

RFC 6875 : The P2P Network Experiment Council's Activities and Experiments with Application-Layer Traffic Optimization (ALTO) in Japan

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 23 février 2013

Date de publication du RFC : Février 2013

<https://www.bortzmeyer.org/6875.html>

Ce RFC décrit une expérience faite il y a quelques années au Japon pour tester l'efficacité des mécanismes de **sélection de pairs** dans un réseau pair à pair. Elle montre que cette sélection peut améliorer l'efficacité des protocoles pair à pair, une leçon utile pour le projet ALTO ("*Application-Layer Traffic Optimization*", RFC 5693¹).

Le réseau pair à pair typique relie des machines, les **pairs**, qui peuvent être très éloignées les unes des autres. Les relations des machines entre elles ne correspondent pas forcément aux réalités de l'interconnexion. En bref, un pair risque de télécharger des données depuis un pair lointain alors qu'il en existait un plus proche. Des tas de propositions ont été faites pour un mécanisme de sélection des pairs, menant théoriquement à une plus grande efficacité. De 2007 à 2008, au Japon, le "*P2P Network Experiment Council*" a réalisé une série de mesures pour comparer théorie et pratique. Une intense répression sous les ordres de l'industrie du divertissement, avec plusieurs arrestations, a réduit l'usage du pair à pair <http://www.sandvine.com/news/global_broadband_trends.asp> dans ce pays depuis, sans toutefois réussir à le réduire à zéro.

Le Japon est souvent un cas à part et, à l'époque en question, le système de pair à pair le plus populaire n'était pas Bit Torrent mais Winny ou Share. Leur trafic total combiné était estimé à 40 % du trafic total. Pour l'opérateur de réseaux, le pair à pair (P2P) a donc un impact important et cela les motive pour chercher des économies, justement ce que promettent les techniques de sélection du pair. Parmi les autres solutions envisagées par les FAI : faire un effort pour placer le contenu très demandé plus près des utilisateurs (les CDN, par exemple), augmenter les prix, ou limiter autoritairement le trafic.

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc5693.txt>

L'approche choisie par le "P2P Network Experiment Council", organisme créé par le gouvernement japonais en 2006, était de tester les solutions de sélection du pair. Pour cela, le conseil a déployé un réseau de serveurs indicatifs (les **oracles**, selon la terminologie généralement utilisée chez les chercheurs de ce domaine), indépendants d'un système de P2P particulier et donnant des indications sur le meilleur pair.

La section 3 décrit plus en détail ce « conseil pour des expériences P2P ». Créé par le ministère des télécommunications, à la suite de débats sur des sujets comme la neutralité du réseau, il est documenté (en japonais) dans le rapport « "Disclosure of the Report [Caractère Unicode non montré ²] Working Group on P2P Networks" » <http://www.soumu.go.jp/menu_news/s-news/2007/070629_11.html> et placé sous l'égide d'une fondation <<http://www.fmmc.or.jp/P2P/about.htm>> (oui, encore un texte dans la langue de Naoki Urasawa). Outre les tests qui font l'objet de ce RFC, le conseil a une activité de promotion des bonnes pratiques en matière de P2P.

La section 4 de notre RFC décrit les expériences menées : au début, il n'y avait aucun doute sur les économies en matière de **serveurs**, lorsqu'on migre vers le pair à pair. Ainsi, l'entreprise Utagoe annonçait <<http://www.utagoe.com/en/technology/grid/live/index.htm>> une réduction de 90 % de son trafic sortant en migrant vers leur système UGLive. Et TV Bank Corp a vu 80 à 96 % de gain avec BB Broadcast <<http://www.tv-bank.com/jp/20081031.html>>. Mais ce sont des gains pour les fournisseurs de contenu. Et pour les opérateurs réseau? Il y a eu de études au Japon sur le trafic total (comme celle de Cho, Fukuda, Esaki et Kato <<http://www.cs.auckland.ac.nz/compsci742s2c/resources/Residential-User-to-User-Traffic.pdf>>). Mais étudier le comportement des applications réellement installées sur les utilisateurs demande des mesures fort intrusives, par exemple avec du DPI.

Le choix du conseil avait donc été d'installer des faux utilisateurs : des PC munis de logiciels P2P modifiés (pour demander à l'oracle) ont été placés dans les réseaux des FAI. On peut ainsi faire varier librement et facilement tous les paramètres de mesure (et, bien que le RFC ne le dise pas, cela évite tout problème lié à la vie privée). Le faux utilisateur est donc un PC avec CentOS, VMware et un Windows par dessus. Le système hôte CentOS peut ainsi observer tout le trafic sans déranger. 60 machines ont été installées chez 40 FAI.

La section 5 décrit l'oracle ("*hint server*"), le serveur qui va distribuer les indications. C'est l'équivalent du « serveur ALTO » dans le protocole ALTO (RFC 6708). Lorsqu'un pair démarre, il indique à l'oracle son adresse IP et des informations comme les caractéristiques de la liaison Internet utilisée. L'oracle, en utilisant des informations sur la topologie du réseau informations fournies par le FAI et une table de routage obtenue via BGP (l'oracle inclut un Quagga), peut alors envoyer au pair une table des meilleurs pairs possibles. Dans ce cas idéal, l'optimisation réalisée est presque parfaite. En pratique, l'oracle n'a pas forcément toute l'information nécessaire : le FAI, l'estimant confidentielle, a pu ne pas donner une information complète.

Pour prendre sa décision, l'oracle s'appuie sur :

- La longueur du chemin d'AS entre le client et le pair potentiel (rappelez-vous que l'oracle participe à BGP),
- La distance géographique entre ces deux pairs possibles. Ne connaissant par la longitude et la latitude de chaque machine, l'oracle fait une approximation : il trouve la préfecture dans une base GeoIP, puis utilise la position physique de la capitale de la préfecture.

Voici un exemple simplifié d'une requête d'un pair à l'oracle, et de la réponse de l'oracle. Le tout est transporté sur HTTP et les données sont encodées en couples {type, valeur} :

2. Car trop difficile à faire afficher par L^AT_EX

```
POST /PeerSelection HTTP/1.1
Host: OracleServerName
User-Agent: ClientName
Content-Type: text/plain; charset=utf-8

v=Version number
ip=IP address of physical interface
port=1978
nat=unknown
ub=256
db=15000
```

Vous avez compris que `ub` et `db` sont les capacités réseau montant et descendantes, et que `nat` indique (si le pair le sait), si le pair est coincé ou non derrière un routeur NAT, qui rend difficile les connexions entrantes. Et voici la réponse :

```
HTTP/1.1 200 OK
Content-Type: text/plain; charset=utf-8

v=Version number
ttl=ttl
client_ip=192.0.2.67
numpeers=11
...
```

La section 6 résume les résultats de l'expérience. Soit un pair situé à Tokyo. Sans l'oracle, de 22 à 29 % des pairs sélectionnés étaient déjà dans le même FAI, et de 19 à 23 % dans le même district (il est important de connaître ces valeurs car, si un contenu est peu populaire et ne se trouve pas du tout dans le même FAI que l'utilisateur, aucun oracle au monde ne pourra indiquer un pair proche). Avec l'oracle, la distance moyenne avec le pair a été réduite de 10 %. (Opinion personnelle : le RFC est beaucoup plus détaillé sur la configuration de l'expérience que sur les résultats obtenus, très brefs.)

Les auteurs estiment que ces résultats montrent l'intérêt d'un système de sélection du pair. Quelles conclusions pour le groupe de travail ALTO <<http://tools.ietf.org/wg/alto>>, qui standardise une technologie analogue? Que l'information publiquement accessible via BGP est une source utile pour optimiser le trafic entre FAI. Que, pour le trafic interne au FAI, il faut une information détaillée de la topologie du réseau de ce FAI, information qui peut être difficile à obtenir. Que les demandes futures seront peut-être d'avoir une sélection plus précise, nécessitant peut-être une structure hiérarchique dans ALTO (des serveurs ALTO utilisant eux-même d'autres serveurs ALTO). Et qu'il serait bon que l'oracle soit un peu plus bavard sur les mécanismes qu'il a utilisé pour trier les pairs possibles, afin que le pair puisse combiner le jugement de l'oracle avec ses propres critères.

Sur l'évaluation des mécanismes de sélection du pair, voir aussi le RFC 6821.