

Systèmes de coordonnées pour Internet

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 1 avril 2008

<https://www.bortzmeyer.org/internet-coordinates.html>

De nombreuses applications Internet ont besoin de connaître la « distance » entre deux machines. Par exemple, un logiciel de pair-à-pair cherche quel pair est le plus « proche » avant de lui demander un fichier. Pour cela, la méthode utilisée est en général à la fois consommatrice en ressources et peu précise. Le déploiement de mécanismes de **coordonnées Internet** pourrait résoudre le problème élégamment.

Aujourd'hui, la métrique la plus courante est le RTT avec l'autre machine et la méthode de mesure la plus fréquente est la force brute : on écrit à tous les partenaires potentiels et le premier à répondre est considéré comme le plus proche et sera désormais davantage utilisé. C'est par exemple ainsi que fonctionne le résolveur DNS BIND ou la plupart des logiciels de pair-à-pair. La métrique n'est pas idéale (il existe des liens Internet à fort débit mais à latence élevée). Et cela consomme beaucoup de ressources réseau (il faut émettre des octets vers chaque pair potentiel) et pour une mesure qui n'est pas forcément très précise.

Les systèmes de **coordonnées Internet** traitent le problème d'une manière radicale. On ne « pingue » plus les pairs potentiels mais on obtient leurs coordonnées dans un espace cartésien à N dimensions. Connaisant ces coordonnées et les siennes propres, déterminer une distance, et donc la proximité d'un futur partenaire, est triviale.

Mais comment déterminer les coordonnées d'une machine ? En testant un petit nombre de machines cibles, les **amers** ("beacons" ou "lighthouses", on peut aussi traduire par **phare** ou **balise**). Contrairement à un pair quelconque, les amers sont en petit nombre (donc on consomme moins de ressources à les tester), stables, et judicieusement placés. Chaque amer a, en testant les autres amers, déterminé ses coordonnées, ce qui permettra à la machine testeuse de déterminer la sienne avec un coût plus faible et une meilleure précision.

Il existe plusieurs systèmes concurrents de coordonnées Internet. À l'heure actuelle, aucun n'est encore réellement déployé et utilisé. L'excellent exposé de Marcelo Pias, "*Internet Coordinate Systems Tutorial*" <<http://www.cl.cam.ac.uk/research/dtg/publications/public/mp431/ast-jon-coord.pdf>> expose les principaux et les compare. Les premiers utilisaient, pour calculer les coordonnées des amers, la minimisation d'erreur par le "*simplex downhill*", alors que les plus récents, comme ICS décrit dans "*Constructing Internet Coordinate System Based on Delay Measurement*" <<http://www.imconf.net/imc-2003/papers/p310-lim11.pdf>> utilisent l'analyse en composantes principales. Voici une liste pas du tout limitative de certains de ces systèmes, outre ICS :

- GNP <<http://www.cs.rice.edu/~eugeneng/research/gnp/>>,
- Vivaldi <<http://pdos.csail.mit.edu/papers/hotnets:vivaldi/>>,

Le projet P4P <<http://www.dcia.info/activities/p4pwg/>>, très à la mode actuellement, est complémentaire de ces systèmes de coordonnées. Il vise à fournir un moyen de distribuer des informations de topologie (qui est connecté à qui) et de politique (le lien de "peering" est préféré au lien de transit) pour que les machines puissent décider judicieusement de leurs pairs. Mais il faut noter que plusieurs présentations de P4P comme "*celle de Pando*" <<http://www.pandonetworks.com/p4p>> sont biaisées dans leur façon de décrire le pair-à-pair actuel, en prétendant que les pairs sont aujourd'hui choisis plus ou moins au hasard, alors que ce choix (qui a l'avantage de simplifier l'implémentation) n'est pas universel.

Une autre approche possible à la question de la détermination d'un pair « proche » est Meridian <<https://www.bortzmeyer.org/meridian.html>>.