The SDSC Storage Resource Broker

University of Maryland
CMSC818S

Presented by: Burcu Karagol-Ayan
burcu@cs.umd.edu

Roadmap

- NPACI Project
- Storage Resource Broker (SRB)
- SRB Features
- Architecture
- Connection Processing
- MCAT
- Federated Servers
- SRB Agent Details
- Client APIs
- Discussion
- MySRB
- Features in SRB
- Conclusion

SDSC NPACI Project

- To develop an infrastructure
  - for high performance distributed computing environment
  - part of NPACI (National Partnership for Advanced Computational Infrastructure)
    - Compute and storage resources
    - 5 compute sites and 10 data cache sites distributed across US
    - Connected via high speed data links
  - Key requirement
    - Support for data-intensive computing by providing
      - High performance I/O to massive data
      - Digital library capabilities for storage, search, and retrieval of scientific data

Data Grid Architecture

- Data Management Environment
  - Distributed data collection
    - Provides a single name space for referencing data stored on multiple storage systems (typically within same administration domain)
    - Capabilities
      - Integrate data collections and associated metadata
      - Handle multiplicity of platforms, resource and data types
      - Seamless access to data and information stored within DGA

Data Grid Architecture - 2

- Digital libraries
  - Integrate remote archival storage systems into a data collection, while providing discovery and manipulation services
  - Capabilities
    - Handle seamless authentication
    - Virtual organization structure for data and information based on a digital library framework
    - Handle dataset scaling in size and number
- Persistent archives
  - Support migration of data collections onto new technologies, while preserving ability to organize, discover, access data
  - Capabilities
    - Replication of data
    - Version control
    - Handle access control and provide auditing facilities

SRB (Storage Resource Broker)

- A middleware part of the infrastructure
- Provides seamless access to data stored on heterogeneous storage resources
  - File systems, DB systems, archival storage systems
- Provides an API:
  - Enables applications executing in distributed NPACI environment to access data stored at the distributed storage sites
  - Provides capability to
    - do information discovery
    - identify data collections of interest
    - select retrieve data items that may be distributed across a wide area network

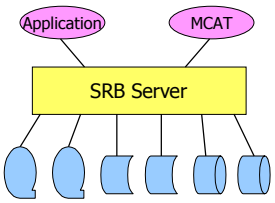
SRB

- Client-server middleware that uses collections to build a logical name space for identifying distributed data
- SRB + MCAT support location transparency by accessing data sets and resources based on their names or physical locations
- SRB provides
  - A logical representation for describing storage systems, digital file objects, collections specific features for use in digital libraries, persistent archive systems, collection management systems
  - Capabilities to store replicas of data, for authenticating users, controlling access to documents and collections, and auditing accesses
  - A facility for co-locating data together using containers
- Store user-defined metadata at collection and object level and provides search capabilities based on metadata

SRB Middleware

- Used to access heterogeneous storage resources in a distributed system by applications
- Employs MCAT
  - A metadata catalog service
  - Manages two types of metadata associated with data collections and system resources
    - Descriptive: describes contents of entire data collections, individual data items
    - System: provides location and access control information
  - Enables SRB to store and retrieve metadata about system entities
  - Provides location transparency
  - Enables attribute-based access to data
    - ⇒Enables information discovery and automated processing of data

SRB Middleware



Goal: improve operating environment for scientific applications that have the requirement to access and process many data sets

DB2, Oracle, Illustra...
HPSS, Unitree... Unix
Distributed Storage Resources

SRB Features

- API for accessing data
- Metadata catalog for organizing data collection attributes
- Data handling system for supporting remote access of data

SRB Features - 2

- **Uniform Storage Interface**
  - Well-defined storage interfaces
  - SRB provides a mapping from defined storage interface to native interface supported by each underlying storage resource
    - via resource-specific drivers

    (implement each interface for each resource)
- **Metadata catalog service (MCAT)**
  - To support attribute-based access to data collections, items, and other system resources
  - Provides a set of APIs for querying and updating metadata catalog

SRB Features - 3

- ***Collection* Hierarchy**
  - The way data stored in SRB is organized
  - Defined as:
    - A collection contains ≥ 0 data items and ≥ 0 sub-collections
    - A sub-collection contains ≥ 0 data items and ≥ 0 sub-collections
    - A data item is a file or binary large object (BLOB)
  - Data items belonging to same (sub-)collection can be stored in physically distributed, heterogeneous storage resources
- **Hierarchical Access Control**
  - Users may control propagation of extensible privileges along the collection hierarchy

SRB Features - 4

- **Tickets**
  - For controlling constrained read access to data
  - Users with appropriate privileges (*control*) may issue them to other users on collections/items
  - Valid either for a fixed amount of time or for a fixed number of uses
  - Registered Users: Necessary metadata is available in MCAT
  - Unregistered Users: Anonymous users of the SRB/MCAT system
- **Physical Storage Resources (PSRs)**
  - For storage resources with a file system interfaces:
    - (*hostname, pathname*) combination
  - For storage resources with database interfaces:
    - (*hostname, db_id, table_id*) triple

SRB Features – 5

- **Logical Storage Resources (LSR) and Replication**
  - A set of one or more PSR's can be combined into a single LSR
  - Client APIs typically refer to LSRs
  - Collections are implemented using LSRs
- **Proxy Operations**
  - Data handling operations that are performed on behalf of client applications by SRB
  - Examples: *move, copy*
- **Federated Operations**
  - A *federation* of SRB servers provides access to distributes storage resources
  - Each SRB server controls a distinct set of PSRs
  - One SRB server can act as a client to another
    - So a client application can access data stored anywhere

SRB Features – 6

- **Authentication and encryption**
  - A variety of authentication systems
  - Password information is maintained by MCAT
  - SRB supports user authentication and data encryption based on SEA system
    - Public-private key mechanisms
    - Symmetric key encryption algorithm
    - Simple key management capabilities
- **Activity Logs**
  - Storage and metadata update operations may be logged using MCAT
  - User may specify, but not obligatory

SRB Features – 7

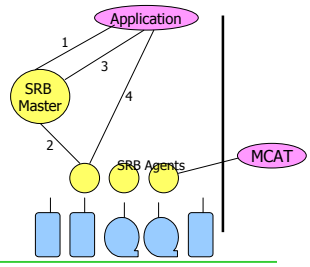
- **Types of Applications**
  - By using SRB middleware, scientific applications can
    - issue ad hoc queries on metadata to identify data of interest
    - then use SRB API's to access data

SRB Architecture

- 1 or more SRB Master daemon processes
  - (*hostname, port number*)
  - Each controls a distinct set of PSRs
  - Monitors its well-known port for connection requests from clients
- Communication via sockets
- SRB Agent
  - Uses MCAT metadata service to obtain necessary system metadata needed for processing client storage requests
- A client may establish
  - only as single connection to a given SRB Master or
  - concurrent connections to multiple, distinct SRB Masters
- A single connection may be used to access multiple data items, possibly stored on different PSRs

Connection Processing

- Connection Processing
  1. Connect request from client to SRB Master
  2. Fork Agent
  3. Return connection info to client
  4. All subsequent communications via Agent



MCAT

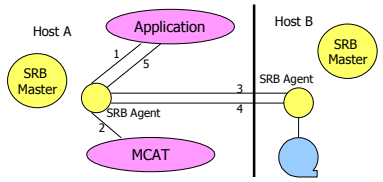
- Data items managed by SRB system are referenced by
  - collection/sub-collection under which they reside
  - their name
- Physical location is distinct from logical collection hierarchy
- MCAT is used
  - to record location information for
    - PSRs
    - Data items
  - To describe contents of collections, data items
- Was implemented in DB2 UDB and Oracle
  - May be used as a central resource or as a local resource

SRB

- SRB is a federated server system, with each SRB server managing/brokering a set of storage resources
- Federated SRB implementation provides
  - Location transparency
  - Improved reliability and availability
  - Logistical and administrative reasons
  - Fault tolerance
  - Integrated data access
  - Persistence

Federated Servers

- In a distributed system, one may choose to control storage resources using different SRB masters due to a variety of technical and administrative reasons.



Federated SRB Operation

SRB Agent Details

- *Dispatcher* module
  - Monitors incoming client requests
  - Dispatches requests to
    - High-Level Request Handler
      - Maps user names and data item names to access control and location information using MCAT
    - or Low-Level Request Handler
      - Expects caller to provide detailed parametric information necessary
  - Returns results to client
- SRB drivers
  - Filesystem drivers (UNIX, HPSS, ADMS, UniTree)
  - DBMS drivers (DB2, Oracle, Illustra, Objectstore)

Client APIs

- Query/Update of metadata
  - Allow applications to manage metadata associated with data collections, data items, users, user groups, storage resources
- Connecting to the server
  - Connect, disconnect, "ticketed connection"
- Creation of data items
  - Does the user have the privilege?
- Open/read/write/delete of data items

Discussion

- SRB system
  - Provides a simplified SQL-like interface to the metadata with a file I/O interface to the actual data item
  - Supports various types of storage resources
    - File system, DB system, archival storage system
  - Implemented for a variety of storage systems and OS platforms

MySRB

University of Maryland
CMSC818S

Presented by: Burcu Karagol-Ayan
burcu@cs.umd.edu

MySRB

- A web-based interface to SRB
- Provides the functionalities
  - Collection and file management
  - Metadata handling
  - Access and display of files and metadata
- Uses https protocol with 128-bit RSA authentication
  - Unique session key with maximum time limit
- Web-browser interface uses a split-window
- Metadata are not just entity-value pairs, have a richer structure

Features in MySRB

- Data movement operations
  - User can
    - Collection-level
      - ingest a file into SRB or
      - create sub-collection through mySRB interface
    - Ingestion-level
      - choose logical resource that will be used for storing
      - or specify a container (file is added to container)
      - (Data type, any metadata)
  - MySRB uses file-browse mechanism of web browsers to identify local file that need to be ingested

Features in MySRB - 2

- Apart from ingesting, a user can
  - Register SRB objects
    - A file that can exist either in a file system, an archival storage system, or as a LOB in a DB system
    - A directory in a file system or an archival storage system
    - A SQL query for a DB resource
    - A URL
    - A method object or virtual data
- Apart from ingesting or registering files into mySRB, user can
  - Replicate
  - Register replicate
  - Ingest replica
  - Copy
  - Move
  - Link
  - Delete
  - Lock, pin, checkout

Features in MySRB - 3

- **Metadata operations**
  - Types of metadata
    - System-defined metadata
      - Created and maintained by SRB
      - User can view them and use them in their search mechanism
    - User-defined metadata
      - Descriptive in nature
      - Made of name, value, unit triplets
      - 4 ways of associating
        - Allows user to associate metadata when ingesting or registering an object, or when creating a new sub-collection
        - Invoke insert metadata function which provides a form for operation (as many times as required)
        - Copy metadata from other SRB objects or collections
        - Extract metadata from an extraction method associated with data-type of file

Features in MySRB - 4

- Type-oriented (domain-oriented) metadata
  - Pre-defined sets of metadata elements that can be associated with SRB objects through their data types or for all SRB objects
- File-based metadata
  - A file in SRB that is associated as a metadata-carrying file for another SRB object
  - Used only for viewing, cannot take part in querying
- Annotations and commentary metadata
  - Useful for associating free-form metadata to a SRB object
  - Can be used for providing notes, comments, errata, queries and answers, annotations, memoranda, …
  - Have a type/location, timestamp, annotation writer's name associated with
  - Can be inserted by any user with a read permission

Features in MySRB - 5

- Metadata can be viewed in two ways
  - User can see both data and metadata at the same time
  - User can select to just view metadata for an object
- Importance of metadata: its queriability
- Operations on metadata
  - Create, view, query
  - Update, copy, delete (user defined metadata and annotations)
  - User registration
  - Access to resource, user, container metadata
  - Ability to navigate collection hierarchy and on-line help

Conclusion

- SRB provides transparent virtualized middleware for sharing data across distributed, heterogeneous data resources separated by different administrative and security domains
- MySRB is a web-based interface to SRB that provides a user-friendly interface to distributed collections brokered by SRB
- Available at *https://srb.npaci.edu/mySRB.html*