HAL
open science

Supra-extensibilité des réseaux P2P

François Baccelli, Fabien Mathieu, Ilkka Norros, Rémi Varloot

To cite this version:

François Baccelli, Fabien Mathieu, Ilkka Norros, Rémi Varloot. Supra-extensibilité des réseaux P2P. 15èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel), May 2013, Pornic, France. pp.1-4. hal-00818551

HAL Id: hal-00818551

https://inria.hal.science/hal-00818551v1

Submitted on 29 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Supra-extensibilité des réseaux P2P †

François Baccelli[1 2 3], Fabien Mathieu[4], Ilkka Norros[5] et Rémi Varloot[1]

[1]*Inria* [2]*ÉNS* [3]*UT Austin* [4]*Alcatel-Lucent Bell Labs France* [5] *VTT*

Une des clés du succès du pair-à-pair (P2P) est la faculté de passer à l'échelle, ou extensibilité (*scalability*) : la capacité de service augmente avec le nombre de participants. Pour la distribution de contenu, l'extensibilité P2P est fortement liée au fait que le principal goulot d'étranglement vient des vitesses d'accès individuelles. Mais avec l'augmentation des accès très haut débit, la congestion pourrait, dans le futur, se déplacer ailleurs. Qu'arriverait-il alors ?

Pour aborder ce problème, nous proposons un modèle stochastique où les pairs sont plongés dans un espace métrique qui détermine les vitesses point-à-point. Notre principal résultat est la description, pour un scénario simple inspiré de BitTorrent, des performances du régime stationnaire. Un phénomène contre-intuitif apparaît : la supra-extensibilité, ou l'augmentation sur-linéaire de la capacité de service ! La supra-extensibilité apparaît naturellement dès que le nombre de pairs permet de choisir des voisins locaux, et continue jusqu'à un point de rupture où le réseau sous-jacent devient subitement incapable de supporter la charge. Enfin, nous construisons à partir du scénario de base un *couteau suisse* d'extensions du modèle pour lesquelles nous donnons des formules de performances.

Une version étendue de cet article sera publiée dans Infocom 2013 [1].

Keywords: P2P, géométrie stochastique, analyse de performances

1 Contexte et modèle

Le scénario étudié est inspiré du protocole d'échange de fichiers BitTorrent [5], où un fichier est segmenté en petits morceaux échangés entre eux par des pairs *leechers* au sein d'un réseau logique. Quand un *leecher* a fini son téléchargement, il peut rester et continuer à contribuer au système. Il est alors *seeder*.

Notre approche est dans la lignée de travaux précédents visant à étudier la performance de tels systèmes en se concentrant sur les problèmes de bande passante [9, 8, 2], mais avec un nouveau modèle de congestion.

Plus précisément, les pairs sont plongés dans un espace D muni d'une distance d. d sera par convention exprimé en mètres, mais son interprétation peut être multiple : distance physique, dans le graphe réseau, dérivée des latences [6], voire distance sémantique. Par souci de simplicité, les résultats sont donnés pour un large tore de dimension 2 approximant le plan euclidien, mais le modèle fonctionne tout aussi bien sur d'autres types espaces adaptés aux réseaux, par exemple les espaces hyperboliques [3].

Des nouveaux pairs arrivent dans le système selon une pluie de Poisson d'intensité λ (exprimé en $m^{-2}.s^{-1}$). On suppose que la congestion ne provient pas de l'accès, mais de la bande passante disponible d'un pair x vers un pair y, qui est déterminée par une fonction $f(d(x,y))$. Des expressions possibles de f sont indiquées sur le tableau 2. En arrivant, chaque pair cherche à récupérer une certaine quantité de données, modélisée par des variables i.i.d. d'espérance finie F. Les données sont récupérées de l'ensemble des voisins à distance au plus R dans D. Une fois son téléchargement terminé, un pair quitte le système.

Nous négligerons ici les problèmes de disponibilité du contenu, et supposerons que toute paire de pairs a toujours des données à s'échanger (une argumentation détaillée est disponible dans [1]).

†Les travaux présentés ici ont été en partie réalisés au *Laboratory of Information, Network and Communication Sciences* (LINCS, `http://www.lincs.fr`).

$f(r)$	Interprétation	$\gamma = 2\pi \int_0^R r f(r) dr$
$\frac{C}{r}$	Comportement TCP Reno lié à la latence (voir [7])	$2\pi CR$
U	Comportement UDP (constant)	$\pi U R^2$
$\frac{C}{r} \wedge U$	TCP avec limitation par flot	$\pi\left(2CR - \frac{C^2}{U}\right)$ [a]
$\frac{C}{r+q}$	TCP avec latence affine	$2\pi C\left(R - q\ln(1+\frac{R}{q})\right)$
$\frac{C}{r} - o$	TCP avec coût de contrôle	$\pi R(2C - oR)$ [b]
$\frac{1}{2}\ln\left(1+\frac{C}{r^\alpha}\right)$	Canal sans fil à bruit blanc gaussien additif	$\frac{\pi^2 C^{\frac{2}{\alpha}}}{2\sin\left(\frac{2\pi}{\alpha}\right)}$ pour $R = \infty$ [c]

TABLE 2: Quelques fonctions de débit f et avec leur force γ associée

[a] Valable pour $C \leq UR$; $C \geq UR$ correspond au cas UDP.
[b] Valable pour $\frac{C}{R} \geq o$; sinon, R devient $\frac{C}{o}$.
[c] Il n'y a pas de formule explicite pour $R < \infty$ et α quelconque, mais pour $\alpha = 4$, nous avons $\gamma = \pi\left(R^2 \log(1+\frac{C}{R^4}) + \sqrt{C}\arctan(\frac{R^2}{\sqrt{C}})\right)$.

Nom	Description	Unités
C	Fonction de débit TCP	$bits \cdot s^{-1} \cdot m$
U	Débit UDP	$bits \cdot s^{-1}$
F	Taille moyenne de fichier	$bits$
R	Rayon d'action	m
λ	Intensité d'arrivée	$m^{-2} \cdot s^{-1}$
W	Temps moyen de téléchargement	s
μ	Débit total moyen d'un leecher	$bits \cdot s^{-1}$
β	Densité de leechers	m^{-2}

TABLE 1: Principales grandeurs du modèle

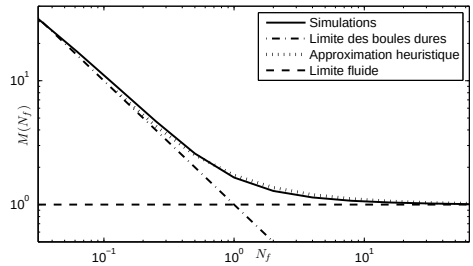

FIGURE 1: $M(N_f)$ pour f de type TCP Reno

2 Résultats principaux

Le processus de vie et de mort des pairs, dans le modèle présenté, est un processus Markovien qui est ergodique pour tout $\lambda > 0$. Autrement dit, il existe un régime stationnaire, sur lequel nous travaillons.

2.1 Analyse dimensionnelle

Pour faciliter l'étude du système décrit, on peut imaginer une limite fluide qui négligerait les fluctuations spatio-temporelles des différentes valeurs. Sous cette simplification, on obtient la densité moyenne de pairs β_f, la vitesse de téléchargement moyenne μ_f et le temps de téléchargement moyen W_f : si l'on note $\gamma := 2\pi \int_{r=0}^R r f(r) dr$, ces valeurs sont

$$\beta_f = \sqrt{\frac{\lambda F}{\gamma}},\ \mu_f = \sqrt{\lambda F \gamma},\ W_f = \sqrt{\frac{F}{\lambda\gamma}}. \tag{1}$$

γ, dont des expressions sont indiquées dans le tableau 2, représente la *force* de f, et quantifie la vitesse de téléchargement point-à-point.

À travers (1) apparaît le principe de supra-extensibilité : le temps de téléchargement décroit avec $\frac{1}{\sqrt{\lambda}}$!

Mais en pratique, l'hypothèse fluide n'est pas forcément justifiée, et (1) ne décrit pas le système proposé dans la section 1. Ceci étant, par un argument d'analyse dimensionnel, le théorème π [4], on sait que le temps de téléchargement moyen W_0 du système *réel* (non fluide) peut s'écrire sous la forme $W_0 = M(N_f)W_f$, où $N_f := \pi R^2 \sqrt{\frac{\lambda F}{\gamma}}$. N_f est une grandeur sans dimension s'interprétant comme le nombre de pairs dans un disque de rayon R dans la limite fluide, et M une fonction (inconnue) qui ne dépend, à f fixée, que de N_f.

M permet de quantifier précisément l'écart entre un système réel et son approximation fluide.

2.2 Résultats théoriques

Un premier résultat important est que $M \geq 1$: la limite fluide donne une borne inférieure du temps de téléchargement réel. Plus encore, cette borne est asymptotiquement exacte : quand N_f tend vers l'infini, M tend vers 1. Ce résultat prouve la supra-extensibilité du système en validant que le comportement est asymptotiquement fluide.

Dans le cas inverse, quand N_f tend vers 0, le comportement du système change radicalement : dès que deux pairs sont à distance inférieure à R, un téléchargement quasi-instantané se produit, de sorte qu'avec

forte probabilité, si l'on prend une photo du système à un instant donné, tous les pairs sont à distance R ou plus les uns des autres. Dans cette limite des *boules dures*, M se comporte en $\frac{1}{N_f}$.

Pour les valeurs intermédiaires de N_f, il n'existe a priori pas de formule close pour M, mais nous avons trouvé une méthode permettant d'obtenir une approximation heuristique. Par exemple, si $f(r)$ est de la forme $\frac{C}{r}$, on peut approcher M par la solution $\hat{M}$ de

$$\hat{M}^2\left(1-\frac{\hat{M}}{2N_f}\ln\left(1+\frac{2N_f}{\hat{M}}\right)\right)=1. \tag{2}$$

Afin de valider ces résultats théoriques, nous avons simulé des systèmes tels que ceux décrits en section 1 en variant les paramètres pour obtenir un échantillonnage de la fonction M. La figure 1 montre la comparaison entre théorie et simulations pour $f(r)$ en $\frac{C}{r}$.

2.3 Réseau physique

On peut se demander si le réseau physique peut supporter la supra-extensibilité. Notre modèle permet de considérer cette question. On suppose maintenant que le réseau physique possède une capacité maximale homogène isotrope Ξ, en $bits.s^{-1}.m^{-1}$ (le réseau physique peut supporter qu'un débit Ξl traverse un segment de longueur l). Le système P2P, lui, génère mécaniquement, dans sa limite fluide un trafic homogène isotrope d'intensité

$$\psi := 4\frac{\lambda F}{\gamma}\int_0^R r^2 f(r)dr. \tag{3}$$

La condition pour que le réseau physique supporte la charge est simplement $\psi \leq \Xi$: f doit posséder un moment fini d'ordre 2 et λ vérifier

$$\lambda \leq \frac{\gamma\Xi}{4F\int_0^R r^2 f(r)dr}. \tag{4}$$

Si la supra-extensibilité tire avantage de l'intensité λ, il existe donc un point de rupture à partir duquel le réseau ne supporte plus la charge. On remarquera que ce phénomène existe aussi dans le modèle extensible traditionnel : le trafic total vaut alors λF en régime stationnaire, ce qui implique également l'existence d'un point de rupture quand λ tend vers l'infini.

3 Extensions

Les résultats précédents autorisent de nombreuses extensions, dont voici un aperçu.

Serveurs permanents Beaucoup de systèmes réels utilisent des serveurs pour améliorer la performance. Dans notre modèle, l'action des serveurs peut être représentée par une *densité de débit* U_C, exprimée en $bits.s^{-1}.m^{-2}$, de sorte que si β est la densité des leechers, un pair reçoit en moyenne un débit $\frac{U_C}{\beta}$.

Pour étudier ce système, il est utile d'introduire le paramètre sans dimension $\chi := \frac{U_C}{\lambda F}$. Deux cas limites apparaissent alors.

– Si $\chi \gg 1$, alors le système fonctionne quasiment en client/serveur, et le temps de téléchargement est

$$W_C \approx \frac{F}{\pi R^2 U_C}. \tag{5}$$

– Si $\chi \ll 1$, alors le temps de téléchargement dans la limite fluide (i.e. pour N_f assez grand) vaut

$$W_{f,C} = W_f\sqrt{1-\chi} \approx W_f. \tag{6}$$

Abandon En s'inspirant de [8], on peut rajouter un processus d'abandon *en cours de route*, décrit par une intensité d'abandon a (en s^{-1}). Le temps de téléchargement dans la limite fluide vérifie alors

$$\frac{1}{W_{f,a}} = \sqrt{\left(\frac{1}{W_f}\right)^2+\left(\frac{a}{2}\right)^2}-\frac{a}{2}. \tag{7}$$

Limitation à l'accès On peut introduire une limitation à l'accès de valeur U (l'upload total d'un pair ne peut dépasser U). Dans ce cas, les équations de supra-extensibilité restent valables dans la limite fluide tant que

$$\sqrt{\lambda F \gamma} \leq U. \tag{8}$$

Au-delà, l'accès devient le point de congestion, et on retrouve un modèle simplement extensible.

Seeders Supposons maintenant que les pairs restent un certain temps dans le système après avoir fini de télécharger, d'une valeur moyenne T_S. Alors, dans la limite fluide, nous avons

$$W_{f,S} = \sqrt{W_f^2 + \left(\frac{T_S}{2}\right)^2} - \frac{T_S}{2}. \tag{9}$$

Les seeders accélèrent le téléchargement. Ils permettent aussi d'échapper à l'effet *boules dures*.

Connectivité limitée Les systèmes P2P réels ne sont généralement pas basés sur un rayon d'action R, mais sur un nombre maximal de connexion L que chaque pair peut établir. Ceci peut être géré par notre modèle. Par exemple, pour f en $\frac{C}{r}$, si l'on suppose que chaque pair se connecte à ses L plus proches voisins, alors dans la limite fluide le temps de téléchargement est

$$W_{f,L} = \left(\frac{F}{2C}\right)^{\frac{2}{3}} \left(\frac{1}{\pi \lambda L}\right)^{\frac{1}{3}}. \tag{10}$$

On constate au passage que la supra-extensibilité existe toujours en dépit du fait que le nombre de voisins, qui tendait vers l'infini dans le modèle initial, est maintenant fixé. Cela ne s'applique en revanche pas à toutes les fonctions f. Par exemple, si f est constante, le système redevient simplement extensible pour des raisons évidentes.

4 Conclusion

Nous avons présenté le concept de supra-extensibilité, qui permet de décrire le comportement d'un système P2P lorsque la vitesse de téléchargement dépend des distances entre pairs, et nous avons constaté que la plupart des modèles et résultats utilisés pour décrire les systèmes simplement extensibles pouvaient être transposés aux systèmes supra-extensibles.

Références

[1] François Baccelli, Fabien Mathieu, Ilkka Norros, and Rémi Varloot. Can P2P Networks be Super-Scalable ? In *INFOCOM*, 2013.

[2] Farid Benbadis, Fabien Mathieu, Nidhi Hegde, and Diego Perino. Playing with the bandwidth conservation law. In *IEEE P2P*, pages 140–149, 2008.

[3] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the Internet with Hyperbolic Mapping. *Nature Communications*, 1(62), Oct 2010.

[4] E. Buckingham. The principle of similitude. *Nature*, 96 (2406) :396–397, 1915.

[5] B. Cohen. BitTorrent specification, 2006. `http://www.bittorrent.org`.

[6] Pierre Fraigniaud, Emmanuelle Lebhar, and Laurent Viennot. The inframetric model for the internet. In *IEEE INFOCOM)*, pages 1085–93, 2008.

[7] T. Ott, J.H.B. Kemperman, and M. Mathis. The stationary behavior of ideal TCP congestion avoidance. *Internetworking : Research and Experience*, 11 :115–156, 1992.

[8] D. Qiu and R. Srikant. Modeling and performance analysis of BitTorrent-like peer-to-peer networks. *ACM SIGCOMM Computer Communication Review*, 34(4) :367–378, 2004.

[9] Gustavo De Veciana and Xiangying Yang. Fairness, incentives and performance in peer-to-peer networks. In *In the Forty-first Annual Allerton Conference on Communication, Control and Computing*, 2003.