Distributed Datastore System

Digital Reasoning Inc.
Larry Mitchell

Introduction

- This presentation describes a distributed datastore/file system that we are developing.
- Who am I?
  - Larry Mitchell
  - Long time Jini/Rio developer
  - Presently engaged at Digital Reasoning as a Jini Architect

Agenda

- Why a new DFS?
- DFS is a collaboration of services
- Common File System
- Virtual File Service
- File Accessor Service
- File Replication Service

Why a new File System?

- Most generic file systems are inherently non distributed
  - Jakarta VFS
- Most Distributed File system are proprietary
  - GFS (google file system)
- The few open source ones are not generic enough.
  - HDFS (Hadoop DFS)
  - AFS (Andrew File System
  - Coda

Common File System

- The common file system is an attempt to create a generic way to describe a hierarchical information store
  - Why genericity?
    - Our current work requires providing a common model of different types of information content for distributed text processing
- Similar but different to jakarta VFS
- Follows the File object interface
  - Why? Most developers are very familiar with it already

Common File System Types

- Regular Java File objects
- NFS
- CIFS/Samba
- Archives (jars, zips, tars, bzips)
- RAM file systems
- /tmp fs
- Hierarchical content such as JNDI
- Web based (HTTP/HTTPS/Web DAV)
- FTP/SFTP fs
- RSS/ATOM sources

Common File System Concepts

- Common element is the CFSFile interface which has a similar signature catalog to java.io.File
- All specific file types are described via URIs
  - eg. smb:// for samba, https:// for https
- A factory creates a representation for a file based on the URI
- The representation is an adapter for the native file type

Virtual File Service

- The virtual file service is the indexing and categorization service for CFS elements.
- FS Representation is a tree or trees.
- There can be multiple VFSs that collaborate on one or multiple roots.
- One or more JavaSpaces are used as a blackboard to describe hierachical file relationships.
- Each VFS defines multiple file system watchers to monitor system changes in the local FS

Virtual File Service (Cont.)

- A preset tree structure can be described via an XML document that describes the form and content of a tree.
- Most of these describe logical nodes which are used to link different file systems together.
- The document can specify which logical nodes are linked to specific physical file systems via the CFS.

File Accessor Service

- The file accessor service provides the means by which CFS file can be accessed and manipulated
- Top level interface is Rio based
- A further stream interface is based on the BEEP protocol and delivered via a jini smart proxy.

File Replication Service

- There are times when files need to be replicated at various points for either fault tolerance or performance reasons
- The file replication system provides a very efficient means of replicating file in the DFS.
- The replication system uses reliable multicast where possible to reduce bandwidth usage.

Summary

- Development is ongoing in the DFS
- At some point we will open source the codebase.
- If you are interested in becoming involved contact me at larry13767@gmail.com