Storage Resource Management for Distributed Data Applications

PI: Arie Shoshani* (LBNL)
CoPIs: Don Petravick (Fermilab), Alex Sim (LBNL)

Summary

Storage Resource Managers (SRMs) are middleware components whose function is to provide dynamic space allocation and file management on shared distributed storage systems. This effort supports the MICS mission in providing the technology needed to manage the rapidly growing distributed data volumes, as a result of faster and larger computational facilities. Over the last year, we deployed our SRMs in real production projects including multiple High Energy Physics experiments as part of the Particle Physics Data Grid (PPDG) project, and the Earth Science Grid (ESG) project, and the Scientific Data Management ISIC. SRMs are used in these projects for space management, robust file movement, and streaming data for analysis. In addition, we continued the coordination of an international effort to standardize SRM interfaces, and developed a powerful second generation standard specification.

One of the goals of the Storage Resource Management project was to develop a standard specification against which multiple implementations can be developed. This approach proved to be a remarkable and unique achievement, in that now there are multiple SRMs developed in various institutions around the world that interoperate.

The SRM functional specification effort was initiated by the Scientific Data Management Group (SDM) at LBNL. Because the SDM group is also participating to deploy SRMs in two of the National Collabratory projects, Earth System Grid and the Particle Physics Data Grid, SRM has developed into an internationally coordinated effort between several DoE laboratories including LBNL, Fermilab and TJNAF, as well as European institutions including CERN and RAL in the UK. This coordinated effort has resulted in the adaptation of the standard specification, and the development of multiple SRM middleware components in various institutions around the world to interface to their specific storage systems. This approach is particularly essential for providing distributed access to complex Mass Storage Systems (MSSs). SRMs were even developed for legacy MSS systems (e.g. at NCAR) which enables them to be accessed from the Grid. The concept of interoperability is illustrated in Figure 1, where the same client can interact with multiple SRMs implemented over existing diverse storage systems. The SRMs can also communicate with each other in a peer-to-peer fashion request and move files.

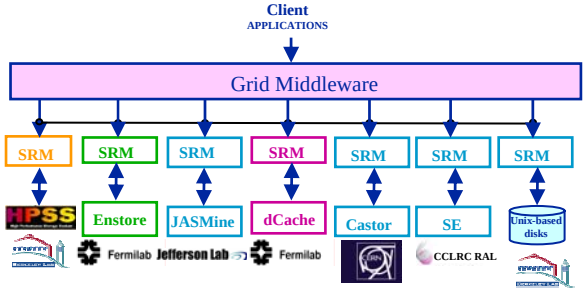

Figure 1: SRMs facilitate access uniformity of diverse distributed storage systems

* 510.486-5171, shoshani@lbl.gov

SRMs have been used in production by several facilities including BNL, NERSC, Fermilab, CERN, and JTNAF. Some are being used for access to files or for storing files in remote systems, and some are used for intensive data movement between storage systems. For example, The BNL to NERSC setup using SRMs that access HPSS, takes advantage of robustness features of the SRMs. They use the Berkeley-SRMs to move about 10,000 files per month (about 1 GB each) in an automated fashion. This arrangement, called a DataMover, is also being used by the Earth Systems Grid to move robustly a large volume of simulation production data from NERSC to NCAR, as well as ORNL to NCAR. The benefits are great reduction in the error rates, and essentially eliminating the human effort previously required. For example, a 50 fold reduction in error rates (from 1% to 0.02%) was achieved in routine file replication of the STAR experiment.

Another example of a successful deployment in the SRM-dCache developed at FNAL. It is widely deployed for use in the CMS project, and it interoperates with the SRM-Castor at CERN. This effort demonstrates the usefulness of SRMs by achieving sustained SRM-to-SRM managed transfers from Castor to FNAL dCache and onto tape at a rate between 40 and 60 MB/s. In addition, high rate, CERN disk to FNAL disk utilizing SRM clients were performed from approximately 150 nodes requested data from CERN. This maxed out at ~700 MB/s.

SRMs have also been used for projects in the Scientific Data Management ISIC. In one application, called a GridCollector, SRMs were used in combination of an efficient indexing method to greatly speed up the analysis of high energy data. In another project, SRMs are used to facilitate data intensive applications that use MPI-IO to transparently access files on remote storage systems, including mass storage systems. This is illustrated in Figure 2.

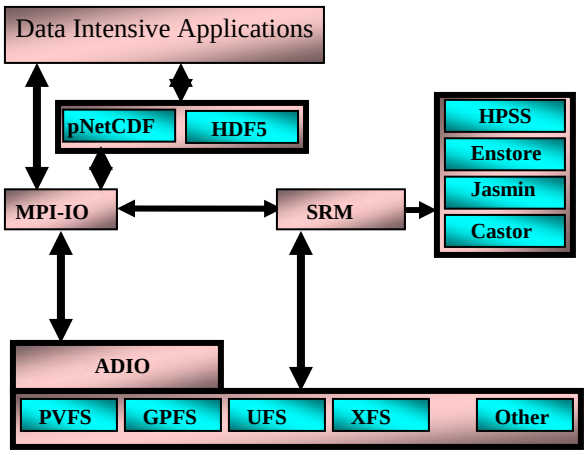

Figure 2: SRMs use for transparent access in data intensive applications.

The dream of having multiple, diverse, distributed storage systems that interoperate was fulfilled by this effort, thanks in great part to the goodwill and commitment of the international participants. The SRM standard is now widely adopted by various international efforts. Although the SRM specification has become a de-facto standard, there is an ongoing effort to standardize this functional specification through the Global Grid Forum (GGF). This activity is facilitated by an international SRM-collaboration that continues to enhance the SRM standard specification based on experience and new requirements.

For further information on this subject contact:
Dr. Arie Shoshani
Lawrence Berkeley National Laboratory
Tel: (510) 486-5171
Email: shoshani@lbl.gov