Managing a Peer-to-Peer Backup System: Does Imposed Fairness Socially Outperform a Revenue-Driven Monopoly?

László Toka and Patrick Maillé

GET/ENST Bretagne
2, rue de la Châtaigneraie CS 17607
35576 Cesson Sévigné Cedex, France
{laszlo.toka, patrick.maille}@enst-bretagne.fr

Abstract. We study a peer-to-peer backup system, where users offer some of their storage space to provide service for the others. The economic model for such a system is different from the ones applicable to peer-to-peer file sharing systems, since the storage capacity is a private good here. We study two mechanisms aimed at incentivizing users to offer some of their capacity: a price-based scheme (here a revenue-driven monopoly) and a more classical symmetric scheme (imposing users to contribute to the service at least as much as use it). We compare the outcomes of such mechanisms to the socially optimal situation that could be attained if users were not selfish, and show that depending on user heterogeneity, a revenue maximizing monopoly can be a worse or a better (in terms of social welfare) way to manage the system than a symmetric scheme.

Keywords: Peer-to-peer networks, economics, incentives, pricing.

1 Introduction

With the convergence of fixed and mobile telecommunication systems, all kinds of digital documents (e.g. videos and audio files, e-mails) are likely to be accessed by different types of devices (mobile phone, personal computer, mp3 player). The storing of all those then documents raises several questions: should there be only one storing location? If so, what happens in case of a crash? If not, how to update documents between several locations? Will it be simple to transfer a document from a storing location to a given device?

In this paper, we suggest that those problems be addressed via a distributed storing system working in a peer-to-peer (P2P) way: using a P2P network infrastructure, a (ciphered) copy of each user's data is stored into the hard drives of other participants in the network. As in peer-to-peer (P2P) file sharing networks, each participant is consequently at the same time a service user and a service provider. Such a service presents a lot of advantages in terms of reliability (data replication within the system provides a protection against failures) and ease of access (each user can access his data from any device connected to the network).

D.J. Veit and J. Altmann (Eds.): GECON 2007, LNCS 4685, pp. 150–163, 2007.

A peer-to-peer backup system has already been proposed and studied in [1], that introduces pStore, a secure distributed backup system based on an adaptive P2P network. pStore exploits unused personal hard drive space attached to the Internet to provide the distributed redundancy needed for reliable and effective data backup. Moreover, support for file encryption, versioning, and secure sharing is provided. Nevertheless, no study on how users would react to such a system is carried out. This paper intents to investigate that particular issue.

Indeed, it seems reasonable to us to assume that each user is selfish, i.e. is only sensitive to the quality of service she experiences, regardless of the effects of her actions on the other users. The framework of *Game Theory* [2] is therefore particularly well-suited to study that kind of interaction among agents: the situation is then studied as a non-cooperative game played among users, where a user strategy is the amount of memory capacity offered to provide service to the other users and the amount of data she stores into the system. Notice that as for other peer-to-peer (P2P) applications, a user valuation for the service depends on the "generosity" of the other users: each user benefits from the others' shared capacity. However, there is no direct incentive to offer one's own capacity to the others, and users are then incentivized to free-ride [3], i.e. benefit from the service without contributing to it: if the sharing efforts do not get some kind of proof of appreciation, nobody has interest to cooperate and the service cannot exist. Therefore it is necessary that some incentive mechanisms be properly designed for the service to actually exist and be valuable for users.

For P2P file sharing systems, there is growing evidence of that need for incentives. For instance, one study of the Gnutella file sharing system showed that almost 70% of the peers only consume resources but do not provide any files [3]. The problem of incentivizing users to contribute in such systems has been the subject of extensive research [4,5,6,7,8].

On the other hand, the existing literature on P2P backup systems mainly focuses on security, reliability and technical feasibility issues [9,10], whereas the incentive aspect received little attention. Notice also that the economic models developed for P2P file sharing systems do not apply to P2P backup services: in file sharing systems, when a peer provides some files to the community, she contributes to the whole system in terms of accessibility. This means that the resource is not dedicated to a certain number of users, but is offered to all the rest of the peers, and in that sense the information stored in the P2P network is a public good. On the contrary, a P2P backup system operates on non-divisible resources, i.e. a certain disk space belongs to one given user (for the time being) and no other peer can access it. The storing resource available on the network is then a private good, and it cannot be managed the same way as a public good from an economic point of view.

The existing models for P2P backup services focus on solutions that do not require financial transactions. Therefore the counter payment for a given service is usually the service in question as well. This approach finally leads to a symmetric scheme where every peer should contribute to the system in terms of service at least as much as she benefits from others [11,12,13].

The incentive part of the scheme proposed in [10] relies on the use of a "probation" period, during which a peer must prove herself reliable before benefiting from the system. A very strict policy based on quotas is suggested in [14]: each peer (identified by her IP address) cannot insert more than a given amount of data into the system. Likewise, the distributed accounting infrastructure proposed in [15] proceeds by simple exclusion of non-cooperating peers from the system, that are detected via an audit.

In this paper, we investigate more flexible solutions, that could still provide peers with incentives to contribute to the system. We focus on the performance of incentive schemes. We propose to study and compare two types of incentive mechanisms that have been suggested in the literature in other - but linked - contexts, like file sharing systems, connection sharing systems, and ad hoc networks: some of those schemes rely on monetary incentives, and some others are based on service degradation for users who do not contribute enough to the service. A particular instance of each type of scheme is considered, and their effects on the overall social welfare are weighted for the particular context of the P2P backup service.

The paper is organized as follows. The model we consider for user preferences is depicted in Section 2, where we also study the maximum reachable value of social welfare yielded by the service. A strict symmetry-based scheme is studied in Section 3, and schemes implementing pricing are investigated in Section 4. The performance of those schemes in terms of social welfare are compared in Section 5. Section 6 presents our conclusions and directions for future work.

2 Model

In this section we describe the model we consider in this paper. We first introduce utility functions that represent user preferences and the decision variables that constitute user strategies. Then we consider the "ideal" situation where users would not behave selfishly and act so as to maximize the total system performance (social welfare). That ideal situation will be used in the next sections as a reference to study incentive schemes.

Note that in this paper, we say that the amount of storage space that is necessary to safely store some data in the system equals the size of those data. This is done without loss of generality: assume that the system introduces a redundancy factor r to improve the data availability on the system, then this is equivalent to replacing C_i^s by C_i^s/r in the user cost functions, or equivalently to replacing C_i^o by rC_i^o in the user valuation function (remark that prices have a different interpretation depending on that choice: they are per unit of physical capacity in the former case, and per unit of "sufficiently available" capacity - taking into account the redundancy - in the latter).

2.1 User Utility Function

We provide here a model for user preferences. The user set is denoted by $\mathcal{I}$, and the perceived utility for a user $i \in \mathcal{I}$ offering capacity C_i^o, storing an amount C_i^s of data in the system and paying a total charge π_i should be a decreasing function of C_i^o and an increasing function of C_i^s. We suggest to use a separable additive function of the utility perceived by a user i, as described in the following definition.

Definition 1. *The utility U_i of a user $i \in \mathcal{I}$ is of the form*

$$U_i\left(C_i^o, C_i^s, \pi_i\right) = V_i(C_i^s) - P_i(C_i^o) - \pi_i, \tag{1}$$

where

- *$V_i(C_i^s)$ is the valuation of user i, i.e. the price she is willing to pay to store an amount C_i^s of data in the system. In this paper we will assume that $V_i(\cdot)$ is positive, continuously differentiable, increasing and concave in its argument, and that $V_i(0) = 0$ (no service yields no value).*
- *$P_i(C_i^o)$ is the opportunity cost of user i for offering capacity C_i^o to the system, i.e. it is the price that she is willing to be paid to devote C_i^o of her disk space to provide service. We assume that $P_i(\cdot)$ is positive, continuously differentiable, increasing and strictly convex, and that $P_i(0) = 0$ (no contribution brings no cost).*

From the valuation and cost functions, we can be derive (by differentiation and taking the inverse functions) two other functions.

Definition 2. *For a user $i \in \mathcal{I}$, we call* demand function *(resp.* supply function*) the function $d_i(\cdot)$ (resp. $s_i(\cdot)$) such that for all $p \in \mathbb{R}_+$,*

$$d_i(p) := \begin{cases} \left(V_i'\right)^{-1}(p) & \text{if } p \leq V_i'(0), \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

$$s_i(p) := \begin{cases} \left(P_i'\right)^{-1}(p) & \text{if } p < \lim_{q \to +\infty} P_i'(q), \\ +\infty & \text{otherwise,} \end{cases} \tag{3}$$

where f' stands for the derivative function of function f.

For a given $p \geq 0$, $d_i(p)$ (resp. $s_i(p)$) is the amount of storage capacity that user i would choose to use from (resp. to offer to) the others if she is charged (resp. paid) a unit price p for it.

Remark that, as intuitively expected, the demand (resp. supply) function is nonnegative and decreasing (resp. increasing) in the unit price.

To carry out a deeper analysis in the next sections, we will assume that the demand functions are of the same form for all users, and only differ through a multiplicative constant. Likewise, we make the same assumption regarding the supply functions.

Assumption A (Common form of supply and demand functions) *There exist a nonnegative and nonincreasing "common" demand function $d(\cdot)$, and a nonnegative and nondecreasing "common" supply function $s(\cdot)$ such that for all user $i \in \mathcal{I}$ there are positive real values a_i and b_i which satisfy*

$$d_i = a_i \times d \tag{4}$$
$$s_i = b_i \times s \tag{5}$$

Moreover,

- $d(0) > 0$ *and* $s(0) = 0$
- $d(\cdot)$ *is strictly decreasing while it takes strictly positive values.*
- $s(\cdot)$ *is strictly increasing (eventually up to a point after which it is constant).*

Notice that the same kind of assumption (i.e. same form of utility functions for all users) is made in [16] in the framework of a P2P file sharing system, for user valuation functions.

Some of our results in the next sections are established for linear demand and supply functions d and v.

Assumption B (Affine demand and supply functions)
- *The common demand function d is affine. More precisely, there exists $\bar{p} > 0$ such that $d(p) = [\bar{p} - p]^+$, where for $y \in \mathbb{R}$, $y^+ = \max(0, y)$.*
- *The common supply function s is linear, i.e. $s(p) = p$.*

Under Assumptions A and B, the demand and supply functions of a user i express as follows:

$$d_i\,(p) = a_i[\bar{p} - p]^+, \tag{6}$$
$$s_i\,(p) = b_i p. \tag{7}$$

This corresponds to quadratic functions for the valuation and cost functions:

$$V_i(C_i^s) = \frac{1}{a_i}\left(-\frac{(C_i^s \wedge a_i\bar{p})^2}{2} + a_i\bar{p}\,(C_i^s \wedge a_i\bar{p})\right) \tag{8}$$
$$P_i(C_i^o) = \frac{1}{b_i}\frac{C_i^{o2}}{2}, \tag{9}$$

where $\wedge$ denotes the min. Finally, we will sometimes consider the following assumption in the case of a large number of users.

Assumption C. *Under Assumption A, the values a_i (resp. b_i) of all users $i \in \mathcal{I}$ are independent and identically distributed. Moreover, a_i and b_i are independent.*

2.2 Social Welfare

A user can choose her own strategy by varying her C_i^s and C_i^o parameters[1]. In this subsection we define social welfare, which will be used later as a performance measure to compare different incentive schemes.

Definition 3. *We call* social welfare *(or welfare) and denote by W the sum of the utilities of all agents in the system:*

$$W := \sum_i V_i(C_i^s) - P_i(C_i^o). \tag{10}$$

Notice that no prices appear in (10). This is because even if we consider a payment-based incentive scheme, we choose to include in social welfare all system agents, eventually including the entity that receives (or gives) payments. The utility of this entity would be its revenue, and all money it exchanges with the users would stay within the system and therefore does not influence social welfare.

Let us have a look at the "optimal" situation that the system can attain (in terms of social welfare maximization). The problem expresses

$$\max_{C_i^s, C_i^o} \left(\sum_i V_i(C_i^s) - P_i(C_i^o) \right), \tag{11}$$

subject to $C_i^o \geq 0$, $C_i^s \geq 0$ for $\forall i$, and

$$\sum_i C_i^o \geq \sum_i C_i^s. \tag{12}$$

This is a classical convex optimization problem, that can be solved by the Lagrangian method.

- The first order conditions imply that for all $i \in \mathcal{I}$, $P_i'(C_i^o) = p$ and $V_i'(C_i^s) = p$, where $p \geq 0$ is the Lagrange multiplier relative to the feasibility constraint (12).
- The complementary slackness condition writes $\min \left[p, \sum_i \left(C_i^o - C_i^s\right)\right] = 0$.

Moreover, p must be strictly positive: otherwise the first order conditions give $C_i^o = 0$ and $C_i^s > 0$ for all i, violating the feasibility constraint (12). We therefore obtain

$$C_i^s = d_i\left(p^*\right), \qquad C_i^o = s_i\left(p^*\right), \tag{13}$$

[1] Since staying online induces a disutility for a user without direct counterpart but improves the quality of the service offered to the others, incentives are needed to honor peers that are online almost all the time. The definition of such incentives is ongoing work, and is out of the scope of this paper. We will not consider it here, assuming that users stay online as much as they can, without trying to minimize the associated costs.

where p^* is the (unique) solution of

$$\sum_i s_i\left(p^*\right) - d_i\left(p^*\right) = 0, \tag{14}$$

and the optimal value of the social welfare is then

$$W^* = \sum_{i \in \mathcal{I}} V_i(d_i(p^*)) - P_i(s_i(p^*)). \tag{15}$$

Figure 1 (displayed in Section 4) gives a graphical interpretation of the maximum social welfare that can be attained by the system. Notice that the Lagrange multiplier can be interpreted as a unit price: if users buy the resource at unit price p^*, and sell their available disk capacity at the same unit price, then the selfish user decisions drive the system to the welfare maximizing solution.

The following result considers our particular assumptions.

Proposition 1. *Under Assumptions A and B, the maximal value W^* of social welfare is*

$$W^* = \frac{1}{2}\bar{p}^2 \frac{\sum_i a_i \sum_i b_i}{\sum_i (a_i + b_i)}. \tag{16}$$

Proof. From (14), we get the social welfare at the ideally fine-tuned unit price:

$$p^* = \bar{p}\frac{\sum_i a_i}{\sum_i a_i + b_i}. \tag{17}$$

Therefore we have $C_i^o = b_i p^* = \bar{p}\frac{b_i \sum_{j\in\mathcal{I}} a_j}{\sum_{j\in\mathcal{I}} a_i + b_i}$ and $C_i^s = \bar{p}\frac{a_i \sum_{j\in\mathcal{I}} b_j}{\sum_{j\in\mathcal{I}} a_j + b_j}$, which gives, after some simplifications,

$$W^* = \frac{1}{2}\bar{p}^2 \frac{\sum_i a_i \sum_i b_i}{\sum_i (a_i + b_i)}, \tag{18}$$

and establishes the proposition.

We therefore have a characterization of the optimal solution. However, as pointed out in the introduction, user selfishness does not lead to this optimal situation when users are not incentivized to offer service to the others. Actually, the unique Nash equilibrium of the game without incentives corresponds to the situation where $C_i^o = 0$ for all i, and the associated social welfare is 0.

In the rest of the paper, we investigate two kinds of incentive schemes, and study their performance in terms of social welfare. We first consider mechanisms without pricing, that simply impose users to provide at least as much memory space as the amount they intend to use (most of the existing related works support this kind of fairness providing approach). Then we turn to payment-based incentive mechanisms where users have to pay for using the service and are paid if they contribute. We finally compare the outcomes of those schemes in terms of social welfare, for some particular types of valuation and cost functions. Since for the two schemes under study the optimal situation cannot be reached in general, we measure the loss of welfare of those schemes with respect to the maximum value.

3 Performance of Schemes Imposing Symmetry

In this section, we follow the ideas suggested in the literature for schemes without pricing. As evoked in the introduction, the principle of those schemes is that users are invited to contribute to, at least as much as they take from, the others. Each user i then chooses C_i^o and C_i^s so as to maximize $V_i(C_i^s) - P_i(C_i^o)$, subject to $C_i^o \geq C_i^s$. As $P_i(\cdot)$ is increasing in C_i^s, no user has an interest to choose a strategy with $C_i^o > C_i^s$. Therefore a user will necessarily choose $C_i^o = C_i^s$. User i maximizes her utility[2] at the point $C_i^s = C_i^o = C_i^*$ where

$$V_i'(C_i^*) - P_i'(C_i^*) = 0. \tag{19}$$

Under our specific assumptions on demand and supply functions, the value of social welfare for such a scheme can be derived:

Proposition 2. *Under Assumptions A and B, the ratio of the social welfare for the symmetric scheme W_{sym} to the maximum social welfare W^* is*

$$\frac{W_{sym}}{W^*} = \left(\frac{1}{\sum_i a_i} + \frac{1}{\sum_i b_i}\right) \sum_i \left[\frac{1}{\frac{1}{a_i} + \frac{1}{b_i}}\right]. \tag{20}$$

Moreover, under Assumption and C, this ratio converges as the number of users tends to infinity

$$\frac{W_{sym}}{W^*} \xrightarrow[|\mathcal{I}|\to\infty]{} \left(\frac{1}{\mathbb{E}[a]} + \frac{1}{\mathbb{E}[b]}\right) \mathbb{E}\left[\frac{1}{\frac{1}{a} + \frac{1}{b}}\right] \tag{21}$$

Proof. We straightforwardly obtain that for all i, $C_i^o = C_i^s = \bar{p}\frac{a_i b_i}{a_i + b_i}$. The corresponding social welfare is then $W_{sym} = \frac{1}{2}\bar{p}^2 \sum_{i\in\mathcal{I}} \frac{a_i b_i}{a_i+b_i}$, and (20) then comes from Proposition 1. The law of large numbers gives (21).

As $f : x, y \mapsto \frac{1}{\frac{1}{x}+\frac{1}{y}}$ is strictly concave, Jensen's inequality implies that $W_{opt} \geq W_{sym}$, and that equality stands if and only if a, b are deterministic, i.e. identical for every user.

4 Performance of Pricing Mechanisms

In this section, we study the influence of introducing a specific pricing scheme for incentivizing users to offer storage capacity, and preventing them from using more capacity than what is available. We consider a simple mechanism: contributors are paid p^o per unit of storage capacity they offer to the system, and service users are charged a unit price p^s when they store their data onto the

[2] Actually, the utility maximization problem for a user is a convex problem, that has the same form as the social welfare maximization problem studied in subsection 2.2, except that there are only two decisions variables (C_i^o and C_i^s) here.

system. Such a mechanism offers users the choice to act as a pure consumer, as a pure service provider, or to both contribute to and benefit from the service. Remark that we will not try here to avoid the presence of a central authority or clearance service: as the model aims to give hints for a commercial application, it is reasonnable to consider the existence of such an entity.

The amount that user i will be charged (this amount can be negative, in which case the user gets paid) is consequently

$$\pi_i = p^s C_i^s - p^o C_i^o. \tag{22}$$

We analyze the model as a full information game, i.e. we assume that the entity that operates the service (the operator) has perfect knowledge of the users and their valuation and cost functions. Therefore, knowing that users will act so as to maximize their utility, it can predict user reactions, and drive the outcome of the game to the most profitable situation for itself. In this sense, the operator acts as the leader of a Stackelberg (or leader-follower) game [2]. We investigate two possibilities: either the coordinator aims at maximizing the user surplus, or it is a revenue-driven monopoly that chooses prices so as to maximize its revenue. In both cases, the feasibility constraint (12) must be satisfied.

Welfare-Maximizing Operator. From our study in subsection 2.2, the theoretically highest level of social welfare can be attained by a payment based scheme. In fact it is reached when the selling and buying prices are the same and equal p^* (see (14)). In that case the operator monetary surplus is null: the operator has no income at all and just acts as a coordinator that redistributes money among users.

Profit-Oriented Monopoly. In this subsection, we assume that the monopoly strives to extract the maximum profit out of the business. The operator therefore faces the following maximization problem.

$$\max_{p^s, p^o} \left(p^s \sum_i d_i\left(p^s\right) - p^o \sum_i s_i\left(p^o\right) \right), \tag{23}$$

subject to $p^s \geq 0$, $p^o \geq 0$ and the constraint (12) that writes $\sum_i s_i(p^o) \geq \sum_i d_i(p^s)$.

This problem is hard to solve for general utility functions, and even under Assumption A since it is not a convex problem. We therefore consider the case where Assumption B holds.

Proposition 3. *Under Assumptions A and B the performance ratio of the social welfare for a profit-maximizing monopoly W_{mon} to the maximum social welfare W^* is*

$$\frac{W_{mon}}{W^*} = \frac{3}{4}. \tag{24}$$

Proof. Under Assumptions A and B, the profit maximization problem becomes a convex problem that can be solved using the Lagrange method for example. However here a simple graphical argument is enough to conclude.

Figure 1 plots two curves: the total demand $D = \sum_i d_i$ and total supply $S = \sum_i s_i$ as functions of the unit price p. First remark that p^o and p^s must be chosen such that $S(p^o) = D(p^s)$: otherwise it is always possible for the operator to decrease p^o (if $S(p^o) > D(p^s)$) or increase p^s (if $S(p^o) < D(p^s)$) to strictly improve its revenue. The operator revenue with such prices is then the area of the rectangle displayed in Figure 1, embedded within a triangle whose area is the maximum value of social welfare. The area of the rectangle is maximum when $S(p^o) = D(p^s) = Q^*/2$ with $Q^* = Dp^*$ and p^* is given in (14). In that case the operator's profit is $W^*/2$, and the total social welfare is $3W^*/4$.

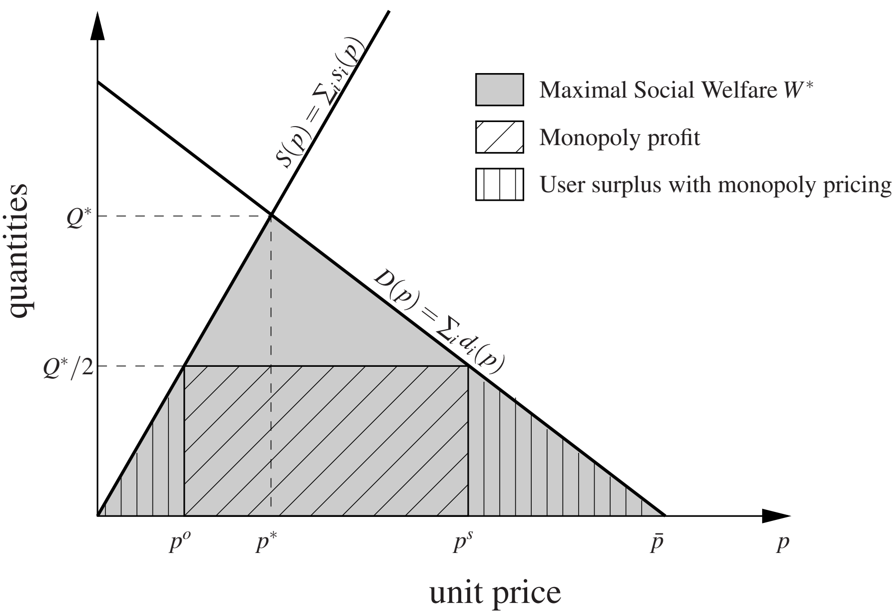

Fig. 1. Maximal social welfare, and optimal choices for a revenue-driven monopoly under Assumptions A and B

5 Does Imposed Symmetry Outperform Monopoly Pricing?

In this section, we compare the two incentive schemes introduced so far in terms of the underlying social welfare at equilibrium. Our point here is not to give necessary and sufficient conditions for one of the two mechanism to provide a larger social welfare than the other for general valuation and cost functions, since the problem becomes much more difficult when Assumptions A and B are relaxed (in particular some optimization problems like revenue maximization are nonconvex and may exhibit local optima). We rather concentrate on our simplifying assumptions and use some examples to highlight situations where one scheme or the other can be better for the overall system.

The following result is a direct consequence of Propositions 2 and 3:

Proposition 4. *Under Assumptions A, B, and C, and for a large number of users, symmetric schemes outperform monopoly pricing if and only if*

$$\left(\frac{1}{\mathbb{E}[a]}+\frac{1}{\mathbb{E}[b]}\right)\mathbb{E}\left[\frac{1}{\frac{1}{a}+\frac{1}{b}}\right]\geq\frac{3}{4}. \tag{25}$$

Proposition 4 highlights in particular the fact that if the population is homogeneous (i.e., a and b are Dirac distributions) then it is better to implement a scheme based on symmetry, since the maximal social welfare can be attained (from Jensen's equality case). On the contrary, user heterogeneity in terms of a and b will make the left-hand side of (25) decrease. If heterogeneity is too important then the left-hand side of (25) may take values below 3/4, which implies that the system (users+coordinator) is better off being driven by a revenue-maximizing monopoly.

At this point of the analysis the distribution of a and b turns out to be the main characteristics of the game. Indeed, these parameters characterize the profile of each user, i.e. her utility and cost functions, and her associated demand and supply functions $(d_i(p), s_i(p))$. In the following we consider two simple examples of distributions (e.g., uniform and exponential) for a and b to illustrate Proposition 4.

Uniform Distribution. We assume here that a (resp. b) is uniformly distributed over $[0, a_{\max}]$ (resp. $[0, b_{\max}]$). In that case we have $\mathbb{E}[a] = \frac{a_{\max}}{2}$, $\mathbb{E}[b] = \frac{b_{\max}}{2}$, and

$$\mathbb{E}\left[\frac{1}{\frac{1}{a}+\frac{1}{b}}\right]=\frac{1}{3}\left(a_{\max}+b_{\max}-\frac{a_{\max}^2}{b_{\max}}\ln(1+\frac{b_{\max}}{a_{\max}})-\frac{b_{\max}^2}{a_{\max}}\ln(1+\frac{a_{\max}}{b_{\max}})\right). \tag{26}$$

The left-hand side of (25) and the plane $z = 3/4$ are displayed on Figure 2 *(left)*. We observe that inequality (25) always holds, thus it is always better for the system to impose symmetry than to introduce a profit-maximizing monopoly.

Exponential Distribution. We now consider the case where a (resp. b) follows an exponential distributions with parameter μ_a (resp. μ_b). Therefore $\mathbb{E}[a] = \frac{1}{\mu_a}$, $\mathbb{E}[b] = \frac{1}{\mu_b}$, and after some calculation we obtain

$$\mathbb{E}\left[\frac{1}{\frac{1}{a}+\frac{1}{b}}\right]=\begin{cases}\frac{1}{3\mu_a} & \text{if } \mu_a=\mu_b,\\ \frac{\mu_a^2-\mu_b^2-2\mu_a\mu_b\ln(\frac{\mu_a}{\mu_b})}{(\mu_a-\mu_b)^3} & \text{otherwise.}\end{cases} \tag{27}$$

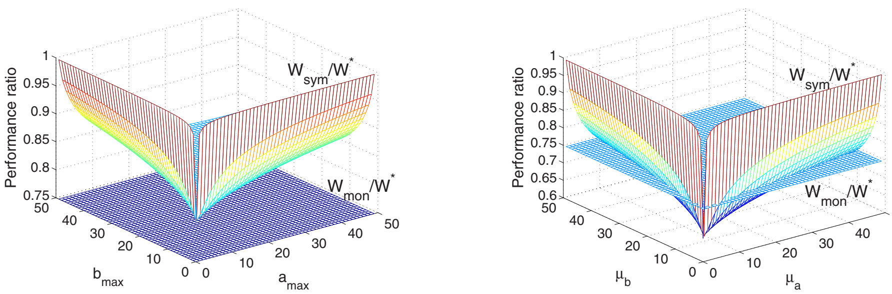

Fig. 2. Social welfare performance ratio of monopoly pricing *(plane surfaces)* and symmetric scheme *(curved surfaces)* as given by Propositions 2 and 3, for uniform *(left)* and exponential *(right)* distributions of a and b

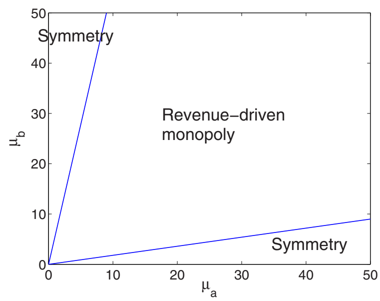

Fig. 3. Best scheme (in the sense of social welfare) for exponential distributions of a and b (μ_a and μ_b are the parameters of the exponential laws)

We again compare both terms of (25) on Figure 2 *(right)*. This time, there is no scheme that always outperforms the other: depending on how much the two variables differ, the monopoly can provide higher welfare than a symmetric scheme. More precisely, when μ_a and μ_b are sufficiently close, then a revenue-maximizing monopoly will drive the system to a situation where the social welfare is larger than what a symmetry-based scheme would have yielded. Figure 5 indicates the scheme that gives the best system welfare, depending on the values of μ_a and μ_b.

6 Conclusions and Future Work

In this paper, we have proposed an economic model for a peer-to-peer backup service. Assuming that users of such a service behave selfishly, we have justified the need for incentive schemes for the system to effectively exist. We have studied and compared two kinds of incentive schemes, namely a symmetry-based scheme (users should contribute to the service as much as they use it) and a pricing-based scheme (introduction of a monopoly that fixes unit prices for buying and selling resource). Under some simplifying assumptions, we have highlighted conditions for one scheme to outperform the other in terms of social welfare. Some examples have shown that user heterogeneity plays a crucial role on the better-suited scheme. Basically, it seems that more user heterogeneity would justify the use of (even profit-driven) pricing.

There remains a lot of work to be done on this subject. First, we have made quite restrictive assumptions on the form of the utility functions to derive some results (we often used quadratic valuation and cost functions). It would be interesting to obtain some results for more general cases. Moreover, we only considered that users had two decision variables, namely the quantity of data they store into the system and the amount of storing capacity they offer. We are currently working toward an extension of the model where users can also choose the proportion of time they are online (a larger availability improves the service offered to the others, but increases the perceived cost of the user), that also needs an incentive scheme.

Acknowledgment. This work has been partially funded by the GET project DisPairSe.

References

1. Batten, C., Barr, K., Saraf, A., Treptin, S.: pStore: A secure peer-to-peer backup system. Technical Report MIT-LCS-TM-632, MIT Laboratory for Computer Science (2001)
2. Fudenberg, D., Tirole, J.: Game Theory. MIT Press, Cambridge, Massachusetts (1991)
3. Adar, E., Huberman, B.: Free riding on gnutella. Tech. rep. Xerox parc (2000)
4. Anagnostakis, K., Greenwald, M.: Exchange-based incentive mechanisms for peer-to-peer file sharing. In: Proc. of 24th International Conference on Distributed Computing Systems (ICDCS'04), Tokyo, Japan (2004)
5. Cohen, B.: Incentives build robustness in bittorrent. In: Proc. of 1st Workshop on Economics of Peer-to-Peer Systems (P2PECON'03), Berkeley, CA (2003)
6. Golle, P., Leyton-Brown, K., Mironov, I., Lillibridge, M.: Incentives for sharing in peer-to-peer networks. In: Proc. of 3rd ACM conference on Electronic Commerce (EC'01), Tampa, Florida, pp. 264–267 (2001)
7. Lai, K., Feldman, M., Stoica, I., Chuang, J.: Incentives for cooperation in peer-to-peer networks. In: Proc. of Workshop on Economics of P2P Systems (2003)
8. Vishnumurthy, V., Chandrakumar, S., Sirer, E.: Karma: A secure economic framework for peer-to-peer resource sharing. In: Proc. of 1st Workshop on Economics of Peer-to-Peer Systems (P2PECON'03), Berkeley, CA (2003)
9. Druschel, P., Rowstron, A.: PAST: A large-scale, persistent peer-to-peer storage utility. In: HotOS VIII, Schloss Elmau, Germany, pp. 75–80 (2001)
10. Lillibridge, M., Elnikety, S., Birrell, A., Burrows, M., Isard, M.: A cooperative internet backup scheme. In: Proc. of 1st Workshop on Economics of Peer-to-Peer Systems (P2PECON'03), Berkeley, CA (2003)

11. Cox, L., Noble, B.: Pastiche: Making backup cheap and easy. In: Proc. of Fifth USENIX Symposium on Operating Systems Design and Implementation, Boston, MA (2002)
12. Cox, L., Noble, B.: Samsara: Honor among thieves in peer-to-peer storage. In: Proc. of 19th ACM Symposium on Operating Systems Principles (SOSP'03), Bolton Landing, NY (2003)
13. Stefansson, B., Thodis, A., Ghodsi, A., Haridi, S.: MyriadStore. Technical Report T2006:09, Swedish Institute of Computer Science (2006)
14. Dabek, F., Kaashoek, M.F., Karger, D., Morris, R., Stoica, I.: Wide-area cooperative storage with CFS. In: Proc. of 18th ACM Symposium on Operating Systems Principles (SOSP'01), Chateau Lake Louise, Banff, Canada (2001)
15. Ngan, T., Nandi, A., Singh, A., Wallach, D., Druschel, P.: On designing incentives-compatible peer-to-peer systems. In: Proc. of 2nd International Workshop on Future Directions in Distributed Computing (FuDiCo II), Bertinoro, Italy (2004)
16. Courcoubetis, C., Weber, R.: Incentives for large peer-to-peer systems. IEEE JSAC 24(5), 1034–1050 (2006)