Proceedings of the 9th International Conference on Applied Informatics
Eger, Hungary, January 29–February 1, 2014. Vol. 1. pp. 213–217
doi: 10.14794/ICAI.9.2014.1.213

A P2P based storage system with reputation points and simulation results*

B. Bakondi, P. Burcsi, P. Györgyi, D. Herskovics, P. Ligeti†, L. Mérai, D. A. Nagy, V. Villányi

Eötvös Loránd University, ELTECRYPT Research Group

Abstract

Within this paper a secure peer-to-peer (P2P) based storage system is presented with additional features and some simulation results. The main motivation for such a system is to provide the functionality of a cloud storage without having to rely on central servers or large datacenters. Users of the proposed P2P storage system are able to store their encrypted files at other users and vice versa, providing availability at the cost of some extra storage space. The system motivates fair behaviour by using reputation points. We investigate the possible optimal user behaviours and the bandwidth/space trade-offs based on simulation results.

Keywords: P2P network, secure storage, distributed hash-table

MSC: 94A60, 91A22

1. Introduction

Decentralized storage has been in use for more than a decade now, e.g. by KaZaA or BitTorrent ([2]). In classical file sharing services however, the required content usually need not be encrypted and usually belongs to or can be read by several users. More recently, P2P storage software for securely storing private user files is also available by few enterprises. Wuala ([5]) addresses reliability issues by using a hybrid storage system: they provide storage space for money or in exchange for the user's space. One's files are then stored on servers and on other users' machines. Symform ([4]) has a similar scheme. Neither of these schemes are known to handle selfish behaviour of users storing the data of others. Space Monkey ([3])

*The research was carried out as part of the EITKIC 12-1-2012-0001 project, which is supported by the Hungarian Government, managed by the National Development Agency, financed by the Research and Technology Innovation Fund and was performed in cooperation with the EIT ICT Labs Budapest Associate Partner Group. (`www.ictlabs.elte.hu`)

†The author was partially supported by the grant OTKA PD-100712.

sells dedicated devices (external hard drives equipped with their software) that implements the storage system in a P2P manner. They still rely on servers for handling user connections. Oualha and Roudier [1] proposed a secure P2P storage system based on a new self-organizing micropayment system and sophisticated verification processes.

2. Proposed P2P Storage System

We propose a simple decentralized P2P storage system and address several potential risks. First, unauthorized data access should be impossible. This can be achieved by using secure encryption schemes before uploading data to the P2P network. Second, data safety (availability) is needed. To ensure this, we introduce a reputation point scheme between any pair of users, instead of a global currency or some consensus based accounting.

The proposed system uses distributed hash table (DHT) for connections. The users have logarithmic number of neighbours and the diameter of the network is also logarithmic.

The users publish their download requests on their own billboard and provide the chunks that a neighbour requests and they have. The upload is very similar: the user simply requests his or her own file. A user only deletes a file when he gets a new file that is more important for him.

Since the storage space of the users is limited, thus each user has to make decisions about which chunks it is willing to store. This decision is taken according to the user's strategy. We propose a strategy that bases the decision on prior behaviour of the neighbour users. The reputation points provide knowledge of the neighbours' prior behaviour, therefore a reputation point is not a global currency, it is just a payment between pairs: if Alice provides the data that Bob needs, then she has better chance to have her data stored by Bob.

3. Simulation

The P2P network can be simulated by a graph. The goal of the simulation is to investigate user strategies and answering the following questions. Does good behaviour pay off in the long run? What is individually a best strategy, and does the individually optimal choice of strategies lead to a globally acceptable/optimal situation?

We make the following simplifying assumptions, in order to get a manageable simulation:

Hypercube. We assume that the nodes are connected in a hypercube pattern. The hypercube shares several properties with a typical P2P network graph: small diameter, logarithmic number of neighbours, short cycles.

Discrete time steps. We assume that requests are announced and served in rounds. In each round, every node announces one single request and looks at its neighbours. Then according to its probabilistic strategy, it fulfils some requests and denies others.

Short memory. The probabilistic strategy only depends on the neighbours' behaviour in the preceding round. It is governed by two parameters p and q, giving the probability of fulfilling a request in the cases when the requesting neighbour behaved nicely/badly in the preceding round.

Individual evaluation. The balance of a round for an arbitrary user u is $\delta t + \phi n_{ff} - r$, where $t = 0$ if none of the neighbours fulfils the request of u, otherwise $t = 1$. n_{ff} is the number of neighbours that fulfil the request, δ and ϕ are conversion parameters and r is the number of the fulfilled requests. We do not consider bandwidth costs in this model.

Strategy tuning. The simulation consists of phases and each phase consists of $numberofRounds$ rounds (a parameter; it was always 100 in our simulations). After every phase, the users modify their strategies based on previous balances. Every user chooses a new strategy for the next phase according to a distribution determined by the average balances of the users who followed that strategy in the previous phase. After every phase we drop some strategies according to a parameter $perc$: if a strategy has an average balance among the last $perc$ percent, we drop it.

4. Settings

During the simulation we used the following settings of the parameters:

- The dimension of the hypercube was always 10.
- Every user chooses a pair (p, q) from $\{0, 0.1, 0.2, \ldots, 1\}^2$ uniformly at random, under the assumption $p \geq q$.
- The conversion parameters are:
  - $\delta \in \{0, 10, 20\}$,
  - $\phi \in \{0, 1, 2\}$.
- The number of the rounds was always 100.
- $perc = 10$ in every case.
- We analysed every settings according to 3 aspects in the first 30 phases:
  - the average values of p and q,
  - the average number of the rounds when there is at least one neighbour who stores the data of a user ('Good Rounds'),
  - the average number of the fulfilled requests.
- We simulated every settings 5 times, the presented results are the mean of these values.

5. Results

We illustrated our results in three diagrams. These results show that the proposed P2P system is operable if it is worth for the users storing at most twice as many data on their devices than they want to store in the system (the $\delta = 0$, $\phi = 2$ case).

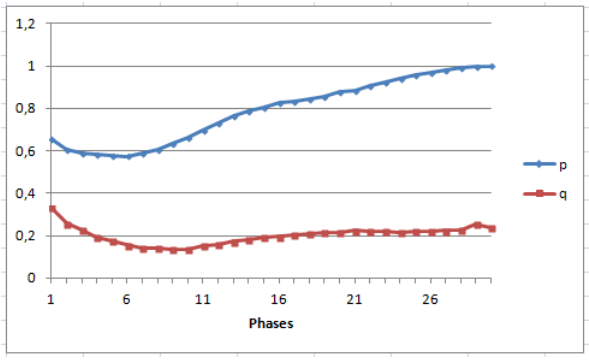

Figure 1: The average value of p and q, in case of $\delta = 0$ and $phi = 2$.

Figure 1 demonstrates that in case of $\delta = 0$ and $\phi = 2$ in the best strategies p is 1, i.e. it is optimal for the user to fulfil the requests of the nicely behaving neighbours. If we increase δ or ϕ, then $p \to 1$ faster, but this is not worth illustrating.

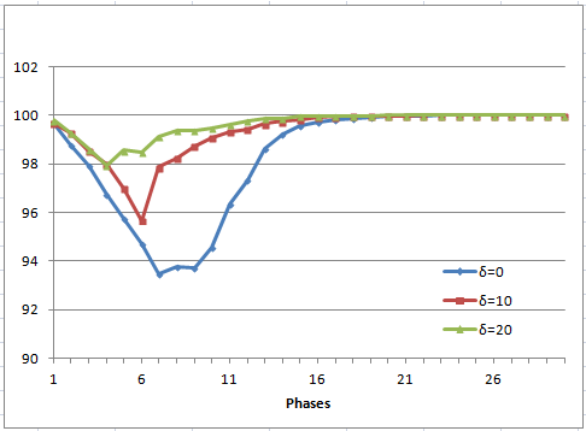

Figure 2: Average number of 'Good Rounds' according to the different value s of δ ($\phi = 2$).

Figure 2 shows that it is very unlikely that there aren't any neighbours who fulfil our request, if $\phi = 2$. However, it is interesting that the average number of the 'Good Rounds' is decreasing in the first few phases. This is due to the fact

that there are a lot of users in these phases with a 'selfish' strategy (a strategy is *selfish*, if the user do not want to fulfil requests) and many other users do not fulfil many of the requests of the selfish users (because q is relatively small).

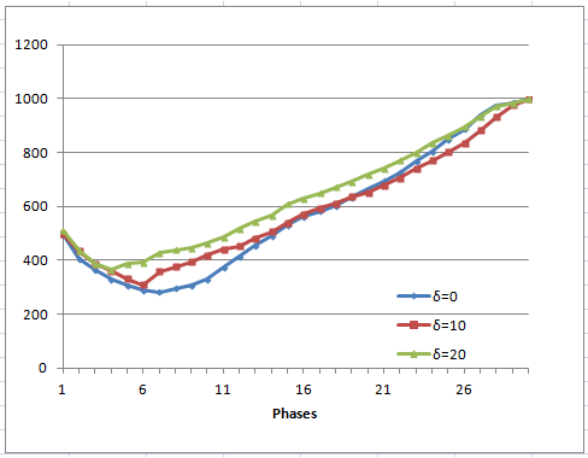

Figure 3: Total number of the fulfilled requests of an average user in eac h phases according to the different values of δ ($\phi = 2$).

The above mentioned property explains the temporary decrease of the total number of the requests, presented in Figure 3. This diagram also shows that the system is not sensitive to δ.

6. Conclusions

We proposed a P2P based data storage system, where the users are motivated to a fair behaviour by reputation points. One of the biggest advantages of the proposed system against other systems is its simplicity. The simulation results are quite positive, and in the future we want to test the proposed system in other models e.g. when the users do not choose values p and q, but determine an upper bound for the size of data that they tend to store and select among the requests in a quite easy way.

References

[1] N. Oualha, Y. Roudier: *Securing P2P storage with a self-organizing payment scheme*, **LNCS 6514** (2011) pp. 155–169.

[2] `http://www.bittorrent.com/`

[3] `https://www.spacemonkey.com/`

[4] `http://www.symform.com/`

[5] `http://www.wuala.com/`