|  |  |
| --- | --- |
| Unique ID of the requirement | #924 |
| Title of requirement | File updates |
| Reference | <https://rt.egi.eu/guest/Ticket/Display.html?id=924> |
| Priority | 4 |
| Submitter | Life Sciences Grid Community (LSGC VRC - <http://www.egi.eu/collaboration/LSGC.html>) |
| Status at TCB | Endorsed (At TCB-9, 15/11/2011) |
| RT ticket belonging to this requirement | File updates: <https://rt.egi.eu/guest/Ticket/Display.html?id=924> |
| Description of the requirement | It should be possible to update an existing file content, including all its replicas (rather than removing and re-creating a file).  We would like to have 'rsync' style file updates and automatic update of all its replicas. |
| Related tickets in the EGI Helpdesk | GGUS ticket:  <https://ggus.eu/ws/ticket_info.php?ticket=76058>  BUG at savannah:  <https://savannah.cern.ch/bugs/?89637> |
| Goals and objectives | To shorten the time and additional work for updating files and synchronizing replicas. |
| Impact | Takes too much time and additional work for updating files and their replicas. |
| Affected services | GFAL/lcg\_util v.1.11.18 |

**SoS, File updates (924)**

**Name:** File updates

**Assessed Requirements**

It should be possible to update an existing file content, including all its replicas (rather than removing and re-creating a file).

It should be possible to update an existing file content, including all its replicas (rather than removing and re-creating a file). We would like to have 'rsync' style file updates and automatic update of all its replicas.

**Executive Summary**

Grid data management has been optimized for a model, which does not allow modification of existing files in a replicated environment. Reversing this decision would be very expensive and is currently not in the scope of the related product team. The reasons are manifold. As the deployed system only loosely coupled replicas, there is now way to guaranty consistency of the same dataset amount different SE’s. Moreover, some SE’s are able storage data on tape. Modifying data after the file has been archived would result in invalidating data on tape, which is a rather expensive operation. As StoRM is build on-top of regular file systems, it would allow for modify/write operations on the SE level. However, in a completely distributed file environment, as provided by EMI, modifying files won’t be possible. For the same technical reasons, Amazon S3 only provides a PUT/GET/REMOVE interface and doesn’t allow modifying existing data.

**Efforts assessment**

Working an a system which allows modifying existing files in a distributed environment is not reasonable with the resource available in our community.

**Milestones and timelines**

N.A.

**Resources**

N.A.

**Risks**

N.A.

**Constraints**

N.A.

**Assumptions**

N.A.