



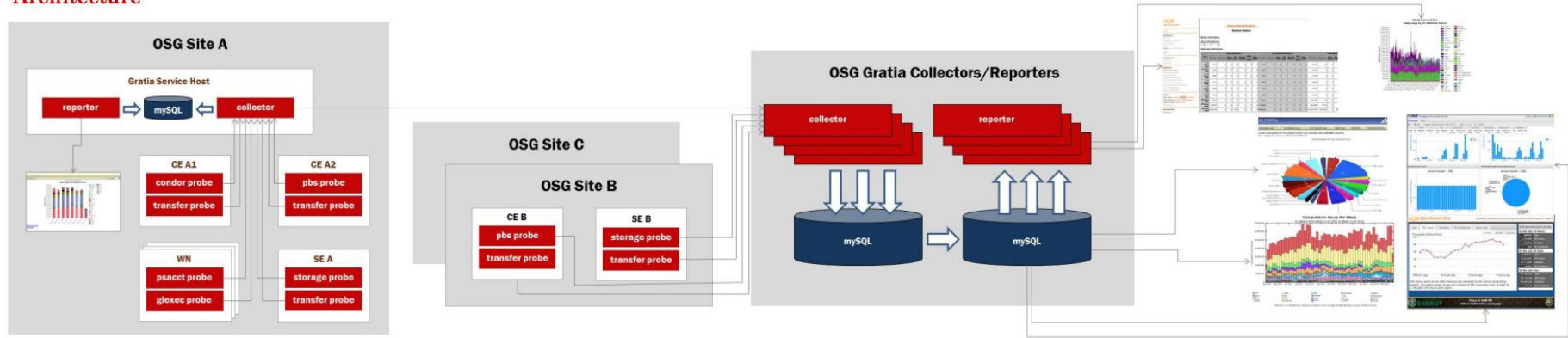
# GRACC

Brian Bockelman, Tanya Levshina, Kevin Retzke, Derek Weitzel  
August 4nd, 2016

# **Part I - Gratia, a beloved relic**

# Gratia Architecture

## Architecture



# Who is using Gratia?

- DOE Annual Report
- WLCG report
- XSEDE report
- OSG ET
- Site admins, VO admins, Experiments PI

# Gratia Overview

Gratia Service consists of several subsystems:

- Collector
- Reporter (WEB UI for admins)
- Database (MySQL)
- 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, slurm)
  - Various Metrics (RSV probes)
  - File transfers
  - Storage Usage and Allocation
  - Cloud accounting (OneNebula and AWS deployed at Fermilab)
- GratiaWeb and OSG Display

Gratia supports multiple collectors. It permits hierarchical forwarding & filtering between collectors. This feature is used at UNL and other sites and allows these sites to filter jobs submitted by local users from OSG users.

# Implementation Details and Stats

- Implementation:
  - Tomcat 6, JMS Queue, Hibernate 4, Java, JSP, python, perl, shell
  - Database (MySQL) 5.6
- Stats
  - OSG collectors since 2005:
    - OSG collector (Started in 2004): 8.1M records in Summary (jobs and pilot jobs); rate ~1M/day; database size 1.1 TB
    - OSG Transfer collector (Started at 2008/04): 45.5M records in Summary (all transfers); database size 1.0 TB
    - 122 sites are reporting 144 batch jobs: 114 Condor; 21 PBS; 4 LSF, 5 SGE; X Slurm
  - 46 campus Grid probes (probe reports actual user jobs):
  - Number of SEs reporting gridftp transfers: 66

# Issues

- Upgrade to hibernate 4 is causing problems with bulk deletion.
- Upgrade to MySql 5.6 has demonstrated “bad” design choices (missing index on auxiliary tables)
- Started to collect Storage and Tape data for AAF project, AWS VM with charges and other new info for HEP Cloud but unable to summarize these because there is no Summary Table for these type of records.
- Any changes in schema - major nightmare
- **You cannot expect that software designed in 2004 will be really adequate in 2016!**

# Part II - Next Generation Accounting



# Generic Requirements

- All historical summary data (job records and transfers) needs to be preserved
- Data should be archived.
- It should be possible to extract the historical data in suitable format in order to upload to the future accounting service.
- Gratia probes should not be changed drastically and new service should be able to deal with older version of probes.
- Should be able to use WebUI to extract the accounting information.

# OSG Requirements

- Daily summary of job wall duration, cpu usage per site, VO, Project, DN (with role), user, exit code for Batch and Payload resources
- Daily transfer summary (size and wall duration) per storage site, vo, user, direction, exit code.
- Ability to aggregate at different levels (site versus cluster versus CE) and rename site aggregation.
- Preserve OSG Display, WLCG and XSEDE reporting.
- Provide WEB UI similar to GratiaWeb.

# Other Institutions Requirements

- Fermilab
  - Ability to add new information:
    - Active Archival Facility project Project is using Gratia to report charges to customers that based on several quantities: tape storage, tape hours used, transfer size to dCache and enstore.
    - HEP Cloud Facility Project is extending the current Fermilab Computing Facility to transparently run on a variety of resources including commercial clouds. It is using Gratia to get historical information about instantiated VMs (wall duration, cpu usage, charges) per vo, instance type, availability zone.
    - FIFE Project is using gratia accounting to facilitate experiments preparation for SPPM meetings, get historical information about users efficiency, success rate, data transfer and dCache pool usage.
- Nebraska
  - All internal HCC usage accounting is done using Gratia. Reports to stakeholders and users often include gratia produced graphs. In order to do so the following information is needed:

# Evaluation of Existing Open Source Grid Accounting Service

- Evaluated:

- XDMoD
- EGI Accounting (APEL)

See details in [google doc](#).

- Researched for the general trends for new development for monitoring and accounting:

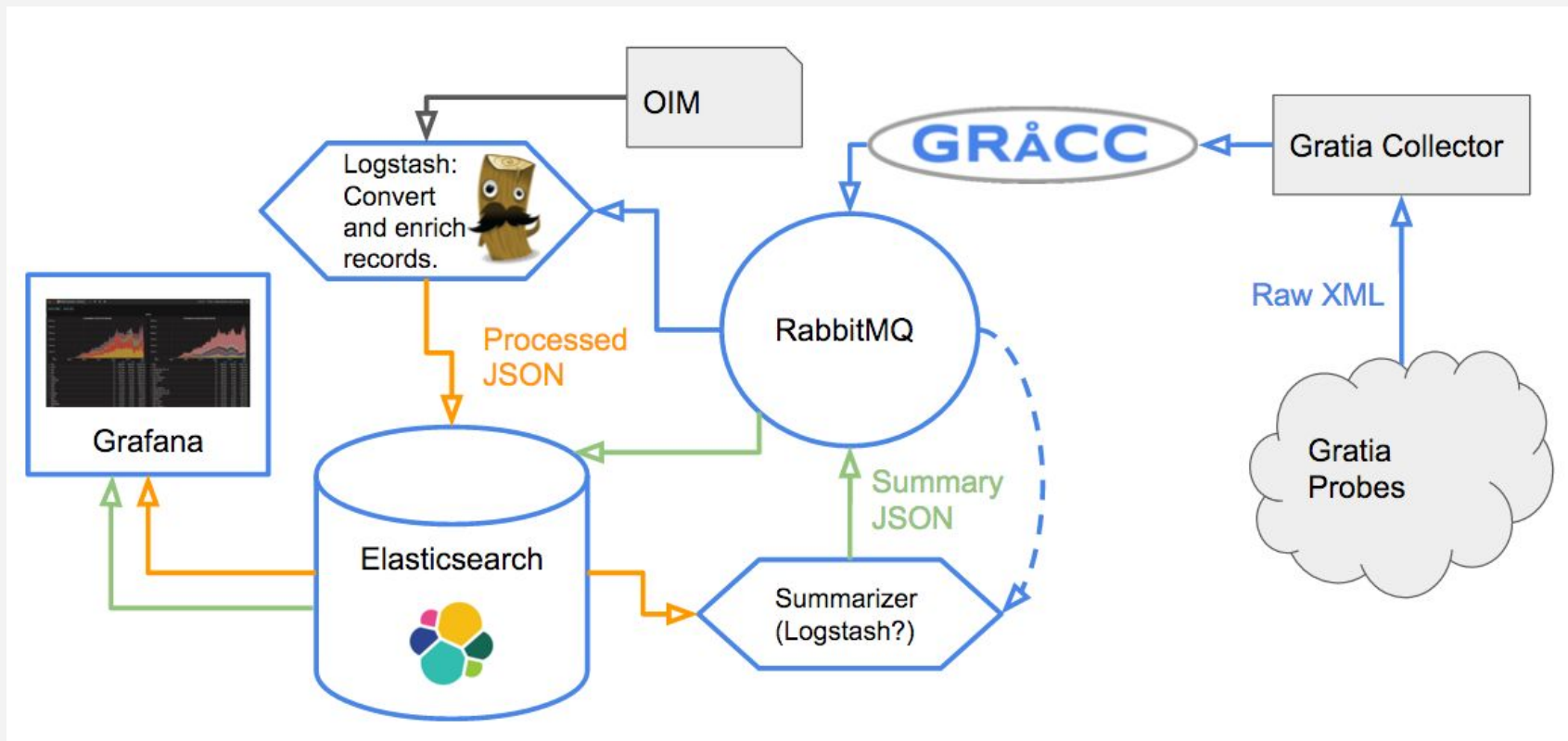
- UChicago
- CERN
- dCACHE
- CMS DAQ
- ALICE Tier-2 the Torino INFN CC

See details in [google doc](#).

## GRACC:

- Next-generation OSG accounting service, based on open-source technology and compatible with Gratia.
- Consists of swappable, independent components, such as Logstash, Elasticsearch, Grafana, and RabbitMQ, that communicate through a data exchange.
- Continues to interface EGI and XSEDE accounting services and provides information in accordance with existing agreements.

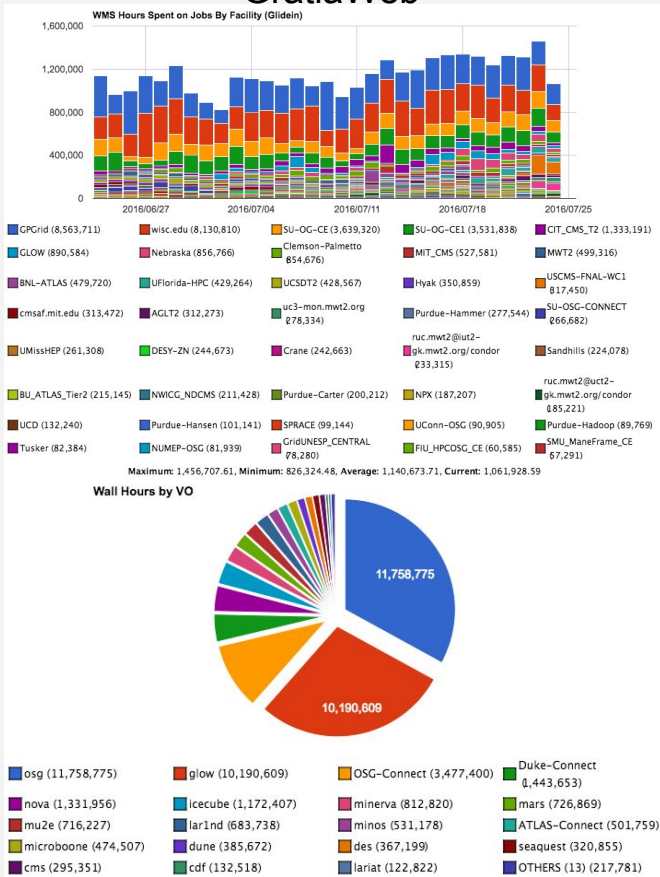
# GRACC Architecture



# Deployment

- ES Cluster at Nebraska (HCC) on “Anvil” cloud
  - head node and 4 data nodes.
  - 8 CPUs and 60GB of RAM.
  - Each data node has a 2TB and a 3TB block device (hard drive) attached to it for ES storage.
  - 4 - 5 TB Ceph-backed network storage
  - Four data/master nodes, one “client” node running collector and agents
- RabbitMQ at Indiana University (GOC)
- Collector, Forwarder (at Fermilab)

## GratiaWeb

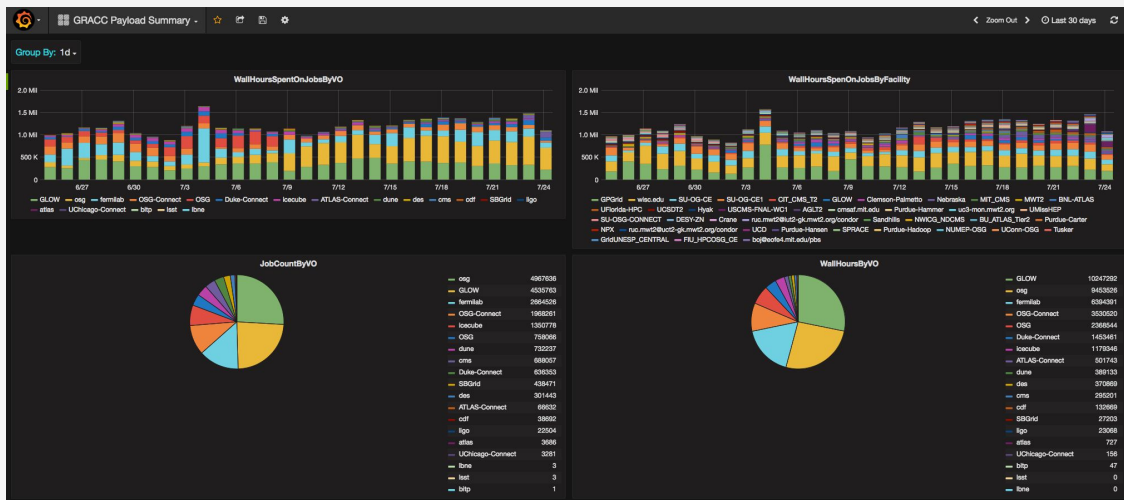


GratiaWeb (Cherry.py, google chart, mysql)

# WEB UI

## GRACC Grafana

- Uses ElasticSearch as data source.
- ElasticSearch contains data coming from all probes.
- Capable to extract time series from Elasticsearch (JobUsage and Transfer Summary Records).
- Grafana allows to add plugins (pie-chart)
- Grafana allows to do templating from ElasticSearch.

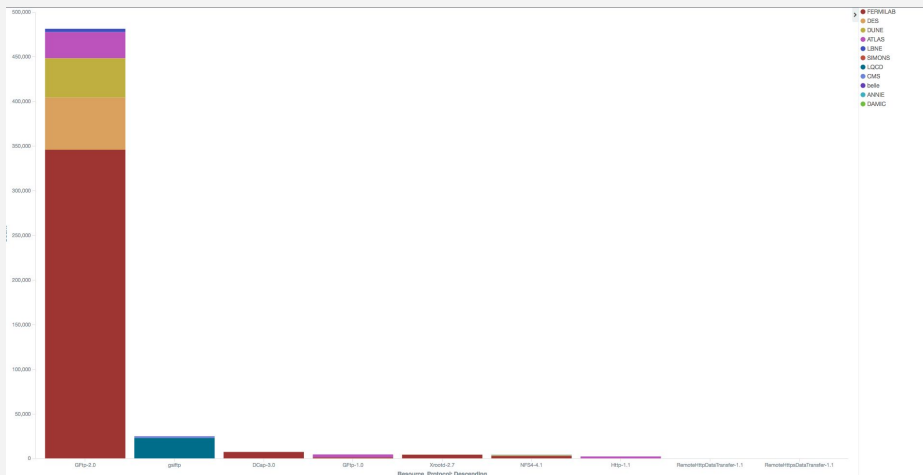


Grafana -> ElasticSearch

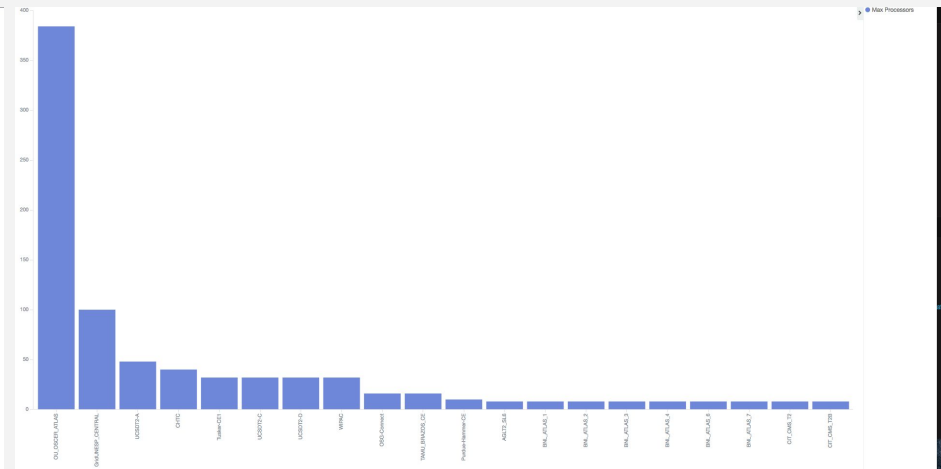


# Analytics

- Use Kibana for analytics. Examples:
  - # of transfers by Protocol by VO
  - Max # of Processors user by job by Site
- Allow access to subset of people.
- Push interesting plots to Grafana.



# of transfers by Protocols by VO



Max # of Processors by Site

# To Do List

- Understand performance, reliability and hardware requirements.
- Develop tape back of raw data
- Come up with an automated procedure of VO, Site, Project Name corrections
- Make sure that we could handle duplicate records.
- Add integration with OIM.
- Modify email reporting, OSG Display, WLCG & XSEDE reporting to query ES.
- Populate Grafana with all dashboards available in GratiaWeb.

# Summary

- It was an easy transition so far: no probes changed!
- It is nice to use well integrated and maintained Open Source Software that is used by many other organizations.
- There are a lot of Gratia features that are still missing.
- Performance, scalability and reliability should be addressed next.