Position Summary - Towards Global Storage Management and Data Placement

Alistair Veitch, Erik Riedel, Simon Towers and John Wilkes
Hewlett Packard Laboratories
{aveitch, riedel, stowers, wilkes}@hpl.hp.com

As users and companies increasingly depend on shared, networked information services, we continue to see growth in data centers and service providers. This happens as services and servers are consolidated (for ease of management and reduced duplication), while also being distributed (for fault-tolerance and to accommodate the global reach of customers). Since access to data is the lifeblood of any organization, a global storage system is a core element in such an infrastructure. Based on success in automatically managing local storage, we believe that the key attribute of such a system is the ability to flexibly adapt to a variety of application semantics and requirements as they arise and as they change over time. Our work has shown that it is possible to automatically design and configure a storage system of one or more disk arrays to meet a set of application requirements and to dynamically reconfigure as needs change, all without human intervention. Work on global data placement expands the scope of this system to a world of distributed data centers.

Data location

Ensuring that data is available in the right location is a key challenge as data and applications go global. Due to speed of light and congestion, network performance will always be a bottleneck. The system will have to transparantly migrate data to have the data that each application needs co-located with the servers that are currently operating on it. Whether data follows a particular user as they travel around the globe; supports a global design team in its daily work; or handles customer data or inventory for a global corporation, the individual data "shadows" of all types of users and applications must be supported efficiently. Such a system can be viewed as a network of "cache" devices – each data center provides a pool of storage that at any one time is caching a subset of the global store. The key problem is deciding when to move data from one to another, when to keep multiple copies, and how many copies to keep – automating data placement such that load is balanced both within and across data centers.

Data replication and consistency

For many applications, the most efficient solution will be to have multiple replicas of the same data. Along with the core requirement of availability in the event of local failures, the makes it necessary to store the same data in multiple global locations and keep it consistent. The ability to adapt consistency levels within the storage system to the varying requirements of individual applications is a key enabler for global data placement. Ideally this would be done transparently, without changes to existing application code, and much of the necessary information and flexibility is available even with storage interfaces designed for local resources.

Mechanisms for maintaining consistency across global sites range from expensive pessimistic approaches with multiple round trips of locking overhead, to low latency optimisitic approaches that allow occasional inconsistencies or require rollback. To evaluate the cost of such schemes, we analyzed traces for a number of applications, including email, software development, databases, and web browsing. At the storage level - after cache accesses have been eliminated - the results do not seem very promising. A high fraction of requests are updates, the ratio of metadata to data is high, and a high fraction of requests are synchronous. However, considering individual applications in isolation, these metrics vary widely, making adaptive consistency that uses different mechanisms as appropriate attractive.[1] We have also begun to quantify how much sharing takes place, and see the fraction of "hard" sharing in a large store is promisingly low.[2]

Security

Data must be secure, especially in a system where facilities are shared amongst many different organizations. This requires strong authentication, authorization and encryption mechanisms, none of which are necessary in the context of local storage systems. Initial analysis shows that large stores quickly encompass large numbers of objects to be protected and principals requiring authentication, posing scalability problems. However, if we consider the number of objects that an individual user handles - their "shadow" on the entire store - and the commonality among access patterns to these objects, the scope of security quickly becomes more tractable. The question then becomes which levels of abstraction to provide for different classes of users and data.

System management and control

Our local management system can determine appropriate placement in the local case with local information. We anticipate a hierarchy in the global setting, with some optimization best done within the data center, and a more global view controlling movement across centers, all informed by the "shadow" that supports a particular coherent data set or user. Such a system must operate at a range of time scales and granularities, and will critically depend on the ability to accurately and efficiently model and predict the behavior of all the components within and the links across the system.

1. Technical memo *HPL-SSP-2001-1*, HP Labs, March 2001.
2. "When local becomes global" *20th IPCCC*, April 2001.

Proceedings of the Eighth Workshop on Hot Topics in Operating Systems (HotOS-VIII '01)
1530-1621/01 $10.00 © 2001 **IEEE**