RFC 1958 : Architectural Principles of the Internet

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 2 janvier 2013

Date de publication du RFC : Juin 1996

https://www.bortzmeyer.org/1958.html

Ce très court RFC posait en 1996 les bases de l'architecture de l'Internet... longtemps après que celui-ci soit largement déployé. C'est une application du principe « faire tourner d'abord, documenter après » qui a été justement un des principes fondamentaux de l'Internet.

Ce document est surtout une synthèse de deux articles essentiels, celui de David Clark, « "*The Design Philosophy of the DARPA Internet Protocols*" <http://ccr.sigcomm.org/archive/1995/jan95/ccr-9501-clark.pdf> » ("*Proc SIGCOMM 88, ACM CCR Vol 18, Number 4, August 1988*") et celui de J.H. Saltzer, D.P. Reed et David Clark, « "*End-To-End Arguments in System Design*" <http://web.mit.edu/saltzer/www/publications/endtoend/endtoend.pdf> » ("*ACM TOCS, Vol 2, Number 4, November 1984*") qui expose très bien les raisons pour lesquelles il ne faut **pas** de réseau intelligent et pourquoi tout le travail difficile doit être fait aux extrémités, dans les machines terminales <https://www.bortzmeyer.org/terminal-host.html>.

Ce RFC 1958 [1] affiches des ambitions modestes : son but était de décrire les principes fondamentaux de l'architecture de l'Internet or, dès le résumé, le RFC note que l'Internet n'a pas été bâti en suivant des grands principes, mais a évolué. Cette capacité à évoluer a fait son succès et tout texte sur les principes ne peut donc être qu'un cliché momentané, pas un texte sacré de principes immuables.

En 1996, date de publication du RFC, la capacité de l'épine dorsale du réseau avait été multipliée par 1 000 depuis les débuts d'ARPANET et son nombre de machines par 1 000 000. Depuis, ces chiffres ont évidemment encore augmenté. Face à de tels changements quantitatifs, il faut s'attendre à ce que les principes qualitatifs aient également changé. Comme le note le RFC « la seule constante de l'Internet, c'est le changement ». On est loin des beaux plans académiques tellement jolis sur le papier mais jamais

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc1958.txt

mis à l'épreuve des faits. L'Internet, comme un être vivant (la comparaison est de moi, elle n'est pas dans le RFC), n'a pas été conçu par un démiurge parfait et c'est pour cela qu'il est si réussi.

La comparaison qu'utilise le RFC est celle d'une ville : de même qu'on ne peut pas détruire une ville pour la refaire « en mieux », on ne peut pas refaire l'Internet de zéro. La ville change tout le temps mais n'a jamais été conçue à partir de zéro (contrairement à ce que voudraient faire aujourd'hui les raseurs de table <https://www.bortzmeyer.org/table-rase-et-john-day.html>).

Le RFC appelle donc à la prudence : il est difficile de prévoir le futur. Et à l'humilité. À noter qu'il ne suit pas toujours ses propres règles. Ainsi, la section 1 parle sans hésiter du « besoin de qualité de service », besoin qui, en 2013, est toujours aussi flou (cf. RFC 5290).

La section 2 commence la liste des principes, après avoir répondu à ceux qui disent, en n'exagérant qu'à moitié, qu'il n'existe pas d'architecture de l'Internet, juste des traditions. Les premiers principes (section 2.1) : la connectivité est un but en soi (contrairement aux experts qui expliquaient doctement qu'il ne servait à rien d'avoir des tuyaux sans contenus, ou aux marketeux qui s'obstinaient à vendre des applications intégrées, et pas juste de la connectivité), l'outil principal d'interconnexion est IP et les fonctions un tant soi peu avancées doivent être dans les extrémités, pas dans le réseau. La connectivité est un but en soi car ce sont les utilisateurs du réseau qui fourniront contenu et applications. (Mon avis est que ceux qui disent le contraire sont ceux qui voudraient verrouiller l'usage de l'Internet en définissant les usages qu'on peut en faire.)

À l'époque de ce RFC, Internet était récemment devenu un réseau mono-protocole, tout fonctionnait avec IPv4. Quelques années auparavant, de nombreux concurrents (de DECnet à UUCP) existaient encore. Depuis, l'Internet est redevenu multi-protocoles, avec IPv6. Notre RFC estime (section 2.2) qu'il ne devrait y avoir qu'un seul protocole au niveau 3, sauf lors d'une transition vers une nouvelle version (le cas actuel) ou si un nouveau protocole est franchement meilleur (personne n'a encore proposé un tel protocole).

La section 2.3 expose et défend le fameux **principe de bout en bout**. Ce principe dit que, toutes les fois où on hésite sur le meilleur endroit pour placer une fonction, on doit la mettre dans les machines terminales et pas dans le réseau. Une des principales justifications de ce principe est qu'il permet à une session en cours de survivre à un redémarrage du réseau, puisque les routeurs ne conservent pas d'**état**. C'est pour cela que le datagramme est meilleur que le circuit.

Bien sûr, les équipements du réseau ont quand même un état : le routeur connait les routes actuelles, par exemple, et il a un cache ARP. Mais cet état doit être auto-réparable : au cas où le routeur redémarre, il doit pouvoir reconstituer son état seul, et que les applications qui l'utilisaient continuent comme si rien ne s'était passé.

Ce principe de bout en bout est également crucial pour la sécurité : une application ne doit pas avoir à faire confiance au réseau (même si le RFC ne le dit pas, l'attaquant peut être le FAI), elle doit pouvoir assurer sa propre sécurité (point traité dans la section 6.2).

Une opinion au passage : ce principe de bout en bout a un autre avantage, bien plus important mais peu mentionné dans le RFC. Ce principe permet l'innovation. Aucun routeur, aucun câble sous-marin n'a dû être modifié pour permettre des inventions comme le Web (apparu quelques années avant ce RFC). Si l'Internet avait été géré par des technocrates français, il aurait fallu réunir une commission avant l'introduction de toute nouvelle application, et le Web n'aurait jamais été déployé. Il n'est donc pas étonnant que le principe de bout en bout soit surtout combattu par ceux qui regrettent le bon temps où le même groupe de technocrates concevait le réseau et toutes les applications.

Depuis la publication de ce RFC, ce principe de bout en bout est aujourd'hui très sérieusement remis en cause par la multiplication des "*middleboxes*", routeurs NAT, pare-feux et autres engins qui, volontairement ou parce qu'ils ont été programmés avec les pieds, bloquent aujourd'hui de nombreuses possibilités du modèle de bout en bout (par exemple, déployer un nouveau protocole de transport comme SCTP est devenu quasiment impossible).

Une des rares incursions de ce RFC dans la politique est en section 2.4 : personne n'est propriétaire de l'Internet, personne ne peut l'éteindre <https://www.bortzmeyer.org/eteindre-internet.html>. Et c'est une bonne chose (même si c'est parfois agaçant).

La section 3 porte sur tout un tas de problèmes de conception d'un grand réseau mondial. Par exemple, il est essentiel de penser au passage à l'échelle (section 3.3). Et de ne pas seulement regarder les fonctions offertes par le réseau mais aussi leur coût (section 3.4), en euros et en performance (un certain nombre de plans alternatifs promettent beaucoup mais oublient de chiffrer les conséquences). Plus généralement, le mieux est l'ennemi du bien : il vaut mieux une solution partielle que d'attendre une solution parfaite et complète (section 3.7).

La longue expérience des praticiens de l'Internet au moment de la publication de ce RFC a aussi mené à d'autres principes : par exemple, que les réglages manuels sont une source d'ennuis sans fin et qu'il faut donc que tout soit automatique (section 3.8). Et il faut éviter les dépendances circulaires (section 3.11), par exemple que le routage dépende du DNS puisque celui-ci dépend du routage... (Des techniques comme Rover <https://www.bortzmeyer.org/rover-bgp.html> ont été critiquées pour cela, pas forcément à juste titre).

C'est dans cette section que le RFC 1958 rappelle un principe souvent cité et souvent discuté, le principe de robustesse (section 3.9). Présent bien avant (je crois que le premier RFC qui le citait explicitement est le RFC 791), il est ici défini comme « soyez strict en envoyant et tolérant en recevant ». L'idée est qu'un programme a intérêt à coller aux moindres détails de la spécification lorsqu'il envoie des données (pour maximiser les chances d'être compris) mais à interpréter cette spécification de manière ouverte lorsqu'il reçoit des données (permettre que le réseau fonctionne est plus important que de pinailler sur la qualité de l'implémentation d'en face).

Et le dernier principe de cette nouvelle section : on ne normalise pas tant qu'on n'a pas deux mises en œuvre distinctes d'un protocole (section 3.14). Notez bien que ce n'est qu'un principe, pas une loi et qu'en pratique, celui-ci est loin d'être toujours respecté.

La section 4 est consacrée aux questions, toujours chaudement disputées, du nommage et de l'adressage. Elle pose des principes comme (section 4.1) d'utiliser des noms plutôt que des adresses (pour leur stabilité <https://www.bortzmeyer.org/pourquoi-le-dns.html>), et n'avoir qu'un seul espace de nommage (section 4.2, qui fait écho au RFC 2826).

Plus rétrograde, le principe 4.3 qui demande que les noms soient uniquement en US-ASCII. Aujourd'hui, où les IDN existent et sont utilisés depuis longtemps, cela sent son époque...

La section 5 porte sur les questions non strictement techniques : acceptation des technologies brevetées (section 5.1), ne pas tenir compte des restrictions à l'exportation (section 5.2, un problème typique des années 1990 où la diffusion des logiciels de cryptographie était bien plus dure qu'aujourd'hui). Il y a aussi un principe (section 5.4) sur l'importance de n'utiliser que des techniques complètement internationalisées, ce qui est franchement contradictoire avec le principe 4.3...