

RFC 6821 : Improving Peer Selection in Peer-to-peer Applications: Myths vs. Reality

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 22 décembre 2012

Date de publication du RFC : Décembre 2012

<https://www.bortzmeyer.org/6821.html>

Le trafic du pair-à-pair peut, on le sait, représenter une bonne part de l'activité d'un réseau, malgré les efforts de l'industrie du divertissement pour diaboliser cette technique. Il y a donc depuis plusieurs années de gros efforts de recherche pour optimiser ce trafic, notamment via l'amélioration de la **sélection des pairs**. Si je veux télécharger la saison 1 de "*Being human*", et que trois pairs ont les données, auquel demander ? Le bon sens répond « au plus proche ». Mais le concept de « plus proche » est plus flou qu'il n'y paraît, et, de toute façon, le logiciel pair-à-pair installé sur ma machine n'a pas forcément accès à toutes les informations nécessaires pour déterminer « le plus proche ». Il existe plusieurs solutions pour résoudre ce problème, mais notre RFC se penche plutôt sur le méta-problème : **la sélection des pairs améliore-t-elle les choses ?**

Tellement de choses ont été dites à ce sujet que l'ingénieur ou l'étudiant débutant qui se penche sur l'optimisation du pair-à-pair peut avoir une impression de grande confusion. Ce RFC choisit donc l'approche « retours aux faits ». Parmi toute la littérature scientifique et technique existante, peut-on trancher sur la question de l'intérêt de la sélection des pairs ?

Je préviens tout de suite que le titre de ce RFC, inutilement sensationnaliste, ne correspond pas à son contenu : le RFC ne dynamite pas de mythes, il examine un certain nombre de questions posées par la sélection des pairs et, pour chacune, en se basant sur des mesures ou des simulations déjà effectuées et publiées, fait une synthèse de leurs conclusions.

A priori, l'intérêt de la sélection est grand : comme les machines, dans un réseau pair-à-pair, ne connaissent pas la topologie sous-jacente (elles ne savent pas si les pairs sont proches ou lointains, s'ils sont joignables par des liens de "*peering*" gratuits ou par du transit payant, etc), elles risquent de ne

pas choisir le meilleur pair. Une des conséquences fâcheuses sera l'utilisation d'inter-connexions lointaines, au lieu de rester dans le réseau du FAI. Il est donc logique que cette idée de sélection « intelligente » des pairs ait fait l'objet de nombreux travaux (résumés dans le RFC 6029¹; sinon, voir les articles « *Can ISPs and P2P systems co-operate for improved performance?* » <<http://www.sigcomm.org/sites/default/files/ccr/papers/2007/July/1273445-1273449.pdf>> » d'Aggarwal, V., Feldmann, A., et C. Scheidler, ou « *Taming the Torrent : A practical approach to reducing cross-ISP traffic in P2P systems* » <<http://www.sigcomm.org/sites/default/files/ccr/papers/2008/October/1402946-1402949.pdf>> » de Choffnes, D. et F. Bustamante, ou encore « *P4P : Explicit Communications for Cooperative Control Between P2P and Network Providers* » <http://www.dcia.info/documents/P4P_Overview.pdf> » de Xie, H., Yang, Y., Krishnamurthy, A., Liu, Y., et A. Silberschatz) et qu'un groupe de travail de l'IETF, ALTO <<http://tools.ietf.org/wg/alto>>, travaille entièrement sur ce sujet (voir ses RFC 5693 et RFC 6708). (À noter que j'avais fait un exposé sur ces techniques en 2010 <<https://www.bortzmeyer.org/thd-meilleur-pair.html>>.) L'évaluation de ces techniques n'est pas évidente, notamment de leur passage à l'échelle lorsque le réseau comprend des dizaines de millions de pairs.

Notre RFC suit le schéma suivant pour synthétiser la littérature existante :

- Décrire une croyance (un mythe, dit le RFC, terme très exagéré),
- Lister les faits (études, mesures, simulations, etc) disponibles,
- Discuter ces données,
- Conclure à la véracité ou à la fausseté de la croyance.

Naturellement, cette synthèse n'est valable qu'aujourd'hui : les progrès de la science pourront changer les conclusions. Ah et, sinon, question terminologie, en l'absence d'une norme unique du pair-à-pair, ce RFC utilise largement le vocabulaire de BitTorrent (section 2), comme le terme d'essaim pour désigner un groupe de pairs ayant les données convoitées.

Place maintenant aux croyances et à leur évaluation. La première : « la sélection des pairs permettra de diminuer le trafic entre domaines ». Les différents essais ou simulations montrent des réductions allant de 20 à 80 % (la variation importante donne une idée de la difficulté à estimer cette réduction). 70 % pour la simulation de Xie, H., Yang, Y., Krishnamurthy, A., Liu, Y., et A. Silberschatz déjà citée. 34 % en sortie et 80 % en entrée pour les mesures du RFC 5632. Et jusqu'à 99,5 % si le trafic est fortement localisé (beaucoup de pairs dans le même domaine) selon « *Pushing BitTorrent Locality to the Limit* » <<http://hal.inria.fr/inria-00534117/en/>> » de Stevens Le Blond, Arnaud Legout et Walid Dabbous.

Bref, cette croyance est tout à fait justifiée (comme la plupart de celles citées par le RFC). On peut vraiment espérer une réduction du trafic entre opérateurs.

Cette réduction profite aux opérateurs, qui voient baisser leur facture d'interconnexion. Une autre croyance fait espérer des gains pour les utilisateurs : « la sélection des pairs se traduira par une amélioration des performances », en clair, on attendra moins longtemps pour charger la dernière ISO d'Ubuntu.

Les simulations de Xie, H., Yang, Y., Krishnamurthy, A., Liu, Y., et A. Silberschatz montrent une diminution du temps de transfert de 10 à 23 %. Celles de « *Applicability and Limitations of Locality-Awareness in BitTorrent File-Sharing* » par Seetharaman, S., Hilt, V., Rimac, I., et M. Ammar (article que je n'ai pas réussi à trouver en ligne et qui ne semble pas avoir été publié « officiellement ») montraient que le gain n'est pas systématique. Les mesures du RFC 5632 indiquent une augmentation du débit de 13

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc6029.txt>

à 85 %. L'expérience de Choffnes, D. et F. Bustamante déjà citée a vu 31 % de gain de débit en moyenne (mais une perte dans certains cas).

La conclusion est donc plutôt « ça dépend ». En général, il y a une amélioration mais dans certains cas (capacité du lien montant faible, peu de pairs disponibles à proximité), on peut voir une dégradation. Donc, la croyance est probablement justifiée mais pas dans tous les cas.

Et l'occupation du lien réseau montant ? Cela pourrait être un effet négatif d'une meilleure sélection des pairs. Ils sont plus proches, donc on envoie plus vite les données, saturant le canal montant, souvent très petit sur les liens asymétriques comme l'ADSL ou DOCSIS. Cette crainte est-elle justifiée ?

Les mesures du RFC 5632 ne montrent pas un tel effet. Théoriquement, cela serait pourtant possible (si la sélection des pairs menait à choisir une machine mal connectée mais proche plutôt qu'une machine lointaine et ayant de fortes capacités réseau, la machine mal connectée verrait son lien montant plus utilisé). Mais, en pratique, cela ne semble pas le cas.

Autre croyance parfois entendue : la sélection des pairs va avoir une influence sur les accords de "peering". Ceux-ci sont souvent fondés sur le volume de trafic échangé. Si la sélection des pairs améliore la localité, le trafic entre opérateurs va baisser, le faisant passer en dessous du seuil qu'exigent certains opérateurs pour "peerer" avec eux. Mais, sur ce point, on ne dispose pas de mesures ou de simulations. On ne peut faire pour l'instant qu'une analyse théorique.

Par exemple, si les deux opérateurs ont une base d'utilisateurs très différente (mettons qu'un opérateur a beaucoup de clients offrant du contenu et l'autre pas du tout), non seulement le trafic entre eux va baisser, mais il peut baisser nettement plus dans une direction que dans l'autre, remettant en cause un accord de "peering" fondé sur la réciprocité.

Curieusement, le RFC affirme que, si les opérateurs sont très similaires (même base d'utilisateurs, mêmes technologies d'accès), il n'y aura sans doute pas beaucoup de changement. Pourtant, comme indiqué plus haut, la seule baisse du trafic entre eux, même symétrique, peut changer les conditions du "peering".

Le RFC rappelle que des opérateurs ont déjà tenté d'injecter des paquets vers un autre opérateur, pour obtenir artificiellement des chiffres élevés de trafic, avant de pouvoir demander un "peering" (« "The art of Peering : The peering playbook" <<http://d.drpeering.net/>> » de Norton). Ce n'est pas très subtil car ce trafic purement unidirectionnel apparaît vite comme suspect. En revanche, avec des techniques de sélection des pairs, le même FAI peu scrupuleux pourrait faire mieux, en redirigeant systématiquement ses utilisateurs vers l'opérateur avec qui il espère négocier un accord de "peering", créant ainsi un trafic nettement plus réaliste.

En conclusion (fondée sur un raisonnement purement théorique, contrairement à la grande majorité des croyances étudiées dans ce RFC), il est probable que la sélection des pairs amène à changer les accords de "peering".

Et sur le transit, y aura-t-il un impact ? Là aussi, on peut imaginer des transitaires peu scrupuleux qui, en utilisant les techniques de sélection de pairs comme ALTO, redirigeront le trafic vers des clients payants, plutôt que vers des opérateurs pairs gratuits. Là encore, on ne dispose pas d'études à ce sujet. Mais, vue l'importance de la question pour le monde des opérateurs, le RFC recommande que l'on se penche sur la question. Des cas comme Sprint contre Cogent <<http://www.pcworld.com>>

com/businesscenter/article/153123/> ou Cogent contre AOL <<http://legalminds.lp.findlaw.com/list/cyberia-1/msg42080.html>> illustrent bien l'extrême sensibilité du problème.

Un effet de bord négatif d'une bonne sélection des pairs avait été envisagé dans l'article Stevens Le Blond, Arnaud Legout et Walid Dabbous ci-dessus (mais les mesures faites ne montraient pas cet effet). Si la sélection marche trop bien, l'essaim de pairs va se fragmenter en plusieurs essaims, un par FAI, avec peu de communications entre eux, ce qui annulerait une partie de l'intérêt du pair-à-pair. En effet, les simulations de Seetharaman, S., Hilt, V., Rimal, I., et M. Ammar ont montré que cela pouvait se produire.

Le RFC conclut que l'effet dépend de l'algorithme utilisé. Si BitTorrent semble assez résistant à ce problème, comme montré par les mesures de Stevens Le Blond, Arnaud Legout et Walid Dabbous, d'autres algorithmes peuvent avoir le problème. Et que la croyance peut donc être justifiée dans certains cas. Attention donc, si vous concevez un algorithme de pair-à-pair à ne pas mettre tous vos œufs dans le même panier en ne sélectionnant que des pairs très proches.

Dernière croyance étudiée, celle que des caches P2P chez les FAI, combinés avec la sélection des pairs, va améliorer les choses. Un problème classique de tous les caches réseau est de les trouver, et de ne les utiliser que s'ils apportent réellement un gain. Avec la sélection des pairs, le problème est théoriquement résolu : le cache est un pair comme un autre et est sélectionné s'il est « plus proche ».

Une étude chez un FAI a montré (« *Are file swapping networks cacheable? Characterizing p2p traffic* » <<http://2002.iwcw.org/papers/18500253.pdf>> » de Nathaniel Leibowitz, Aviv Bergman, Roy Ben-Shaul et Aviv Shavit) que le cache pouvait potentiellement gérer 67 % du trafic pair-à-pair. Cette croyance est donc plausible.