Freenet

Will Holcomb

CSC 203 - November 4, 2002

Freenet is a piece of software designed to provide "totally free" communications across the Internet. It is a large scale peer-to-peer network that pools member computers so as to create a virtual information store open to anyone to publish any information of any kind. It is designed to be impossible to monitor or control aside from simply disallowing its usage.

Freenet is different from other peer-to-peer systems like say gnutella, in that you are not sharing your files with other people, rather you are providing some resources to the global freenet and how those resources are used is dependent on the behavior of people on the net.

The backbone of freenet is a set of "persistent" nodes that are always connected. Anyone can start up a persistent node and tell it to join into the global freenet based on a list of nodes that can be retrieved off on the net. Once a node joins it takes whatever space it was allowed and creates an encrypted cache where it will store all of it's information. The cache is not available to the person running the node (at least not easily). Information may be inserted into the network at a given node, but as the information is requested at different places it will be copied to other nodes.

Freenet nodes know very little about the world. When searching with a traditional p2p client the search goes out and what comes back is the IP address of the computer that has the information that you are looking for. A transfer is then set up between those two endpoints and you get your file. With freenet anonymity is very important, so there is no traceability either to who made the request of who provided the information. When a request goes through the freenet, each node only knows to ask the nodes around it and when the file is coming back each node copies the file to its cache and then sends it to the node that requested it.

This means that as a request comes in a node doesn't know if the request originated with the asking node or if it has been floating around the freenet for a hundred queries. There is a time to live (TTL) value, which is decremented with each hop, but it is randomly set initially and may be randomly changed by any node, so it cannot be used in any useful way.

There is an advantage to the way that freenet works beyond providing anonymity. Nodes know what is in their cache at a given time and as information is requested it gets better spread through more caches and so the popularity of a piece of information is directly proportional to its availability (as opposed to normal networks with a centralized host where the inverse is true).

There is also several disadvantages as well. A big one is there is a size constraint on the amount of information that may be stored in a node. That means that if enough stuff comes through then some things will have to get pushed out. If you have your website published on your persistent node then there is no way to keep it from being pushed out if it is not requested enough. The accepted solution is to write a bot to regularly reinsert information to make sure it stays present, but this is quite a pain compared to keeping a regular website up.

Another issue goes at the basic philosophy on which freenet is founded. Most people would agree that anonymous publishing is important in say, a country where an oppressive government will come hunt down and jail anyone publishing dissenting views. Very few people would say though that anonymous publishing is valuable to protect child pornographers from being apprehended. Freenet protects them both equally however, and it is pretty much impossible to create a system that could effectively protect one and not the other. How does the value weigh against the cost?

Resources:

- The Freenet Project - http://freenetproject.org/
- A Distributed Decentralised Information Storage and Retrieval System by Ian Clarke - http://freenetproject.org/freenet.pdf