SRB in Action
Introduction

• What is SRB
• CCLRC and SRB
• Using SRB in CMS
• Using SRB in the e-Minerals Mini-Grid
• Future Projects
• Questions
Managing Data

• Historically data has been **STORED** rather than **MANAGED**
• Problems arising from this include
  – Scaling
  – Distribution
  – Access Control, Authentication, Security
  – Data Migration
  – Data Curation
What is SRB?

- Storage Resource Broker (SRB) is a software product developed by the San Diego Supercomputing Centre (SDSC).
- Allows users to access files and database objects across a distributed environment.
- Actual physical location and way the data is stored is abstracted from the user.
- Allows the user to add user defined metadata describing the scientific content of the information.
How SRB Works

- 4 major components:
  - The Metadata Catalogue (MCAT)
  - The MCAT SRB Server
  - The SRB Server
  - The SRB Client
The MCAT Database

- MCAT database is a metadata repository that provides a mechanism for storing information used by the SRB system.
- Includes both
  - Internal system data required for running the system
  - Application (user) metadata regarding data sets being brokered by SRB.
The MCAT Server

- At least one SRB Server must be installed on the node that can access the MCAT database. This is known as the MCAT SRB Server.
- MCAT SRB Server works directly against the MCAT database to provide SRB Services.
- All other SRB Servers interact through the MCAT.
The SRB Server

- The SRB Server is a middleware application that accepts requests from clients and obtains/queries/manages the necessary data sets.
- It queries the MCAT SRB Server to gather information on datasets and supplies this back to the SRB client.
SRB Clients

- Provides a user interface to send requests to the SRB server.
- 4 main implementations of this:
  - Command line
  - MS Windows (InQ)
  - Web based (MySRB).
  - Java
- Web Services (MATRIX) in beta
  - Precursor to GRID Services
Concepts

• Location: A physical node running an SRB Server
• Physical Resource: A storage area managed by an SRB Server
• Logical Resource: One or more Physical Resources – can be distributed
• Collection – Data abstraction of resources
SRB in Detail

Application

SRB

C, C++, Linux I/O
Unix Shell
Java, NT Browsers
Prolog
Python
Web

Archives
HPSS, ADSM, UniTree, DMF
HRM
File Systems
Unix, NT, Mac OSX
Databases
DB2, Oracle, PostgreSQL

Resource, User
User Defined
MCAT

User Defined
Application Meta-data
Dublin Core

Third-party copy
Remote Proxies
DataCutter

Michael Doherty
RAL
Administration

- Users/Locations/Resources must be managed
- Two methods for doing this
  - Java MCAT Admin Tool
  - Command line tools
CCLRC and SRB

- The Data Management Group in CCLRC started working with SRB in November 2002 after a fact finding mission to the USA.
- There was an immediate requirement for a storage based product that allowed the addition of searchable metadata.
- Generated lots of internal interest, which led to a number of projects with SRB.
SRB Example: CMS

- Largest project using CCLRC SRB services at present is the CERN CMS experiment.
- SRB Chosen for ‘Pre-Challenge Production’, producing data for Data Challenge 2003/2004 (DC03/DC04)
- Need to prove data can be transferred, replicated and stored at LHC rates
- Need to coordinate data in the DC
CMS: Why SRB

• Not originally planned to SRB for DC
  – Planned RLS based system
• Nothing available CMS-wide on the timescale (end of year) for DC2003
• Needed stable and supported product for 6 months continuously
MCAT: CCLRC Database Service

- MCAT requires professionally run database
- Two IBM x440 clusters, one based at Daresbury Laboratory and the other at Rutherford Appleton Laboratory.
- The clusters connect to their own 1TB RAID 5 storage arrays via a independent fibre channel Storage Area Networks (SAN).
- Run Oracle Real Application Clusters software on Redhat Advanced Server for high availability/scalability RDBMS
- CMS MCAT hosted by 2 nodes
- Can load balance
Atlas Datastore (ADS)

- The Datastore currently has an on-line, nominal capacity of 1 Petabyte
- Based on an STK 9930 (PowderHorn) tape robot.
- The ADS team recently migrated from IBM 3590 (10GB) tape drives to STK 9940B (200GB) drives.
- CCLRC have written a custom SRB driver for this for CMS
• Implemented Storage System Driver
• Implement (most) of the 16 standard calls that implement the driver layer such as copy, move, delete and create.
• Some functions have no equivalent in ADS

Michael Doherty
RAL
CMS Results (So far)

• How long does it take to register and replicate 500GB between to widely separated locations?
  – *Tested 200GB which took approximately 6 hours, almost completely network limited*

• How long does it take to register and replicate 50k 10kB files?
  – *SRB has a bulk file registration mode which they have clocked at 400 files per second. Registered and replicated 1000 files in a few seconds*

• What is the maximum sustainable transfer rate out of a single server and what is the maximum rate a server can accept data from three servers?
  – *About 80-90% of network speed for 5 streams x number of servers*

• How many files can be registered in a day?
  – *No inherent limits*

• How many parallel streams can a server accept?
  – *Unknown, very small load on CPU with 10 streams*
SRB Example: e-Minerals

- [Environment from the Molecular Level]

- UK e-science project for modelling the atomistic processes involved in environmental issues
e-Minerals Requirements

• Data Management Requirements
  – Scientists want to store input and output files from simulations in different locations
  – manage their own files/data via the web
  – give access to other project members
  – give temporary access to others
Michael Doherty  
RAL

Architecture

Web Browser

Daresbury App Server
- MySRB
- SRB Server
- Oracle Client

Daresbury Database server
- MCAT

Eminerals MiniGrid

UK Grid

Application server runs SRB software

Database server holds locations of files

Cambridge  
SRB Resource

Reading  
SRB Resource

Bath  
SRB Resource
Future Projects

• E-Materials Project: A mini grid similar to E-minerals above.

• National Crystallographic Service (Southampton University): For both data management and storage in the ADS.

• British Atmospheric Data Centre (BADC) archive. For both data management and storage in the ADS.

• NERC Data Grid. Details to be confirmed.
Future Directions?

- 3 major areas that CCLRC are commencing work with in collaboration with SDSC.
  - Web/Grid Services
  - High Performance Testing
  - Database Advice
  - Training
Summary

- Have established links with SRB community and SDSC
- Implemented SRB on projects
- Set up test systems for new projects
- Can help community with
  - SRB Test Systems
  - SRB Productions Systems
  - SRB Local Support
- Will contribute to future versions
Questions

- ?