Utilization of Storage Resource Managers by Compact Muon Solenoid Large Hadron Collider Experiment

Donald Petravick, FNAL
Fermi National Accelerator Laboratory, MS 120
Box 500, Batavia IL 60510
Tel: +1-630-840-3935
Fax: +1-630-840-2783
e-mail: petravick@fnal.gov

Timur Perelmutov
Fermi National Accelerator Laboratory, MS 120
Box 500, Batavia IL 60510
Tel: +1-630-840-5547
Fax: +1-630-840-2783
e-mail: timur@fnal.gov

Storage Resource Managers (SRMs) are middleware components whose function is to provide dynamic space allocation and file management on shared storage components on the Grid[1]. SRMs support protocol negotiation and reliable replication mechanism. The SRM specification standardizes the interface, thus allowing for a uniform access to heterogeneous storage elements.[2,3,5] The SRM standard allows independent institutions to implement their own SRMs. SRMs leave the policy decision to be made independently by each implementation at each site. Resource Reservations made through SRMs have limited lifetimes and allow for automatic collection of unused resources thus preventing clogging of storage systems with "forgotten" files.

The storage systems can be classified on basis of their longevity and persistence of the data they store. Data can be considered to be temporary and permanent. For example disc caches might allow for spontaneous deletion of the files, while deletion of the file stored in a robotic tape storage can be very problematic. To support these notions, SRM defines three types of files and spaces: Volatile, Durable and Permanent. Volatile files can be removed by the system to make space for new files upon the expiration of its lifetime. Permanent files are expected to exist in the storage system for the lifetime of the storage system, unless explicitly deleted by the user. Finally Durable files have a both the lifetime associated with them and a mechanism of notification of owners and administrators of lifetime expiration but can not be deleted automatically by the system and require explicit removal.

SRM interface consists of the five categories of functions: Space Management, Data Transfer, Request Status, Directory and Permission Functions. Among the functions worth mentioning are srmReserveSpace which creates advanced space reservation with a user specified lifetime, and identified by a unique space token. The space token can later be utilized for storing files via srmPrepareToPut, srmPrepareToGet and srmCopy functions. For example srmPrepareToPut will take the list of files, file sizes, list client supported transfer protocols and space tokens. SRM interface utilizes Grid Security Infrastructure (GSI) for authentications. SRM service is a Web Service implementation of a published WSDL document. Fermilab SRM is based on and is an integral part of the dCache Distributed Disk Cache[6] coupled with Enstore Tape Storage System.

The Compact Muon Solenoid (CMS) experiment is designed to explore the full range of physics at the high-energy frontier up to TeV mass/energy scales made available for the first time at CERN's Large Hadron Collider (LHC).[7,8] CMS raw data originate at the CERN laboratory, near Geneva, Switzerland. Collaborators worldwide will analyze this data. Fermilab will play the role of a Tier One center in the CMS experiment. When transferring data from CERN, data will be exchanged between different petabyte class storage systems (CASTOR at CERN, and Enstore at Fermilab). Data will be transferred between Fermilab and North American Tier Two centers, which typically are large university data centers. The Storage Resource Manager (SRM) protocol was selected as a management protocol for both of these transfers, and the development of the protocol is seen by CMS as an essential abstraction allowing for more uniform software and system development.

Among the benefits to CMS of adaptation of the SRM as a major interface to the storage systems are the following: ability to use a single client to access multiple storage systems in multiple site; security; reliable replication services, that implement pacing of transfer to prevent disk thrashing and network clogging by queuing and limiting the number of simultaneous transfers. Another benefit of SRM is its ability to address storage resources independently of a transfer protocol and a specific user. This allows the SRM URLs to be used as physical file names for storage resources in CMS replica catalogs.

References

[1] Storage Resource Management: Concepts, Functionality, and Interface Specification, *Arie Shoshani*, Future of Data Grids Workshop, Berlin, 2004

[2] SRM Joint Functional Design, Version 1.0 http://sdm.lbl.gov/srm/documents/joint.docs/SRM.joint.func.design.part1.doc

[3] SRM joint methods specification version 1.0 http://sdm.lbl.gov/srm/documents/joint.docs/srm.v1.0.doc

[4] http://sdm.lbl.gov/srm-wg

[5] The Storage Resource Manager Interface Specification, version 2.1, Edited by *Junmin Gu*, *Alex Sim*, *Arie Shoshani*, available at http://sdm.lbl.gov/srm/documents/joint.docs/SRM.spec.v2.1.final.doc.

[6] Dcache Chep2003 presentation. Michael Ernst, Patrick Fuhramann, Tigram Mkrtchyan, available at http://www.dcache.org/manuals/chep2003.pdf

[7] Distributed Data Management in US-CMS, Michael Ernst, available at http://www.uscms.org/s&c/reviews/doe-nsf/2003-01/docs/USCMS-Grid-Storage.pdf

[8] A Scalable Storage Solution for U.S. CMS based on Cluster File Systems and Object Storage Devices Michael Ernst, Donald Petravick, available at http://www.uscms.org/s&c/reviews/doe-nsf/2003-01/docs/USCMS-OSD.pdf

[9] Security for Grid Services. V. Welch, F. Siebenlist, I. Foster, J. Bresnahan, K. Czajkowski, J. Gawor, C. Kesselman, S. Meder, L. Pearlman, S. Tuecke. Twelfth International Symposium on High Performance Distributed Computing (HPDC-12), IEEE Press, to appear June 2003, available at http://www.globus.org/research/papers/gauth02.pdf