

GRATIA REVIEW

Quote of the Day

“Gratia is the lens through which we can see and understand OSG”

Chander S. Sehgal , **OSG** Project Manager

Outline

3

- Status
- Effort
- Architecture Overview
- Ongoing work
 - ▣ Packaging
 - ▣ Improvements & New development
 - ▣ Production Upgrade
 - ▣ Testing and Support
 - ▣ OSG/WLCG Integration Status
 - ▣ OSG/XSEDE Integration
- Stakeholders requests/future work
- Risk
- Summary

Status

4

- Multiple instances of Gratia Service are installed at
 - Fermilab (5 Fermi Collectors, 4 OSG Collectors)
 - Nebraska, UCSD, Purdue – site collectors
 - BNL, UofChicago – ITB collectors
 - GOC - RSV collector
- 15 different types of probes that collect batch jobs, storage allocation and usage, transfers, metrics and cloud records
 - More than 200 active probes are currently reporting to OSG collector
 - These probes are running on more than 100 clusters
- Email reports are sent daily, weekly and monthly to various VO admins and management.
- All parts of gratia software (collector & reporter, email reports and probes) are packaged as rpms and are included in OSG Software 3.x distribution.
- Multiple releases of software have been made since October 2011. They include new features, improvements as well as bug fixes.
 - Collector and reporter releases – 4. Upgraded at UNL and installed in UofChicago.
 - Probes releases - 8. All sites that have installed OSG 3.x are using probes from the release.
 - Email reports releases - 1. Is planned to install at UNL
- Gratia records are one of the data sources for OSG Display, gartiawebtools UI, XSEDE, and EGI Accounting Service.

Efforts

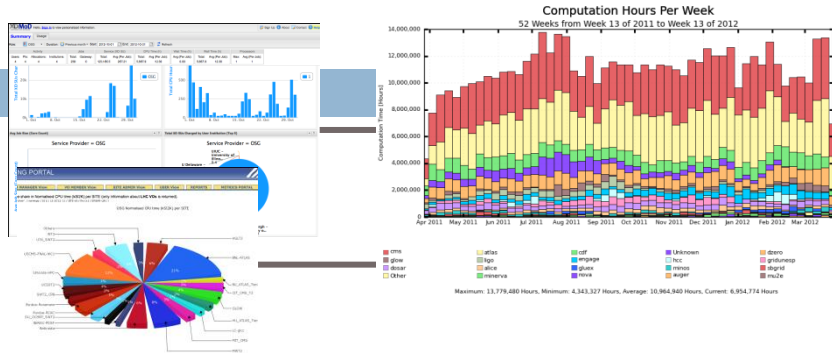
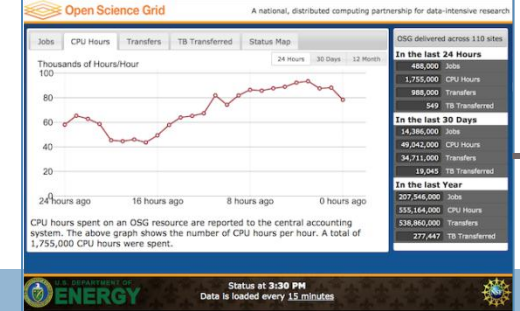
Name	Affiliation	Expertise	Efforts	Funded by
Tanya Levshina	Gratia development team (project leader) /Fermilab	collector, probes, packaging, documentation, support	0.5 FTE 0.1 FTE	Fermilab ExTENCI
Hyunwoo Kim	Gratia development team /Fermilab	email reporting, probes, packaging, testing	0.2	Fermilab (Started in March 2012)
John Weigand	Gratia development team (consultant) /Fermilab	mySQL, collector, installation scripts, support OSG/WLCG Integration	0.25 ?	CMS (Gratia funding for 2013 is unknown) CMS
Parag Mhashilkar	DOCS/Fermilab	Cloud/openNebula API	0.1	Fermilab
Philippe Canal	ADSS/Fermilab	Expert help: main architect and developer	0.1	Fermilab (no funding for 2013)
Gratia Operation Team	GCC/Fermilab	Running gratia services	0.20	Fermilab & OSG
Suchadra Thapa	OSG Software Team	sge, lsf-pbs, rsv probes	?	OSG
Neha Sharma	OSG Software Team	transfer and storage probes	0.05	OSG
Brian Bockelman	HCC/OSG Metrics	Condor probe, common probe libraries, support	0.1	HCC & OSG
Derek Weitzel	HCC/OSG Campus Grid	Condor probe for campus grid	0.03	HCC & OSG
Ashu Guru	HCC/OSG Metrics	Gratia web UI	0.15	HCC & OSG



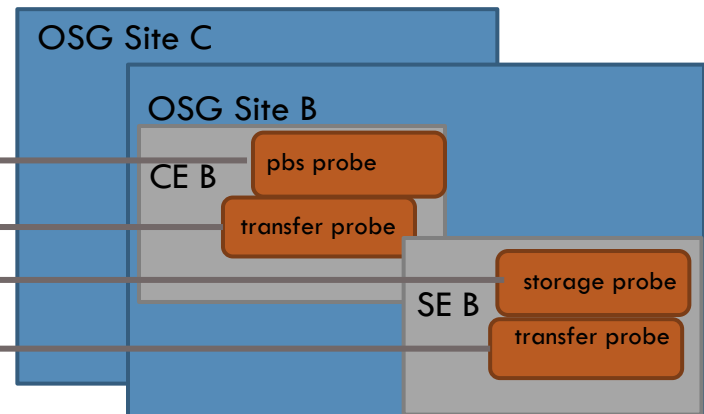
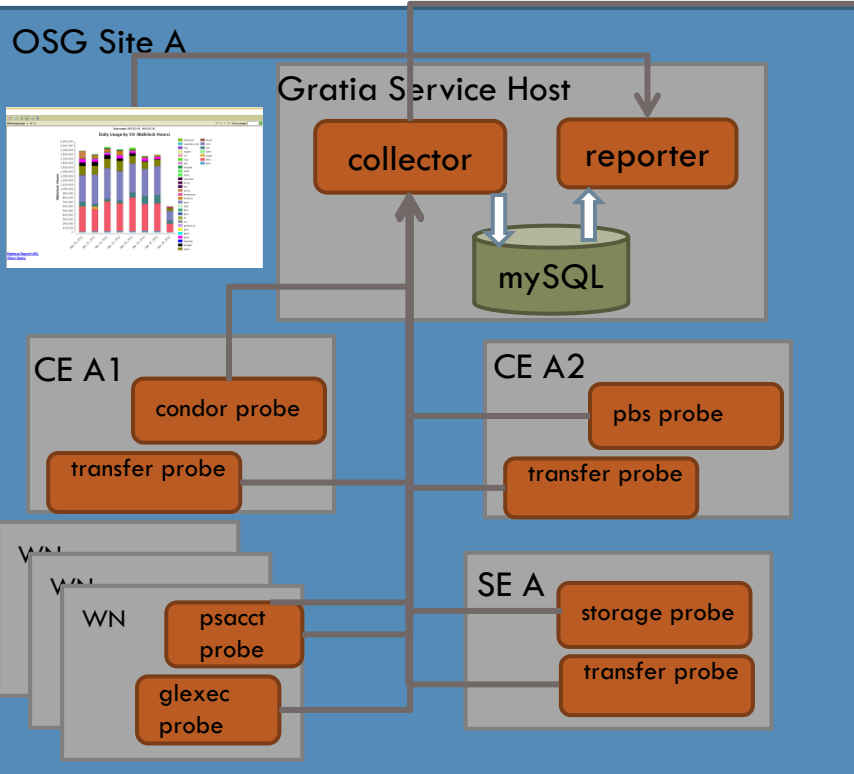
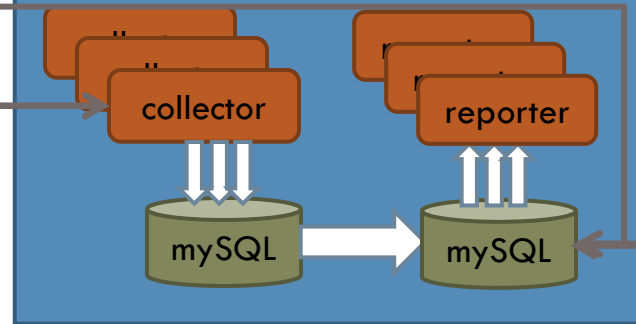
Overview

Gratia Architecture

7



OSG Gratia Collectors/Reporters



Gratia Review

2012-12-03

Gratia Overview & Statistics

8

- Gratia Service consists of several subsystems:
 - Collector
 - Reporter (WEB UI for user and admins)
 - Database
 - Email reports generating scripts
- Information is generated by various probes and sent to Gratia collectors via Gratia API. Probes collect information about:
 - Batch and glide-in jobs (condor, lsf, pbs, sge)
 - Linux process accounting (deployed at Fermilab)
 - Various Metrics (RSV probes)
 - File transfers
 - Storage Usage and Allocation
 - Cloud accounting (deployed at Fermilab)
- Gratia supports multiple collectors. It permits hierarchical forwarding & filtering between collectors. This feature is used at UNL and UCSD and allows these sites to filter jobs submitted by local users from OSG users.

Gratia Statistics

9

- The maximum rate for processing observed during the last year is 200,000 records per hour in production.
- Number of records processed for OSG and OSG-Transfer collectors since 2005:
 - ▣ gratia jobs in ~ 870 million
 - ▣ gratia transfers ~2 billion
- Number of Sites reporting batch jobs ~133 (Site could have several entry points)
 - ▣ Condor probes ~111
 - ▣ PBS probes ~37
 - ▣ LSF probes ~2
 - ▣ SGE probes ~5
- Number of SEs reporting gridftp transfers ~49 (SE could have multiple gridftp doors)
 - ▣ gridftp-transfer probes ~112
 - ▣ dcache-transfer probes ~4
 - ▣ Very few T2 and T3 sites setup and enabled transfer probe.

Implementation Summary

10

Software implementation details:

- ▣ Tomcat
- ▣ JMS Queue
- ▣ Hibernate
- ▣ Database (MySQL)
- ▣ Java (43, 358 lines of code), JSP (3,560), python (36,832), perl(1,858), shell (6,420)
- ▣ Business Intelligence and Reporting Tools (BIRT). An open source Eclipse-based reporting system.
- ▣ Build tools: Makefile
- ▣ Packaging: rpm
- ▣ Repository: SVN (<http://sourceforge.net/projects/gratia/>)
- ▣ Ticket tracking: JIRA (<https://jira.opensciencegrid.org/browse/GRATIA>)



Ongoing Work

Code Support Responsibilities

12

#	Name	Purpose	Maintainer	Stakeholders & Customer
1	gratia-probe rpms	Probes installation package	Gratia Dev Team	OSG & Fermi
2	gratia-service rpm	collector & reporter package	Gratia Dev Team	OSG & Fermi
3	gratia-email-reporting	gratia email reports package	Gratia Dev Team	OSG & Fermi
4	gratia core code	collector, reporter, email reports, installation and upgrade scripts, configuration template	Gratia Dev Team	OSG & Fermi
5	probe common libs	Probe common libraries	Gratia Dev Team	OSG & Fermi
6	osg-configure (gratia.ini)	configuration and scripts for gratia probes to be installed/enabled automatically on CE	Software Team	OSG

Probes List & Support Responsibilities

13

#	Name	Purpose	Maintainer	Stakeholders & Customer
1	condor	monitors job submitted via condor	Software team	OSG & Fermi
2	pbs-lsf	monitors job submitted via pbs/lsf	Software team	OSG & Fermi
3	sge	monitors job submitted via sge	Software team	OSG & Fermi
4	gridftp-transfer	monitors transfers by parsing gridftp logs	Software team	OSG & Fermi
5	dcache-transfer	monitors transfer submitted to dcache (scan billing database)	Software team	OSG & Fermi
6	xrootd-transfer	gathers information about xrootd transfer from xrootd monitor	Software team	OSG
7	hadoop-storage	gathers storage area and raw disk utilization	Software team	OSG
8	xrootd-storage	gathers storage area and raw disk utilization	Software team	OSG
9	dcache-storage	monitors storage area and pools utilization (pull info from Information provider)	Software team	OSG & Fermi
10	psacct	collects Linux accounting	Gratia team	Fermi
11	glexec	monitors glexec logs	Gratia team	Fermi
12	metric	collects RSV reports	Software team	OSG
13	bdii-status	collects bdii info	Software team	OSG
14	onevm	collects ONEVM usage information	Gratia team	Fermi

Software Packaging

14

- RPM Packaging
 - ▣ Started in October, 2011
 - ▣ Followed OSG Software guidelines
 - ▣ Currently all software is packaged as rpms (probes, collector & reporter, email reports).
 - ▣ Most time has been spent on “gratia-service” packaging – need to figure out right dependencies, clean up libraries locations, modified various hardcoded locations, do a lot of modifications for sl6/tomcat6 in installation scripts. We still need to upgrade to the newest libraries (e.g hibernate).
 - ▣ Used the OSG Software framework (osg-build, koji) to build and release software, provided documentation.
 - ▣ Worked with the OSG Software team to include gratia probes installation and configuration as part of CE package.
 - ▣ Tested on FermiCloud VMs – **crucial help to do the testing of various releases and OS!**

This work is mostly done by the members of Gratia Dev Team.

Improvements

15

- Most improvements have been done for the probes:
 - ▣ Restructuring of common libraries (done by Bockelman)
 - ▣ Re-implementation of condor probe (done by Bockelman)
 - ▣ Implementation of common start/stop service script for all probes (done by Gratia Dev Team)
 - ▣ Modification of gridftp probe to handle ipv6 entries, multiple fixes for netlogger (done by N. Sharma/Software Team)
 - ▣ Modification of pbs-lsf probe to handle correctly new PBS log format (done by the Gratia Dev Team)
 - ▣ Modification of glexec probe to handle old and new glexec log format (done by the Gratia Dev Team)
- Separation of service properties file to service-authorization and service-configuration files.(requested by HCC, done by the Gratia Dev Team)
- Minor code cleanup and bug fixes have been done for email reporting by the Gratia Dev Team. The work has been requested by HCC.

New Features (I)

16

- Several modifications were made to gratia probe common libs to handle Campus Accounting. Probes can map a local user properly if the relevant information is specified in the probe configuration. This configuration is now used on OSG-XSEDE, SAGA, Virginia Tech, Nebraska submission nodes. (done by D.Weitzel)
- Project Name addition to summary of Job Usage and transfer records requested by XSEDE/OSG. The implementation of this feature required the following steps:
 - Database schema change
 - Addition to Administration WEB UI to allow administrators to modify reported Project Name (similar to VO Name correction)
 - Addition of Project Name reporting to common lib probe
 - Modification of upgrade script
 - UNL has been upgraded to the newest gratia-service rpms
 - We have installed a new gratia service on the FermiCloud VM and started forwarding relevant probes from the OSG gratia collector to this collector.
 - A. Guru has started to work on gratiaweb tool modification to include queries by ProjectName by using data collected in this collector

Work has been done by the Gratia Dev Team

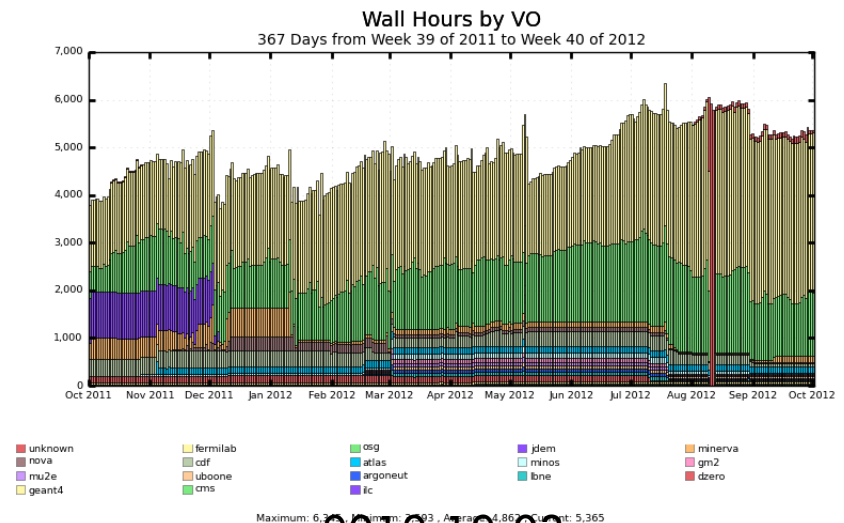
New Features (II) – Cloud Probe

17

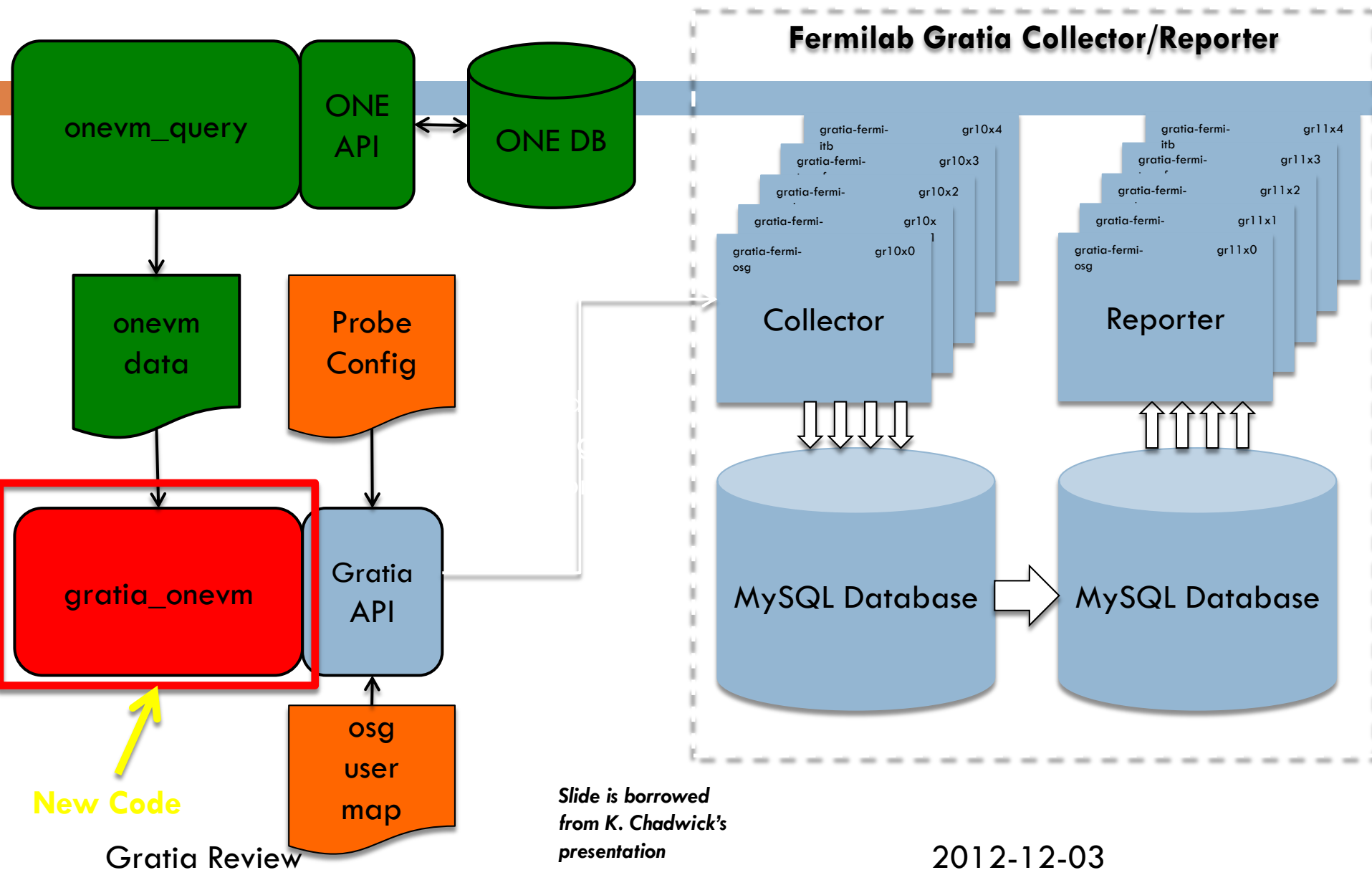
Design and implementation of gratia cloud accounting probe requested by Fermi Cloud project.

- All information about VMs owner, physical machine, states, start and stop times, memory and cpu usage could be acquired from interfaces provided by Open Nebula.
- Gratia Job Usage Record structure is currently sufficient to represent VM accounting information.
- No changes have been required to the core Gratia Service at least during the initial phase to store the records.
- Gratia webtool queries have been modified to display relevant information.

Work has been done T.Levshina, P. Mhashilkar,
S. Ma (summer student)



Cloud Probe Architecture



OSG Gratia Collectors Upgrade at Fermi

19

- FermiGrid Services host multiple production OSG Gratia instances. The latest major upgrade was done in January 2012. The preparation for upgrade and testing took a lot of time (about 6 months) due to various reasons (see http://cd-docdb.fnal.gov/cgi-bin/RetrieveFile?docid=4599;filename=gratia_rca_history%203.pdf;version=1).
- The current installation of the Gratia in production is highly customized and not very well documented. It will be challenging to fully understand it and replace it with the standard rpm distribution.
- We are working with Fermi Gratia operation team to come up with a plan for gratia upgrade. The upgrade will include
 - ▣ Installation of gratia software from rpm
 - ▣ Database schema upgrade
- So far we were able to demonstrate that we can restore database from backup (7 hours for OSG Gratia db), install gratia from rpm and do upgrade of the database (~20 minutes). The next step will be the upgrade of the test system and roll back.

Testing

20

- The previous testing infrastructure doesn't work with restructured code. We would like to use OSG Software Team framework to build software and Fermi Cloud VMs to do the testing. A lot of work needs to be done in order to automate that process.
- The Gratia Dev Team is doing all the testing of newly released gratia-service, gratia-email-reporting rpms and share responsibilities with the OSG Software Team to test probes. It would be ideal if these tests could be automated too.
- sha1/sha2 testing is done. The test was successful.
- ipv4/ipv6 testing has yet to be done. We have found and fixed some bugs related to ipv6 addresses in gridftp-transfer probe.

Support

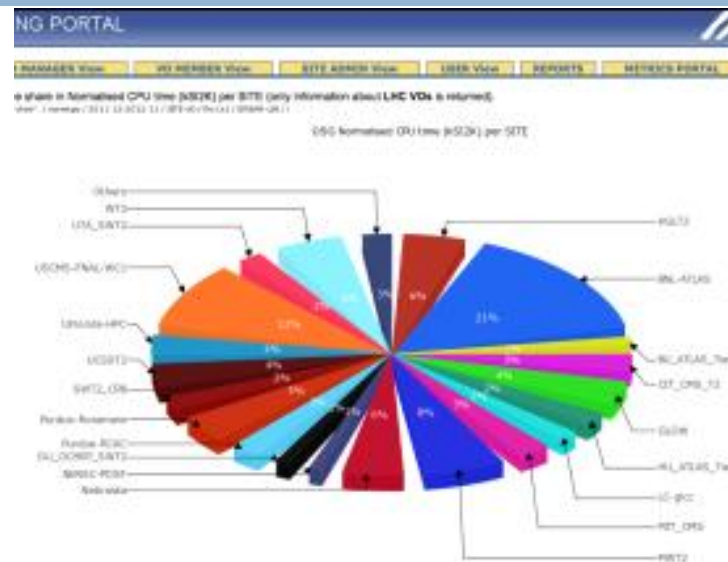
21

- Support responsibilities are shared between Fermi Gratia Operation, Software and Gratia Dev Teams
- The tickets are created at
 - ▣ Jira
 - ▣ GOC
- Most tickets are about:
 - ▣ Installation problems (decreased drastically after several releases)
 - ▣ Probe misconfiguration (we have started troubleshooting doc on twiki)
 - ▣ Modification of batch logs structure (new version of pbs, sge)
 - ▣ Ipv6 addresses parsing
 - ▣ Various questions related to configuration, backfilling the data, etc

OSG/WLCG Accounting Integration

22

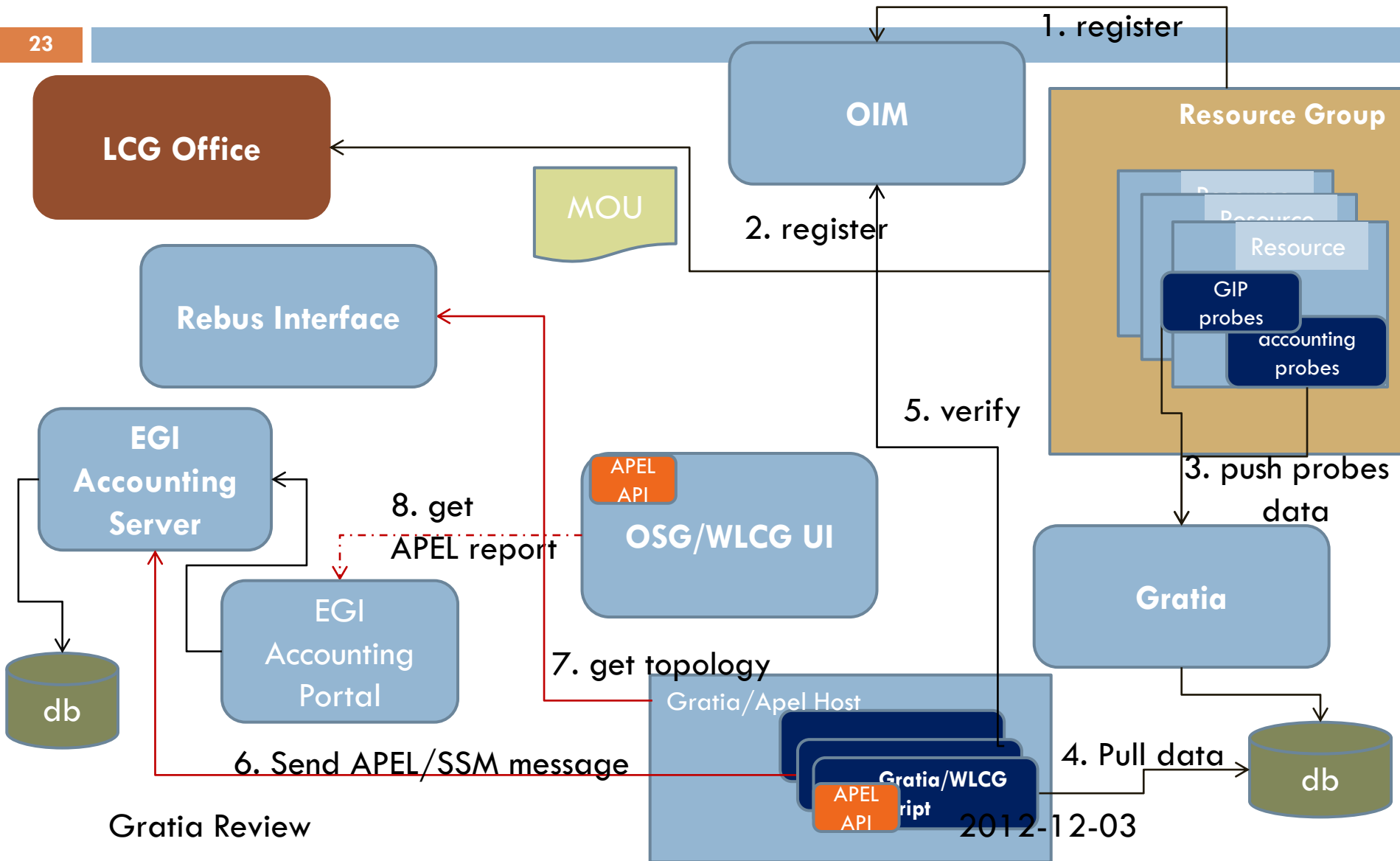
- Accounting records for US-LHC Tier1 and Tier2 facilities collected by Gratia are forwarded to the EGI accounting system (APEL), in accordance with signed MOU agreements.
- A service outside the OSG software infrastructure runs at Fermilab, parses and analyzes accounting records for a resource, scales wall time and cpu usage, and then forwards this data to the APEL Server at the EGI GOC.
- The work is currently funded by CMS and executed by J. Weigand.
- Currently we are reporting the data from 33 sites (ATLAS -18, CMS – 13, ALICE -2)
- The daily cron job checks the presence of the data for each site and notifies us about missing data. If data is missing a corresponding ticket is created at GOC.



This OSG data can be viewed in the EGI Accounting Portal – OSG view

Gratia/APEL & SSM Workflow

23



Gratia/APEL & SSM Work Status

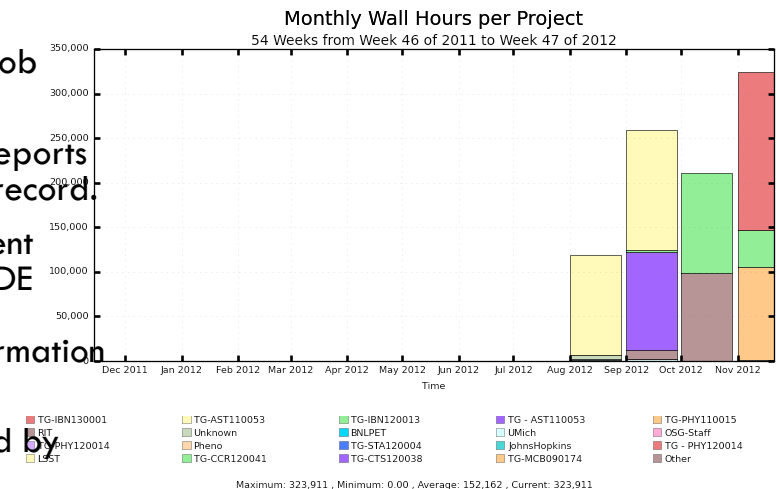
24

- The new interface (SSM) to APEL has been introduced during the summer of 2011.
- Development has started in October, 2011.
- We have provided feedback and participated in early testing. In production since April, 2012
- Work on rpm packaging has started in September, 2012
- We are coordinating efforts to move this service to GOC.

OSG as an XSEDE Service Provider

25

- OSG is an XSEDE Service Provider since April, 2012. OSG provides a Virtual Cluster that shields XSEDE users from the necessity to be aware of the OSG infrastructure. Jobs submitted into this Virtual Cluster will be executed on machines at several remote physical clusters.
- In order to successfully submit a job, a user should specify, among other attributes, a valid Project name in a condor job description file.
- The condor gratia probe running on the submission node reports usage records to Gratia. It includes Project name into the record.
- XSEDE tracks resource usage in its own account management system (AMIE). The script, that is running on the OSG-XSEDE submits information from Gratia collector, summarizes this information by grouping it by Project, and pushes this information to XSEDE accounting system using AMIE API.
- We have implemented the summarization of Gratia record by Project Name.
- The separate collector is installed on FermiCloud VM that collects and summarized these records. This will be dropped as soon as production Gratia collector is upgraded.
- The work is underway to modify gratiaweb queries in order to display this information (A. Guru)





Future Work/Stakeholder Requests

OSG/WLCG Integration

27

- Plans for FY13:
 - ▣ Coordinate transition of Gratia-APEL Service to GOC
 - ▣ Maintain communication with APEL team
 - ▣ Modify code to use SSM2 when it is requested
 - ▣ Change the message schema according new EGI requirements
- Future work:
 - ▣ Discuss/coordinate with US LHC VOs modification of storage probes according to new OGF SAR requirements, implement changes. It is on my list but I don't know if OSG HEP experiments want to report storage to WLCG. Potentially it could mean a lot of work.
 - ▣ Many international VOs use resources across the EGI and OSG infrastructure. We have a request from some of them (ILC, D0, SBGrid/WeNMR) to investigate if we can share the accounting information between OSG and EGI.
 - ▣ Participation in EGI accounting Task Force activities (meetings ~4 times a year)

Integration with XSEDE

28

□ FY13

- ▣ Continue to coordinate/test modification to gratiaweb.
- ▣ Provide instructions/help for xsede-portal installations (e.g similar to SAGA) with gratia condor probe and configuration, so Project could be reported correctly.
- ▣ Upgrade OSG production collector to the newest software.

□ Future work

- ▣ OSG-XSEDE proposal “Usage Analysis and Accounting Service Extensions” is currently in the works. It may require our group involvement if it is approved.

New Requests from UNL

29

- Coordinate design & implementation of memory accounting - development will be done mostly at UNL (Jira/Gratia 66, 67).
- Coordinate design & implementation of network - development will be done mostly at UNL; postponed for now (Jira/ Gratia 31).
- Incorporate new probes into distribution (e.g slurm-meter), provide appropriate testing.
- Reduce usage of root privileges in Gratia probes (Jira/Gratia 29).

More Requests

30

- Discuss/understand evolving cloud accounting records standards, report to EGI cloud accounting information using APEL/SSM (requested by K. Chadwick)
- Better understand/improve gratia accounting collection from campus submission nodes (requested by C. Sehgal/OSG). We are not getting any requests from Campus Grid Team so we are not aware if there are any issues.

Production Upgrade/Testing

31

□ Fermilab Upgrade

- ▣ Understand Fermilab Gratia installation and customized scripts used only by operation team.
- ▣ Prepare detailed upgrade plan from custom installation to rpm packages and database schema upgrade
- ▣ Test the upgrade, provide appropriate documentation for FermiGrid
- ▣ We hope that with this approach the next upgrade will be less painful!

□ Test infrastructure

- ▣ Coordinate testing of the probes with OSG Software
- ▣ Understand how to automate periodic & comprehensive testing

Risk

32

- The outcome of this review is very important to us. The failure to define stakeholders , validate requirements, and set priorities could lead to waste of efforts and prevent us from addressing the real problems.
- Efforts:
 - ▣ Core team is small and the efforts are limited
 - ▣ If J. Weigand's funding for Gratia core development is not approved, we will loose mySQL expert.
- OSG/WLCG integration software should be transition to GOC. We need GOC cooperation to do it successfully.
- We have to rely on collaboration with the OSG Software team to maintain probes and provide support.
- Failure to establish collaboration with the OSG Campus Grid team could cause underreporting campus usage of OSG resources.
- The failure to prepare for inclusion of public/commercial cloud resources accessible to OSG could impact the completeness of accounting data we are gathering.

Summary

33

- We are adapting the infrastructure to meet the new technological challenges of campus infrastructure, cloud computing, and interoperability.
- New probes development are in progress.
- Modifications of the old probes are ongoing.
- We have to develop comprehensive production upgrade procedure.
- Integration with other Grids is important (EGI, XSEDE). Good communication is crucial for integration success!
- New requests keep coming: we need to understand how to setup priorities and where we can look for additional efforts.