OSG STORAGE OVERVIEW

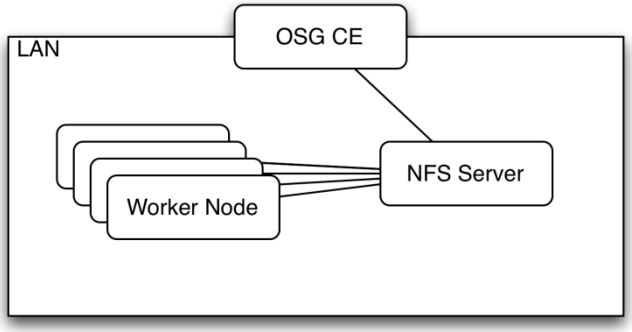
Tanya Levshina

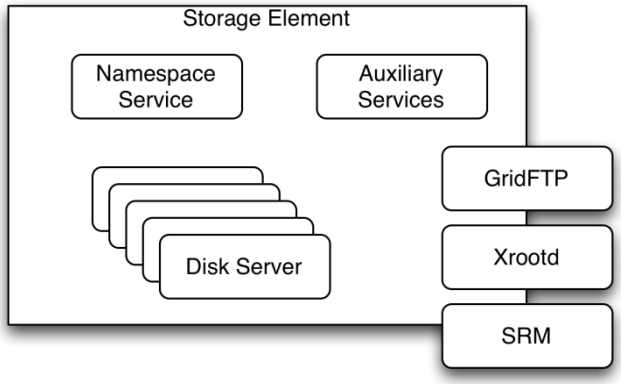
Talk Outline

- OSG Storage architecture
- OSG Storage software
  - VDT cache
  - BeStMan
  - dCache
  - DFS:
  - SRM Clients
  - Auxiliary software
- Statistics
- OSG Storage Group
- Summary

OSG Storage Architecture

- Classic Storage Element
  - POSIX-mounted storage
  - Mounted and writable on the CE.
  - Readable from the worker nodes
  - Not-scalable under heavy load
  - High-performance FS is not cheap
  - Space management is not trivial
- Storage Element
  - SRM endpoint
  - Provides GridFTP Load balancing
  - Transfers via GridFTP servers
  - May provide internal access protocols (xroot, Posix)





Pictures from B. Bockelman's presentation at OSS2010

Virtual Data Toolkit

VDT provides:

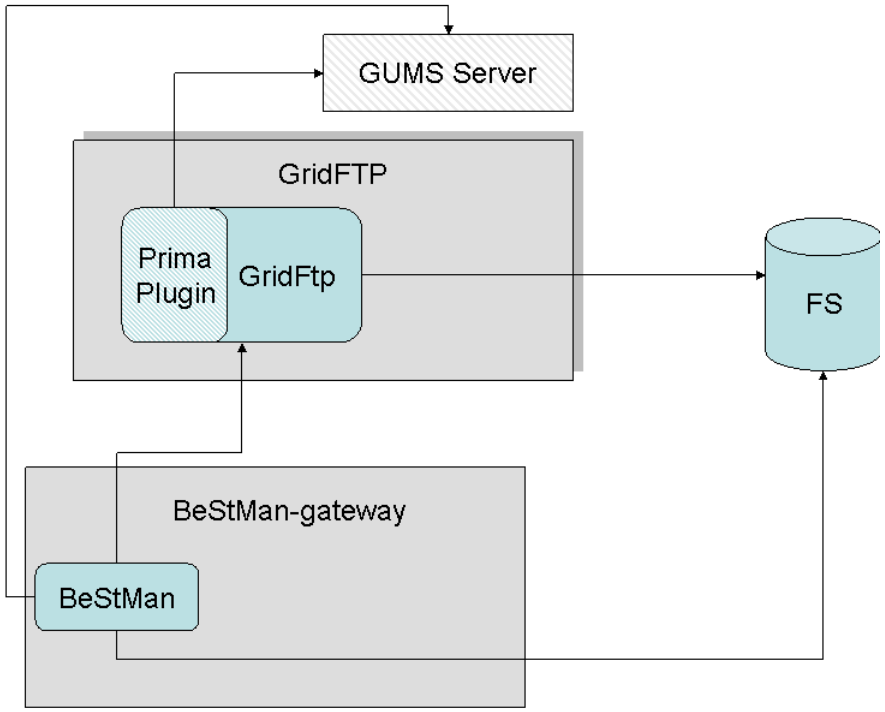
- A standard procedure for installation, configuration, services enabling, startup and shutdown
- Simplified configuration scripts
- All packages in one cache:
  - BeStMan
  - GridFTP
  - CA certificates, CRL installation, update
  - Log rotation scripts
  - Probes
- Straightforward upgrade procedure

BeStMan 2

- **Berkeley Storage Manager (BeStMan) 2**
  - **Retains all functionalities of the previous BeStMan**
    - SRM v2.2 implementation – interoperable and compatible to other implementations
    - Works on existing storages with posix-compatible file systems
      - Adaptable to special file systems and storages with customized plug-in
    - Supports multiple storage partitions
      - Supports pre-defined static space tokens
    - Easy adaptability and integration to special project environments
    - Supports multiple transfer protocols
      - Supports load balancing for multiple transfer servers
    - Supports grid-mapfile or GUMS server
    - Supports Gateway Mode for faster performance
      - Scales well with some file systems and storages, such as Xrootd and Hadoop
  - **Improvements from the previous BeStMan**
    - Jetty based web server container
      - Better performance in http connection handling through configurations
    - Updated dependency libraries – both server and clients
    - New packaging but the same setup process as the previous package

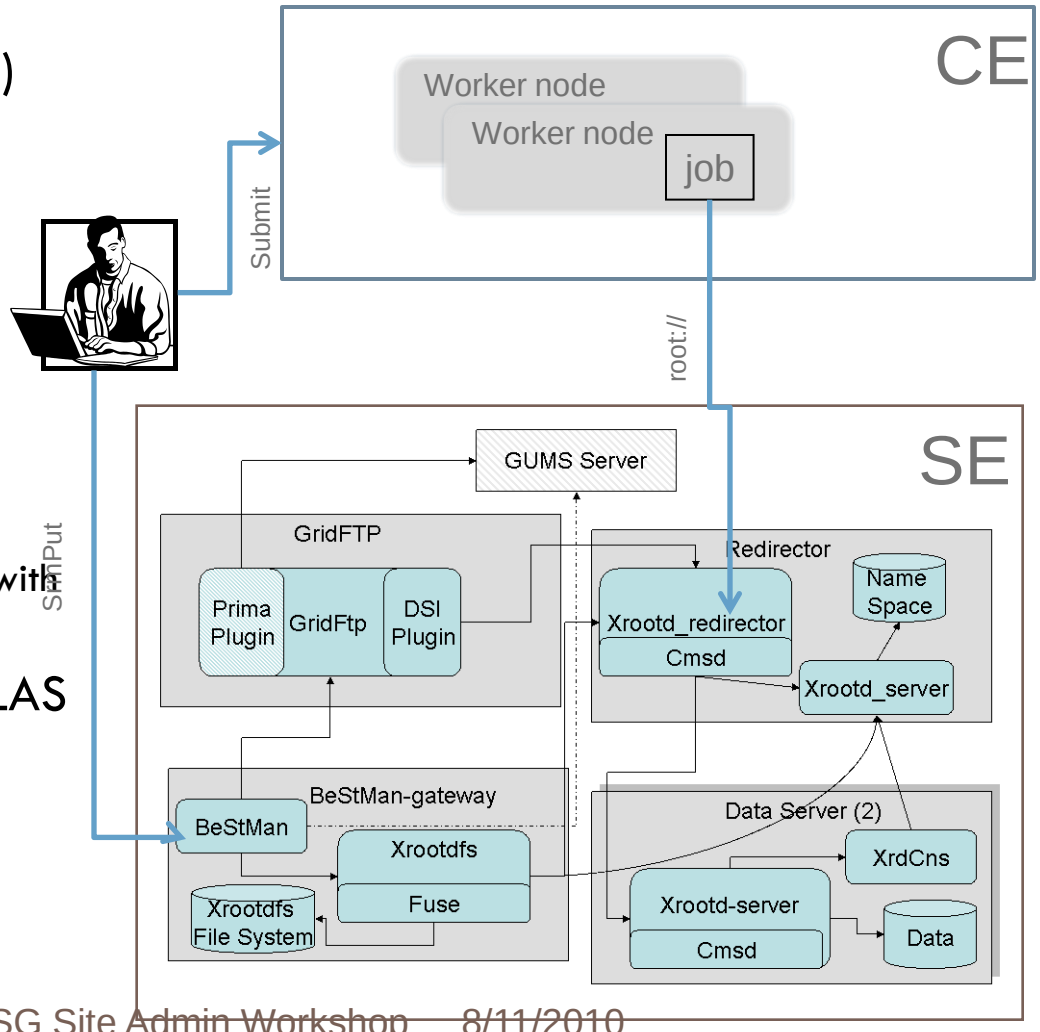
BeStMan-gateway

- Generic SRM v2.2 load balancing frontend for GridFTP servers
- Light-weight implementation of SRM v2.2 for POSIX file systems
  - srmPing,
  - srmLs
  - srmRm
  - srmMkdir
  - srmRmdir,
  - srmPrepareToPut (Status, PutDone),
  - srmPrepareToGet (Status, ReleaseFiles)
- Designed to work with any Posix-like file systems
  - NFS, GPFS, GFS, Lustre, XrootdFS, HDFS
- Doesn't support queuing or disk space management
- Hands-on installation will follow



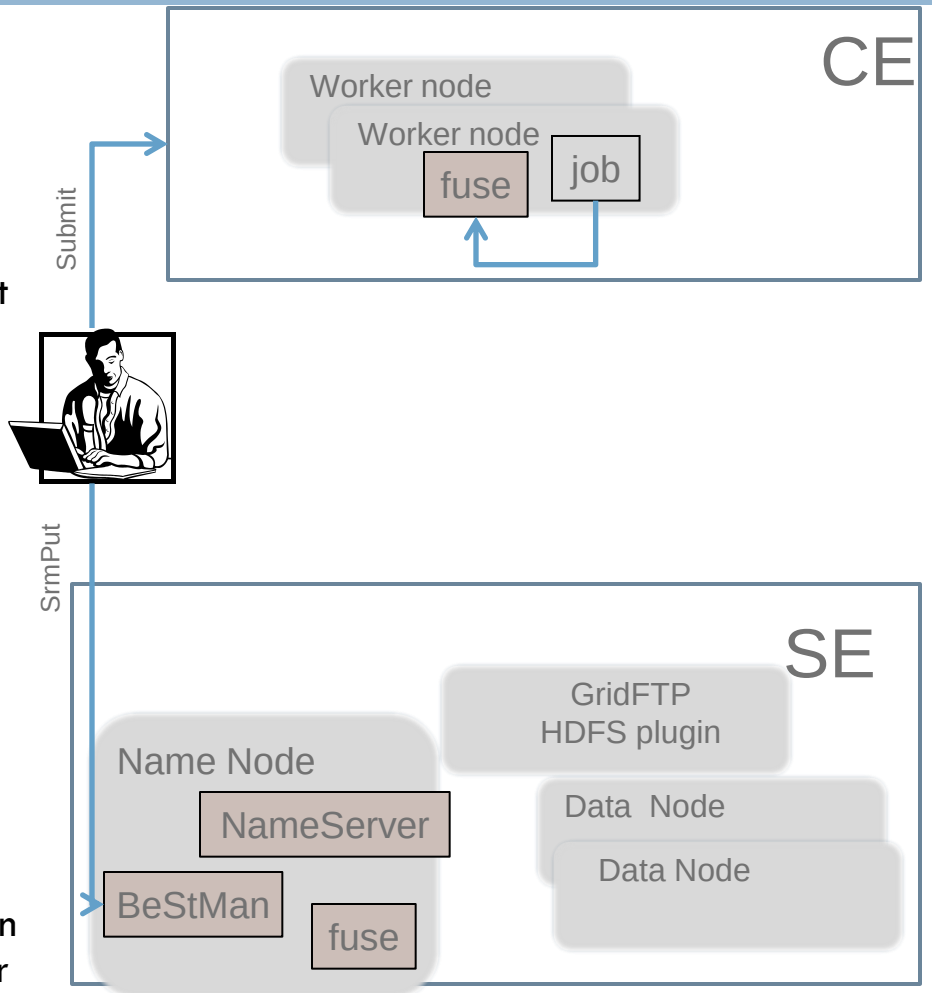
BeStMan-gateway/Xrootd

- Xrootd (developed at SLAC, contribution from CERN, others)
  - is designed to provide access
    - POSIX-like
    - via root framework (root://)
    - Native commands (xrdcp,…)
  - Allows cluster globalization
  - Allows unix-like user/group authorization as well as X509 authentication.
  - Requires FUSE, XrootdFS to hook with BeStMan, GridFTP DSI plugin
- Currently is used by many ATLAS and ALICE T2 sites , recommended for all Atlas T3
- Can be installed from VDT (pacman)



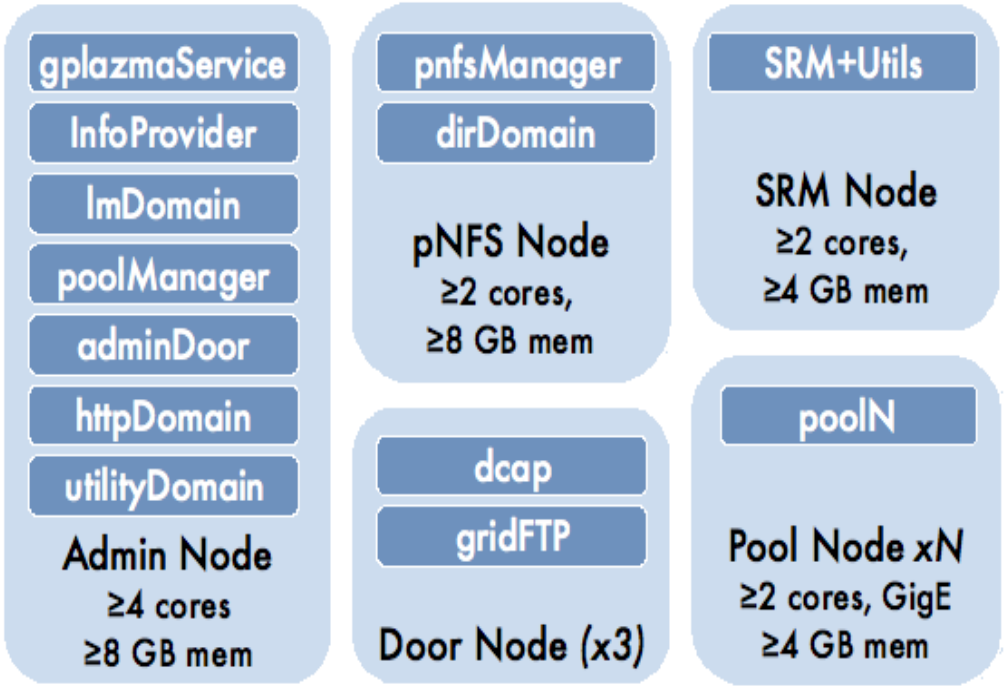
BeStMan-gateway/HDFS

- Hadoop DFS is developed in the Apache project.
  - Creates multiple replicas of data blocks
  - Distributes them on data nodes throughout a cluster
  - Consists of two major components:
    - Namenode: central metadata server.
    - Datanode: file servers for data
  - Allows replication
  - Runs on commodity hardware
  - unix-like user/group authorization, but no strict authentication
  - Requires FUSE to hook with BeStMan, GriFTP –HDFS plugin
- BeStMan/HDFS and all auxiliary software can be installed from rpms (hands-on tutorial later today)



dCache

- dCache is a distributed storage solution developed at DESY, Fermilab and NGDF
- dCache supports requesting data from a tertiary storage system
- Full SRM 2.2 implementation
- nfs-mountable namespace
- Multiple access protocols
- Replica Manager
- Role-based authorization
- Information Provider
- Probably, requires more administration then T3 may provide
- Available from dcache.org and VDT with auxiliary software and installation/configuration script



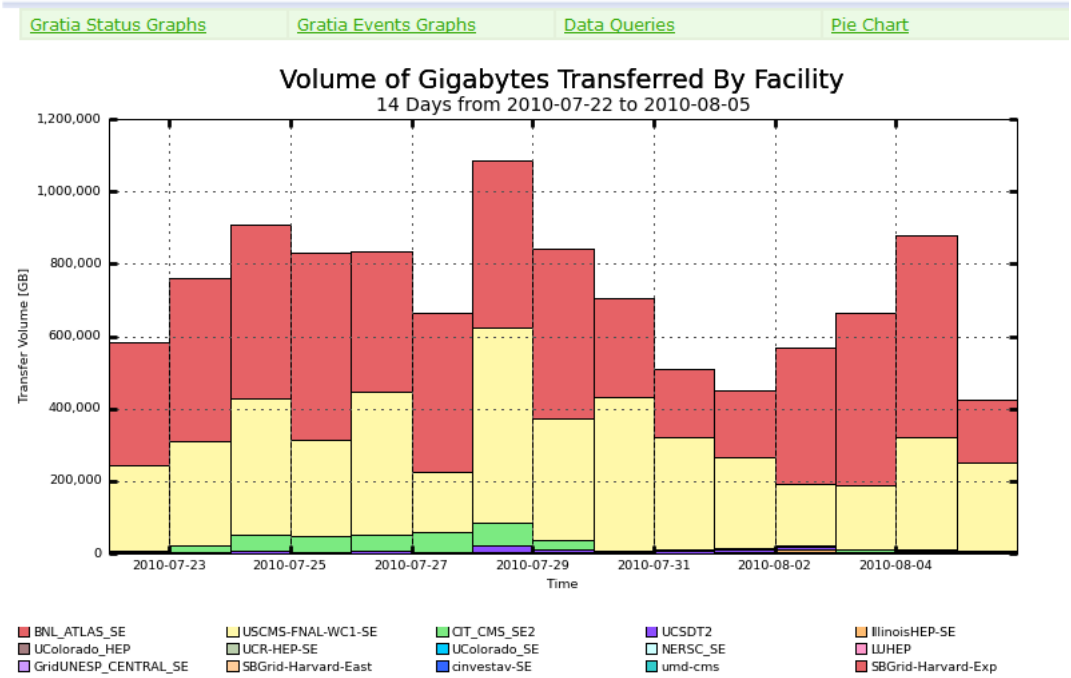
Picture from Ted Hesselroth's
(from presentation: "Installing and Using SRM-dCache"

SRM Clients

- Available from VDT (OSG-Client, wn-client)
- SRM-Fermi-Client commands
  - developed and maintained at Fermilab
  - access any Storage Element that complies with the SRM 1 or 2 specification
- SRM-LBL-Client commands
  - developed at LBNL,
  - access any SRM v2.2 based storage components
- LCG-utils is a suite of client tools for data movement written for the LHC Computing Grid.
  - based on the Grid File Access Library,
  - access any SRM v2.2 based storage components
  - May use logical file names and require a connection to a BDII-based catalog for some commands file copies and deletions, which take endpoints based on the SRM URL.

Gratia transfer probes

- Included in BeStMan, dCache VDT Cache
- Reports to Gratia Accounting System
- Generates accounting information about file transfers, source, destination, size of the file and owner

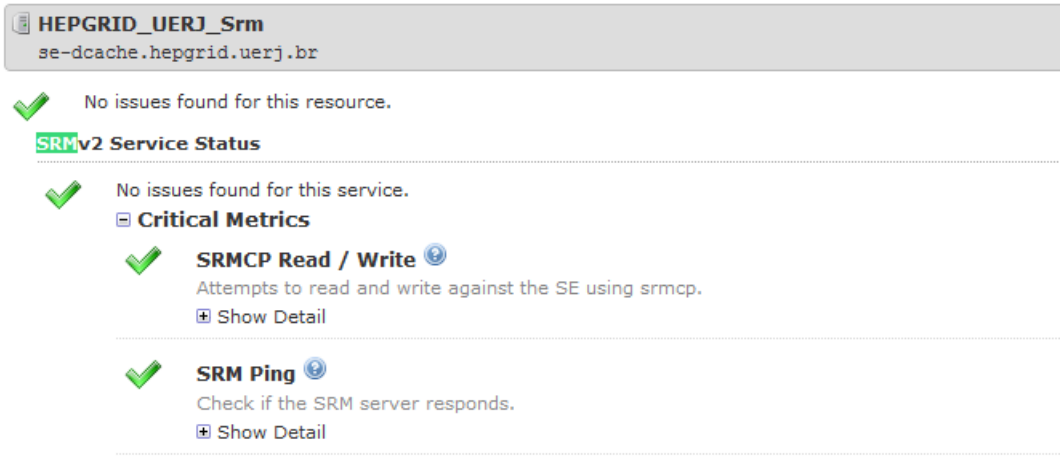


http://t2.unl.edu/gratia/xml/facility_transfer_volume

RSV Storage probes

- The Resource and Service Validation (RSV) provides monitoring infrastructure for an OSG site admin.
  - Client
  - Collector/Server
  - Periodic Availability Reports
- Storage RSV probes:
  - Current probes:
    - srm-ping,
    - srm-copy
  - Coming soon: srmtester suite

MyOSG



OSG SE Statistics

These are the unofficial statistics based on BDII:

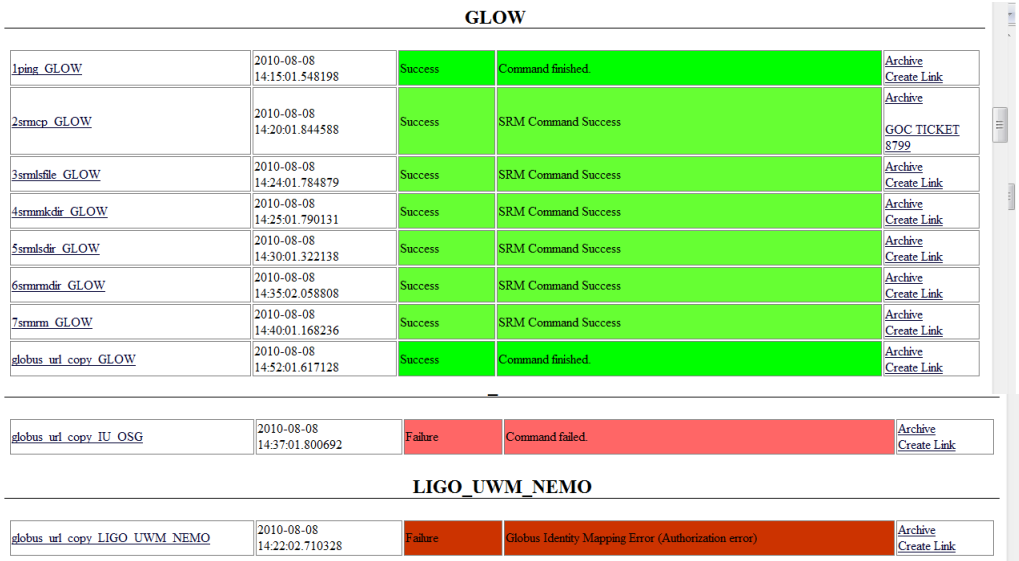
- Number of sites providing Storage Elements: 49
- Number of sites running dCache: 12
- Number of sites running BeStMan-gateway: 37
  - HDFS 6
  - Xrootd 3
  - Lustre 3
  - REDDNet 1
  - All other sites: Local disk, NFS?
- Number of sites reporting Gratia GridFTP Transfer Probes: 15 (daily transfer ~170000 files, 800 TB)

OSG Storage Group

- Group members (all part time):
  - Ted Hesslroth (dcache, discovery tools)
  - Tanya Levshina (OSG Storage coordinator)
  - Abhishek Rana (hadoop)
  - Neha Sharma (support, dcache, probes,test suites)
  - Alex Sim (bestman develper and support)
  - Douglas Strain (rsv probes, xrootd, pigeon tools)
- Packages certification
- Test suites development, test stands
- Auxiliary software development
  - Gratia and RSV probes
  - Discovery tools
  - Pigeon tools (not in VDT yet)
- Documentation
- Support for site administrators
  - GOC Tickets creation/monitoring
  - Liaison to developers groups
- Active mailing list: osg-storage@opensciencegrid.org

Discovery and Pigeon Tools

- Discovery tools provide a convenient way to discover storage elements and related information (surl, end path, available space) for a particular VO by queering BDII information.
- Pigeon tools (created on top of Discovery tools) help a non-owner VO to debug site problems with Public Storage allocated for this VO
  - Runs periodically
  - Allows to see detailed errors
  - Allows to generate and monitor GOC ticket
  - Keeps archive
- Will be available as RSV probes



GLOW

1ping_GLOW	2010-08-08 14:15:01.548198	Success	Command finished.	Archive Create Link
2srmcp_GLOW	2010-08-08 14:20:01.844588	Success	SRM Command Success	Archive GOC TICKET 8799
3srmlsfile_GLOW	2010-08-08 14:24:01.784879	Success	SRM Command Success	Archive Create Link
4srmmkdir_GLOW	2010-08-08 14:25:01.790131	Success	SRM Command Success	Archive Create Link
5srmlsdir_GLOW	2010-08-08 14:30:01.322138	Success	SRM Command Success	Archive Create Link
6srmrmdir_GLOW	2010-08-08 14:35:02.058808	Success	SRM Command Success	Archive Create Link
7srmrm_GLOW	2010-08-08 14:40:01.168236	Success	SRM Command Success	Archive Create Link
globus_url_copy_GLOW	2010-08-08 14:52:01.617128	Success	Command finished.	Archive Create Link

globus_url_copy_IU_OSG	2010-08-08 14:37:01.800692	Failure	Command failed.	Archive Create Link

LIGO_UWM_NEMO

globus_url_copy_LIGO_UWM_NEMO	2010-08-08 14:22:02.710328	Failure	Globus Identity Mapping Error (Authorization error)	Archive Create Link

Storage Documentation On OSG Twiki

- Release Documentation:

  https://twiki.grid.iu.edu/bin/view/ReleaseDocumentation

- Main Storage Page:

  https://twiki.grid.iu.edu/bin/view/ReleaseDocumentation/Storage

  - End User Guide
  - Site Admin Guide

- Tier-3 specific documentation:

  https://twiki.grid.iu.edu/bin/view/Tier3/WebHome

- OSG Storage Group Meetings

  https://twiki.grid.iu.edu/bin/view/Storage/MeetingMinutes

Summary

- There is plethora of available storage software solutions
- Each solution has some pros and cons
- Tier-3 coordinators are trying to come up with the most comprehensive solution that satisfies:
  - The needs of experiments
  - Hardware availability
  - Available efforts for installation, support and maintenance
- VDT provides means to improve and simplify installation and configuration
- OSG Storage group is ready to help!

Announcement

- OSG Storage Forum
  - University of Chicago
  - September 21-22, 2010
  - General discussion of various storage solutions (new features, major improvements), scalability and performance.

http://indico.fnal.gov/conferenceDisplay.py?confId=3377