Incentives in Peer-to-Peer and Grid Networking *

Torsten Ackemann*, Richard Gold†, Cecilia Mascolo* and Wolfgang Emmerich*

* Department of Computer Science
University College London
Gower Street
London WC1E 6BT, UK
{T.Ackemann|C.Mascolo|W.Emmerich}@cs.ucl.ac.uk

† Department for Computer Systems
Uppsala University
Box 325
SE-75105 Uppsala, Sweden
rmg@docs.uu.se

Abstract

Today, most peer-to-peer networks are based on the assumption that the participating nodes are cooperative. This works if the nodes are indifferent or ignorant about the resources they offer, but limits the usability of peer-to-peer networks to very few scenarios. It specifically excludes their usage in any non-cooperative peer-to-peer environment, be it Grid networks or mobile ad-hoc networks. By introducing soft incentives to offer resources to other nodes, we see an overall performance gain in traditional file-sharing networks. We also see soft incentives promoting the convergence of peer-to-peer and Grid networks, as they increase the predictability of the participating nodes, and therefore the reliability of the services provided by the system as a whole. Reliability is what is required by Grid networks, but missing in peer-to-peer networks.

1 Introduction

In recent years, a new paradigm for networking architectures has been introduced, or rather re-introduced. This moves from the dominant client-server architecture, where one or more centralized servers serve a number of clients, to a system where any node can take either roles. This model is referred to as *peer-to-peer*.

Peer-to-peer became popular quickly, not so much because of the technology behind it, but for the services it offers. The early peer-to-peer systems were file sharing applications, used to exchange music files and later also other types of files. In these applications, people would let their computers participate in the network, so that every user could download the other user's audio files, and vice versa. After two years of operation of the first of those platforms, Napster, millions were using it.

Obviously, the reason for the popularity of peer-to-peer networks is the availability of a huge number of files. This was due to the high availability of popular files as they were present on a large percentage of users' hard drives. The inherent redundancy of people sharing their own private collections meant that it was statistically likely that there would be someone located nearby who would have the file that is wanted. The upshot of this attribute is that the more users there are downloading a particular file, the more available it will become for all users of the network. Due to the inherent scalability of this approach, the average download speed remained relatively constant, even during times of massive growth of the peer-to-peer network.

Solutions have emerged for speeding up searching, most of them using distributed hash tables, but despite this, much leaves still to be desired by the actual download speed in those systems, which is very often far from the line speed maximum, especially considering that the latest generation of file sharing software allows the limitation of the bandwidth available to the peer-to-peer network. Moreover, if a download is slow, there is nothing the user can do to improve the speed.

A number of factors play a role in download speed for a particular file to a particular host, not least the hardware network limitations of the participating nodes. But besides these hard limits, the performance can be influenced by the decisions of the peers. The average download performance for a given file will increase the more nodes offer the file, and the more of their up-link bandwidth those nodes offer to the peer-to-peer network. This is especially important since a huge percentage of participants in file sharing networks are so called "freeloaders", who offer very few files or no files at all. And since the total upload speed of all nodes combined equals the total download speed, these freeloaders effectively lower the average performance of the network.

*Torsten Ackemann is supported by a research studentship of Kodak Ltd.

There are systems which have implemented anti-freeloader mechanisms, notably the now defunct MojoNation [16] and eDonkey2000 [6], but these systems are concerned with fighting freeloaders to achieve “fairness”, and not with improving the average performance and reliability of the system.

So what it takes to make the nodes offer more files and more resources, and subsequently improve the average download speed, is an *incentive* for them to do so. The difference to the freeloader prevention mechanism is that here, the nodes are given incentives to share as many files and as much bandwidth as possible, instead of punishing those who offer nothing or very little.

Along with the increased *performance* by introducing incentives comes an increased *reliability*. When there is a reason to share files, users become predictable, and thus it is not pure chance anymore whether a given file is available in the network. This predictability of the peers makes it possible to use peer-to-peer architectures in non peer-to-peer areas which have a peer structure but are dominated by either centralized approaches or by the requirement of a cooperative environment, specifically Grid networks and mobile ad-hoc networks.

In this paper, we will investigate the existing types of incentives, and how they affect the overall service quality, be it download speed or service reliability, in the different peer-oriented networks. The networks we look at range from pure peer-to-peer networks to Grid networks and mobile ad-hoc networks. We will then present as an example a way of improving the average download speed in the Gnutella network, based on our findings.

Section 2 will give an overview of incentives mechanisms used in peer-oriented networks and a classification of incentives. In Section 3, we apply our analysis results and describe a way to enhance an existing file sharing network in order to improve the average download speed, using bartering. Finally, we will discuss our conclusions and future work in this area in Section 4.

2 Incentives in Existing Systems

2.1 Examples of the Use of Incentives

Incentives in File Sharing

There is a multitude of different file sharing networks, and some include a notion of fairness and mandatory sharing, while the majority do not.

Most of the current peer-to-peer systems share the same use case, which is concerned with individual users using either one computer at home or at work. This means that every node in an overlay network is being seen as one computer controlled by one individual, where this individual does not control any other node in the peer-to-peer network.

All what the users do in this scenario is downloading relatively small amounts of data, for example a number of audio files. The quality of the service they receive, which is the download speed in this case, is usually not very important to them. Even if the download of a file is slow, the user is unlikely to cancel it. The tolerance to the service quality is generally high. However, users would certainly like to speed up the downloads if they could do so easily.

Since most current peer-to-peer systems do not include any kind of incentive for the users to offer files, there are only three reasons why users share files. The first reason is some kind of idealism, that people offering files are needed to keep the system working. The second reason is indifference: The files are on the users’ hard drive already, and the up-link capacity is usually underused anyway. The third reason is ignorance i.e., the user is not aware that uploads are taking place from his machine.

Incentives in other Peer-to-Peer Networks

There is a rising number of peer-to-peer networks which are not for file sharing, such as the Internet Indirection Infrastructure [14], or the Jabber instant messenger [10]. These peer-to-peer networks are based on a cooperation model of resource usage. There are no incentive models in place in any of these networks. Any node performs any task given to it by other nodes, be it searching, uploading, or forwarding of messages.

Incentives in other Peer-Oriented Networks

Grid networks are currently a huge trend in supercomputing and distributed computing. Here, the unused CPU capacity of a number of computers is being used for a computational cluster, by running a client application on each computer, which makes the unused CPU cycles available to applications. The analogy here is that of a power grid, where unused generated electricity just disappears, if it is not consumed.

This model has many analogies to the peer-to-peer model, where many independent computers are connected through a virtual network, in order to use each other’s services. However, in this case the service is the provision of CPU time, or more generally resources, and not the provision of particular files.

On the organisational layer, another difference to peer-to-peer networks is that often, many machines belong to one entity i.e., all the computers of a computing centre are administered by the same person. Grid networks are usually formed only within one organisation, or as a collaboration of a small number of organisations [8].

The “incentive” for organisations to offer the usage of their Grid networks to other institutions is that of externally billing the other party for the resource usage. Alternatively, organisations combine two or more Grids to a larger one and are using the resulting, larger Grid in turns. In any case, the incentives are completely external to the system.

Notable exceptions are the highly distributed computational Grids such as SETI@home [13]. In these systems, many private users participate out of idealism, and the reason that the organisers of these Grid applications cannot use one of the "professional" Grids is precisely because they cannot provide enough of their incentive i.e., money. This model only works for a few highly prominent tasks, which need to require vast amounts of resources to make the effort of promoting them worthwhile.

There is obviously a gap between the plain billing model and the idealism driven model. This is where an incentive based approach would fit in, which provides a model which is decentralized, based on internal incentives, rather reliable with respect to the service quality, and scalable in terms of the amount of resources needed for a task.

The users of the network, in this case companies or organisations, require a reliable and predictable environment in order to make use of the network. This is very different to the private user downloading an audio file, where the download speed, and often even the quality of the audio file, is usually not so important.

Mobile ad-hoc networks also have a peer-oriented structure. The mobile nodes are organizing themselves and forward messages for each other. Again, there are no incentives for the nodes to actually do so, everything depends on the cooperation of the participating nodes. Current research into Mobile Ad-Hoc Networks [11] relies on the assumption that nodes collaborate by forwarding traffic for others because they wish their own traffic to be forwarded.

Some research has already been done to encourage cooperation between nodes in such networks. For example, the Nuglets [2] work uses tamper-proof hardware in order to overcome the problem of not having a single trusted entity in such a decentralized network. This is an expensive solution which is probably fairly fragile given the susceptibility of such hardware to being circumvented [1].

2.2 Incentives Taxonomy

We will now have a closer look at which incentives and similar mechanism have been used, and categorize the incentive models into two groups. The first group is those which are based on no or weak incentive models, where weak means that there is no actual representation of an incentive in the network. The second group comprises the strong models, where some kind of service offering is mandatory.

We can identify three different weak models:

relying on the user this is the "classic" model, where the users are not required to offer services. Most of the prominent file sharing tools can be subsumed under this point, namely Napster, Gnutella and Freenet. The same holds true for all other peer-to-peer applications such as Jabber [10] and PeerCast [12].

higher good in this model, the users are only offering a service, without being able to consume any services. This model is only being used for computational Grids and computational peer-to-peer networks, e.g. SETI@home [13].

common interest this model is similar to the previous one, except that the result of the shared services have a direct advantage for the participants. For example, owners of the digital VCR TiVo organised themselves to crack the password which enables access to the backdoor functions of the device [5].

Besides the weak models, there is also a number of strong models. We identified three different types of strong incentive models:

micro currency here, a virtual currency is established between all nodes. Services between nodes (such as download, searching, disk space) are then paid in this currency. There needs to be a central authority of some kind, which issues this currency, and means to avoid inflation need to be in place as well. In fact, a whole micro-economic model is needed. The now defunct MojoNation [16] used this model.

external billing as opposed to the other incentive models we discussed is only concerned with accounting within the network. The actual billing takes place outside the network. This model is what is predominantly used in Grid networks, often without a detailed accounting.

force sharing the users are invariably offering files (or part of files) themselves, should they download files from other users. The only systems which are using this model are eDonkey2000 [6] and its successor Overnet [17]. An upload/download ratio is in place i.e., for each kB/s of maximum download speed, a certain amount of kB/s of upload speed has to be made available. The relationship between the two is not linear though, with unlimited download speed above a certain upload speed (10 kB/s currently).

The first three weak models are not really helpful in increasing the fairness and performance of the peer-to-peer system, since they require a motivation external to the network. While this works in some highly specific scenarios, it cannot be relied on, and the offered resources will very likely not exceed a minimal amount in non-cooperative networks.

Of the strong models, the micro currency approach is certainly the most complex one. Many additional economical problems are being introduced with a virtual currency. Security then also becomes a major concern. All in all, it is doubtful whether its advantages outweigh the numerous disadvantages. There is also the problem that having a currency leads to concrete prices for all services, which in turn undermines the "soft incentives" concept. Tschudin showed a potential way out of this, by giving each node its own currency [15].

External billing, the second strong model, also introduced a number of new problems. Most notably, it requires an external billing structure in place, which relies on proper accounting within the peer-to-peer network. The accounting in turn requires a unique ID for each node for proper authentication and authorisation. There also has to be a defined price in a real currency for each service, which again breaks the "soft incentives" concept.

The "force sharing" model has none of these problems. No logging is required, which makes in scalable, and the services do not have a concrete price, leaving the "soft incentives" concept intact. However, in the case of eDonkey2000 [6], it only encourages offering a minimum amount of resources, as there is no reward for offering additional resources.

In the next section, we will present an architecture which combines a "force sharing" model with "soft incentives" on top of the already existing Gnutella network.

3 Applied Incentives for File Sharing

The next step after the incentives taxonomy is now to actually apply the incentive-based approach in a real network. Out of the three possible network types, namely Grid networks, mobile ad-hoc networks and peer-to-peer networks, we chose to start with latter. This is due to the fact that there are already widely used implementations of large scale peer-to-peer networks, namely file-sharing networks. One of the most popular ones, which is also completely decentralized and for which open-source clients exist, is the Gnutella network [9].

Our goal is to extend Gnutella with a soft incentives model i.e., users should get rewarded with an increased download speed for offering additional up-link bandwidth.

3.1 Requirements

There are a number of essential requirements. First of all, our improvements have to be scalable, or rather not break the scalability of Gnutella. Most importantly this means that we cannot save state in the nodes, so we cannot keep track of other nodes, other than for the current transaction. The additional temporary state as well as the additional communication necessary between nodes should be kept to a minimum.

Having an incentive-free peer-to-peer system works fine to a certain extent, and it is clearly a good foundation for a file sharing network. Yet, we want to increase the average download speed. We want to do that by providing incentives for users to offer additional files or resources, and by attracting additional users. On the other hand, we do not want to drive users away who do not offer files, since in the long run every user of the system will make the network grow, if only by increasing the user base by word of mouth. It might also only be a temporary condition of a user not to offer anything, e.g. users new to the network usually do not offer anything until they downloaded files themselves.

Thus, it makes sense to always have a minimal service being provided, even for freeloaders. Whenever users want to improve their download speed though, they can do so by simply increasing the service quality they offer, by either increasing the upload bandwidth, or by increasing the number of files they offer. This is somewhat similar to the concept of super-peers as introduced in Gnutella, but here users get rewarded for making their nodes super-peers.

Some of the currently implemented systems are very concerned with keeping the system fair. The way this fairness is achieved is that each service or resource is metered or defined to an exact internal currency equivalent (which might be a virtual currency or just an upload/download ratio), and for each unit of service usage the correct amount of this currency has to be "paid", in some form. Without being able to compensate for the service, a node is thus unable to use it.

While fairness is a good idea, we think that it is not necessary for every service exchange to be fair, and that this fine-grained definition of fairness is, in fact, actually limiting the overall performance of the system. Instead of a "quid pro quo" compensation, we argue that almost any kind of roughly equivalent compensation is enough. A collaboration does not necessarily have to be fair, as long as all parties profit from it. In the long run, the system must be fair though, in order to be accepted by the users. By fairness we mean that there needs to be a balance between consumed and offered resources.

Fairness is not only important as part of the system design, but it is also vital to make circumvention of these fairness mechanism impossible or at least very hard. More concrete, it should be on an architectural level impossible for nodes to get more resources than the other nodes are willing to give to them.

In a peer-to-peer network, all readily available resources tend to be used (logging into any peer-to-peer network and seeing the up-link being saturated within a minute or two proves this point). Yet, the download speed is often far from optimal i.e., from saturating the down-link. The performance bottlenecks seem to be either the performance of the used links of the underlying network, or the lack of nodes which host a specific requested file (or the inability to find all those nodes hosting the file).

Experience with peer-to-peer networks suggests that the participants are willing to share resources which they do not need anyway, and which cannot be saved for future use, namely unused CPU cycles, and up-link bandwidth. Users seem not to be comfortable with offering disk space for random public usage, except for offering downloaded files which are on their hard disks, of course. Losing relatively small amounts of useful resources, such as the CPU time for processing search requests (Gnutella), or the link capacity for downloading another data block (SETI@home), seem to be tolerated though. Thus, efficient replication strategies as suggested by [3] might not be suitable for the inherently multi-domain private user use case.

Since there is no control over the peers, the only way to stimulate the offering of additional files and bandwidth is to provide the users with an incentive to do so. The only existing system-inherent incentive, is to improve the quality the users receive themselves. Therefore, a direct connection between the *received* quality and the *offered* quality is needed.

Of course, in this specific case where we build on top of Gnutella, we also have to be careful not to break compatibility with nodes running unmodified Gnutella clients.

Another important aspect is the use of *feedback loops* in the system. As we said before, all the currently used incentive-based approaches are just freeloader prevention mechanisms. The nodes get denied service for not sharing files, instead of being awarded for sharing more files. The nodes are especially not being rewarded for sharing more than the average. Thus, nodes tend to share only just as much as needed to receive the maximally realistic performance of the network. This results in a *negative feedback loop*, where less and less resources per node are being offered. Instead, the system should implement a *positive feedback loop*, where nodes are rewarded for offering additional resources, not punished for offering less resources. That means that the more a node offers, the better the received performance gets.

3.2 Bartering Rings

One way of providing a direct feedback between downloading and uploading is to include a trading mechanism into the peer-to-peer network. If two nodes discover that they are downloading files from each other, they can increase their mutual resource usage 1).

These direct trading mechanism have been discussed already in [4], and there are practical implementations of this concept, e.g. in eMule [7]. However, this approach showed some problems, namely the lack of finding suitable trading partners in the first case, and the necessity of keeping a log of all the past transactions with all peers in the second case.

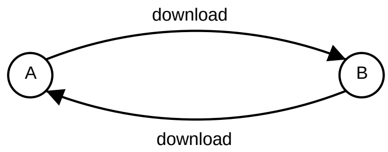

Figure 1: Trading between two Nodes

In most cases, it is very hard to find two nodes which actually need each other's service. While this may be possible if the nodes need generic resources such as computing power or storage space, as it is the case in current Grid networks, we imagine it to be very rare that two nodes are interested at the same moment in a file which the other node has and vice versa.

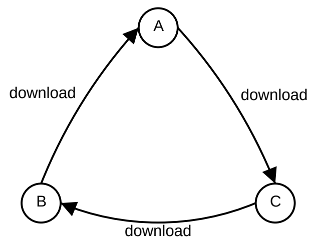

Figure 2: Bartering Ring with 3 Nodes

We generalize the concept of a direct trading mechanism to include a larger number of hosts. The more nodes are involved, the more likely they have content of mutual interest. So instead of node A receiving a file from node B and in exchange send another file to node B (Figure 1), these two parts of the transaction are decoupled, and the receiving and the transmitting end of one's nodes transaction are not the same. Instead, all nodes form a downloading ring. In a ring of three nodes, node A might receive a file from node B, while node B receives one from node C, and node C from node A (Figure 2).

In general, these rings are cycles in a directed graph of downloads. When these rings are found, the participating nodes can then improve the service quality (download speed or download queue position) they perceive from their upload ring neighbour by improving the service quality they provide to their download ring neighbour. Ideally, this improved quality should not go at the expense of the other downloading nodes (those which are not part of this particular ring), but by allocating more resources of the underlying network to the peer-to-peer network. Inevitably, this is the case to some extent though, but there will still be an incentive to increase the actual hard limit of the percentage of the bandwidth allocated to the peer-to-peer application.

Since we build on top of the existing Gnutella, we try to reactively find these rings in already existing download relationships. Given the structure of the Gnutella network, a prototype implementation of this is relatively straightforward, as all that is needed is to send all the uploading nodes a list of the downloading nodes. This is enough to find 3 node rings. For rings with more nodes, it is necessary for the nodes to forward the uploading nodes lists they receive to their downloading nodes.

4 Conclusions and Future Work

After we introduced bartering rings, the next step is to verify the validity of this approach. We plan to do this by collecting search queries and responses from the Gnutella network with a modified client, and then analyzing this data with respect to how many of these rings of different sizes actually exist. This will give us an idea of the potential of the concept. We also plan to run simulations for different replica distributions, in order to find out about the performance increase our system could provide. This is necessary in order to quantify the performance gain by providing incentives to share additional files, as opposed to additional up-link bandwidth.

If this analysis shows our concept to be promising, we will enhance a Gnutella client to actually support our principles. Then, we can investigate in the dynamic behaviour and how the users influence the system with their decisions. We will also see whether the bartering rings only lower the download speed for those who do not offer enough, or whether it actually makes users use more of their bandwidth for the peer-to-peer application.

Once we established the concept of bartering, we can extend it to be used in Grid networks, and later also in mobile ad-hoc networks. Along with the use in Grid networks goes a generalisation from file sharing to a generic service model. Quite likely though, the development of different incentive-based architectures might required, since bartering rings are not very useful for generic resources. In mobile ad-hoc networks on the other hand, performance is more critical than reliability.

Performance of the bartering rings approach could also be improved by using a proactive approach i.e., to actively look for these rings in the download initialisation phase already, and possibly start downloading the file from another node if this helps to form a ring.

Furthermore, we plan to investigate whether an increased availability of files will provide a increased statistical reliability for the file availability. Having a provably reliable peer-to-peer network would enable new use cases.

To conclude, we argue that the area of sharing incentives in peer-to-peer networks has been neglected up to now. Many peer-to-peer networks implement the concept of sharing incentives in one way or another, but the concept and implementations used were not very systematic, and have only been implemented for defensive reasons of freeloader prevention and introducing "fairness".

Properly recognizing the concept of incentives, and applying it to different peer-oriented networks, ranging from classical file sharing to Grid networks, will potentially enhance both the quality and the reliability of the services in those networks.

References

[1] R. Anderson and M. Kuhn. Tamper Resistance - a Cautionary Note. In *Proceedings of the Second Usenix Workshop on Electronic Commerce*, pages 1–11, Nov 1996.

[2] L. Buttyan and J. Hubaux. Stimulating Cooperation in Self-Organizing Mobile Ad Hoc Networks. Tech. Rep., 2001.

[3] E. Cohen and S. Shenker. Replication Strategies in Unstructured Peer-to-Peer Networks. In *Proceedings of ACM SIGCOMM 2002*, Aug 2002.

[4] B. Cooper and H. Garcia-Molina. Peer-to-peer data trading to preserve information. *Information Systems*, 20(2):133–170, 2002.

[5] DClient. http://edo.lcs.mit.edu/dclient.

[6] eDonkey2000. http://www.edonkey2000.com.

[7] eMule. http://www.emule-project.net.

[8] I. Foster and A. Iamnitchi. On Death, Taxes and the Convergence of Peer-to-Peer and Grid Computing. http://people.cs.uchicago.edu/ anda/papers/foster_grid_vs_p2p.pdf, 2002.

[9] Gnutella Community Portal. http://www.gnutella.com.

[10] Jabber. http://www.jabber.org.

[11] IETF MANET working group. http://www.ietf.org/html.charters/manet-charter.html.

[12] PeerCast. http://www.peercast.org.

[13] SETI@home. http://setiathome.ssl.berkeley.edu.

[14] I. Stoica, D. Adkins, S. Zhuang, S. Shenker, and S. Surana. Internet Indirection Infrastructure. In *Proceedings of ACM SIGCOMM 2002*, Aug 2002.

[15] C. Tschudin. Talk on Computational Economies for Auto Mobile Code. http://www.icsi.berkeley.edu/t̃schudin/ps/ucsb-talk9802.ps.gz, 1998.

[16] B. Wilcox-O'Hearn. Experiences Deploying a Large-Scale Emergent Network. In *Proceedings for the 1st International Workshop on Peer-to-Peer Systems (IPTPS '02)*, March 2002.

[17] Overnet. http://www.overnet.com.