<sup>1</sup> est le cahier des charges produit.

Il est très court : il est plus facile de spécifier le problème que de décrire les solutions. Le problème (les services difficiles à fournir tout en passant à l'échelle) a déjà été présenté par le RFC 4984. Reste à concevoir les solutions et, pour cela, à décrire les exigences auxquelles doit se plier la future architecture, et leurs priorités respectives. Logiquement, le cahier des charges s'écrit au début et est publié longtemps avant que le travail ne se termine mais, ici, cela n'a pas été le cas. Les conclusions du travail sur la nouvelle architecture sont déjà sorties (RFC 6115). Simplement, le cahier des charges avait bien été écrit au début mais il a été modifié et raffiné au fur et à mesure de l'avancement du travail et d'une meilleure compréhension du problème, pour n'être publié officiellement que maintenant.

Chaque exigence listée dans ce RFC a une priorité : requise, fortement souhaitée ou souhaitée, et les exigences avec leurs priorités sont résumées dans un tableau dans la section 3.11. D'autre part, les principes traditionnels de l'Internet, tels qu'explicités dans le RFC 1958 s'appliquent toujours (notamment le principe de bout-en-bout qui indique que les deux machines terminales qui communiquent doivent faire l'essentiel du travail, puisque ce sont elles qui doivent décider, par exemple des priorités de trafic).

La « liste au Père Noël » est en section 3. D'abord, compte-tenu de la croissance continue de la table de routage globale (cf. les rapports sur la croissance de la table BGP <<http://bgp.potaroo.net>>), il est **fortement souhaité** (section 3.1) que la nouvelle architecture permette de séparer la croissance

---

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc6227.txt>

de l'Internet et celle de la table de routage. Actuellement, un nouvel utilisateur qui veut être « multi-homé » proprement ajoute une entrée BGP dans la table de routage globale, qui risque donc de devenir proportionnelle au nombre de tels utilisateurs. Ce n'est clairement pas souhaitable. Notez que le RFC est ambitieux : on ne parle pas de faire croître la table de routage moins vite que le nombre d'utilisateurs, mais de ne pas la faire croître du tout lorsque ce nombre d'utilisateurs augmente.

Autre exigence, concernant l'ingénierie de trafic, c'est-à-dire la possibilité d'envoyer le trafic sur tel ou tel chemin selon des critères qui ne sont pas ceux du système de routage classique, par exemple pour forcer le passage par un lien moins coûteux. Une méthode courante aujourd'hui est d'injecter dans la table de routage globale des préfixes d'adresses IP plus spécifiques. Cela augmente la taille de ladite table. Il est donc **fortement souhaité** qu'une meilleure solution soit trouvée (section 3.2).

Même chose pour le "*multi-homing*", déjà cité : il est **fortement souhaité** qu'il puisse aussi passer à l'échelle (section 3.3).

Une bonne partie des problèmes de l'architecture actuelle de l'Internet vient de la confusion de deux rôles dans l'adresse IP : localisateur (indique un point d'attachement dans l'Internet) et identificateur (nomme une machine ou une interface). Il est depuis longtemps reconnu (cela se trouvait déjà dans l'IEN1 <<http://www.postel.org/ien/pdf/ien001.pdf>>) que cette confusion est dommageable <<https://www.bortzmeyer.org/separation-identificateur-localisateur.html>>. Il est donc **souhaité** qu'on sépare ces deux rôles (section 3.4).

Un certain nombre de gens (note au passage : je n'en fais pas partie) considèrent qu'il est essentiel qu'une machine terminale, voire un réseau entier, puisse se déplacer dans l'Internet, changer de fournisseur, etc, sans remettre en cause son adresse IP et les sessions existantes. Par exemple, une connexion SSH devrait rester intacte si mon "*smartphone*" passe du Wi-Fi chez moi avec Free, au 3G chez Bouygues, voire si je passe de Free à la maison à Paris à un accès Verizon en Californie. Comme ces deux scénarios entraînent un changement de l'adresse IP, la connexion SSH cassera. Les mécanismes existants pour permettre ces changements d'attachement sans casser les connexions en cours sont la mobilité IP (RFC 5944), avec introduction d'une machine relais, le "*home agent*", ou bien l'injection de préfixes dans la table de routage lorsque la machine se déplace. Les deux solutions sont imparfaites (la seconde impacte sérieusement la table de routage globale) et on **désire** donc une meilleure solution pour la mobilité (section 3.5). Le mieux serait qu'elle sépare complètement la mobilité du routage.

Autre problème du système de routage : comme il est très difficile de rénumérotter un grand réseau (cf. RFC 5887), beaucoup d'organisations font tout pour garder des adresses IP à elles, ce qui impacte le système de routage (l'agrégation des préfixes devient plus difficile). Le RFC note bien qu'une rénumérotation complètement automatique est utopique, il estime néanmoins (section 3.6) qu'on peut abaisser sérieusement le coût de rénumérotation (cf. RFC 4192) et que c'est **fortement souhaitable**.

Autre **souhait fort** : que la nouvelle architecture soit conçue de manière modulaire, de manière à ce qu'on puisse en utiliser seulement une partie et à ce que les différents composants de la solution ne soient pas excessivement couplés (par exemple, si la solution introduit un tunnel, la détermination de la PMTU devrait se faire par les moyens standard, sans introduire un nouveau mécanisme pour les machines terminales).

Ce n'est pas tout que les paquets arrivent : encore faut-il qu'ils arrivent vite. La qualité du chemin trouvé par le système de routage compte donc beaucoup. Il est donc **fortement souhaité** (section 3.8) que ce dernier trouve une route rapidement (convergence efficace), n'en change tout le temps (routes stables) et que les routes trouvées soient proches de l'idéal (faible élongation - "*stretch*" - cette dernière étant définie comme le rapport entre la longueur du chemin effectif et la longueur du chemin le plus

court possible, cf. « *"A Trade-Off between Space and Efficiency for Routing Tables"* <<http://web.sau.edu/lilliskevinm/wirelessbib/PelegUpfal.pdf>> » de Peleg et Upfal).

Autre préoccupation importante, la sécurité (section 3.9). Il est **requis** que la nouvelle architecture ait une sécurité au moins équivalente à l'ancienne. C'est un des deux seuls points où l'exigence est **requis**. Les cyniques diront que celle-ci n'est pas trop difficile à atteindre, vu le niveau de sécurité existant.

Enfin, une dernière exigence, elle aussi **requis**, en section 3.10 : la déployabilité. Il ne sert à rien de concevoir une splendide architecture si celle-ci reste dans les cartons. Or, des expériences malheureuses comme celle d'IPv6 <<https://www.bortzmeyer.org/ipv6-et-l-echec-du-marche.html>> ont montré que déployer des nouveautés dans l'Internet était très difficile. Il faut notamment que le nouveau système soit déployable de manière incrémentale : les premiers adoptants doivent pouvoir l'installer, et même en tirer des bénéfices immédiats. Sinon, chacun attendra que l'autre fasse le premier pas et on n'avancera pas. (Il n'est évidemment pas question de repartir de zéro, vus les investissements réalisés.)

Voilà, il n'y a maintenant plus qu'à trouver des solutions correspondant à ce cahier des charges : le RFC 6115 résume les résultats de cette quête.