AGnuS: The Altruistic Gnutella Server

Daniel Hughes, Ian Warren, Geoff Coulson
Computing Department, Lancaster University, Lancaster, UK
d.r.hughes@lancaster.ac.uk | iw@comp.lancs.ac.uk | geoff@comp.lancs.ac.uk

Abstract

The first generation of peer-to-peer file sharing systems followed the traditional client-server paradigm. However, legality and scalability issues have driven the development of decentralized file sharing protocols; the most popular of these being Gnutella. To date, such systems have been unable to match the Quality of Service (Qos) offered by centralized architectures. AGnuS improves QoS on Gnutella by increasing file availability, improving network friendliness and increasing file quality. This is achieved by layering caching, load balancing, content-based routing and filtering services on top of the core Gnutella protocol.

1. Introduction

Napster and other file sharing clones use a hybrid model of interaction involving servers. However, scalability problems and legal issues have made the development of semi-centralized services unviable. The most successful decentralized system is Gnutella[1]. However, it suffers from a number of QoS issues:

Scalability: The broadcast search mechanism employed by Gnutella does not scale well; bandwidth consumption rises exponentially with the number of queries. For this reason, queries are assigned a time to live (TTL) value that limits how far they propagate through the network. This technique limits the number of results that nodes receive in response to queries.

The Tragedy of the commons: Rational users with limited resources will always try to maximize the benefit they accrue from using the network, donating few of their limited resources. This behavior, known as "Tragedy of the Commons"[2], reduces QoS for each user of the network.

Uneven resource distribution: Adar and Huberman[3] suggest that 50% of all files on Gnutella are served by just 1% of nodes. This makes the actual architecture of the Gnutellanet closer to a client-server model than a true peer-to-peer system. Consequently, the resulting network suffers from single points of failure, problems associated with flash crowds[4], and resource unavailability.

Poor quality files: Gnutella networks may distribute two varieties of poor quality files: masquerading files that simply waste bandwidth, and malicious files that also threaten system security.

2. AGnuS

AGnuS[5] was designed to address some of the shortcomings of Gnutella. AGnuS is a specialized Gnutella host which improves QoS both across the Gnutellanet and for the system's primary user. It does this while remaining compatible with the core Gnutella protocol and hence maintains compatibility with the thousands of existing nodes that conform to the Gnutella protocol.

In particular, AGnuS offers four QoS improvements. First, AGnuS nodes locate a greater number of files, when compared to a standard Gnutella node, in response to any given Query request. Second, AGnuS attempts to prevent a node's available bandwidth being exhausted by processing incoming query requests. Third, AGnuS provides support to prevent users from downloading low quality files that include masquerading and malicious files. Finally, the collective set of mechanisms implemented in AGnuS leads to a reduction in the time taken to download files matching any given search term.

	Increased file availability	Improved network friendliness	Increased file quality	Improved file acquisition time
Caching	✓			✓
Load balancing		✓		✓
Content based routing	✓			✓
File filtering		✓	✓	✓

Table 1: QoS improvements and enabling mechanisms

Table 1 identifies the four QoS enhancing mechanisms employed by AGnuS. The table also reveals that these mechanisms are mutually supportive in terms of the above QoS improvements they address.

Load balancing: The load balancing mechanism routes a greater volume of messages to those peers that are more able to process requests. This ensures a node's bandwidth

is not exhausted by responding to incoming query messages. The mechanism also reduces variation in bandwidth consumption across nodes, providing more consistent and predictable response times to users.

Content based routing: The content based routing mechanism directs search requests to those nodes that are most likely to yield matching files. The consequent benefits of this mechanism include increased file availability and improved file acquisition time.

Caching: The caching mechanism aims to distribute popular files more evenly over the Gnutellanet, resulting in increased file availability and a reduction in the time required to locate files. In addition, this mechanism reduces the effects of flash crowds and tends to restore the peer-to-peer architecture one associates with Gnutella.

Heuristic filtering: Heuristic filtering improves file quality by removing low quality files from further consideration by AGnuS nodes. This avoids the caching of poor quality files and reduces bandwidth consumption.

3. Performance

In this section we evaluate AGnuS in terms of the QoS improvement claims introduced in Section 2.

Increased file availability: To evaluate this claim, we conducted a test using an AGnuS node and a standard Gnutella node. The two node types were connected to a Gnutellanet for a 24 hour period. At 2 hour intervals, each node initiated a search for the most popular current file on Gnutella and the number of corresponding QueryHit messages was recorded. The AGnuS node consistently located more files than its standard Gnutella counterpart. The average number of hits received per search was 76 for the standard Gnutella node and 128 and for the AGnuS node, giving a 68% improvement in file availability for AGnuS nodes.

Improved network friendliness: Here we focus on the effectiveness of AGnuS for regulating a node's bandwidth consumption. Three one-hour tests were conducted using an AGnuS node and a standard Gnutella node. The load on an AGnuS node's peers was found to be much less volatile, with the average standard deviation in host load being reduced from 194 queries per minute to just 91. Thus AGnuS nodes gave a 56% reduction in load variation when compared to a standard Gnutella node, reducing the likelihood of nodes on low-bandwidth connections being flooded.

Increased file quality: We tested an AGnuS node with popular search terms spanning multiple file types. We ran the tests three times with the node connected to a different point in the network for each test. We then repeated the tests using a standard Gnutella node. In all cases, AGnuS' filtering mechanism significantly increased the proportion of quality files that are downloaded. Specifically, the number of quality files downloaded increased by: 22% for audio files, 82% for video files, 16% for program files and 70% for text files.

4. Conclusions

In this paper, we have reported on AGnuS, an enhanced Gnutella node which remains backwards compatible with the Gnutella protocol. AGnuS improves the QoS experienced by both individuals and the Gnutellanet, and is thus likely to be acceptable to rational users.

Using four mutually supportive mechanisms, AGnuS offers the following QoS improvements over standard Gnutella nodes:

- *Increased file availability*. Caching enables files to propagate over the Gnutella network while content based routing allows search requests to be directed to hosts most likely to satisfy them.
- *Improved network friendliness*. Load balancing reduces the probability that individual nodes will be flooded by query requests. Furthermore this mechanism provides individual users with more predictable response times for search requests. The heuristic file filtering mechanism contributes to improved network friendliness by preserving bandwidth.
- *Increased file quality*. In addition to reducing bandwidth consumption, the heuristic file filtering mechanism discards low quality files.
- *Improved file acquisition time*. Improved resource distribution, more effective use of bandwidth and a greater range of nodes from which to download files all contribute to a reduction in file acquisition time.

5. References

[1] **The Gnutella Protocol Specfication**, available at http://dss.clip2.com/GnutellaProtocol04.pdf

[2] G. Hardin, **The Tragedy of the Commons.** *Science volume 162, pp. 1243-1248,* 1968.

[3] E. Adar and B. A. Huberman, **Free Riding on Gnutella.** *Fist Monday Oct. 2000.*

[4] L. Zhang et al**, Adaptive Web Caching,** *Proceedings of the 1997 NLANR Web Caching Workshop*, April 1997

[5] D. Hughes, **AGnuS: The Altruistic Gnutella Server for Improving QoS**, Technical report, Computing department, Lancaster University, 2003.