Efficient and Scalable Client-Clustering for Proxy Cache

Kyungbaek Kim, Woo Jin Kim, and Daeyeon Park

Department of Electrical Engineering & Computer Science,
Division of Electrical Engineering,
Korea Advanced Institute of Science and Technology (KAIST),
373-1 Kusong-dong Yusong-gu, Taejon, 305-701, Korea
{kbkim, wjkim}@sslab.kaist.ac.kr and daeyeon@ee.kaist.ac.kr

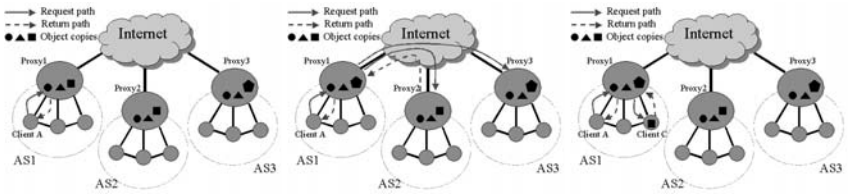
Abstract. Many cooperated web cache systems and protocols have been proposed. These systems, however, require expensive resources, such as external bandwidth and proxy cpu or storage, while inducing hefty administrative costs to achieve adequate client population growth. Moreover, a scalability problem in the cache server management still exists.
This paper suggests peer-to-peer client-clustering. The client-cluster provides a proxy cache with backup storage which is comprised of the residual resources of the clients. We use DHT based peer-to-peer lookup protocol to manage the client-cluster. With the natural characteristics of this protocol, the client-cluster is self-organizing, fault-tolerant, well-balanced and scalable. Additionally, we propose the Backward ICP which is used to communicate between the proxy cache and the client-cluster, to reduce the overhead of the object replication and to use the resources more efficiently.
We examine the performance of the client-cluster via a trace driven simulation and demonstrate effective enhancement of the proxy cache performance.

1 Introduction

The recent increase in popularity of the Web has led to a considerable increase in the amount of Internet traffic. As a result, the Web has now become one of the primary bottlenecks to network performance and web caching has become an increasingly important issue. Web caching aims to reduce network traffic, server load, and user-perceived retrieval delay by replicating popular content on caches that are strategically placed within the network.

By caching requests for a group of users, a proxy cache can quickly return documents previously accessed by other clients. Using only one proxy cache has limited performance, because the hit rate of the proxy is limited by the cache storage and the size of the client population. That is, if a cache is full and needs space for new documents, it evicts the other documents and it will retrieve the evicted documents from the Internet for other requests. In Figure 1(a), if the square object is evicted, the proxy cache obtains it from the Internet. But if the near proxy cache has a square object like that in Figure 1(b), Proxy 1 can obtain it from Proxy 2 and reduce the latency and the Internet traffic. According to this procedure, multiple proxies should cooperate with each other in order to increase the total client population, improve hit ratios, and reduce document-access latency; that is the cooperative caching.

M.M. Freire, P. Lorenz, M.M.-O. Lee (Eds.): HSNMC 2003, LNCS 2720, pp. 83–92, 2003.



(a) Initial Situation (b) After Square is evicted from Proxy 1 (c) Client C stores the evicted Square

Fig. 1. Request and Response path when client A requests the Square object

Various cooperative caching systems have been proposed in [2], [3], [4]. However, these techniques need high bandwidth, expensive infrastructure and high administrative cost. ICP-based cooperative caches communicate with other caches that are connected by busy core-links, which are the inter-proxy links, to find and obtain requested objects in other caches. Even if the requested objects are not in these caches, they spend bandwidth of core-links in order to find the objects. Some cooperative caches use the proxy cluster, as a single large cache so as to be overprovisioned to handle bursty peak loads. However, this approach still needs too much administrative cost for the frequent variation of clients. For example, a growth in client population necessitates increasing the cluster size and updating the cluster information.

In this paper, we suggest a new web caching system which uses the residual resources of clients. In Figure 1(c), not only the proxy cache but also the clients are responsible for storing objects; the proxy cache stores more popular objects and the client-cluster stores evicted objects from the proxy cache. That is, the client-cluster is used as a backup storage for the proxy cache. In this case, Client A can get the square object from Client C, which is inside the network, not outside of it. This behavior reduces the usage of core-links and improves the performance of the proxy cache, in terms of the hit rate, the byte hit rate and the reduced latency. Furthermore, the size of the backup storage of the proxy increases as more clients use the proxy. According to this feature, this approach reduces the administrative cost and makes the proxy cache more scalable.

The client-cluster is composed of the clients' residual resources. Since the clients join and leave dynamically, in order to use its storage efficiently, the client-cluster must be self-organizing and fault tolerant and the load of each client should be balanced. To meet these requirements, we manage the client-cluster by using Distributed Hash Table (DHT) based peer-to-peer protocol. By using this protocol, all clients receive roughly the same load because the hash function balances load with high probability. Additionally, the proxy cache does not need to gather the client information and we reduce administrative cost.

This protocol is responsible for the routing of the object, but it needs to cope with updating the object whenever clients join or leave. Typically, we can replicate the object. However, this approach leads to extremely large traffic overhead and wasted storage. To reduce this overhead, we suggest the *Backward ICP* which is responsible for storing

and finding objects in a manner similar to replication. A proxy saves objects to a client-cluster and gets objects from it by using this protocol.

This paper is organized as follow. In section 2, we describe cooperated web caching and peer-to-peer lookup algorithm briefly. Section 3 introduces the detail of the peer-to-peer client-clustering. The simulation environment and the performance evaluation are given in section 4. We mention other related works in section 5. Finally, we conclude in section 6.

2 Background

2.1 Cooperated Web Caching

The basic operation of the web caching is simple. Web browsers generate HTTP GET requests for Internet objects such as HTML pages, images, mp3 files, etc. These are serviced from a local web browser cache, web proxy caches, or an original content server - depending on which cache contains a copy of the object. If a cache closer to the client has a copy of the requested object, we reduce more bandwidth consumption and decrease more network traffic. Hence, the cache hit rate should be maximized and the miss penalty, which is the cost when a miss occurs, should be minimized when designing a web caching system.

The performance of a web caching system depends on the size of its client community. As the user community increases in size, so does the probability that a cached object will soon be requested again. Caches sharing mutual trust may assist each other to increase the hit rate. A caching architecture should provide the paradigm for proxies to cooperate efficiently with each other. One approach to coordinate caches in the same system is to set up a caching hierarchy. With hierarchical caching, caches are placed at multiple levels of the network. Another approach is a distributed caching system, where there are only caches at the bottom level and there are no other intermediate cache levels.

Internet Cache Protocol (ICP) [2] is a typical cooperating protocol for a proxy to communicate with other proxies. If a requested object is not found in a local proxy, the proxy sends ICP queries to neighbor proxies; sibling proxies and parent proxies. Each neighbor proxy receives the queries and sends ICP replies without the existence of the object. If the local proxy receives an ICP reply with the object, it uses that reply. Otherwise, the local proxy forwards the request to the parent proxy. ICP wastes expensive resources; core-link and cache storage. Even if the neighbor caches do not have the requested object, ICP uses the core-links between proxies, which are used for many clients and are bottlenecks of the network bandwidth. Another protocol for cooperated caching is the Cache Array Routing Protocol (CARP) [3], which divides the URL-space among an array of loosely coupled caches and lets each cache store only the objects whose URL are hashed to it. For this feature, every request is hashed and forwarded to a selected cache node. In this scheme, clients must know the cache array information and the hash function, making the management of CARP difficult. Additionally, there are other issues such as load balancing and fault tolerance.

Another problem of CARP, as well as ICP, is scalability of management. Large corporate networks often employ a cluster of machines, which generally must be over-

provisioned to handle burst peak loads. A growth in user population creates a need for hardware upgrades. This scalability issue cannot be solved by ICP or CARP.

2.2 Peer-to-Peer Lookup

Peer-to-peer systems are distributed systems without any centralized control or hierarchical organization, where the software running at each node is equivalent in functionality; this includes redundant storage, selection of nearby servers, anonymity, search, and hierarchical naming. Among these features, lookup for a data is an essential functionality for peer-to-peer systems.

A number of peer-to-peer lookup protocols have been recently proposed, including Pastry [5], Chord [6], CAN [7] and Tapestry [8]. In a self-organizing and decentralized manner, these protocols provide a distributed hash-table (DHT) that reliably maps a given object key to a unique live node in the network. Because DHT is made by a hash function that balances load with high probability, each live node has the same responsibility for data storage and query load. If a node wants to find an object, a node simply sends a query with the object key corresponding to the object to the selected node determined by the DHT. Typically, the length of routing is about O(log n), where n is the number of nodes. According to these properties, peer-to-peer systems balance storage and query load, transparently tolerate node failures and provide efficient routing of queries.

3 Peer-to-Peer Client-Clustering

3.1 Overview

As we described in the previous section, the use of only a proxy cache has a performance limitation because of potential growth in client population. Even if proxy caches cooperate with each other to enhance performance, high administrative cost and scalability issues still exist. To improve the performance of the cache system and solve the scalability issues, we exploit the residual resources of clients for a proxy cache. That is, any client that wants to use the proxy cache provides small resources to the proxy and the proxy uses these additional resources to maintain the proxy cache system. This feature makes the system resourceful and scalable.

We use the residual resources of clients as a backup storage for the proxy cache. While a conventional proxy cache drops evicted objects, our proxy cache stores these objects to the backup storage, which is distributed among the client-cluster. When a client sends a GET request to a proxy cache, it checks its local storage. If a hit occurs, it returns the requested object; otherwise, it sends a lookup message to the backup storage and this message is forwarded to the client that has responsibility for storing the object. If the client has the object, it returns the object to the proxy; otherwise, the proxy gets the object from the original server or other proxy caches. This interaction between the proxy cache and the backup storage decreases the probability of sending requests outside the network, reduces the usage of inter-proxy links, and increases the performance of the proxy cache.

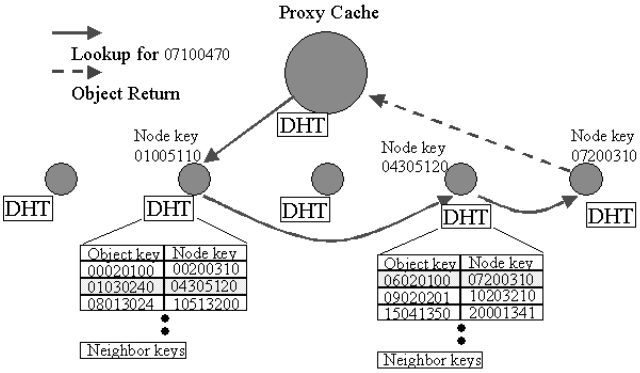


Fig. 2. Basic Lookup Operation In the Client-Cluster. In this Figure, Total Hop count is 3.

3.2 Client-Cluster Management

In our scheme, a proxy cache uses the resources of clients that are in the same network. Generally, if a peer wants to use other peers, it should have information about those. This approach is available when the other peers are reliable and available. However, the client membership is very large and changes dynamically. If the proxy cache manages the states of all clients, too much overhead is created to manage the client information and complex problems such as fault-tolerance, consistency and scalability arise. In consideration of these issues, we establish the proxy cache such that it has no information for the clients and the client-cluster manages itself.

We design the client-cluster by using DHT(distributed hash table) based peer-to-peer protocol [5], [6], [7], [8]. To use this protocol, each client needs an application whose name is *Station*. A Station is not a browser or a browser cache, but a management program to provide clients' resources for a proxy cache. A client can not use resources of a Station directly, while a proxy cache sends requests issued from clients to Stations in order to use resources of a client-cluster. When a Station receives requests from a proxy cache, it forwards requests to another Station or checks whether it has the requested objects. Each Station has a unique node key and a DHT. The unique node key is generated by computing the SHA-1 hash of the client identifier, such as an ip address or an ethernet address, and the object key is obtained by computing the SHA-1 of the corresponding URL. The DHT describe the mapping of the object keys to responsible live node keys for efficient routing of request queries. It is similar to a routing table in a network router. A Station uses this table with the key of the requested object to forward the request to the next Station. Additionally, the DHT of a Station has the keys of *neighbor Stations* which are numerically close to the Station, like the leaf nodes in PASTY or the successor list in CHORD.

The basic operation of the lookup in a client-cluster is shown in Figure 2. When a proxy cache sends a request query to one Station of a client-cluster, the Station gets the object key of the requested object and selects the next Station according to the DHT

and the object key. Finally, the *home Station*, which is a Station having the numerically closest node key to the requested object key among all currently live nodes, receives the request and checks whether it has the object in local cache. If a hit occurs, the home Station returns the object to the proxy cache; otherwise, it only returns a null object. In Figure 2, the node whose key is 07200310 is the home Station for the object whose key is 07100470. The cost of this operation is typically O(log n), where n is the total number of Stations. If 1000 Stations exist, the cost of lookup is about 3, and if 100000 Stations, the cost is about 5. Since the RTT for any server in the Internet from one client is 10 or 100 times bigger than that for another client in the same network, we reduce the latency for an object by 2 or 20 times when we obtain the object in the client-cluster.

The client-cluster can cope with frequent variations in client membership by using this protocol. Though the clients dynamically join and leave, the lazy update for managing the small information of the membership changes does not spoil the lookup operation of this protocol. When a Station joins the client-cluster, it sends a join message to any one Station in the client-cluster and gets new DHT and other Stations to update their DHT for the new Station lazily. On the other hand, when a Station leaves or fails, other Stations, which have a DHT mapping with the departing Station, detect the failure of it lazily and repair their DHT. According to this feature, the client-cluster is self-organizing and fault-tolerant.

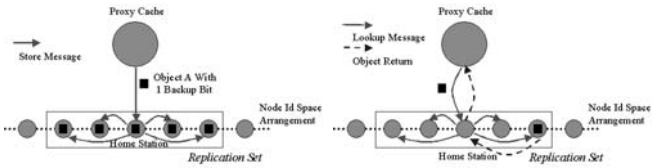
The proxy cache stores the evicted objects to a particular Station in the client-cluster by using this lookup operation. All Stations have roughly the same amount of objects, because the DHT used for the lookup operation provides a degree of natural load balance. Moreover, the object range, which is managed by one Station, is determined by the number of live nodes. That is, if there are few live nodes, the object range is large; otherwise, it is small. According to this, when the client membership changes, the object range is resized automatically and the home Stations for every object are changed implicitly.

As described, the routing information and the object range are well managed by this protocol. Consequently, after updating the information for variation in the client membership, future requests for an object will be routed to the Station that is now numerically closest to the object key. If the objects for the new home Station are not moved, subsequent requests miss the objects. According to these misses, the performance of a client-cluster decreases remarkably. We can replicate the objects to neighbor Stations to prevent such misses. This approach ensures the reliability of the objects, but leads to serious traffic overhead and inefficient storage usage. To reduce this overhead and use the storage efficiently, we store and lookup objects using the *Backward ICP*. This is described in the next section.

3.3 Backward ICP

The Backward ICP, which is a communication protocol between the proxy cache and the client-cluster, is similar to the ICP used between the proxy caches. However, the Backward ICP uses a local area network rather than an inter-proxy link.

There are two types of messages in the Backward ICP, as shown in Figure 3. One is a *Store* message and the other is a *Lookup* message. A Store message is used to store evicted objects from a proxy cache. The proxy cache sends a Store message and



(a) Store Message (b) Lookup Message

Fig. 3. Two types of Backward ICP Message

the evicted object to the home Station and the home Station replicates the objects to the replication set, which is composed of neighbor Stations. Before sending a Store message for an evicted object, the proxy cache checks the *Backup bit* of the evicted object. This Backup bit is used to prevent duplicated storage of an object that is already in the client-cluster. If the Backup bit is set to 1, the proxy cache knows that the client-cluster has this evicted object and drops this object immediately. If the bit is set to 0, the proxy cache backs up the evicted object to the client-cluster. When the proxy cache gets the object from the client-cluster, this bit is set to 1. When the object is refreshed or returned from the original server, this bit is set to 0.

A Lookup message is used to find objects in the client-cluster. When the proxy cache sends a Lookup message to the home Station, this Station returns the object to the proxy cache if it has the requested object. Otherwise, if a miss occurs, it sends Lookup messages to the replication set simultaneously and waits for a response from any Station. If the object is somewhere among the replication set, the home Station stores this object and returns this to the proxy cache; otherwise, it returns a null object. Following this, the home Station replicates the object to the replication set, except the responding Station.

This protocol replicates objects only at the time when they are stored or a lookup miss occurs. It reduces traffic overhead incurred by object replications. Moreover, it uses storage efficiently by giving more opportunities to retrieve popular objects. The first time when any object is stored, the object is replicated to increase the probability of accessing the object. As time goes by, popular objects are requested more than other objects and they are replicated again to increase the probability.

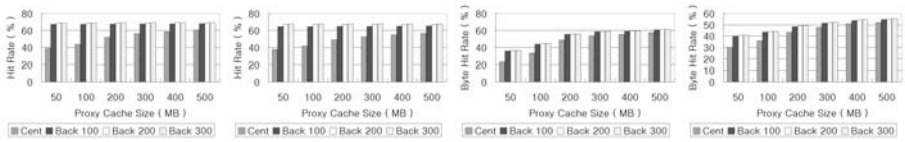
4 Performance Evaluation

4.1 Traces Used

In our trace-driven simulations we use traces from KAIST, which uses a class B ip address for the network. The trace from the proxy cache in KAIST contains over 3.4 million requests in a single day. We have run our simulations with traces from this proxy cache since October, 2001. We show some of the characteristics of these traces in Table 1. Note that these characteristics are the results when the cache size is infinite. However, our simulations assume limited cache storage and ratios including hit rate and byte hit

Traces	Measuring day	Requests Size	Object Size	Request #	Object #	Hit Rate	Byte Hit Rate
Trace 1	2001.10.08	9.02GB	3.48GB	699280	215427	69.19%	63.60 %
Trace 2	2001.10.09	11.66GB	1.38 GB	698871	224104	67.93%	57.79%

Table 1. Traces used in our simulation



(a) Hit rate for Trace 1 (b) Hit rate for Trace 2 (c) Byte Hit rate for Trace 1 (d) Byte Hit rate for Trace 2

Fig. 4. Hit rate and byte hit rate comparison between only proxy cache(cent) and client-cluster(back-n)

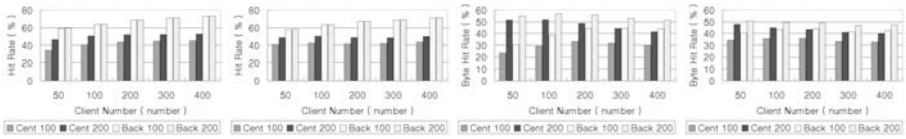
rate cannot be higher than *infinite-hit rate* and *infinite-byte hit rate*, which are the hit rate and the byte hit rate when the infinite cache is used.

4.2 Hit Rate and Byte Hit Rate

Figure 4 shows a comparison of the *hit rate* and the *byte hit rate*. By the hit rate, we mean the number of requests that hit in the proxy cache as a percentage of total requests. A higher the hit rate means the proxy cache can handle more requests and the original server must deal with proportionally lighter load of requests. The byte hit rate is the number of bytes that hit in the proxy cache as a percentage of total number of bytes requested. A higher byte hit rate results in a greater decrease in network traffic on the server side.

In the figures, cent means using only a proxy cache and back n means using the client-cluster with n hundreds clients. The hit rate of only the proxy cache is greatly affected by the cache size, but the hit rate of using the client-cluster achieves nearly an infinite-hit rate without any relationship to the proxy cache size. This is achieved by the plentiful resources provided by the clients. That is, though the proxy cache size is limited, the storage of the client-cluster is sufficient to store evicted objects and the proxy cache gets almost all requested objects from the client-cluster.

For the byte hit rate, we can obtain a similar result as that for the hit rate. However, in this case, using the client-cluster does not yield infinite-byte hit rate, particularly with a small proxy cache size. The reason for this result is the different byte size of the object range, which is roughly the same for each client, because of the different sizes of the objects. Thus some clients that usually have large objects cannot store many objects, and the hit rate and the byte hit rate decrease. In particular, large size objects whose size is bigger than that of one client storage, which is the Station's storage, 10MB, are not stored on the client-cluster and the byte hit rate decreases remarkably.



(a) Hit rate for Trace 1 (b) Hit rate for Trace 2 (c) Byte Hit rate for Trace 1 (d) Byte Hit rate for Trace 2

Fig. 5. Hit rate and byte hit rate comparison with various client number

Client #	Mean Req.	Max Req.	Dev.	Mean Byte Req.	Max. Byte Req.	Dev.
100	1024	1369	2.2	13422KB	316805KB	11.1
200	602	733	2.4	6711KB	315158KB	12.1
300	401	510	2.5	4474KB	314197KB	12.9

Table 2. Summary of Client Loads for Trace 1 with the 200MB proxy

4.3 Client Size Effect

In this section, we show scalability of the client-clustering. We assume every 100 clients makes 0.35million requests and simulate with variable client number.

In Figure 5, the hit rate when only the proxy cache is used does not increase markedly. Even in Trace 2, the hit rate decreases. However, when a proxy cache uses the client-cluster, the hit rate increases by 30-40% over that when only the proxy cache is used. Additionally, as the client number increases, the hit rate increases accordingly. For the byte hit rate, when the proxy cache uses the client-cluster, the byte hit rate increases by 20-30% over that when only the proxy cache is employed.

According to these results, when client population grows, using only a proxy cache should take on administrative cost to provide sufficient service to clients. However, using the client-cluster does not need any management cost to handle the growth in client population. Consequently, the client-cluster is scalable.

4.4 Client Load

We examine the client loads, which include the request number, storage size, stored objects, hit rate, etc, to verify that the client-cluster balances the storage and request queries. Table 2 shows a summary of the request number and the sizes of the requested objects. Each client receives roughly the same load, and when the client number increases the load of each client decreases. The properties of DHT-base peer-to-peer protocols account for these findings. For the byte request, we again see the effect of the different sizes of the objects, which we strongly believe account for the performance degradation.

5 Related Works

A similar proposal for our approach appeared in Squirrel [9], which described a decentralized web browser cache. Squirrel fully distributes the web caches storage among the browser cache of clients. Hence, when the availability of clients is asymmetric, some clients decrease the total performance of the Squirrel network. Additionally, all contents are distributed and it is hard to manage the objects according to the characteristics of web objects. In our scheme, a web object is assigned to the proxy cache or the client-cluster according to the popularity of the object, which optimizes the overall performance of the proxy cache.

6 Conclusions

In this paper, we propose and evaluate peer-to-peer client-clustering, which is used as a backup storage for the proxy cache. The proxy cache with this client-cluster is highly scalable and more efficient, and has low administrative cost. Even if the clients take the load, this load has been verified on a range of real workloads to be low. Moreover, the utility of the client-cluster can be improved by managing objects according to their properties such as size, popularity and update frequency. We can extend the usage of the client-cluster to other proxy systems. If a proxy performs demanding jobs such as encoding/decoding and complex calculation for many clients, it can use the residual resources of the clients to accomplish these tasks.

References

1. P.Rodriguez, C.Spanner, and E.W.Biersack, Web caching architectures:Hierarchical and distributed caching. In proceedings of the 4th International Web Caching Workshop, 1999.
2. A.Chankhunthod, P.B.Danzig, C.Neerdaels, M.F.Schwartz and K.J.Worrell, A hierarchical internet object cache. In proceedings of the 1996 Usenix Technical Conference, January 1996.
3. J.Cohen, N.phadnis, V.valloppillil and K.W.Ross, Cache Array Routing Protocol v1.0. http://www.ietf.org/internet-drafts/draft-vinod-carp-v1-03.txt, September 1997.
4. A.Wolman, G.M.Voelker, N.Sharma, N.Cardwell, A.Karlin and H.M.Levy, On the scale and performance of cooperative Web proxy caching. In proceedings of the 17th ACM symposium on Operating Systems Principles, December 1999.
5. A.Rowstron and P.Druschel, Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems. In proceedings of the International Conference on Distributed Systems Platforms, November 2001.
6. I.Stoica, R.Morris, D.Karger, M.F.Kaashoek and H.Balakrishnan, Chord: A scalable peer-to-peer lookup service for Internet applications. In proceedings of ACM SIGCOMM 2001, August 2001.
7. S.Ratnasamy, P.Francis, M.Handley, R.Karp and S.Shenker, A Scalable Content-Addressable Network. In proceedings of ACM SIGCOMM 2001, August 2001.
8. B.Y.Zhao, J.Kubiatowicz and A.Joseph, Tapestry: An Infrastructure for Fault-tolerant Wide-area Location and Routing. In UCB Technical Report UCB/CSD-01-114, 2001.
9. S.Iyer, A.Rowstron and P.Druschel, Squirrel: A decentralized peer-to-peer web cache. In proceedings of Principles of Distributed Computing'02, 2002.