Efficient File Distribution in a Configurable, Wide-area Distributed File System

Irene Zhang

Background

WheelFS is a wide-area distributed file system designed to be a general storage solution for wide-area applications. WheelFS simplifies wide-area applications by helping them cope with the challenges of sharing data over the wide-area network. WheelFS is applicable to a wide range of distributed applications because applications can fit the behavior of WheelFS to their requirements using *semantic cues*.

Problem Statement

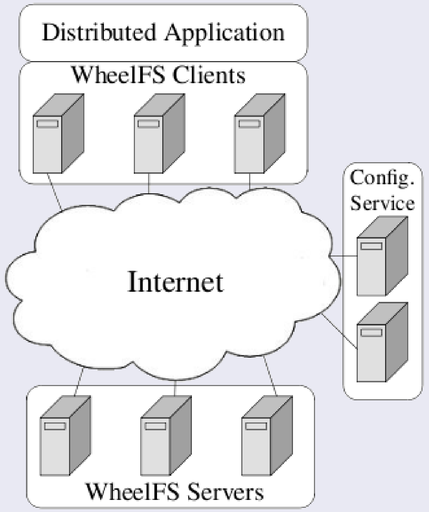
Many applications require their storage layer to be able to efficiently distribute files. The limited bandwidth and high latency of wide-area links make it difficult for most file systems to serve large files quickly and scale to handle many simultaneous requests for a file. Our goal is to ensure that WheelFS can provide efficient file distribution for wide-area applications.

Solution

Incorporate ideas from file systems and distributed applications to improve the performance and scalability of file distribution in WheelFS:

- **Prefetching.** WheelFS allows applications to control prefetching of entire files or directories with the .WholeFile cue.
- **Cooperative Caching.** WheelFS clients keep local caches and can serve files out of their cache. Applications can choose to fetch files from other clients using the .Hotspot cue.

WheelFS Overview

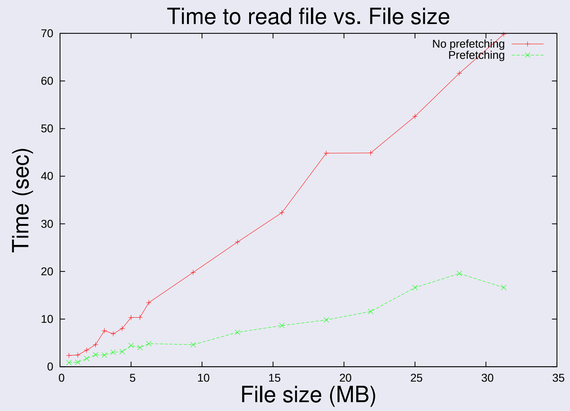


A WheelFS deployment consists of WheelFS servers, that store data, and WheelFS clients, that access stored data on behalf of applications. These clients and servers are scattered across the wide-area network. In addition a configuration service, that tracks which servers store each file, runs on nodes at several sites.

Prefetching Design Overview

WheelFS performs prefetching asynchronously in the background using prefetch threads. Prefetch requests for files are placed in a FIFO prefetch queue. For each request, a prefetch thread will send a request for the file to the WheelFS servers and then store the result in the client's local cache for the application to use later. WheelFS limits the number of outstanding requests to the prefetch request window size.

Prefetching Evaluation: Latency



Latency of reading a file over a 100ms RTT, 6 mbps link using a request window of five pending requests at a time.

Prefetching Evaluation: Throughput

Throughput for file reads with different numbers of simultaneously pending requests over a 100ms RTT, 6 mbps link.