


Status and plans of
dCache and SRM projects
at Fermilab



Priorities

- Ensure success of our stakeholders' dCache by improving reliability, maintainability as needed
- Collaborate with DESY and NDGF in the development and release process for mutual benefit
- Develop new features needed by Fermilab Experiments, US CMS-T1, WLCG and OSG
- Improve automation and monitoring to reduce operational burdens
- Various levels of support for Public, CDF, US CMS-T1, OSG and WLCG dCache systems.

Resilient Manager

- Stable operations in US CMS T1
- More development, including integration with Space Management, only if required by experiments, OSG or WLCG

Addendum to the WLCG Usage Agreement

- Operations on Space reservations are protected by ACLs
- Support for “Read Space Management” - space reservation with `SrmPrepareToGet` and `SrmBringOnline`
- Extension of the set of supported Storage Classes to include T1D1
- We are working on implementation of the short term solution and waiting for feedback from experiments on the rest of the features included in new WLCG Usage Agreement, after the experiments gain production experience

Short Term Solution for WLCG experiments

- Usage of T1D0 in conjunction with srmBringOnline instead of T1D1
 - srmReleaseFiles by name without request id
- Protection of space reservation release operation only

SRM BringOnline functionality

- PinManager retains PinRequest Creator's Authorization info
- Authorize extend and release operations on pins
- Allows release of all pins by file names only without unintended consequences - no releases of other people's pins

SRM LS and RM

- SRM Blocking operations cause great performance impact
- Reduce the performance impact
- Investigate scheduled execution

SRM Horizontal Scalability

- Collaboration with BNL
- Multiple SRM Web Service front-ends perform CPU and memory intensive GSI Authentications
- One or more SRM Back-ends handle the requests
- Vendor neutral JMS technology for messaging

SRM monitoring and Error reporting

- Improve monitoring and reporting of transfers, especially remote GrifFTP transfer failures
- More descriptive error reports about SpaceManager related errors, especially for transfer operations
- Space Manager Web Monitoring
- Query page for debugging specific Remote Transfer failures

gPlazma Plans

New User Mapping Schema

- Motivation
 - Multiple rootpaths per session can lead to inconsistent placement.
 - File sharing access via common uid -ownership is ambiguous.
- Mapping Schema
 - Each DN maps to one rootpath.
 - Each DN maps to one uid (recommended unique).
 - Each Group/Role maps to a gid or list of gids.
 - File sharing access is through group permissions.
- Mapping Functions
 - grid-uidmap, grid-gidmap files.
 - Internal or External service, such as and LDAP server.
 - Dynamically assign new uids, etc to new users.
 - Consistent mapping for returning users.
 - Site-wide uniformity of assignments.
- Previous methods still supported, with the above schema restriction.

gPlazma Plans

New XACML Authorization Plugin

- Motivation
 - Avoid writing a new plugin for each new authorization service to be utilized.
- Solution
 - OASIS Standard, XACML: **eXtensible Access Control Markup Language**.
 - Grid Collaboration: EGEE, OSG VO Services Project, Globus, and Condor
 - Defined specific schema to be used for grid authorization.
 - Adapting existing client and server software.
 - gPlazma, PRIMA, GT, VOMS, GUMS, SCAS.
- Any service supporting the specification will be usable by gPlazma.

Namespace

- Scale the namespace to support CMS T1 rate of namespace operations and number of entries
 - Performance improvements in Pnfs as a Stop Gap measures
 - Chimera as a long term solution
- Registration of namespace entry deletions
 - Risk free automatic cleanup of Precious data

Berkley DB Based PNFS

PostgreSQL 16 tasks, 16384 files/directories

SUMMARY: (of 3 iterations)

| <i>Operation</i> | <i>Max</i> | <i>Min</i> | <i>Mean</i> | <i>Std</i> |
|----------------------------|-----------------|-----------------|-----------------|------------|
| ----- | --- | --- | ---- | ---- |
| <i>Directory creation:</i> | <i>55.344</i> | <i>38.516</i> | <i>49.245</i> | <i>7.</i> |
| <i>Directory stat :</i> | <i>1105.956</i> | <i>1013.797</i> | <i>1049.322</i> | <i>40.</i> |
| <i>Directory removal :</i> | <i>90.188</i> | <i>68.425</i> | <i>81.317</i> | <i>9.</i> |
| <i>File creation :</i> | <i>187.439</i> | <i>153.335</i> | <i>165.913</i> | <i>15.</i> |
| <i>File stat :</i> | <i>1049.236</i> | <i>1013.307</i> | <i>1030.509</i> | <i>14.</i> |
| <i>File removal :</i> | <i>286.152</i> | <i>256.363</i> | <i>270.032</i> | <i>12.</i> |

Berkley DB 16 tasks, 16384 files/directories

SUMMARY: (of 3 iterations)

| <i>Operation</i> | <i>Max</i> | <i>Min</i> | <i>Mean</i> | <i>Std</i> |
|----------------------------|-----------------|-----------------|-----------------|------------|
| ----- | --- | --- | ---- | ---- |
| <i>Directory creation:</i> | <i>295.922</i> | <i>239.936</i> | <i>262.260</i> | <i>24.</i> |
| <i>Directory stat :</i> | <i>2800.140</i> | <i>2743.410</i> | <i>2771.290</i> | <i>23.</i> |
| <i>Directory removal :</i> | <i>356.093</i> | <i>296.062</i> | <i>333.736</i> | <i>26.</i> |
| <i>File creation :</i> | <i>335.957</i> | <i>284.474</i> | <i>317.327</i> | <i>23.</i> |
| <i>File stat :</i> | <i>2784.884</i> | <i>2706.608</i> | <i>2738.489</i> | <i>33.</i> |
| <i>File removal :</i> | <i>411.862</i> | <i>406.396</i> | <i>409.017</i> | <i>2.</i> |

Internal Consistency Scans

- Petabyte scale systems will routinely observe data corruption
- Proliferation of custom scripts for checksum verification, namespace and companion consistencies
- Internal dCache scans can detect and fix variety of the problems and/or alert the administrators

Summary

- No radical changes or ambitious new function implementations
- Concentration on stability, scalability, code quality and testing should result in reduced operational and support burden for everyone