



# Summary of Monitoring BOF

James Casey, CERN

WLCG Operations Meeting, 25<sup>th</sup> January 2007





# Aims of the BOF



- Present the three new WLCG Working Groups to the WLCG community
  - System Administration
  - Grid Service Monitoring
  - System Analysis
- Show some early work
  - Monitoring Questionnaire
- Gather feedback
- Get more participation from site administrators
  - That means you !
    - What do you want?
    - What have you got that you want to share?

# WLCG Collaboration Workshop (Tier0/Tier1/Tier2)

22-26 January 2007

CERN

[Home](#) > [Timetable](#) > Session details



## Monitoring BOF session

Presentation of new WLCG Monitoring efforts. This will cover both the interface of site fabric and grid-level monitoring and the 'System Analysis' done by ARDA to correlate experiment monitoring with system monitoring




**Place:** CERN  
Room: **160 1-009**

**Dates:** Tuesday 23 January 2007 16:00

**Conveners:** **CASEY, James**  
**ANDREEVA, Julia**  
**NEILSON, Ian**  
**Dr. FORTI, Alessandra**  
**JOUVIN, Michel**

**Material:**  [Meeting Notes](#)

[Contribution List](#) [Time Table](#)

Tuesday, 23 January 2007	
16:00	[80] <b>Overview of System Management group</b> by Dr. Alessandra FORTI (University of Manchester) (160 1-009: 16:00 - 16:20)  <a href="#">slides</a>
	[79] <b>Overview of Grid Service Monitoring group</b> by Ian NEILSON (CERN) (160 1-009: 16:20 - 16:40)  <a href="#">slides</a>
	[81] <b>Overview of System Analysis group</b> by Julia ANDREEVA (CERN) (160 1-009: 16:40 - 17:00)  <a href="#">slides</a>
17:00	[82] <b>Discussion &amp; Questions</b> (160 1-009: 17:00 - 18:00)

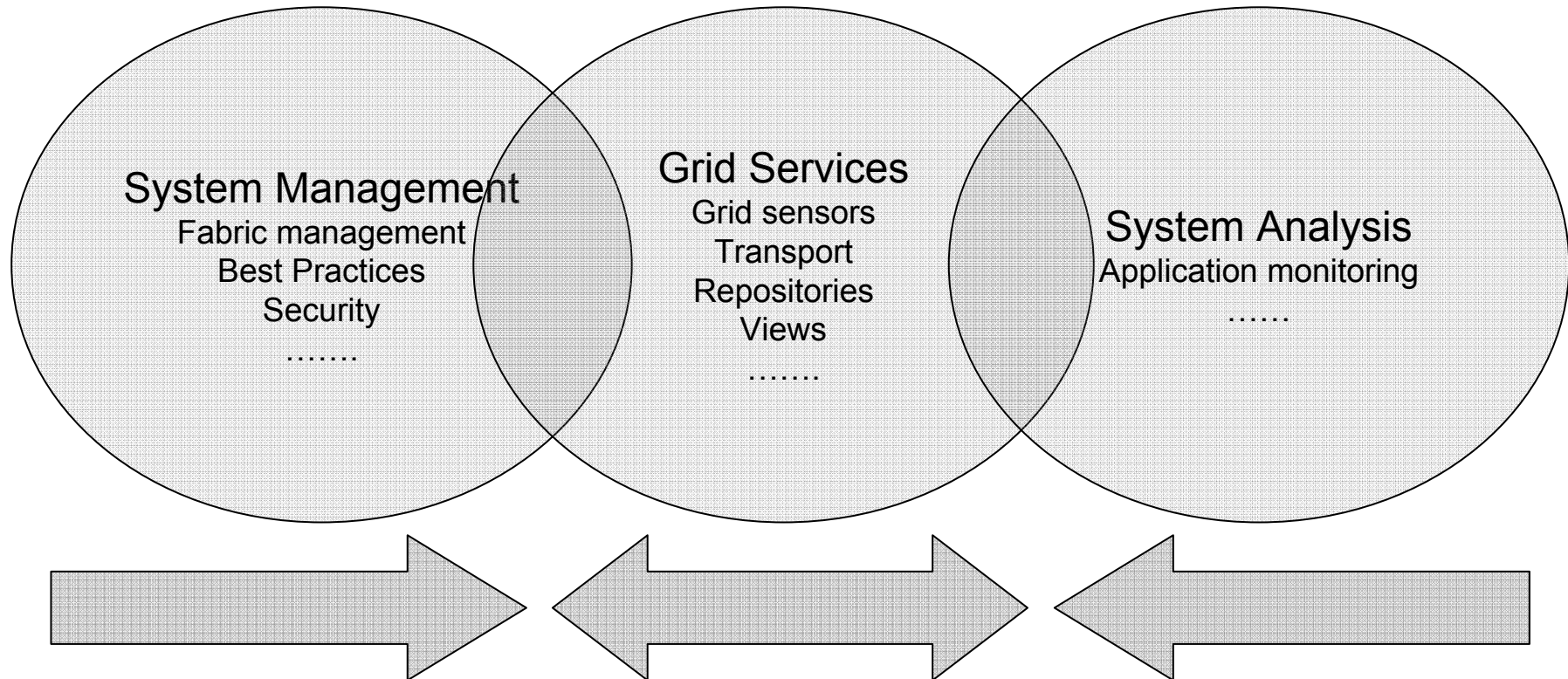


# Group Mandates



# WLCG Monitoring Working Groups

- 3 groups proposed by Ian Bird [LCG-MB, Oct 06](#).
  - *Goal to improve the reliability of the grid*



# **System Management Working Group**

**Alessandra Forti**

**WLCG workshop**

**CERN, 23 January 2007**

# Background

- Ian Bird at the Fall 2006 Hepix and at the WLCG Management board
  - <https://indico.fnal.gov/materialDisplay.py?contribId=34&sessionId=8&materialId=slides&confId=384>
  - <http://indico.cern.ch/materialDisplay.py?contribId=s0t14&sessionId=s0&materialId=slides&confId=a063271>
- 3 groups have been created to set up a comprehensive monitoring framework to improve the robustness of grid sites.
  - **System Management WG: system management and fabric monitoring tools and cookbook**

# Mandate: Intro

- One of the problems observed (by EGEE and LCG) in providing a reliable grid service is the reliability of the local fabric services of participating sites.
- The SMWG should bring together the existing expertise in different area of fabric management to build a common repository of tools and knowledge for the benefit of HEP system managers' community.
- The idea is not to present all possible tools nor to create new ones, but to recommend specific tools for specific problems according to the best practices already in use at sites.
- Although this group is proposed in order to help improve grid sites reliability, the results should be useful to any site running similar local services.
- Two areas should be improved by the group: tools and documentation.



# Mandate: Goals

- Improve overall level of grid site reliability, focussing on improving system management practices, sharing expertise, experience and tools
- Provide a repository
  - Management tools
  - Fabric monitoring sensors
  - HOWTOs
- Provide site manager input to requirements on grid monitoring and management tools
- Propose existing tools to the grid monitoring working group as solutions to general problems
- Produce a Grid Site Fabric Management cook-book
  - Recommend basic tools to cover essential practices, including security management
  - Discover what are common problems for sites and document how experienced sites solve them
  - Document collation of best practices for grid sites
- Point out holes in existing documentation sets
- Identify training needs
  - To be addressed in a workshop or by EGEE for example?

# Group Organisation

- Chairs:
  - Alessandra Forti (University of Manchester)
  - Michel Jouvin (LAL)
- The group organisation is a big question mark at the moment as it depends very much on the number of people and quality (ie dedicated time) of participation.
  - To be sustainable in the long term it has to be light weight and loosely bound, i.e people joining and leaving according to their availability. However this might not be feasible at the beginning when the initial structure has to be setup and a smaller core of dedicated people among the loosely bound are needed.

# Preliminary list of areas and tools

- **System Management Areas**
  - Filesystems: ext(2,3), XFS, NFS, AFS, dcache, DPM
  - Networking: Interfaces, IPs, Routers, Gateways, NAT
  - Databases: mysql, Oracle, ldap, gdbm
  - Processes: system, users monitoring
  - Servers: http, dhcp, dns, ldap, sendmail or other, sshd, (grid)ftp rfio
  - Batch systems: LSF, Torque, Maui, BQS, Sun Grid Engine, Condor
  - Security: login access pool accounts, certificates management and monitoring, non required services, ports list backups, monitoring(file systems, processes, networking), log files (grid services included)
- **Common Fabric Monitoring and Management Tools**
  - Monitoring: Ganglia, Nagios, Ntop, Home grown, SAM, GridICE, Lemon
  - Management: Cfengine, Npaci rocks, Kickstart, Quattor
  - Security: iptables, rootkit, tripwire, nmap, ndiff, tcpdump, syslog, yummit
  - Grid Configuration: Yaim , Quattor



# Grid Service Monitoring Working Group

Monitoring WG BOF, January  
2007

James Casey/Ian Neilson



# Grid Services Monitoring WG

- Mandate

- “....to help improve the reliability of the grid infrastructure....”
- “.... provide stakeholders with views of the infrastructure allowing them to understand the current and historical status of the service. ...”



# Grid Services Monitoring WG

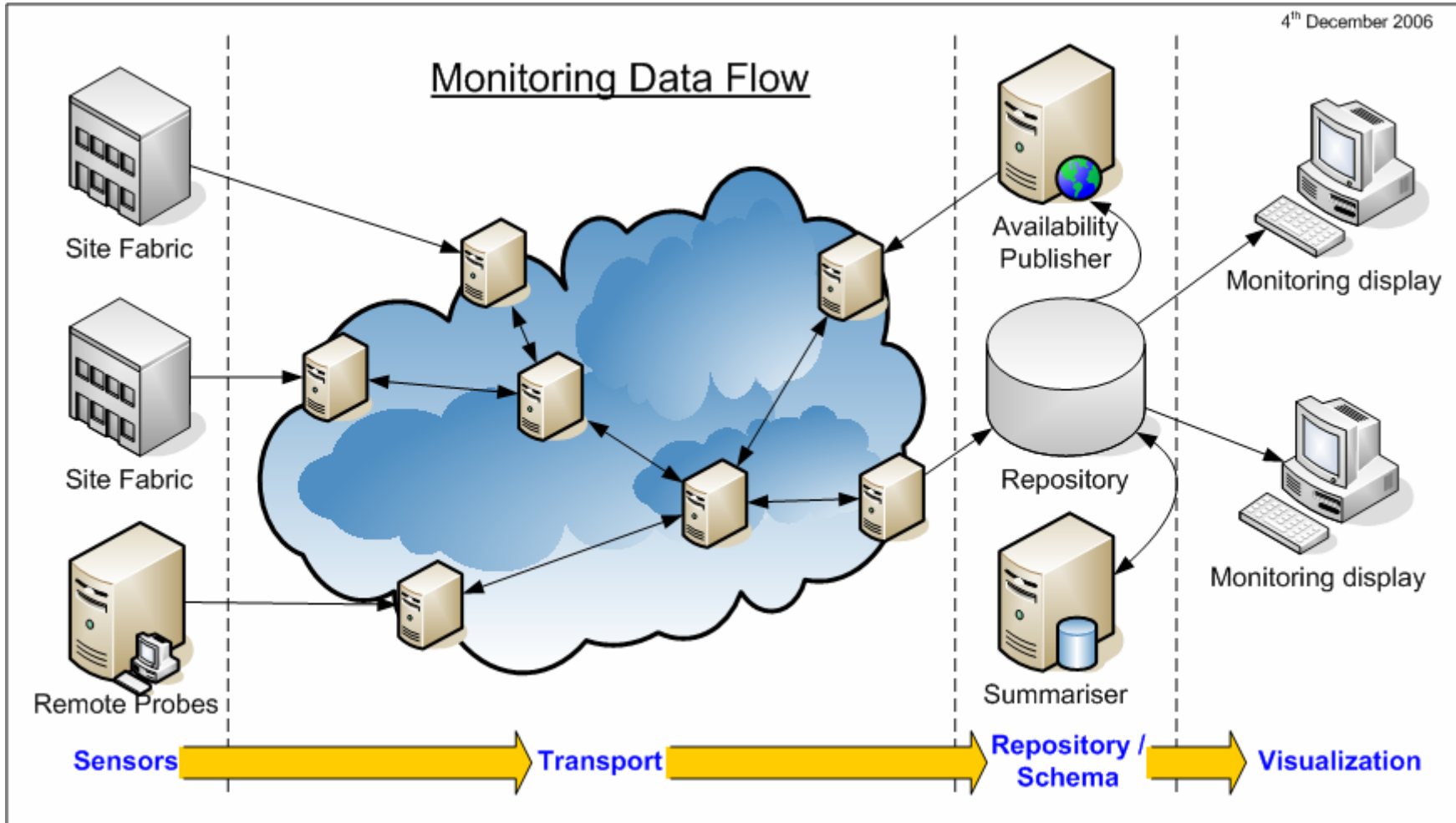
- Mandate

- to develop more monitoring tools
  - *unless a specific need is identified*
- to replace existing fabric management systems





# Monitoring Data Flow



Monitoring BOF, 23<sup>rd</sup>  
Jan 2007





# WG Structure

- 2 coordinators
- “core” team of ~10 across domains
- 4 domain sub-groups
  - Sensors
  - Transport
  - Repository
  - Views



# Immediate Tasks

- “What do you have and what is needed?”
  - questionnaire to site administrators (Dec 06)
- Per-service sensor definition
  - Plain english
  - Sensor ‘architecture’
- Characterise monitoring data traffic
  - → transport requirements
- Repository schema
  - Understand relationship between multiple DB’s
- Include security requirements

*Monitoring BOF, 23<sup>rd</sup>*

*Jan 2007*



# Timeline

- Dec 06
  - Background research
  - Establish core group
- Feb 07
  - Establish sub-groups
  - Agree interfaces and workplan
- April/May 07
  - Prototype instrumented services to local FM
  - Remote metrics to local FM
- end-Summer 07
  - Demonstrated improvement in reliability of grid

# Mandate: Interaction with GSWG

- Some of the areas covered by this group overlap with the Grid Services Monitoring Working Group ones particularly the local fabric monitoring area.
- The two groups are required to work in close contact and boundaries and division of responsibility should be discussed between the groups.
- The SMWG should act as a bridge between the system managers and the developers in the GSMWG giving feedback for what concerns monitoring tools and concerns



Enabling Grids for E-science

# Overview of System Analysis Working Group

*Julia Andreeva CERN,*

*WLCG Collaboration Workshop , Monitoring BOF session  
23 January 2007*





Goal



- As stated in the mandate the goal is to gain understanding of application failures in the grid environment and to provide an application view of the state of the infrastructure
  - *Application view in this context means the comprehensive picture of the experiment activities on the Grid , i.e. combining Grid-related and application specific information and allowing to detect and address problems of various nature*
  - *This work is the continuation of what had been started in the Experiment Dashboard project*
  - *The view of the experiments activities on LCG can be achieved by taking into account the progress done in the monitoring area by all involved parties:*
    - developers and providers of the Grid services*
    - developers and providers of the existing monitoring tools*
    - experiments themselves, in particular developers of work load management tools and data management systems*
- Summarize experience gained by the LHC experiments in achieving this goal and provide input to grid service monitoring and management



## In practical terms



- We are not planning to introduce a new monitoring system
- In close collaboration with the experiments development work aimed to provide the application view of the infrastructure had been started by the Experiment Dashboard project and dashboard development will continue following the output of the Working Group
- But the scope is wider. Overview what experiments had achieved in the area of application monitoring, define common patterns and identify common problems to address them in collaboration with two other monitoring groups



# Example of combined (Grid - application) monitoring



any user
any site
any ce
any submissiontool
any dataset
any application
any rb
analysis
any grid
<input type="checkbox"/> unk <input type="checkbox"/> pend <input type="checkbox"/> run <input type="checkbox"/> term
<input type="checkbox"/> done <input type="checkbox"/> canc <input type="checkbox"/> abort <input type="checkbox"/> g-unk
<input type="checkbox"/> succ <input type="checkbox"/> fail <input type="checkbox"/> a-unk
<input type="checkbox"/> donesuccess
2006-12-09 19:44:55
to
2006-12-10 19:44:55
sort by user



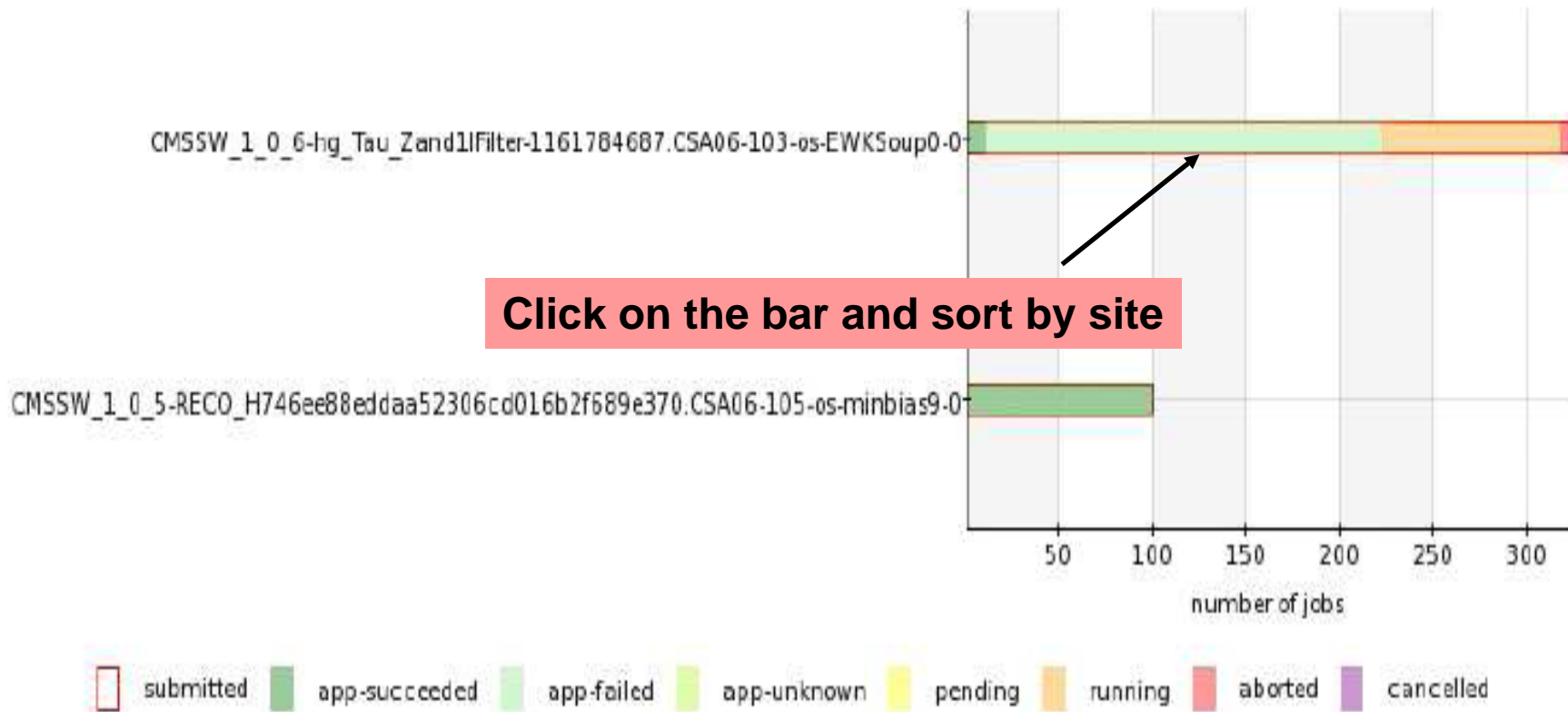




# Example of combined (Grid-application) monitoring



jobs per dataset

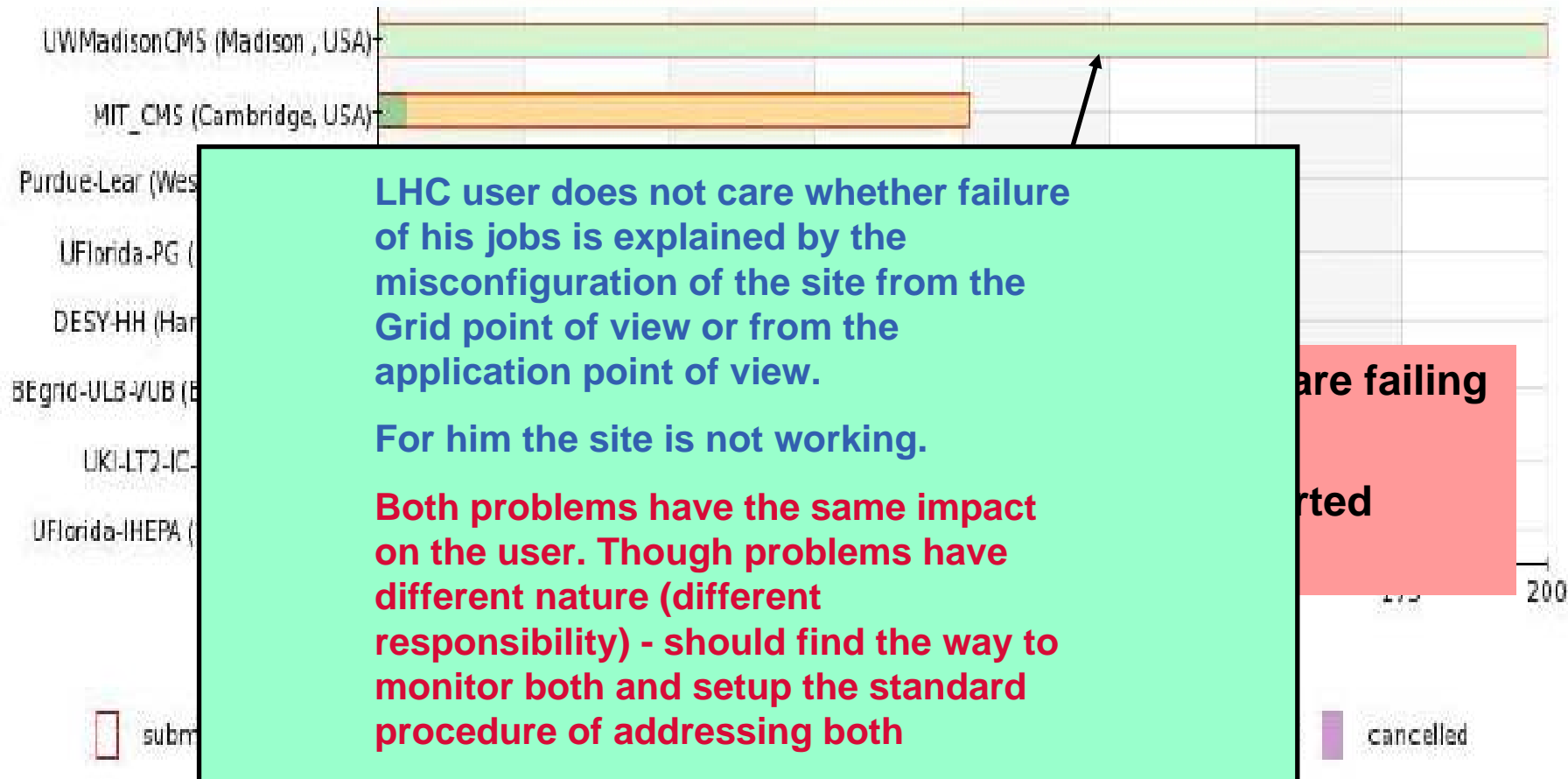




# Example of combined (Grid-application) monitoring



jobs per site



LHC user does not care whether failure of his jobs is explained by the misconfiguration of the site from the Grid point of view or from the application point of view.

For him the site is not working.

Both problems have the same impact on the user. Though problems have different nature (different responsibility) - should find the way to monitor both and setup the standard procedure of addressing both



# How to achieve



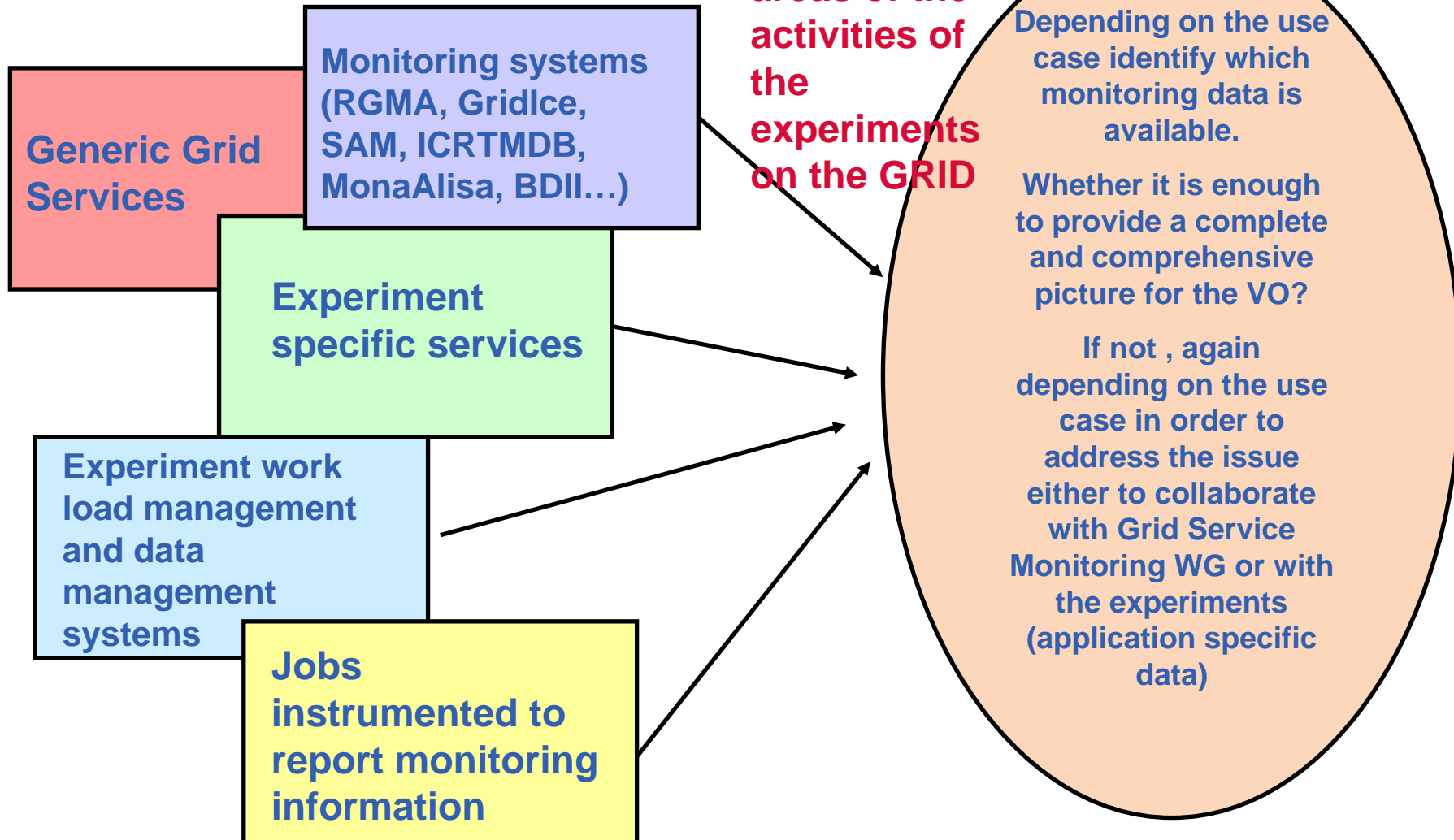
- Identify current experiment use cases related to main areas of activities - job processing, data management, DB replication.
- For every use case analyze present experience and requirements of four LHC experiments
- For every use case analyze the existing sources of monitoring data and understand whether they are sufficient, whether bits of information coming from various sources can be complementary to each other, how they can be correlated. Identify information holes (if any).
- Where possible identify common approach and implementation for the application monitoring (experiment dashboard, SAM tests with VO specific content)
- Provide input to the Grid Service Monitoring Working Group regarding identified problems or suggestions for the improvements of the grid service monitoring
- Based on information collected in the experiment dashboard identify VO-site related failures of the real user jobs. Coupled with SAM tests with VO-specific content. This is the area where System Analysis WG has to collaborate with two other monitoring groups. Come with the suggestion of the procedure how these problems should be followed up.



# Analysis of the information flow of the monitoring data



## Information sources





# Analysis of the current VOs experience and requirements



- Experiments have different requirements for the high level monitoring depending on the way their workload management systems and data management systems are organized :
  - *different level of centralization (example job submission)*
  - *different progress already done in the experiments regarding a given activity (example Phedex with very advanced monitoring system)*
  - *variety of the platforms used by the experiments*
  - *different technology for a similar task (example – DB replication)*
- Still a lot of common issues regarding job processing, data transfer, data access, usage of the distributed DBs.
- Nothing to enforce, but to identify where possible the ways to implement monitoring in a common way:
  - *via the experiment dashboard*
  - *using existing monitoring frameworks (SAM)*
  - *by defining common problems/requirements/suggestions to the middleware developers, providers of the monitoring tools, via Grid Service Monitoring working group (example – improving error reporting for the Grid related failures)*
- Share experience and ideas related to the application monitoring between LHC experiments
- Different activities on the Grid are very much dependent on each other (example- job processing efficiency is dependent on data distribution, data publishing, data access). Analyze these dependences in order to understand how VO monitoring data can be better presented/navigated.



# Core group membership



- Chaired by

**Julia Andreeva**

- Group is focused on the needs of the main WLCG customers – LHC experiments. LHC experiments should be the key players in the group. One representative per experiment, plus VO experiment experts will be invited to take part the meetings on the specific subject

**Dietrich Liko (ATLAS)**

**Latchezar Betev (ALICE)**

**Stefano Belforte (CMS)**

**To be confirmed (LHCb)**

- Experiment dashboard made a good start. Output of the work of the System Analysis WG will define directions for the further dashboard development.

**Benjamin Gaidioz**

**Pablo Saiz**

**Ricardo Brito Da Rocha**

- MonAlisa monitoring system is widely used by several LHC experiments for the application level monitoring

**Iosif Legrand**

- ROC at CERN started work to follow site problems for LHC VOs

**Diana Bosio**

- LCG Experiment Integration Support

**Roberto Santinelli**

Monitoring tools developers will be invited to take part in the meetings related to a specific topic





## Expected outcome of the work



- Further development/improvement of the Experiment Dashboard following the output of the working group
- Make sure that via experiment dashboard, SAM and experiment specific monitoring systems LHC experiments are provided with the monitoring framework where LHC VO user depending on his role and use case can find necessary monitoring data
- Suggest the procedure to address VO-related problems at the sites which should improve overall level of site reliability from the point of view of the LHC experiments
- Provide input for the Grid Service Monitoring Working Group for the issues/requirements related to Grid Service monitoring and collaborate with it in order to find the solution



# Grid Services Monitoring WG Site Survey

Results to 17 Jan 2007





# Questionnaire

- 1) What local fabric monitoring system do you use?:
  - a) GridICE/Lemon
  - b) Nagios
  - c) Other (please specify)
  - d) None.
- 2) Which Grid level sensors do you use?:
  - a) which services are monitored
  - b) what values/metrics are measured
- 3) Who provided the sensors?
- 4) Is your fabric monitoring part of any regional/off-site monitoring framework?
  - a) who are you linked with
  - b) generally, how is this implemented
- 5) When you learn that something is wrong with the services at your site, what is the most frequent way you are informed?
  - a) looking in the local fabric or Grid monitoring system
  - b) getting a trouble ticket
  - c) getting a mail/telephone call from VOs/users
  - d) other (please specify)..
- 6) Briefly describe what you see as your top 3 monitoring priorities to help improve your service reliability/availability



# Summary of Returns 1

- 34 responses analysed up to 17 Jan 2007
  - Not so easy to summarise sometimes so numbers don't always add up!
- Local monitoring frameworks in use
  - Sites using multiple frameworks
    - a) Nagios: 22
    - b) GridICE/Lemon: 10
    - c) Other: =majority as (a or b) + Ganglia: 13
    - d) None : 3
- Grid Services Monitored
  - 12 sites monitoring some Grid services
    - Most commonly CE+SE
  - Non-Grid default Nagios sensors in use



## Summary of Returns 2

- How problems get reported
  - Most common from local monitoring : 21
  - Support Ticket : 10
  - Looking at SAM/GSTAT : 4
  - Direct from User/VO : 3
- Sites reported being in regional infrastructures : 10
  - Not clear from the reports how these are implemented.



---

# Feedback and summary





# Feedback



- ~80 people attended the BOF, and there was good discussion
- We still don't know what proportion are running fabric monitoring systems
  - We assume site *\*will\** run a fabric monitoring system
- On the whole site admins want data back into their sites
  - But some sites are very happy with the central repository e.g. CIC Portal
  - Coordination with work by CIC portal developers is important



# Feedback



- Sites go to their client monitoring system first and see what the problems are from the experiment perspective
  - But then need the local information to do the troubleshooting and debugging
- Connection and drilling down between systems is important for debugging



# Feedback



- Experiments want to see the data in their experiment specific view e.g 'by cloud' for ATLAS
  - Dashboard is fine for them
- Still not sure on how much regional monitoring is used/needed
  - Some regions do it now, will more want to ?



## For more information



- Twiki for all groups:
  - <https://twiki.cern.ch/twiki/bin/view/LCG/LCGMonitoringWorkingGroups>
- Mailing lists:
  - [wlcg-system-management-wg@cern.ch](mailto:wlcg-system-management-wg@cern.ch)
  - [wlcg-monitoring-wg@cern.ch](mailto:wlcg-monitoring-wg@cern.ch)
  - [wlcg-system-analysis-wg@cern.ch](mailto:wlcg-system-analysis-wg@cern.ch)





EGEE





# Extra Slides

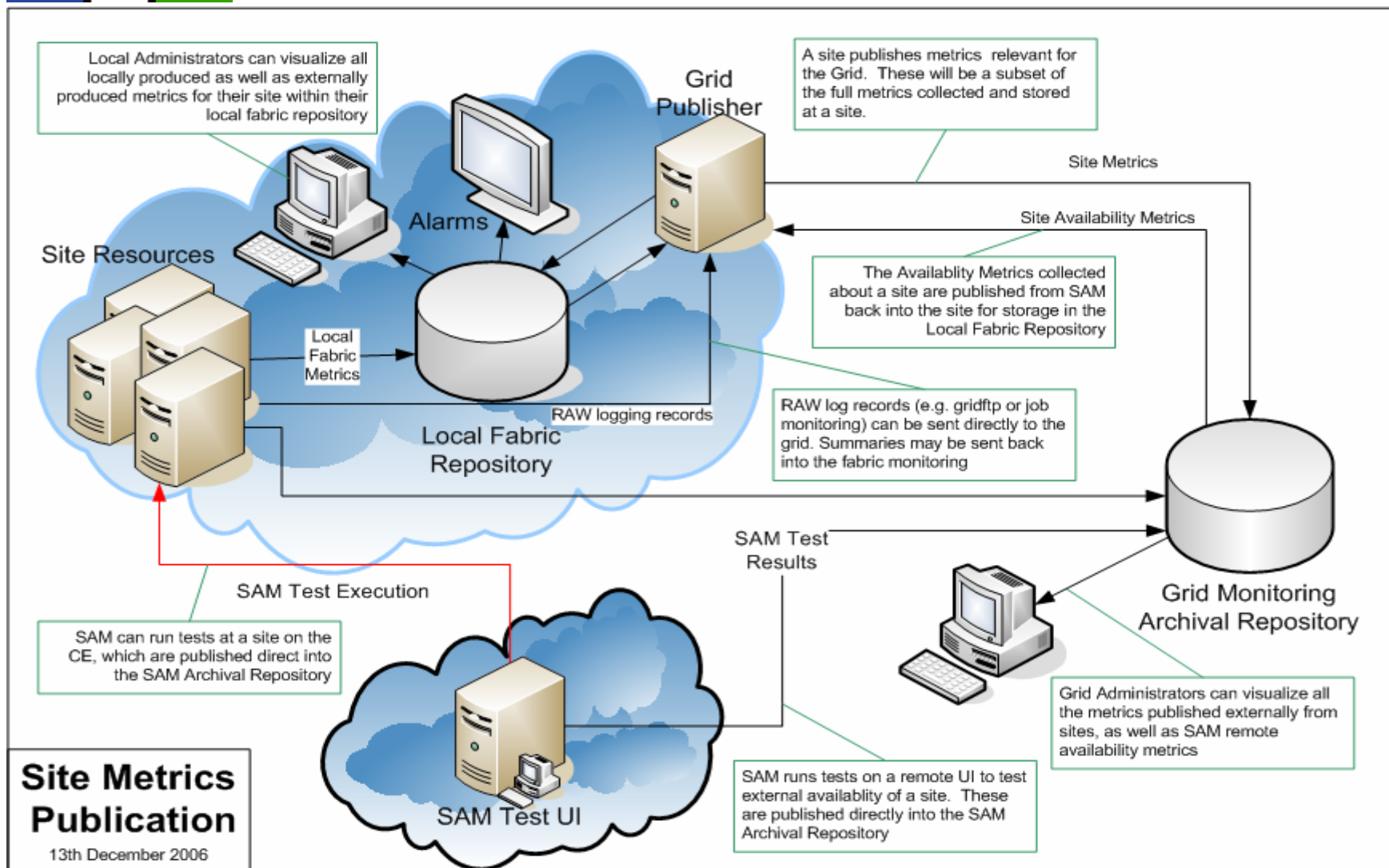


# Further Information

- Group mandate link:
  - <https://uimon.cern.ch/twiki/bin/view/LCG/SystemManagementWGMandate>
- Mailing list for the group:
  - [wlcg-system-management-wg@cern.ch](mailto:wlcg-system-management-wg@cern.ch)
- If you want to contribute contact:
  - [Alessandra.Forti@cern.ch](mailto:Alessandra.Forti@cern.ch)
    - It would be useful to know your areas of expertise.



# Site Metrics Publication





# Priorities

- Priorities
  - Quite difficult to summarise but keywords are....
    - single view - common interface - global view
    - unified tools - repository
    - more/deeper diagnostics
    - more flexible – alarm levels
    - improved/reliable/redundant SAM
    - hardware/network monitoring
  - Also non-monitoring replies
    - Working/debugged middleware
    - Reliable hardware
    - Experience/knowledge transfer



# Follow job failures at the site for a given VO



V  
=

SiteName (click on any site)	Successful jobs	Failed jobs	Percentage
<b>CERN-PROD</b>	<b>197</b>	<b>626</b>	
ce106.cern.ch:2119/jobmanager-logsif-grid_cms	30	109	21.58%
ce106.cern.ch:2119/jobmanager-logsif-grid_2nh_cms	12	21	36.36%
ce105.cern.ch:2119/jobmanager-logsif-grid_2nh_cms	16	20	44.44%
ce105.cern.ch:2119/jobmanager-logsif-grid_cms	10	118	7.81%

Jobs	# jobs	Successful?	Error message
<a href="#">See all the jobs...</a>	1	Ignored	Job successfully submitted to Globus
<a href="#">See all the jobs...</a>	56	No	Got a job held event reason: Globus error 101: the user proxy expired job is still running
<a href="#">See all the jobs...</a>	18	No	cannot retrieve a job held event reason: Globus error 131: the user proxy expired job is still running
<a href="#">See all the jobs...</a>	15	No	Got a job held event reason: Globus error 131: the user proxy expired job is still running
<a href="#">See all the jobs...</a>	13	No	Got a job held event reason: Globus error 131: the user proxy expired job is still running
<a href="#">See all the jobs...</a>	4	No	Got a job held event reason: Globus error 150: the job manager could not lock the state lock file
<a href="#">See all the jobs...</a>	2	No	Cannot download file from file

Submitted###Waiting###Ready###Done#Got



# Backup slides. Dashboard Architecture

