Using the Small-World Model to Improve Freenet Performance

Hui Zhang, Ashish Goel, and Ramesh Govindan
Information Sciences Institute and Department of Computer Science
University of Southern California
Marina del Rey, CA 90292-6695

A Peer–to–peer networked system is a collaborating group of Internet nodes which overlay their own special–purpose network on top of the Internet. Such a system performs application–level routing on top of IP routing. These systems share some of the same characteristics as the Internet in that they can grow to be quite large, may need to utilize distributed control and configuration, usually employ a naming scheme that allows them to address a node without knowing its exact whereabouts, and possess a routing mechanism that allows each node to meaningfully communicate with the rest of the system. Such systems can have interesting and unanticipated properties. We taxonomize these systems into two categories: structured systems where the assignment of a key to the node on which its corresponding data is stored is determined by the structure of the key space, and unstructured systems where no such assignment exists. Systems like CAN [5], CHORD [6], and OceanStore [7] are examples of the former, and Freenet [1] is an instance of the latter.

In this paper we study the impact of route-cache replacement policy on the performance of the Freenet [1] in terms of two metrics- request hit ratio and average hops per successful request. We find that the hit ratio (the likelihood of finding the datum associated with a key within a fixed number of network hops) in Freenet is crucially dependent on the local policy used to manage the cache of data (called *datastore* in Freenet) and the routing table. A standard LRU–like cache replacement policy can result in significantly low hit ratio under high load (i.e. when the number of files stored in the network is high). This is an important finding. While memory is cheap these days, Freenet-like peer-to-peer systems will always have limits on cache sizes. Thus, examining the performance degradation of such systems under high load becomes important, especially, for example, in the context of understanding the immunity of these systems to denial-of-service attacks.

To understand this particular performance degradation in Freenet, we look at Freenet's internal routing tables when LRU is used and observe that the entries in a routing table are not highly clustered. This is undesirable since the success of the Freenet routing algorithm relies on a high clustering of the routing table entries. We then observe that the routing tables can be shaped by changing the cache replacement policy at each node; this does not include any changes to the Freenet routing protocol. Additionally, we use intuition from the small world model [2][3] which says that the routing distance in a graph is small if each node has pointers to its geographical neighbors (causing clustering) as well as some randomly chosen far away nodes. Based on the intuition from the small-world models and the theoretical result in [4],, we propose a simple clustered cache replacement scheme with a small amount of randomization.

We then perform a simulation study to compare the two schemes under high loads. Our proposed cache replacement scheme results in a significantly higher hit ratio compared to LRU, and also a significantly small number of average hops per request. The proposed scheme and LRU result in similar values for the average number of hops per successful request. It is interesting that such a small change in the system can lead to a significant improvement.

We also develop an idealized model of Freenet under our cache replacement scheme based on the model in [4]. We prove that the expected message delivery time in this idealized model is $O(\log^2 n)$, provided we have $O(\log^2 n)$ amount of memory per node. Here n is the number of nodes id the system. More structured systems such as CAN or CHORD [5][6]achieve a delivery time of O(log n) with O(log n) memory. It is surprising that an unstructured system such as Freenet can come so close to the structured systems, even in a highly idealized model. We believe that a theoretical understanding of unstructured systems in more realistic models is an important open problem.

References

[1] Freenet. http://freenet.sourceforge.net/
[2] J. Kleinberg. Navigation in a small-world. Nature, 406, 2000.
[3] D. J. Watts and S. H. Strogatz. Collective dynamics of `small-world' networks, Nature 393, 440--442 (1998).
[4] J. Kleinberg. The small-world phenomenon: an algorithmic perspective. Cornell Computer Science Technical Report 99-1776, 2000.
[5] S. Ratnaswamy, P. Francis, M. Handley, R. Karp, and S. Shenker. A scalable content- addressable network. ACM SIGCOMM, 2001.
[6] I. Stoica, R. Morris, D. Karger, F. Kaashoek, and H. Balakrishnan. Chord: A peer-to-peer lookup service for internet applications. ACM SIGCOMM, 2001.
[7] J. Kubiatowicz, D. Bindel, Y. Chen, S. Czerwinski, P. Eaton, D. Geels, R. Gummadi, S. Rhea, H. Weatherspoon, W. Weimer, C. Wells, and B. Zhao. Oceanstore: An architecture for global-scale persistent storage. Proceedings of the Ninth international Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2000), November 2000.