

# Les conséquences politiques des choix techniques dans l'Internet

Stéphane Bortzmeyer <stephane@bortzmeyer.org>

17 avril 2008

## Gouvernance de l'Internet, les acteurs

- ① Les États, notamment les USA
- ② Les auteurs de logiciel, par exemple Microsoft, Cisco, GNU...
- ③ Les opérateurs réseaux et autres prestataires
- ④ Les organismes de normalisation comme l'IETF ou le W3C
- ⑤ Les organismes de régulation comme l'ICANN ou les RIR  
(*Regional Internet Registry* comme le RIPE-NCC en Europe)

Tous font de la politique, même s'ils se gargarisent souvent de termes comme **décisions purement techniques** ou bien **politiques consensuelles**.

## Débats politiques

Il y a des questions dont tout le monde sait qu'elles sont politiques :

- Normalisation de OpenXML et Open Document à l'ISO,
- Le pouvoir du gouvernement US, via l'ICANN, sur la racine du DNS,
- La Neutralité du Réseau,
- La protection des données personnelles (est-ce que Google et Facebook connaissent mes préférences sexuelles et qu'en font-ils?),
- Le pouvoir exorbitant des titulaires de propriété intellectuelle, par exemple via les brevets ou les marques,

Mais d'autres sont ignorées, reléguées au rang de « questions techniques ».

## L'architecture est la loi

Lawrence Lessig :

À Coney Island, le choix de routes étroites, où les autobus ne pouvaient pas passer, garantissait que la plage reste blanche, car uniquement accessible aux voitures individuelles. Ceci, sans mettre de panneaux *Whites only* comme le faisait le Sud.

## L'allocation d'adresses IP

Petits rappels techniques :

- ① Les adresses IPv4 ont une taille de 32 bits et les adresses IPv6 de 128 bits.
- ② Une adresse est composée d'un **préfixe** de longueur variable et d'un identificateur de la machine. Exemple :  
2001:DB8:425::/48 est un préfixe de 48 bits et  
2001:DB8:425:1:0:0:CAFE:1 est une adresse dans ce réseau.
- ③ On peut **agréger** des préfixes qui ont le même préfixe (plus général) en commun. Par exemple 2001:DB8:425::/48 et 2001:DB8:DEAD::/48 peuvent être agrégés dans 2001:DB8::/32. Cela permet de réduire la taille des tables de routage.
- ④ Diverses raisons (le « facteur H ») font qu'il y a des trous dans l'allocation. Il y a en théorie 4 milliards d'adresses IPv4 mais beaucoup moins en pratique.

## Politique d'allocation

- ① Les adresses IPv4 ont longtemps été données sur simple demande. Les pays du Nord se sont ainsi servis largement (40 % des allocations ont été faites avant la mise en place des RIR).
- ② Les restrictions ont été mises en place lorsque les pays du Sud ont commencé à se connecter. Les nouveaux arrivés ont reçu des sermons sur la nécessité de préserver la ressource commune... une fois que les premiers arrivés l'avaient largement entamée.
- ③ Aujourd'hui, les adresses IPv4 s'épuisent (fin prévue vers 2011).
- ④ Qui doit avoir les dernières adresses ? En fonction des besoins, ce qui permettra aux « riches » de s'enrichir encore ? Ou en fonction de la justice ?
- ⑤ Le marché étant la solution à tout, faut-il un marché des adresses IP, ce qui permettrait au Nord de faire payer le Sud ?

## Politique d'allocation d'adresses IP

Les documents comme le RFC 2050 ou comme les politiques d'allocation que les RIR affichent sur leur site ne mentionnent en général pas ces aspects.

Contrairement à l'allocation de noms de domaine, celle des adresses IP est souvent perçue comme « purement technique » ou « neutre ». Les utilisateurs ne sont pas représentés dans les RIR, qui ne comportent que des opérateurs.

Au fur et à mesure de l'épuisement, on peut au contraire parier que ce problème va devenir de plus en plus aigu.

Le déploiement d'IPv6 aurait résolu le problème (on se dispute moins lorsque les ressources sont abondantes) mais il est bien trop tardif pour empêcher la crise.

## Internationalisation de l'Internet

Un très vaste chantier, qui a de nombreux aspects, tout le monde peut participer :

- Logiciels internationalisés (ce qui est largement fait),
- Logiciels localisés (ce qui est surtout fait pour les « grandes » langues),
- Protocoles internationalisés (ce qui est fait à 95 %, on attend juste les adresses de courrier internationalisées et le transfert de texte Unicode dans FTP),
- Contenu en langues locales (dépend des utilisateurs),
- Déploiement des solutions techniques existantes (ce qui nous amène aux IDN).

## Internationalized Domain Names

Les noms de machines étaient traditionnellement restreints à un sous-ensemble d'ASCII.

IDN permet de mettre des noms de domaine en Unicode (café.org ou lætitia-casta.com).

La norme date de 2003 et les premières implémentations aussi. Mais des années de débats acharnés ont été nécessaires !

Encore faut-il le déployer ! (La racine et plusieurs TLD comme "fr" n'ont pas IDN.)

## Petits rappels techniques sur les IDN

- 1 La limite aux lettres d'US-ASCII, aux chiffres et au tiret n'est pas dans le DNS mais dans les règles de nommage des machines.
- 2 Le poids de l'existant fait que personne n'a jamais considéré sérieusement de changer cette règle.
- 3 IDN travaille donc en convertissant localement en ASCII, dans l'application.
- 4 IDN s'appuie sur Unicode, unique norme sur le sujet mais très riche et complexe.
- 5 La question est ultra-passionnelle.

## Une révision nécessaire ?

Les IDN de 2003 n'étaient même pas encore déployés dans la racine que l'ICANN réclamait déjà une réforme.

Deux buts : gagner du temps mais aussi faire passer le **même** que les IDN étaient « dangereux », avec de vagues allusions à des problèmes sans rapport comme le hameçonnage.

L'ICANN a donc obtenu le début du processus de création d'un groupe de travail de réforme d'IDN, dont un des éléments de la charte est de déplacer le choix des caractères acceptables depuis les registres locaux vers l'IETF, qui se lance désormais dans des débats de fous sur l'acceptabilité du tsek tibétain.

## Deux approches opposées

Il n'est pas vraiment nécessaire de fournir des identificateurs dans toutes les écritures. Après tout, les pilotes d'avion acceptent de lire le balisage des aéroports, qui est en alphabet latin partout dans le monde

Argument souvent cité par ceux dont l'alphabet provincial est justement l'alphabet latin.

Le but de l'internationalisation est que l'Internet soit exactement aussi facilement utilisable par un Chinois et par un Américain.

(Ted Hardie)

## Si l'Internet avait été conçu en Chine...

Imaginez-vous face à des URL en chinois...

Et les chinois vous répétant que c'était plus simple et que c'était un standard...

Choisir des identificateurs en ASCII (même pas tout l'alphabet latin) limite donc l'accès d'une partie de la population mondiale.

Aujourd'hui, presque toutes les normes Internet sont complètement internationalisées. Cela ne veut pas dire qu'elles sont déployées !

## La future architecture de l'Internet

Une fois passé le problème de l'épuisement des adresses IPv4, l'Internet va rencontrer un second obstacle.

La table de routage globale grossit beaucoup, et, surtout, elle change trop vite, poussée par le *multihoming* et la mobilité.

Les recherches sur une nouvelle architecture se poursuivent...

Les voix qui se font entendre se réclament souvent d'ensembles flous comme « la communauté Internet » sans expliciter dans quel groupe elles se situent.

## Adresses PI ou PA ?

Les opérateurs souhaiteraient que les adresses IP soient de type PA (*Provider Aggregatable*). Cela diminue la consommation de mémoire dans leurs routeurs, diminue les changements de la table de routage et rend plus difficile le changement d'opérateur.

Exemple radical : les propositions « 8192 entrées dans la table de routage » des débuts d'IPv6. Le RIPE-NCC n'attribue toujours que des adresses PA en IPv6.

Les utilisateurs souhaiteraient que les adresses IP soient de type PI (*Provider Independent*). Cela rend le *multihoming* plus facile et cela permet de changer facilement de fournisseur, puisqu'on n'a plus à renuméroter.

## Cahier des charges

La nouvelle architecture n'est pas encore décidée. Elle n'a même pas de cahier des charges (RFC 4984 et *Internet-Drafts draft-irtf-rrg-design-goals* et *draft-narten-radir-problem-statement*).

Elle reposera peut-être sur la séparation de l'identificateur et du localisateur, permettant de réconcilier PI et PA.



## Petit détour technique

La séparation de l'identificateur et du localisateur consiste à séparer l'adresse IP actuelle en deux, un **identificateur** unique, stable, permanent et indépendant du fournisseur (comme les adresses IP PI), d'une part et un **localisateur** dépendant du fournisseur actuel (comme les adresses IP PA).

La proposition la plus aboutie est HIP mais il y a aussi LISP, Six/One, etc. Certains protocoles s'en inspirent comme SHIM6 ou SCTP.

C'est une très vieille idée et un sujet de recherche appliquée, loin de tout déploiement à l'heure actuelle.

Les demandes des utilisateurs ne font pas explicitement partie du projet.

## Autre exemple, le NAT

La traduction d'adresses IP (NAT, *Network Address Translation*), fait faire des économies à l'opérateur (qui n'a pas à déployer IPv6) mais se traduit par un surcoût pour le développeur (qui doit programmer des techniques élaborées de traversée de NAT) et pour l'utilisateur (beaucoup de protocoles marchent mal à travers le NAT).

Qui doit payer ?

## Conclusion

- ① L'architecture est la loi (*The code is the law*),
- ② Les questions techniques sont trop sérieuses pour être laissées aux techniciens,
- ③ La participation de toutes les parties prenantes est nécessaire.