

Middleware Support for Locality-aware Wide area Replication

Sai Susarla and John Carter
{*sai, retrac*}@*cs.utah.edu*

UUCS-04-017

School of Computing
University of Utah
Salt Lake City, UT 84112 USA

November 18, 2004

Abstract

Coherent wide-area data caching can improve the scalability and responsiveness of distributed services such as wide-area file access, database and directory services, and content distribution. However, distributed services differ widely in the frequency of read/write sharing, the amount of contention between clients for the same data, and their ability to make tradeoffs between consistency and availability. Aggressive replication enhances the scalability and availability of services with read-mostly data or data that need not be kept strongly consistent. However, for applications that require strong consistency of write-shared data, you must throttle replication to achieve reasonable performance.

We have developed a middleware data store called *Swarm* designed to support the wide-area data sharing needs of distributed services. To support the needs of diverse distributed services, *Swarm* provides: (i) a failure-resilient proximity-aware data replication mechanism that adjusts the replication hierarchy based on observed network characteristics and node availability, (ii) a customizable consistency mechanism that allows applications to specify allowable consistency-availability tradeoffs, and (iii) a contention-aware caching mechanism that monitors contention between replicas and adjusts its replication policies accordingly. On a 240-node P2P file sharing system, *Swarm*'s proximity-aware caching and replica hierarchy maintenance mechanisms improve latency by 80%, reduce WAN bandwidth consumed by 80%, and limit the impact of high node churn (5 node deaths/sec) to roughly one-fifth that of random replication. In addition, *Swarm*'s contention-aware caching mechanism outperforms RPCs and static caching mechanisms at all levels of contention on an enterprise service workload.