WLCG and SRM: some background

SRM v2.2 deployment Workshop - Edinburgh, UK
12-13 November 2007

Flavia Donno
IT/GD, CERN

Outline

- Storage Services in WLCG
- The classic SE
- The SRM v1.1 based SE
- The requirements for a Storage protocol
- The Storage Resource Manager v2.2
- Protocol definition activities
- The study of the SRM v2.2 specification
- Available implementations and test suites
- The GLUE Schema
- The GSSD working group
- Conclusions

Storage Services in WLCG

- *Storage Services* are *crucial* components of the *Worldwide LHC Computing Grid* (WLCG) infrastructure spanning more than 200 sites and serving computing and storage resources to the High Energy Physics LHC communities.
- Up to *tens of Petabytes of data* are collected every year by the 4 LHC experiments at CERN.
- It is crucial to *efficiently transfer* data to *Tier-1s* that contribute with their storage and computing power to the reconstruction step.
- An important role is also covered by the *Tier-2s* that provide experiments with the *results of the simulation*. Such results need to be transferred to Tier-1s and safely stored on permanent media.

Storage Services in WLCG: the Classic SE

- *The Classic SE* : an optimized FTP server with Grid authentication and authorization.
  - The first Storage Server in the Grid based on Globus GridFTP
  - Very simple solution that included simple and complex tape-based systems
- What are the capabilities of such a service ?
  - No possibility to query the service itself about its status, space available, etc. (one has to rely on the Information System)
- How are data accessed on such a storage server ?
  - Protocols supported: NFS/file, rfio, root, gsiftp
  - Discovery of related information
    - Different root directory for GridFTP and NFS or rfio
- What about growing file systems according to needs ? Or more in general managing space ?
  - Sometimes very hard
  - No explicit support for tape backend (pre-staging, pool selection, etc.)

Storage Services in WLCG: the SRM v1.1 SE

- The need for a standard interface for storage services in Grid was recognized.
  - International collaboration: LBNL, FNAL, CERN, JLAB
  - Provide basic functionality required
  - SRM v1.0 proposed. SRM v1.1 provides a Copy operation
  - SRM v1.1 implemented by all major storage providers: CASTOR, dCache, DPM
- The main functions:
  - Get, getRequestStatus, pin, unpin
  - Put, setFileStatus, Copy
  - getProtocols, AdvisoryDelete, FileMetaData
- Main features:
  - Asynchronous operations
  - Support for bulk requests
  - Protocol negotiation
- Main problems:
  - Missing reference implementation/No clear specs
  - Advisory delete
  - No space management
  - No explicit quality of storage management
  - No abort operations
  - No staging operations
  - ...

Requirements definition by dates

- In *June 2005* the *Baseline Service Working Group* published a report:
  - http://lcg.web.cern.ch/LCG/peb/bs/BSReport-v1.0.pdf
  - A Storage Element Service is mandatory and high priority.
  - The experiment requirements for a Grid storage service are defined
  - The full set of recommended feature available by *February 2006*
  - Experiments agree to use only high-level tools as interface to SRM
- The report was based on the early experience acquired with SRM v1.1 and v2.1 (never deployed).
- *Mumbai workshop* (*CHEP2006*): the experiments had learned more about what was needed and changed their requirements.
- In *May 2006* at FNAL the WLCG SRM Memorandum of Understanding (MoU) was agreed on:
  - http://cd-docdb.fnal.gov/0015/001583/001/SRMLCG-MoU-day2%5B1%5D.pdf

Basic requirements

- Support for Permanent files (and volatile copies)
- Support for Permanent Space
- Space Reservation : static only per VO.
  - If dynamic space reservation is available, allow for the possibility of releasing the allocated space
- Permission Functions only on directories based on VOMS group/roles
- Directory Management Functions
- Data Transfer and File Removal Functions
- File access protocol negotiation
- VO-specific relative paths

The Storage Resource Manager SRM v2.2

- The *Storage Resource Manager* (SRM) is a ***middleware component*** whose function is to provide ***dynamic space allocation*** and ***file management*** on shared storage components ***on the Grid***.
- More precisely, the SRM is a ***Grid service*** with several different implementations. Its main specification documents are:
  - A. Sim, A. Shoshani (eds.), The Storage Resource Manager Interface Specication, v. 2.2, available at http://sdm.lbl.gov/srm-wg/doc/SRM.v2.2.pdf.
  - F. Donno et al., Storage Element Model for SRM 2.2 and GLUE schema description, v3.5 available at: http://glueschema.forge.cnaf.infn.it/uploads/Spec/V13/SE-Model-3.5.pdf

The SRM Interface in a nutshell

- The SRM Interface Specification lists the *service requests*, along with the *data types* for their arguments.

- Function signatures are given in an implementation-independent language and grouped by functionality:
  - *Space management* **functions** allow the client to reserve, release, and manage spaces, their types and lifetimes. Support for different qualities of storage space.
    - Reserve/Release Space, ChangeSpaceForFiles, ExtendFileLifeTimeInSpace
  - *Data transfer functions* have the purpose of getting files into SRM spaces either from the client's space or from other remote storage systems on the Grid, and to retrieve them.
    - PrepareToPut/StatusOfPutRequest/PutDone
    - PrepareToGet/StatusOfGetRequest
    - BringOnline
    - Copy
    - ReleaseFiles, AbortRequest/Files, ExtendFileLifeTime
  - Other function classes are *Directory*, *Permission*, **and** *Discovery* **functions**.
    - Ping, Ls, Mkdir, Rm, Rmdir, SetPermission, CheckPermission, etc.

SRM 2.2 protocol definition activities

- *Long protocol definition process*: many discussions on the concepts of spaces, file copies management, name space handling.
- *Storage services providers* had to be heavily involved in the discussions.
- From May to December 2006 *weekly meetings* to follow the development.
- *Formal static and dynamic protocol models* defined: identifying unanticipated behaviours and interactions.
  - F. Donno, A. Domenici, *A Model for the Storage Resource Manager,* ISGC 2007, International Symposium on Grid Computing, Taipei, March 26 . 29, 2007

SRM 2.2 protocol definition activities

- Need for *testing suites to validate the protocol* against real use cases and to verify the *compliance of the implementations* against the spec.
- The testing campaign motivated the developers to *reconsider many of the initial assumptions and decisions*, leading to solutions that seem to better satisfy the needs of the users.
- The *LBNL testing suite* written in Java was the only one available till September 2006, run manually. Since January 2007 running automatically every day.
  - http://sdm.lbl.gov/srm-tester/v22-progress.html
  - http://sdm.lbl.gov/srm-tester/v22daily.html
- In September 2006 CERN took over from RAL the development of the *S2 SRM 2.2 testing suite*, enhancing it with a complete test set and with publishing and monitoring tools.
- *Reports to WLCG Management Board* to monitor progress.

Study of SRM 2.2 specification

- In September 2006 *very different interpretations* of the spec
- *6 Releases of the SRM v2.2* specification document: July, September, December 2006 and January(2x), April 2007
- Study of the spec (*state/activity diagrams*): many unspecified behaviours.
- A list of about *50 open issues* has been compiled in *September 2006*.
- Last *30 points discussed* and agreed during the *WLCG Workshop* in *January 2007*. Other major points delayed to SRM 3.0.

- *The study of the specifications*, the discussions and testing of the open issues have helped *insure coherence* in the protocol definition and *consistency between SRM implementations*.

https://twiki.cern.ch/twiki/bin/view/SRMDev/IssuesInTheSpecifications

Storage providers involvement: the available implementations

- SRM v2.2 implementations are available today for the following Storage Services:
  - *CASTOR2* : Hierarchical Storage Server (HSS). Developed by CERN and RAL. SRM v2.2 support in v.2.1.4.
  - *dCache* : HSS developed by DESY and FNAL. SRM 2.2 support in v1.8.
  - *DPM* : disk-only developed by CERN. SRM v2.2 support in v1.6.5 in production.
  - *StoRM* : disk-only developed by INFN and ICTP. SRM v2.2 interface for many filesystems: GPFS, Lustre, XFS and POSIX generic filesystem. SRM v2.2 support in v1.3.15.
  - *BeStMan* : disk-based developed by LBNL. SRM v2.2 support in v2.2.0.0.

Need for tests: the black box testing

- Functional testing used for:
  - *Validation*: against the explicit and implicit user requirements
  - *Consistency*: checking for inconsistency, incompleteness, or inefficiency
  - *Verification*: correctness of implementations with respect to the specification
  - *Performance*: … and reliability
- *Exhaustive test* can demonstrate the correctness of System Under Test (SUT). But it is impracticable.

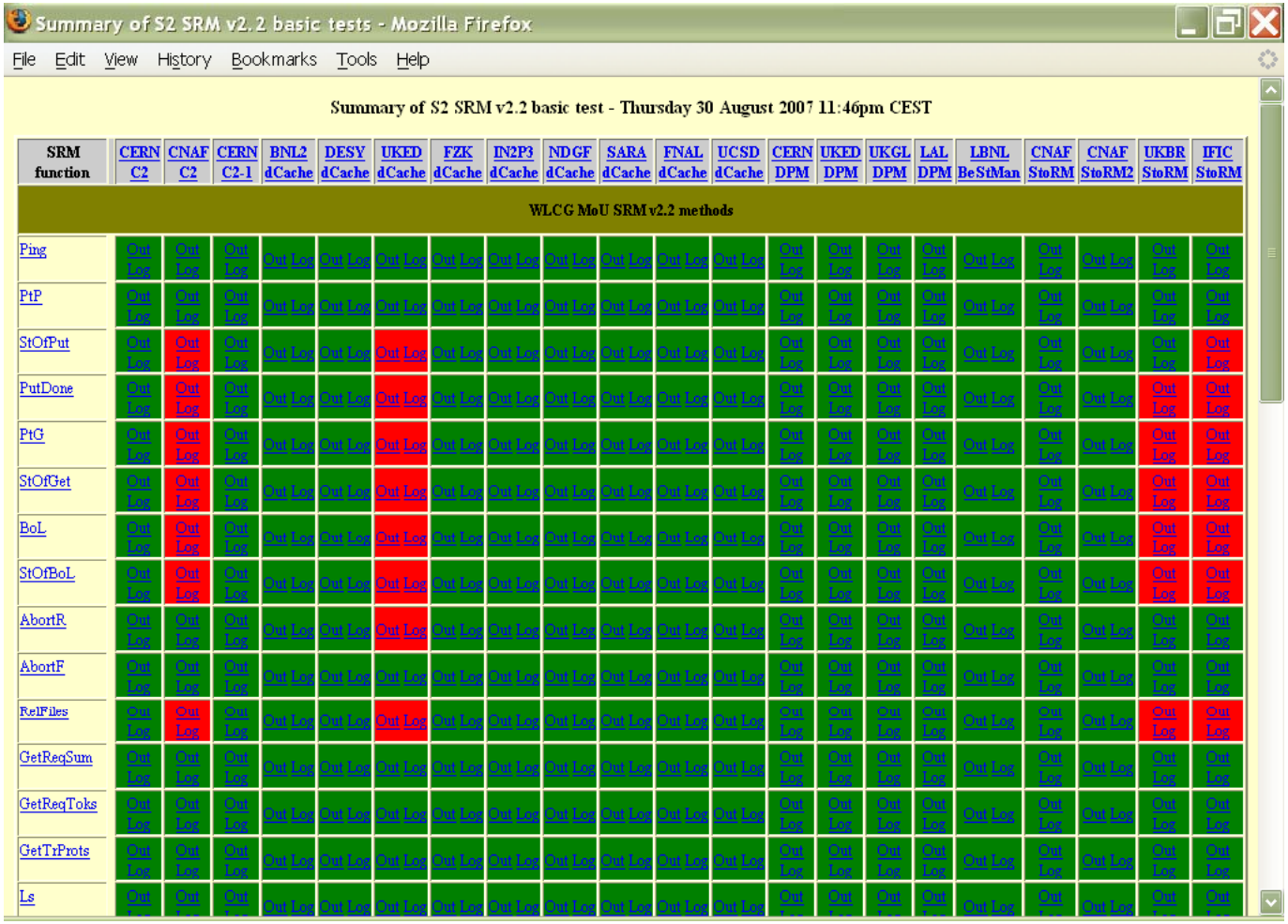
The S2 test families

- The S2 test suite families.
  - **Availability**: Ping and full put cycle
  - **Basic**: basic functionality, checking only return codes and passing of basic input parameters
  - **Usecases**: testing boundary conditions, exceptions, real use cases extracted from the middleware clients and experiment applications.
  - **Interoperability**: servers acting as clients for put and get requests, cross copy operations
  - **Exhaustive**: Checking for long strings, strange characters in input arguments, missing mandatory or optional arguments. Output parsed.
  - **Stress**: Parallel tests for stressing the systems, multiple requests, concurrent colliding requests, space exhaustion, etc.
  - **Bdii**: Checks that the information published in the information system is correct
  - **Lcg-utils**: Equivalent to current SAM tests adapted to SRM v2.2
- S2 tests cron job running *5 times per day* on *21 SRM v2 endpoints*/5 storage flavours with *different configurations*.
- The S2 test suite has been integrated in the *WLCG certification process*
- *S2* has been *used* as well *by the developers* to check their implementations and *distributed to OSG* for Tier-2 validation.
- Test families *continuously expanded* to cover new scenarios, new decisions taken at the protocol level, and check for discovered bugs or race conditions.
- *Integration with WLCG high-level tools* is done.
- The S2 test families will **soon** be part of **SAM framework**

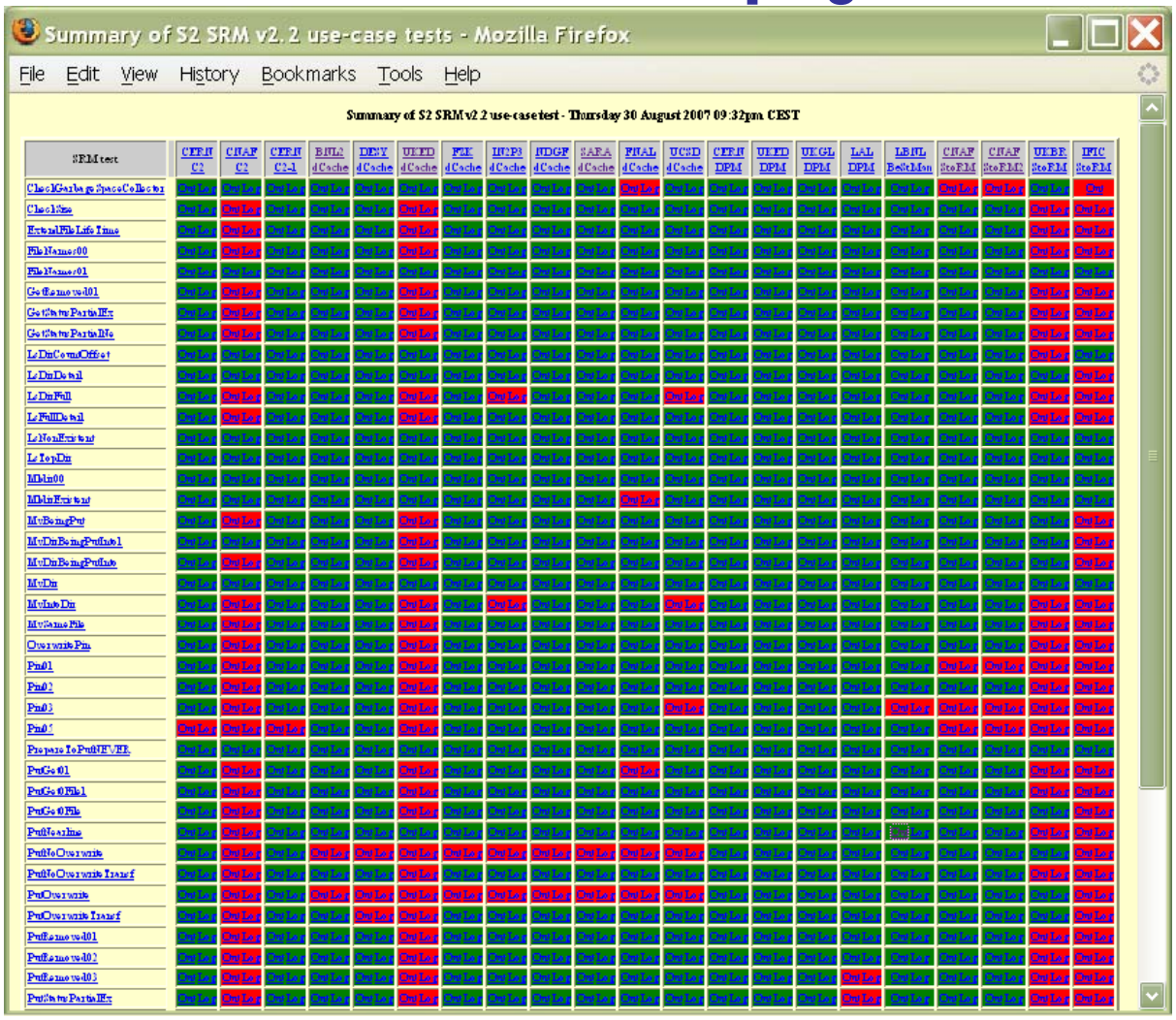
Tests executed

- *Results published* on a web page. Latest and history available:
  - https://twiki.cern.ch/twiki/bin/view/SRMDev
- Test results and issues are *discussed on srm-tester and gssd lists*
  - https://hpcrdm.lbl.gov/mailman/listinfo/srmtester
  - https://mmm.cern.ch/public/archive-list/s/storage-classes-wg

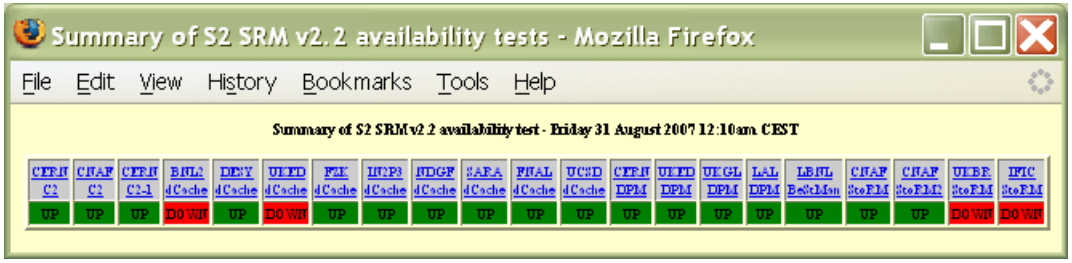
The S2 result web-pages: Basic tests

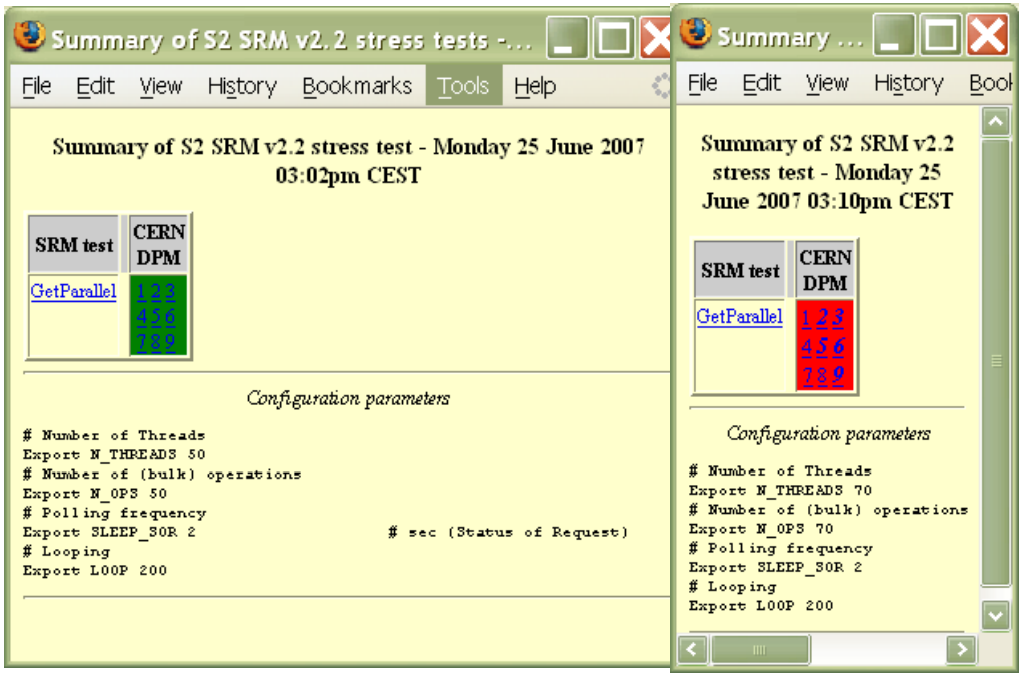


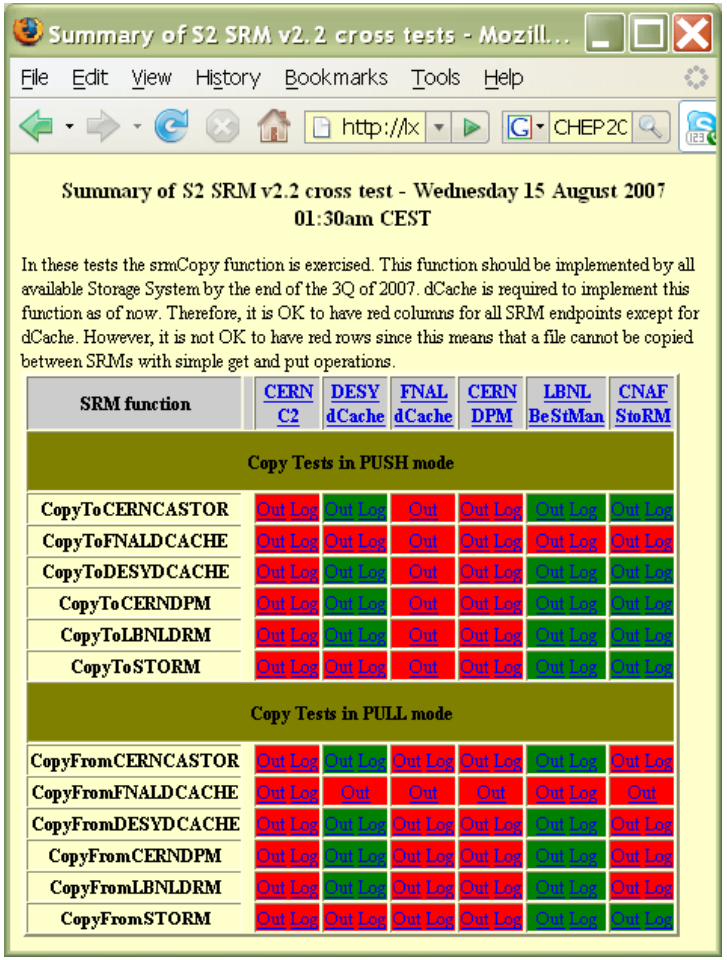
The S2 result web-pages: Use-case tests



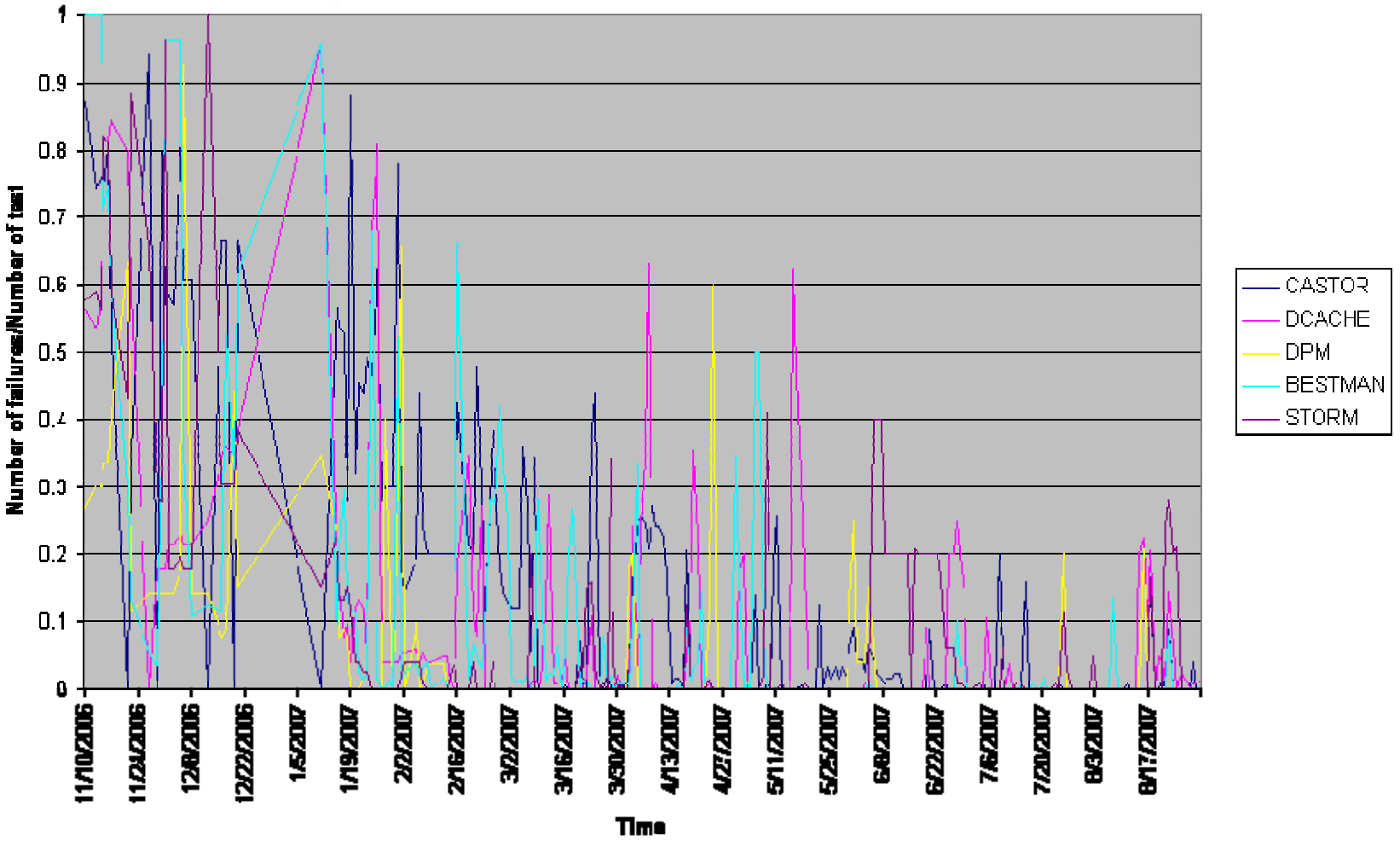
The S2 result web-pages: Availability, interoperability and stress







Test results



Other testing efforts: the SRM clients

- Other testing groups to:
  - Check the status of the high-level SRM v2 clients such as GFAL, lcg-utils, FTS
  - Transparent access to storage independent of the protocol in use (SRM v1 vs SRM v2).
  - Experiment data access patterns.
  - Correct handling of VO specific proxies with VOMS groups/roles.
  - Compliance to user requirements and functionalities needed
  - Availability of required bindings (C++, Python, Perl, etc.) and support for mandatory platforms (SL4 32-bit mode, 64-bit mode later).
- Outcomes:
  - User documentation.
  - User education and training.
  - Preparation to the integration of experiment frameworks with the SRM v2 environment.

GLUE Schema

- The WLCG information system publishes details about grid services in a format that is specified by the GLUE schema.
- GLUE 1.3 includes information needed for SRM v2.2, such as the space information.
  - http://glueschema.forge.cnaf.infn.it/Spec/V13
- The modelling and testing exercises have helped in the specifications of GLUE 1.3 for Storage Services.
- The current experience has also been taken as a base for the design of the new GLUE 2.0 model for Storage.

Grid Storage System Deployment (GSSD)

- https://twiki.cern.ch/twiki/bin/view/LCG/GSSD
- Mailing list: storage-class-wg@cern.ch

- Mandate:
  - Coordinating the deployment of SRM v2.2 in WLCG with sites, experiments, and developers.
  - Establishing a migration plan from SRM v1 to SRM v2 to ensure smooth operations during the transition.
  - Coordinating the provision of the necessary information by the Storage Providers in order to publish and monitor the status of storage resources, guaranteeing the availability of the requested resources and their correct usage.
  - Ensure transparency of data access and the functionalities required by the experiments.
  - Organize support for storage in WLCG
  - Provide a forum where storage issues are discussed and solved

Conclusions

- The SRM specification definition and implementation process has evolved in a *world-wide collaboration effort* with developers, independent testers, experiments and site administrators.
- Much *clearer* description of SRM *specifications*.
- Well *established* and agreed *methodology* to verify the status of the implementations and validate sites.
- *Clear plan* to make sure that experiment requirements are covered.
- The modelling and testing exercises have helped in the specifications of Storage Services in the *Grid Information System*.
- The GSSD working group is in charge of the *deployment plan* for SRM v2.2 in production on the WLCG infrastructure, to be ready for LHC operations.