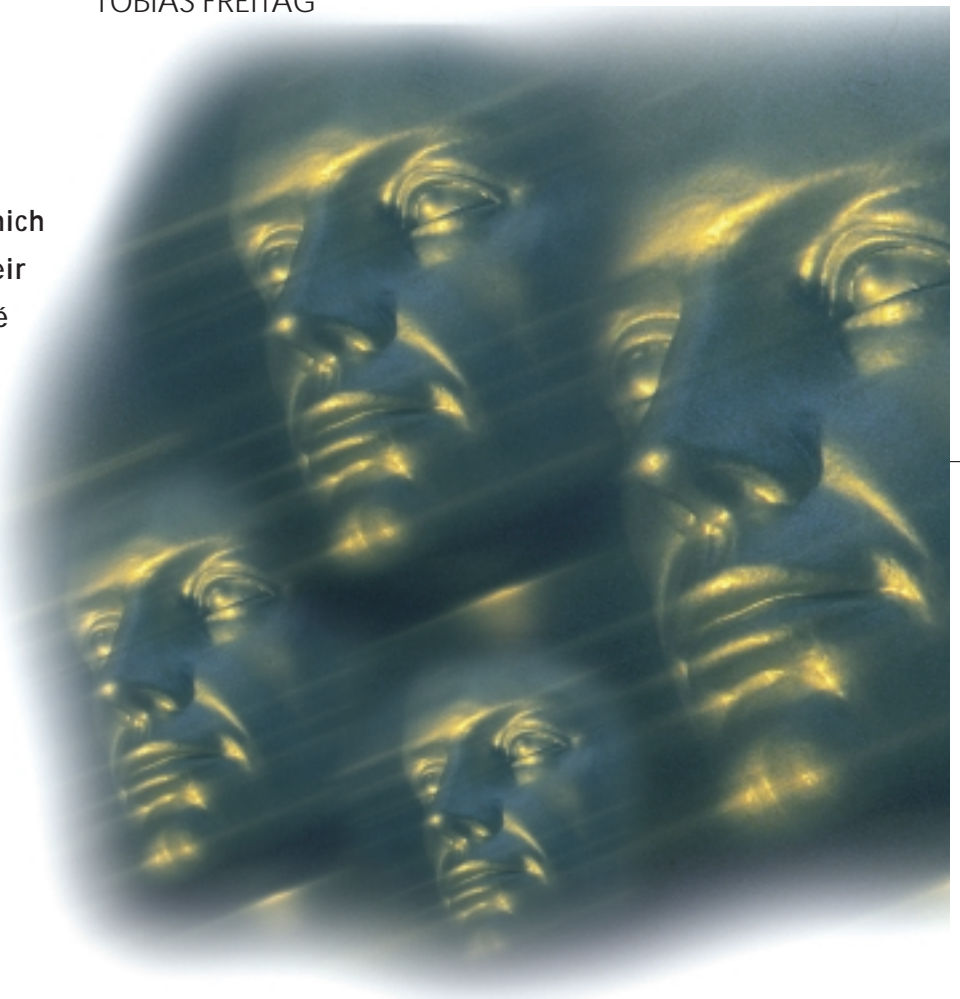


# Distributed File sharing, part I

# COLLECTIVE MEMORY

TOBIAS FREITAG

The whole world as a file system which everyone can access and can reach their home directory even from an internet café in Timbuktu – this, or something similar, is what the developers of Freenet, Gnutella and Co. imagine the world could be with their software. This two-part article aims to show how close they are to achieving this goal.



It's a logical idea: a globalised world should have a global data archive. The idea is realised in each of the many projects connected with it, usually by a small program that, in contrast to normal file transfer programs, not only transfers data from other computers but is also able to provide it.

Users themselves are solely responsible for what kind of data they transfer. Often, only the basic system is stipulated. This is the case with probably the best-known operator of a file sharing network. Napster, a company set up two years ago by an American student and now used by almost 5 million people, only allows MP3 files to be transferred. Unfortunately, it currently faces legal proceedings – initiated by the heavy metal band Metallica among

others – which aim to prevent the service from transferring copyrighted music. This is because users wishing to build up a music collection with the help of Napster or the other networks (see table 1) do not actually take any notice of copyright.

Until now the court case has had the opposite effect of that intended. According to Media Metrix's estimates, the number of users rose from 1.1 million in January to 4.9 million in July of this year making Napster the fastest growing application ever registered on the internet. Nevertheless, the question remains whether the service will be closed in the coming months. For precisely this reason the *gnutella* project was formed on a website hosted by AOL.

## REPORT

## FILE SHARING



Shawn Fanning, just nineteen years old, founded Napster together with his uncle

In contrast to Napster, this service aims to allow users to transfer all kinds of files and not to rely on one central server to produce and manage the index of all the distributed data. Instead, a "real" network is in operation. Each participant produces their own index, and search requests are forwarded to the next clients, which in turn forward these to their network neighbours. The reasoning behind this is that if there is no central server, a court cannot shut down the network.

But this is where the system reaches its limits. When it seemed that Napster was about to be shut down, thousands of former Napster users flooded to the "rival" Gnutella, whose capacities were put to the test for the first time. And it seems that the concept behind the protocol used had not been properly thought through. The web publisher Clip2 published a report according to which Gnutella had now come up against a "modem barrier". The many search requests forwarded by each client have slowed the throughput to such an extent that it is no longer possible to download properly, and this affects not only modem owners but also all the other participants, as their search requests are also forwarded, and so delayed, by the "weak" modem users.

### Classified information

Though Gnutella claims to be open source and part of GNU, the reality appears to be a little different. Even when development began, when AOL still accommodated the Gnutella website directly on one of its servers, the developers postponed the publication of the source code until the distant future. They said the code would be released once they had reached version 1.0. But after more than half a year's development work they are only up to version 0.56.

Even if the developers could finally make up their mind to make their program open source, they fear that AOL could intervene. After all, the software was developed by the company Gnullsoft, backed by employees of the company Nullsoft (particularly well-known in Windows circles for its MP3 player *Winamp*), which in turn was bought up some time ago by AOL. The crucial detail: AOL, following the merger with Time Warner if approved by the competition regulators, will be one of the major players in the music business.

### Freenet

Freenet is quite clearly open source and completely open to anyone wishing to participate. The project was first conceived by Ian Clarke while he was studying for his doctorate and now has a few "nodes" or installations. The system was immediately implemented in Java for preview purposes, and a graphical user interface is under development. However, the client isn't yet suitable for serious use and the indexing and searching with what are called keys doesn't even seem to be able to satisfy several developers.

The Freenet nodes don't just form an interface with the released and exported resources. They also automatically replicate data where it is frequently requested. However, as they only index some of the data available on the network, it is possible that data for which there is insufficient demand will disappear from the network completely.

Therefore, this approach is not suitable for the "Eternity Service" (based on a paper by Ross J. Anderson) which is said to keep data for all eternity. The Eternity Service attempts to combine the advantage of longevity with that of modern data communications by storing the data on servers scattered around the world. There seems to be a need for it, particularly in view of the ever decreasing lifespans of the storage media used today to store all kinds of data regardless of whether it is a medical report or birth certificate.

However, Anderson actually had quite different ideas about his vision. He draws a comparison with the first translations of the Bible, which were one of the main reasons for the Reformation, the subsequent social upheaval and everything that followed. He says that, then as now, a technical innovation – in this case the printing press – ensured that the acquired information could not be suppressed or lost. He therefore requests that the system protect the anonymity of its users and not allow any government or other institution the opportunity to take the information away from the world.

### ... and all the rest

The idea of sharing something among equals on the peer-to-peer network goes even further. Just as users of file search services only share hard disk

Table 1: Common file-sharing networks

Name	Founded in	OpenSource	Average data volume available	Users*	Globbering search	Special features
Freenet	June 1999	yes	GByte	c. 200	no	encrypted transfer possible
Gnutella	1997	planned	20 TByte	200-800	in part	
Napster	May 1999	no	4 TByte	300; 4.9 million in total (as at July 2000)	no	central server, only MP3 files permitted.
Scour	September 1997	no	20 TByte	70	no	central server
*All hosts that can be searched simultaneously when operating normally				see text	" depends on the support provided by the host searched	

space with other users, unused computer time could be redistributed. However, the projects which have focussed on this subject to date, such as Seti@HOME or Distributed.net, don't distribute computer capacity but use it for a special purpose. Distributed.net, for example, is currently attempting to crack the RC5 encryption algorithm.

The GnuSpace project hosted on Sourceforge intends to redistribute the CPU time to the user. Unfortunately, the project is still in its infancy and so usable programs can't be expected for some time.

There is no shortage of ideas about where to go with the concept of "distributed resources". The best example was O'Reilly's Peer-to-Peer Summit where Hank Barry, CEO of Napster, Andy Herzfeld from Eazel and representatives of IBM, Microsoft, Red Hat and Gnutella among others came together in the middle of September. Also participating in the event was Stanford professor Lawrence Lessig, who has made a name for himself with the sociology papers he has published on computers and society in particular. This informal meeting was also concerned with investigating the potential of distributed systems for technology and society and countering the idea that this kind of technology would only make it easier for pirates to make copies.

The participants agreed that the use of dormant memory capacity, computer time and transfer bandwidth through peer-to-peer networks and

### Info

#### **Eternity Service implementation:**

<http://www.cypherspace.org/~adam/eternity/>

**Another implementation:** <http://www.kolej.mff.cuni.cz/~eternity/>

#### **The Eternity Service, Ross J. Anderson:**

<http://www.cl.cam.ac.uk/users/rja14/eternity/eternity.html>

**GnuSpace Project:** <http://gnuspace.sourceforge.net>

#### **O'Reilly's Peer-to-Peer Summit:**

<http://www.oreillynet.com/pub/a/linux/2000/09/22/p2psummit.html>

technologies can only be useful. Due to the increase in the amount of data, it will be increasingly important in future to be able to categorise data sensibly. XML is only the beginning, and it must be made much easier for users to furnish newly produced data with the correct meta data straight away, they said.

This technology obviously harbours a considerable degree of dynamism. By the time you read this a verdict will have been reached in the Napster case, and there may be other networks too. There will certainly be new clients and implementations. We will go into more detail on the latter next time. ■

**Table 2: Useable filesharing clients/programs**

Network	Linux clients	Homepage	Special features
Freenet	official Java client	<a href="http://freenet.sourceforge.net/">http://freenet.sourceforge.net/</a>	Java client
Gnutella	gnubile	<a href="http://www.gnutelladev.com/source/gnubile/">http://www.gnutelladev.com/source/gnubile/</a>	Linux client of developers, can be used for uploads
	gnujatella	<a href="http://gnujatella.sourceforge.net/">http://gnujatella.sourceforge.net/</a>	Java client
	gnut	<a href="http://www.mrob.com/gnut/">http://www.mrob.com/gnut/</a>	console client
	gtk-gnutella	<a href="http://gtk-gnutella.sourceforge.net/">http://gtk-gnutella.sourceforge.net/</a>	
	hagelslag	<a href="http://TieFighter.et.tudelft.nl/hagelslag">http://TieFighter.et.tudelft.nl/hagelslag</a>	console client, GUI planned
Napster	gnapster	<a href="http://jasta.gotlinux.org/gnapster.html">http://jasta.gotlinux.org/gnapster.html</a>	
	gnome-napster	<a href="http://gnome-napster.sourceforge.net/">http://gnome-napster.sourceforge.net/</a>	
	gtk-napster	<a href="http://www.geocities.com/xilliator/">http://www.geocities.com/xilliator/</a>	
	iNapster	<a href="http://members.optusnet.com.au/~iwade/inapster/">http://members.optusnet.com.au/~iwade/inapster/</a>	Web interface with Napster
	javaNapster	<a href="http://www.mp3s4u.f2s.com/jnapster/">http://www.mp3s4u.f2s.com/jnapster/</a>	Java client, search requests possible on several servers at the same time
	jnap	<a href="http://www.perham.net/mike/jnap/">http://www.perham.net/mike/jnap/</a>	Java client
	jNapster	<a href="http://members.nbc.com/harikris_v/">http://members.nbc.com/harikris_v/</a>	Java library for Napster access
	jnerve	<a href="http://jnerve.sourceforge.net/">http://jnerve.sourceforge.net/</a>	Java server
	knapster	<a href="http://knapster.netpedia.net/">http://knapster.netpedia.net/</a>	
	Linux Napster Client	<a href="http://www.gjs.net/~nite/">http://www.gjs.net/~nite/</a>	console client
	Lopster	<a href="http://lopster.sourceforge.net/">http://lopster.sourceforge.net/</a>	can be used for uploads
	MyNapster	<a href="http://mynapster.sourceforge.net/">http://mynapster.sourceforge.net/</a>	Web interface with Napster
	TekNap	<a href="http://www.teknap.com/">http://www.teknap.com/</a>	
XNapster	<a href="http://www.xnapster.com/server.html">http://www.xnapster.com/server.html</a>	Opennap/Napster server	
Scour	gsx	<a href="http://freshmeat.net/projects/gsx/homepage/">http://freshmeat.net/projects/gsx/homepage/</a>	
	JavaScour	<a href="http://freshmeat.net/projects/javascour/homepage/">http://freshmeat.net/projects/javascour/homepage/</a>	Java client
	JScour	<a href="http://freshmeat.net/projects/jscour/homepage/">http://freshmeat.net/projects/jscour/homepage/</a>	Java command line
	Scour Media Agent	<a href="http://freshmeat.net/projects/scourmediaagent/homepage/">http://freshmeat.net/projects/scourmediaagent/homepage/</a>	Perl wrapper for SMB download from Scour Windows server