



Open Science Grid

VO Application Monitoring on the Open Science Grid

Suchandra Thapa

Computation Institute

University of Chicago



Open Science Grid

Look at 4 VO monitoring systems

- ATLAS (PanDA)
- CMS (ARDA)
- GADU (GNARE)
- D0 (SAMGrid and JIM)

June 25, 2007

HPDC Grid Workshop 2007 -
Monterey, California



Open Science Grid

ATLAS

- All US jobs on OSG are handled by PanDA
- Started development in Aug 2005 and went into production in Dec 2005
- Handles 100K-200K jobs at US ATLAS sites and ~4 times that including non-US sites

June 25, 2007

HPDC Grid Workshop 2007 -
Monterey, California



Open Science Grid

Panda Job Submissions

- All jobs get sent to a global task queue
 - Users can submit jobs using a python client interface
 - Automated submissions from analysis tools
 - Regional production jobs also submitted
- Queue is used to track all active jobs and send new jobs to idle nodes

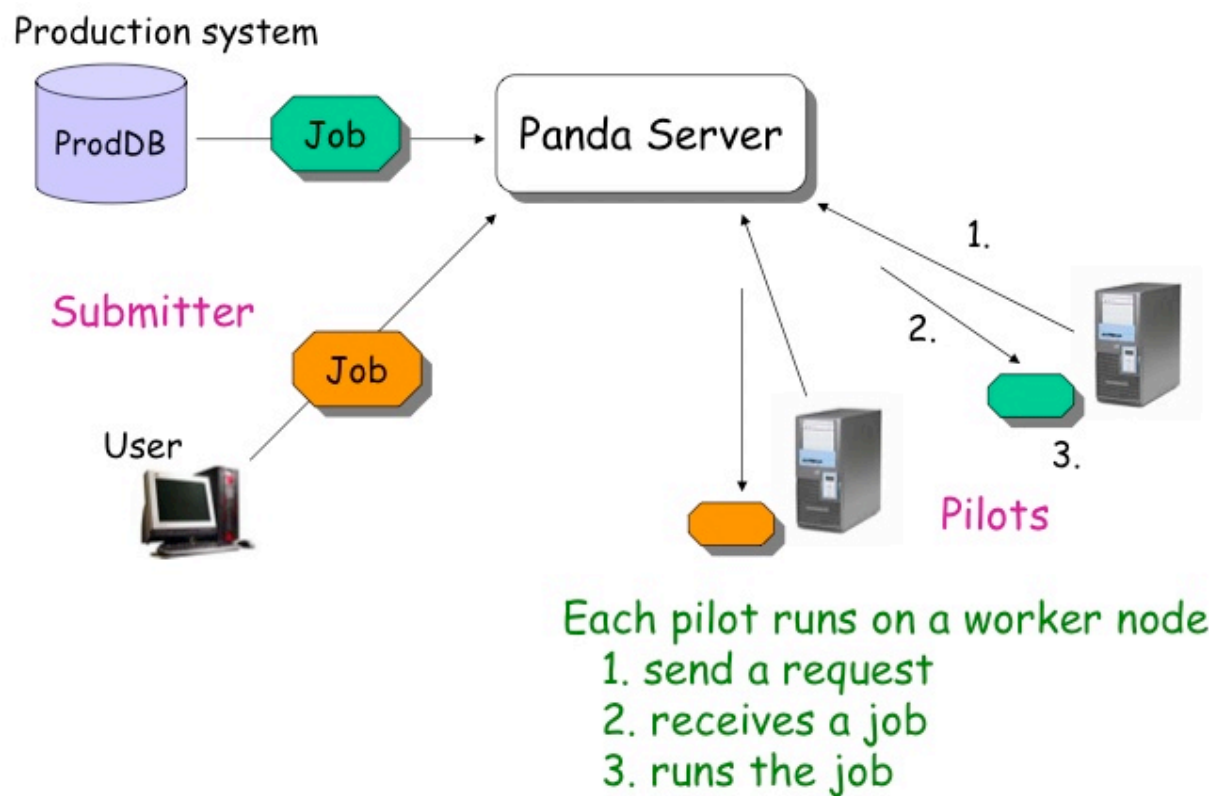


Job Execution

- Participating cluster are sent batches of pilot jobs
- When pilots start running
 - Contacts global task queue to see next task to run
 - Check to make sure data has already been staged in
 - If data is present, pilot gets necessary task from queue and starts computations
 - Heartbeat communications are initiated with the central server every 30 minutes
 - When completed, data is staged out and job is marked as completed in global task queue
 - If any errors occur or if data is not present, pilot immediately quits



Panda Job Flow





Job State and Reporting

- Pilot jobs report to PanDA dashboard every 30 minutes with job state
- Job state file also created on system with recovery information at each of these reports
- If a pilot crashes, then state file is not updated
- After 60 minutes, other pilot jobs on the cluster will notice the stale state file and report this to the server and attempt error recovery



Open Science Grid

PanDA Dashboard

Panda Production Operations Dashboard

http://gridui02.usatlas.bnl.gov:25880/server/pandamon/query?

Documents Tools Validation Styles Transformations Views Stats S.O.A. Authoring Help

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options Q N

Configuration
Dashboards: [Production](#) [DDM](#) [AutoPilot](#) [Sites & Grids](#) [Analysis](#) [Physics data](#) [Usage & Quotas](#) [Plots](#) [ArdaDash](#)
0 min old [Update](#)

Panda monitor
[Quick guide](#), [twiki](#)

User info
[Jobs - search](#)
Recent [running](#), [activated](#), [waiting](#), [assigned](#), [defined](#), [finished](#), [failed](#) jobs
Select [analysis](#), [production](#), [test](#) jobs
Quick search
Job
Dataset
Task
File

Summaries
Blocks: days
Errors: days
Nodes: days
[Daily usage](#)

Tasks - search
[Generic Task Req](#)
[EvGen Task Req](#)
[CTBSim Task Req](#)
[Task list](#)
[Task browser](#)

Datasets - search
[Dataset browser](#)
[New datasets](#)
[Aborted MC datasets](#)
[Panda subscriptions](#)
[All subscriptions](#)

Datasets Distribution
[DDM Req](#)
[Req list](#)
[AODs](#)
[RDOs](#)
[DB Releases](#)
[Validation Samples](#)

Panda Production Operations Dashboard
Panda shift [guide](#) [calendar](#) [mailing list](#)

Servers: Panda:OK Panda-dev:OK Logger:OK DQ2:offline
Tasks assigned to OSG
Jobs updated >12 hrs ago: **activated:80** **running:none**
Jobs updated >36 hrs ago: **transferring:11005**

Space available at sites:

Site	GB	As of
AGLT2	5782	06-22 06:28
BU_ATLAS_Tier2	663	06-22 06:22
BU_ATLAS_Tier2o	676	06-21 18:51
SLACXRD	14141	06-22 06:58
UTA_SWT2	2573	06-22 06:56

Pilot job requests per hour, last 3 hours

	Production An
AGLT2	113
ANALY_BNL_ATLAS_1	69
ANALY_LONG_BNL_ATLAS	39
BNL_ATLAS_1	39
BNL_ATLAS_DDM	48
BU_ATLAS_Tier2	231
BU_ATLAS_Tier2o	229
IU_ATLAS_Tier2	4
MWT2_IU	126
MWT2_UC	146
OU_OCHEP_SWT2	229
OU_OSCER_ATLAS	212
SLACXRD	229
UC_ATLAS_MWT2	230
UC_Teraport	64
UTA_SWT2	319

Production job summary, last 12 hours (Details: [errors](#), [nodes](#))

Site	Nodes	Jobs	Latest	defined	assigned	waiting	activated	running	holding	transferring
All	1314	13194	06-22 12:19	0	100	3	83	193	84	12474
AGLT2	15	50	06-22 11:59	0	0	0	0	1	0	8
BNL_ATLAS_1	169	460	06-22 12:14	0	0	0	80	118	70	0
BNL_ATLAS_2	0	0		0	0	0	0	0	0	0
BU_ATLAS_Tier2	26	269	06-22 11:57	0	15	0	0	6	0	248
BU_ATLAS_Tier2o	71	2825	06-22 05:38	0	1	0	0	0	0	2824
IU_ATLAS_Tier2	20	26	06-19 06:26	0	0	0	0	0	0	26
MWT2_IU	35	247	06-22 06:20	0	21	0	0	0	0	226
MWT2_UC	34	1687	06-19 07:27	0	0	0	0	0	0	1687
Unassigned	1	2	06-22 07:16	0	0	0	0	0	0	0

Done

June 25, 2007

HPDC Grid Workshop 2007 -
Monterey, California



Open Science Grid

CMS Remote Analysis Builder

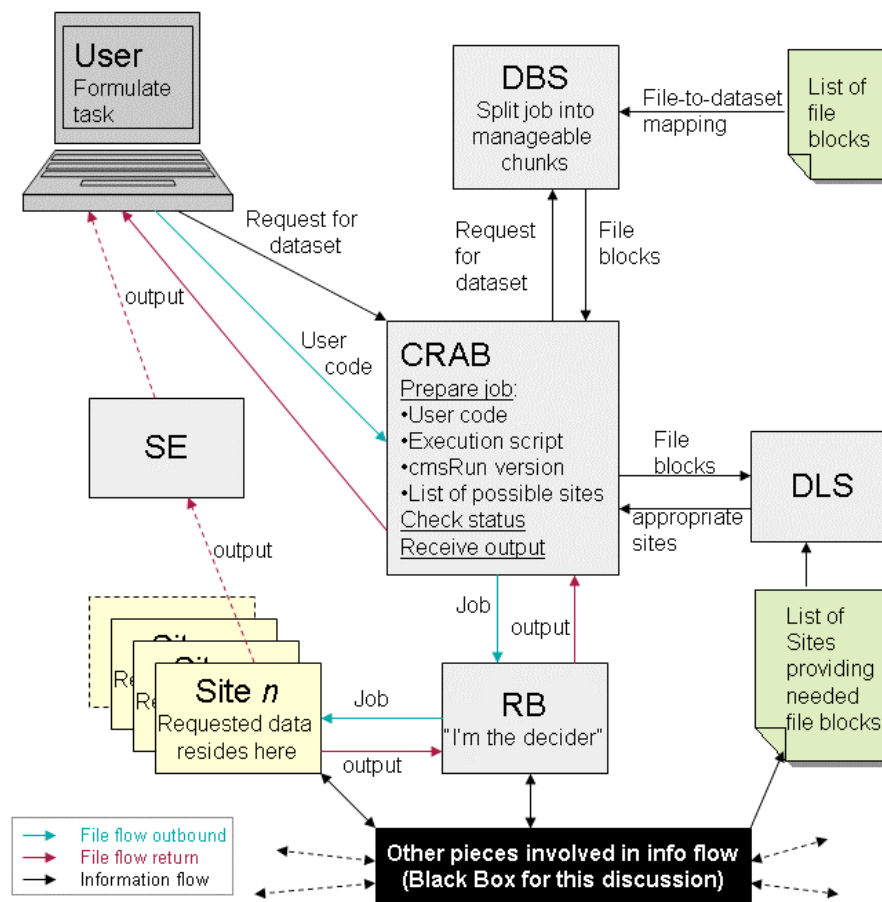
- User submits job and specifies data set using client interface
- CRAB contacts central server to decide how to split job and where data sets are located
- Job and job information sent to resource broker (RB)
- RB determines the best sites to run the job
- CRAB responsible for monitoring status and obtaining results

June 25, 2007

HPDC Grid Workshop 2007 -
Monterey, California



CMS Workflow





Open Science Grid

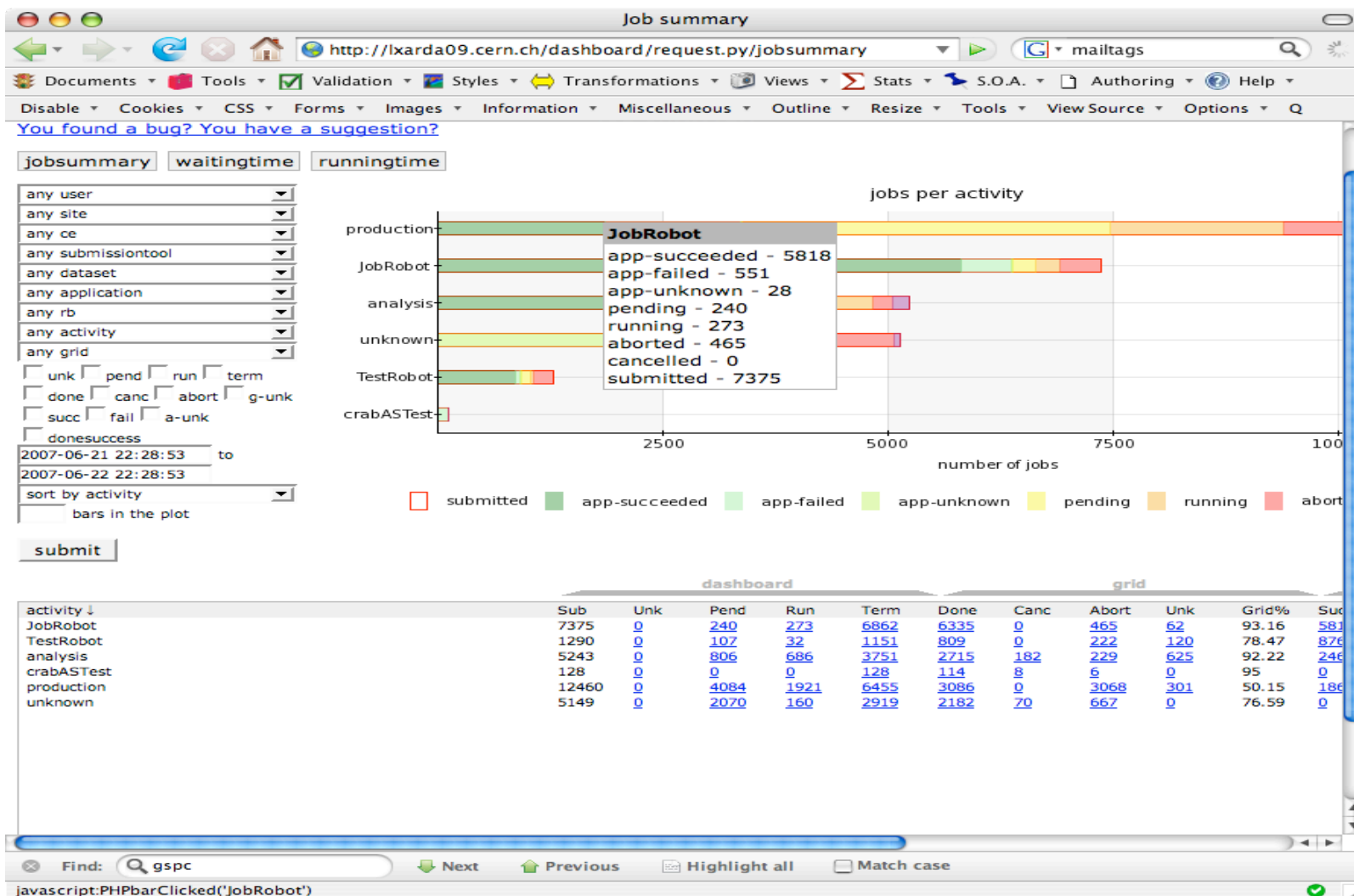
CMS Monitoring

- Uses CMS job bookkeeping and monitoring system
- Obtains information from workflow monitoring system using the MonaLISA protocol
- Information displayed on the ARDA dashboard
- Users can also use CRAB to get status information on their jobs



Open Science Grid

Arda Dashboard



June 25, 2007

HPDC Grid Workshop 2007 -
Monterey, California



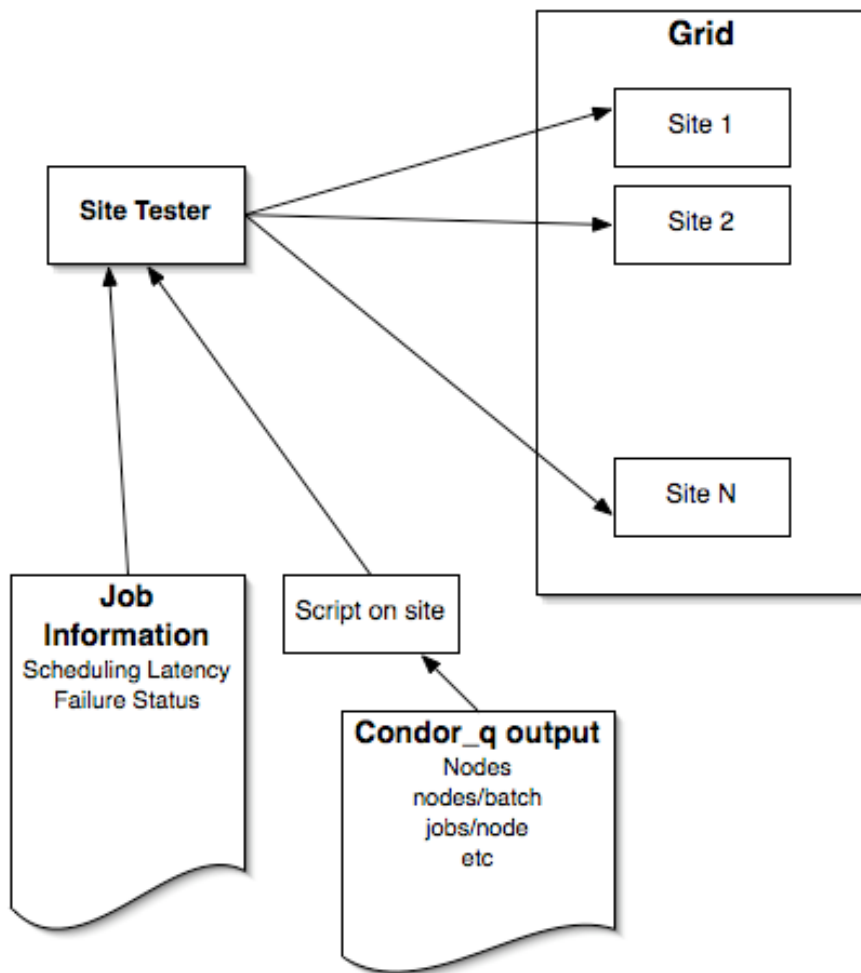
Open Science Grid

GADU Monitoring

- Monitoring and site selection done by Site Selector
- Testing done by spending probe jobs to sites and monitoring responses
- Site Selector keeps a list of jobs running on each site as well as information on site details



GNARE Probes





Open Science Grid

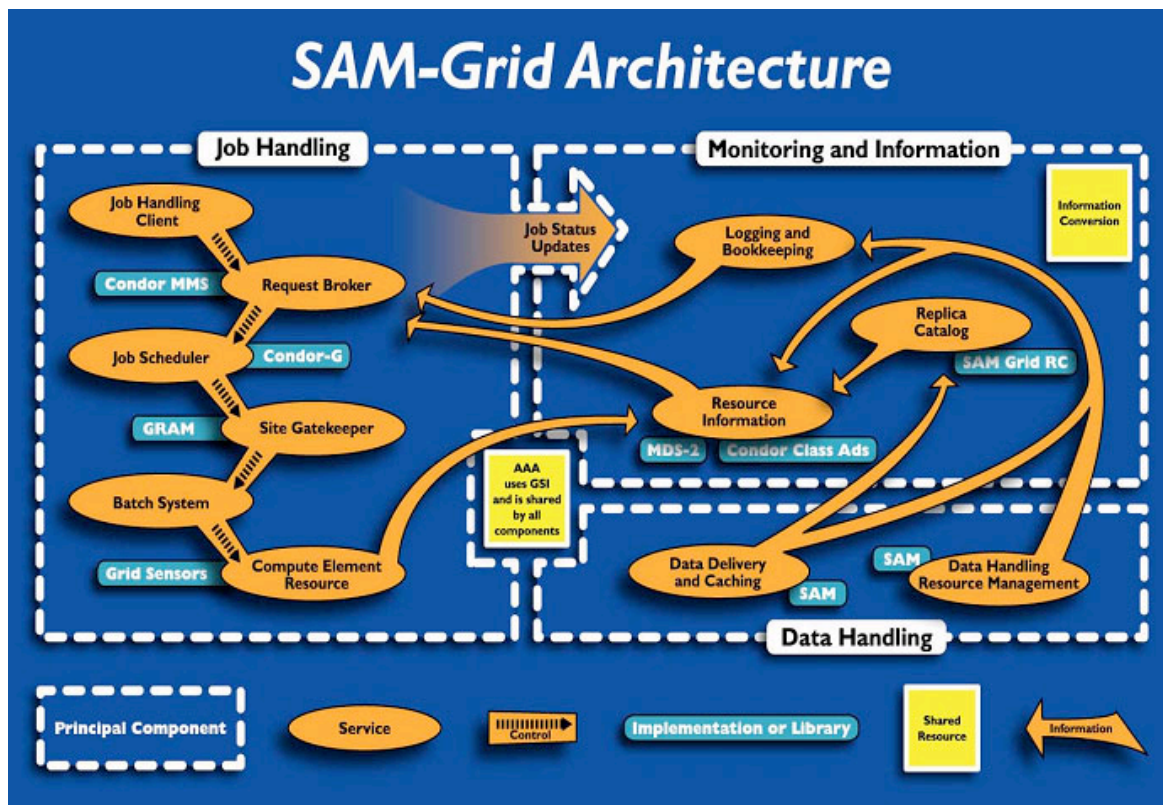
D0 Monitoring for SAM-Grid

- Done by Job Information Monitor (JIM)
- Obtains information different sources
 - Condor classads
 - MDS information service



Open Science Grid

SAM Architecture



June 25, 2007

HPDC Grid Workshop 2007 -
Monterey, California



MDS System

- Based on OpenLDAP and OpenSSL
- Pull service - server must query MDS system on each site to get information
- Scripts run locally on the cluster obtain and populate local database
- Central service later pulls and collates information from all sites



Condor Classads

- Modified condor classads to add additional information
- Classads collected on a central server which will act as a **Resource Selection Server** (ReSS)
- ReSS is a push server with each CE sending data to the central service at specified intervals (~10 min)
- Work on integrating this into OSG using CEMon to central server at Fermilab



Open Science Grid

Samgrid monitoring

SAM-GRID INFORMATION AND MONITORING SYSTEM

http://samgrid.fnal.gov:8080/

Documents Tools Validation Styles Transformations Views Stats S.O.A. Authoring Help

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options S

Launching the Monitoring System:

Please click at the map to monitor the execution sites.
Get information about the [submission sites](#).
Get information about the [advertised sites](#).

Participating Experiments:
● D0 ● CDF

This web portal is best viewed in Mozilla Firefox with a 1280 x 1024 or higher screen resolution.
Send comments about website to: sam-admin@fnal.gov
For more information, see www-d0.fnal.gov/computing/grid
[Security](#) [Privacy](#) [Legal](#)

Find: classad Next Previous Highlight all Match case

Done

June 25, 2007

HPDC Grid Workshop 2007 -
Monterey, California



Summary

- PanDA - uses heartbeat information reported from running jobs to monitor sites
- CRAB - uses information from workflow management system to monitor and track jobs
- GNARE - uses probe jobs sent to sites to monitoring sites
- SAMGrid - uses services running on sites to obtain monitoring information



Open Science Grid

Future of Monitoring?

- More comprehensive systems provided by OSG
- Gratia, CEMon, Monalisa, GIP, etc.
- Central resource databases run by Grid Operations Center
- Hopefully will allow more robust and accurate monitoring as well as provide something that a VO can use and build on



Links

- ATLAS -
<https://twiki.cern.ch/twiki/bin/view/Atlas/Panda>
- CMS -
<https://twiki.cern.ch/twiki/bin/view/CMS/WorkBook>
- GADU -
<http://compbio.mcs.anl.gov/gnare/doc.cgi>
- D0 - <http://www-d0.fnal.gov/computing/grid/>



Open Science Grid

Thanks

Special thanks to: Robert Gardner, Marco Mambelli,
Frank Wuertwein, Dinanath Sulakhe, Parag
Mhashilkar

June 25, 2007

HPDC Grid Workshop 2007 -
Monterey, California