GRÅCC Grid Accounting

(or, why grid accounting systems have to have such ~~terrible~~ great names)

Kevin Retzke
OSG All-Hands Meeting
07-MAR-2017

Grid Accounting

Who ran jobs where, when, using what resource, and for what purpose?

Site Admins want to know: Who’s been using my resources, and how well?

VO and **Project Coordinators** want to know: What’s my job throughput?

Stakeholders want to know: Where’s the money going?

Gratia (ca. 2006)

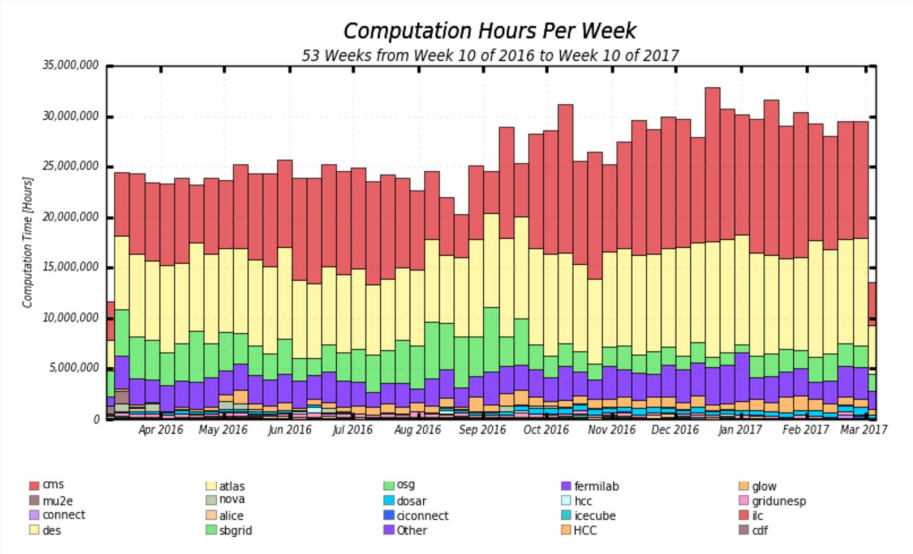
Probes run within each site's batch system, collecting job usage information (5W1H) at regular intervals. *Usage Records* are transmitted to a central *Collector*.

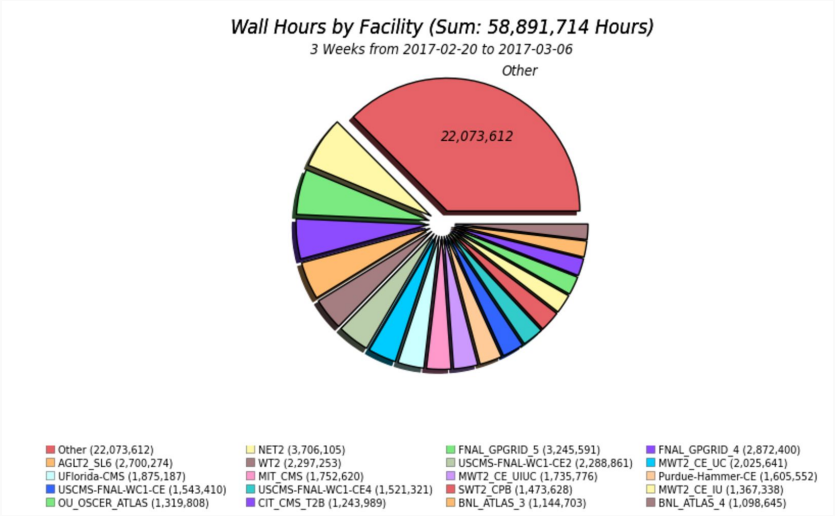
Probes exist for many batch systems: HTCondor, PBS, LSF, SLURM, etc. Also for grid storage and transfer systems.

Collector was developed as joint FNAL-OSG project; MySQL-backed Java web application running under Tomcat.

GratiaWeb

Python/matplotlib web application for graphing Gratia accounting data.





Gratia is unable to keep up with today’s rapidly-evolving grid.

GRACC

Design Requirements

Flexible document-based storage

Support for modern visualization and reporting tools

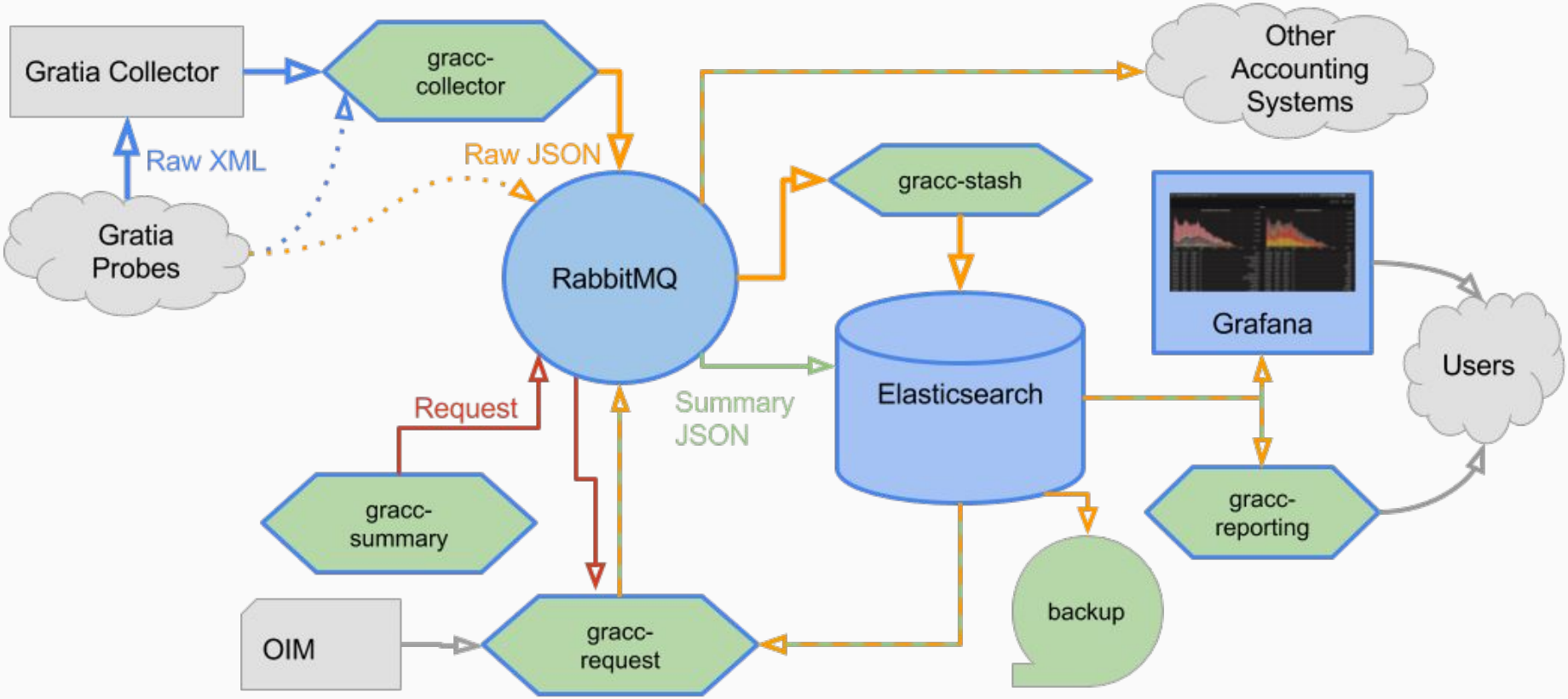
Able to support multi-dimensional long-term analytics

Compatibility with existing accounting data and Gratia probes

Modular architecture designed to evolve and scale

Leverage open-source tools and limit custom development

GRACC Architecture (v1)



Interfaces

Grafana

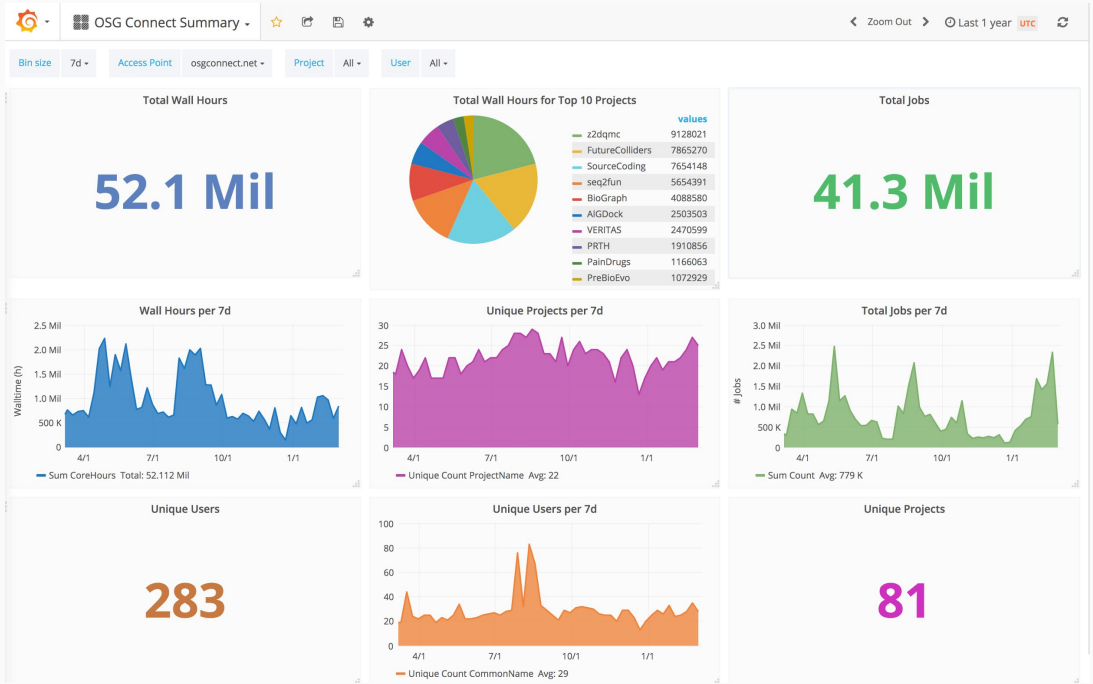
Web-based dashboarding app, primarily focused on time-series data.

Widely used in system and service monitoring.

Dashboards are easy to design and tweak. Best used for constrained “at-a-glance” information, not in-depth analytics.

Two years of success as primary user interface to Fifemon batch monitoring system at FNAL.

Ever-growing support for Elasticsearch.



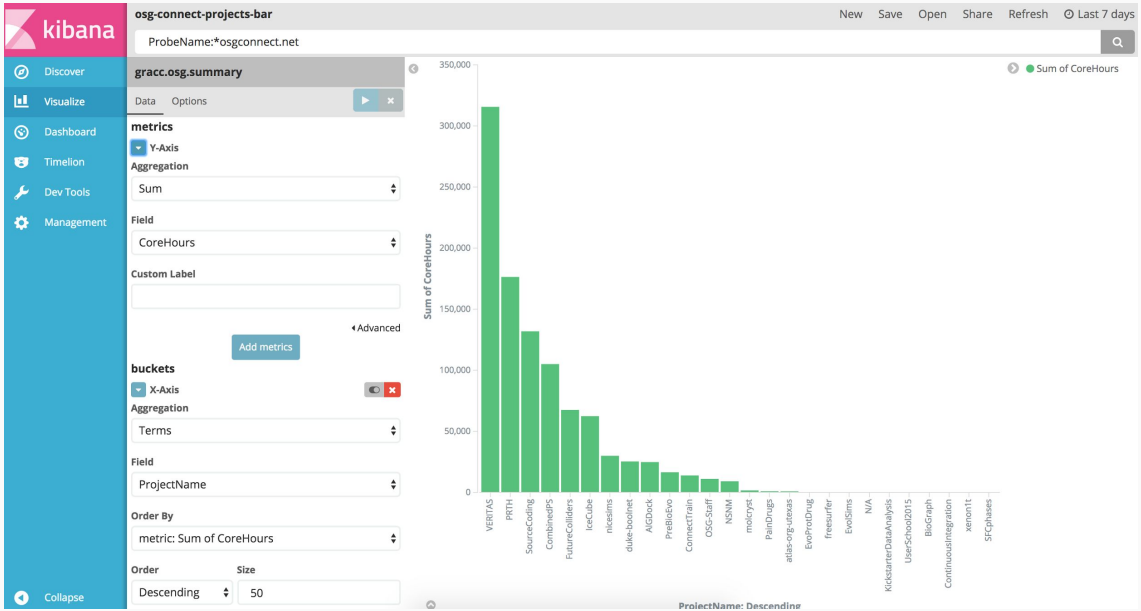
Kibana

Native Elasticsearch exploration and visualization tool.

Support for many visualization types, beyond time-series bar/line graphs and pie charts.

Best used for ad-hoc analytics; data can be easily filtered, sliced, and explored.

Visualizations are easily shared, and can be combined into dashboards.



Reporting

Regular email reports to interested parties

Some carry-over from Gratia, but many have been scrapped (until someone notices)

VOName	SiteName	ProbeName	ProjectName	Wall Hours
IU-CSIU	osg	condor:csiu.grid.iu.edu	SPLINTER	606,252
MIT_SUBMIT	osg	condor:SUBMIT.MIT.EDU	CpDarkMatterSimulation	9,655
MIT_SUBMIT	osg	condor:SUBMIT.MIT.EDU	AMS	117
SPTConnect	SPT-Connect	condor:scott.grid.uchicago.edu	spt-all	3,489

Open Science Grid

VOs Usage of OSG Sites: 2017-02-01 00:00:00 - 2017-02-28 00:00:00

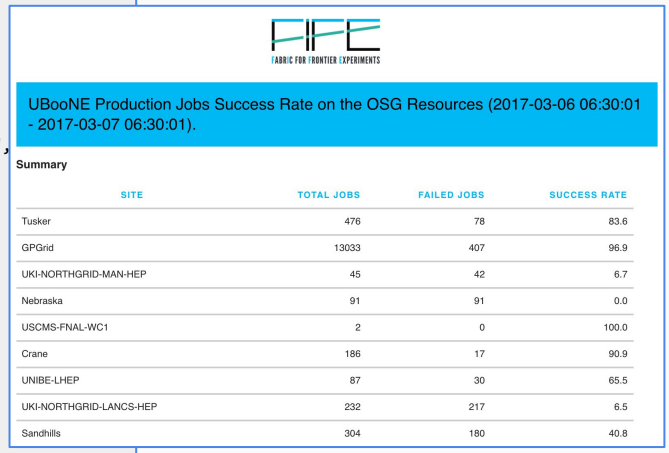
Site	Total	glow	hcc	osg	sbgrid	Opportunistic Total	Percent Opportunistic	Prev. Month Opp. Total	Percentage Change Month-Month
AGLT2	5,466,756	0	0	251,665	0	251,665	5	348,518	-28
Boston University ATLAS Tier2	6,019,843	0	0	29,608	0	29,608	0	9,128	224
Brookhaven ATLAS Tier1	9,217,568	56,662	10	97,713	0	154,385	2	342,374	-55
CCHDV	57,942	0	0	57,858	71	57,929	100	38,338	51
CHPC	48,333	0	0	48,333	0	48,333	100	0	100
COMET	137,599	0	0	103,192	0	103,192	75	162	63,560
Caltech CMS Tier2	4,000,527	217,301	53	255,815	16,053	489,221	12	394,384	24
Clemson IT	302,329	34,921	0	252,588	3,963	291,471	96	162,892	79
FIU HPC	12,573	0	0	12,573	0	12,573	100	15,974	-21
FNAL HPC	19,850	0	0	0	0	0	0	0	0
FNAL USCMS Tier1	11,980,877	120,599	0	254,935	737	376,271	3	13,238	2,742
FZU	3,883	0	0	0	0	0	0	0	0
FermiGrid	11,736,558	557,226	0	880,529	4,645	1,442,400	12	2,861,294	-50
Florida Tech	81,094	26,140	0	29,149	0	55,289	68	39,526	40

Roll your own

Elasticsearch read-only endpoint available for custom reporting.

- Grafana (your own)
- Python
- cURL
- RFC 1149
- etc.

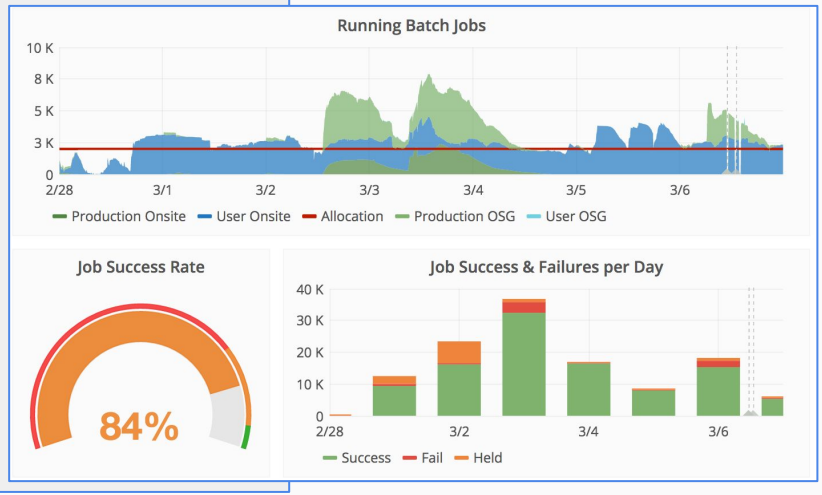
```
{
  "size": 0,
  "query": {
    "filtered": {
      "query": {
        "query_string": {
          "query": "ResourceType:BatchPilot",
          "analyze_wildcard": true
        }
      },
      "filter": {
        "bool": {
          "must": [
            {
              "range": {
                "@timestamp": {
                  "gte": 1462742321424,
                  "lte": 1462828721424,
                  "format": "epoch_millis"
                }
              }
            }
          ],
          "must_not": []
        }
      }
    }
  },
  "aggs": {
    "site": {
      "terms": {
        "field": "SiteName",
        "size": 5,
        "order": {
          "_count": "desc"
        }
      }
    }
  }
}
```



UBooNE Production Jobs Success Rate on the OSG Resources (2017-03-06 06:30:01 - 2017-03-07 06:30:01).

Summary

SITE	TOTAL JOBS	FAILED JOBS	SUCCESS RATE
Tusker	476	78	83.6
GPGrid	13033	407	96.9
UKI-NORTHGRID-MAN-HEP	45	42	6.7
Nebraska	91	91	0.0
USCMS-FNAL-WC1	2	0	100.0
Crane	186	17	90.9
UNIBE-LHEP	87	30	65.5
UKI-NORTHGRID-LANCS-HEP	232	217	6.5
Sandhills	304	180	40.8



Schema

ResourceType: Batch vs Payload?

Batch usage records are collected from the batch systems on each site. For *most* OSG usage these are **pilot** jobs from the GlideinWMS factory.

~~BatchPilot~~ **Payload** usage records are collected from the submit nodes on the VO frontend batch system.

Rule of thumb: **Sites** should focus on **Batch** records, while **VOs** should focus on **Payload** records.

Summary Records - `gracc.osg.summary`

ResourceType	“Batch” or “Payload”
@timestamp / EndTime	Date job finished (summary records are per-day)
CoreHours	WallDuration * Processors / 3600
CpuDuration_[user/system]	CPU usage in seconds
Count / NJobs	Number of jobs included in summary
VOName	Corrected VO name
ProjectName	Corrected project name
CommonName / DN	Distinguished Name of job submitter
OIM_*	Site and Project information from OIM

Use Sparingly!

Raw Records - gracc.osg.raw-%{YYYY.MM}

Same fields as summary, plus lots more individual job details:

VOName	“Raw” VO name
ProjectName	“Raw” project name
SiteName	Site name reported by probe
Host_description	(Payload only) Site on which job ran
MachineName	Host on which job ran
Network	Data transferred (if reported, units typically bytes, see Network_storageUnit)
Memory	Memory usage (if reported, units typically bytes, see Memory_storageUnit)
Resource_*	Extra information sent by the probe

NO OIM INFO!

Tour

Grafana Basics

Grafana Menu

Personal settings
Light or dark theme
Homepage (must be “starred” first)

Dashboard List

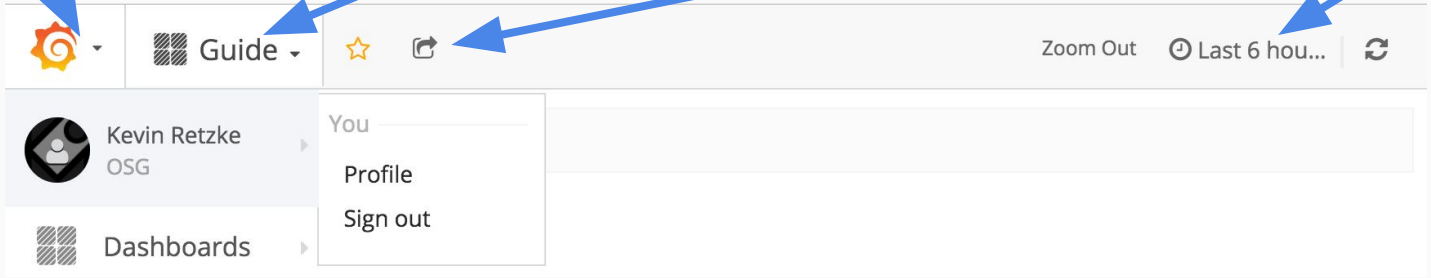
Browse/search dashboards by name or tag

Share

URL to share current dashboard including variables and time range.

Timepicker

This affects time range shown on all graphs (unless overridden, e.g. home page)



GRÅCC Grid Accounting

GRACC Service Status

Quick Links

Site Summary | VO Summary | OSG Connect | OSG Projects | Opportunistic

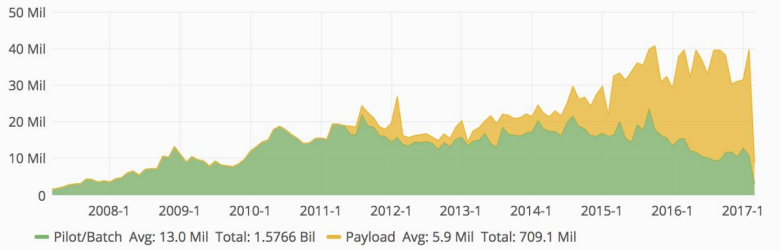
Payload Jobs | Pilot/Batch Jobs | Transfers | About the OSG | About GRACC | Help

Total Core Hours per Month

Last 10 years

Total Jobs per Month

Last 10 years



Active Virtual Organizations

Last 30 days

65

Active Sites

Last 30 days

54

Represented Fields of Science

Last 30 days

24

Top 10 VOs

Last 30 days

VOName	Total Wall Hours
cms	47.84 Mil
atlas	44.78 Mil
osg	12.21 Mil
fermilab	12.14 Mil

Top 10 Sites

Last 30 days

Site	Total Wall Hours
FermiGrid	13.34 Mil
MWT2 ATLAS UC	12.75 Mil
FNAL USCMS Tier1	12.60 Mil
Brookhaven ATLAS Tier1	10.35 Mil

Top 10 Fields of Science

Last 30 days

Metric	Total Wall Hours
High Energy Physics (LHC)	92.62 Mil
Unknown	15.20 Mil
Medical Sciences	2.55 Mil
Neuroscience	1.99 Mil

Site Summary

Zoom Out | Last 30 days UTC

Bin size: auto | Site: SU OrangeGrid | Job Type: Batch | ? Batch vs Payload Jobs

Total Wall Hours

3.50 Mil

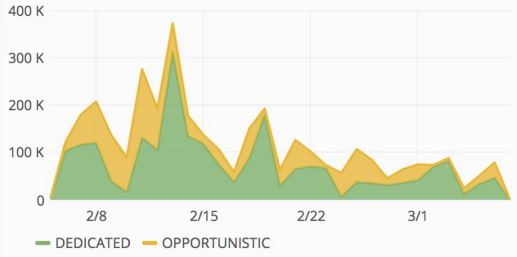
Total CPU Hours

2.512 Mil

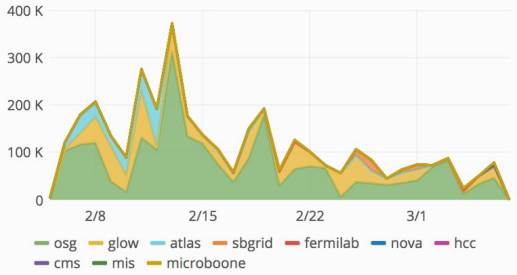
Total Jobs

506 K

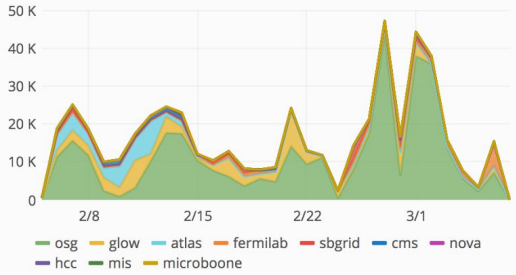
Wall Hours by Usage Model by 1d



Wall Hours by VO per 1d



Jobs by VO per 1d



VOs

VOName	Total Wall Hours
osg	2.24 Mil
glow	883.50 K
atlas	270.21 K
sbgrid	64.96 K
fermilab	26.13 K
nova	6.12 K
hcc	5.60 K
cms	1.38 K
mis	425.20
microboone	5.77

Opportunistic VOs

VOName	Total Wall Hours
glow	883.50 K
atlas	270.21 K
sbgrid	64.96 K
fermilab	26.13 K
nova	6.12 K
hcc	5.60 K
cms	1.38 K
mis	425.20
microboone	5.77

Fields of Science

Field of Science	Total Wall Hours
Unknown	1.26 Mil

Site Summary

Zoom Out | Last 30 days UTC

Bin size: auto | Site: SU OrangeGrid | Job Type: Payload | ? Batch vs Payload Jobs

Total Wall Hours

2.870 Mil

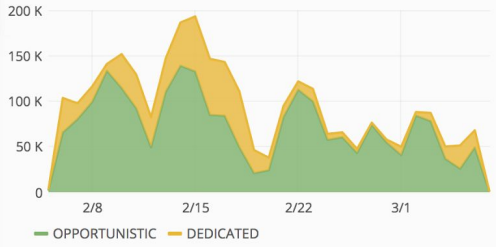
Total CPU Hours

1.607 Mil

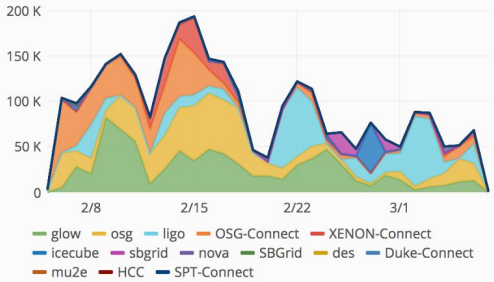
Total Jobs

2.556 Mil

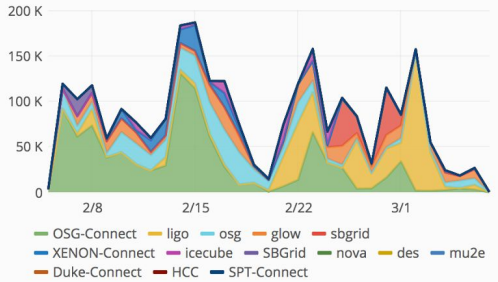
Wall Hours by Usage Model by 1d



Wall Hours by VO per 1d



Jobs by VO per 1d



VOs

VOName	Total Wall Hours
glow	782.73 K
osg	689.57 K
ligo	563.96 K
OSG-Connect	519.14 K
XENON-Connect	150.73 K
icecube	80.59 K
sbgrid	56.23 K
nova	21.32 K
SBGrid	5.44 K
des	238.37
Duke-Connect	117.11

Opportunistic VOs

VOName	Total Wall Hours
glow	782.73 K
ligo	563.96 K
OSG-Connect	519.14 K
XENON-Connect	150.73 K
icecube	80.59 K
sbgrid	56.23 K
nova	21.32 K
SBGrid	5.44 K
des	238.37
Duke-Connect	117.11
mu2e	15.01

Fields of Science

Field of Science	Total Wall Hours
Unknown	946.55 K
Gravitational Physics	563.96 K
Medical Sciences	230.11 K
Astrophysics	190.63 K
Physics	104.93 K
Chemistry	70.48 K
Engineering	36.82 K
Physics and astronomy	16.61 K
Biophysics	12.81 K
Computer and Information Science and Engineering	3.96 K
Training	1.27 K

VO Summary

Zoom Out | Last 30 days UTC

Bin size: auto | VO: glow | Job Type: Payload | ? Batch vs Payload Jobs

Total Wall Hours

2.254 Mil

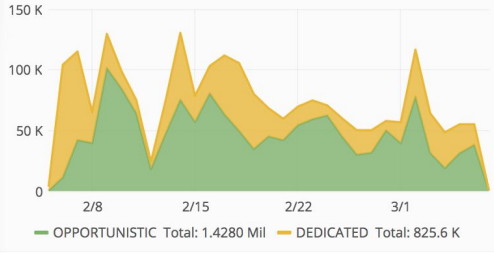
Total CPU Hours

430 K

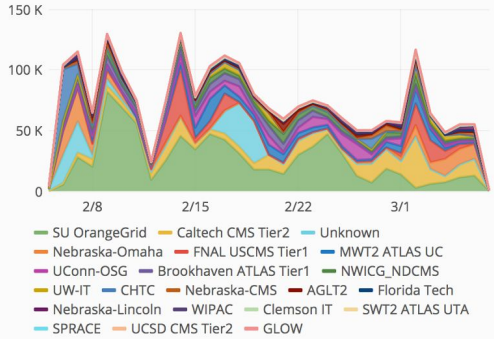
Total Jobs

1.261 Mil

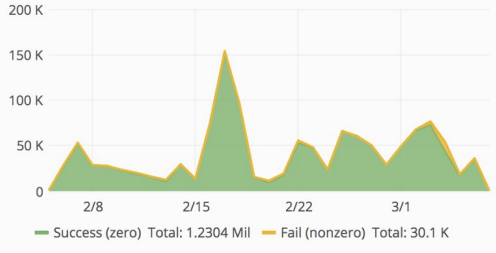
Wall Hours by Usage Model by 1d



Wall Hours by Site per 1d



Jobs by Exit Code per 1d



Sites

Site	Total Wall Hours
SU OrangeGrid	782.73 K
Caltech CMS Tier2	266.90 K
Unknown	153.11 K
Nebraska-Omaha	144.75 K
FNAL USCMS Tier1	134.42 K
MWT2 ATLAS UC	128.88 K
UConn-OSG	119.44 K
Brookhaven ATLAS Tier1	116.21 K
NWICG_NDCMS	102.29 K

Opportunistic Sites

Site	Total Wall Hours
SU OrangeGrid	782.73 K
Caltech CMS Tier2	266.90 K
MWT2 ATLAS UC	128.88 K
NWICG_NDCMS	102.29 K
Nebraska-CMS	52.95 K
Florida Tech	28.83 K
Nebraska-Lincoln	21.37 K
WIPAC	12.92 K
Clemson IT	12.38 K

Payload Submission Nodes

Submission Node	Total Wall Hours
condor:submit-5.chtc.wisc.edu	1.89 Mil
condor:submit-3.chtc.wisc.edu	344.51 K
condor:submit-4.chtc.wisc.edu	12.82 K
condor:deepdivesubmit.chtc.wisc.edu	8.73 K
condor:cecc7test.hep.wisc.edu	0.00

VO Summary

Zoom Out | Last 30 days UTC

Bin size: auto | VO: glow | Job Type: Batch | ? Batch vs Payload Jobs

Total Wall Hours

3.94 Mil

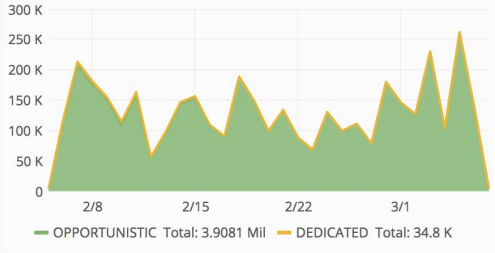
Total CPU Hours

1.766 Mil

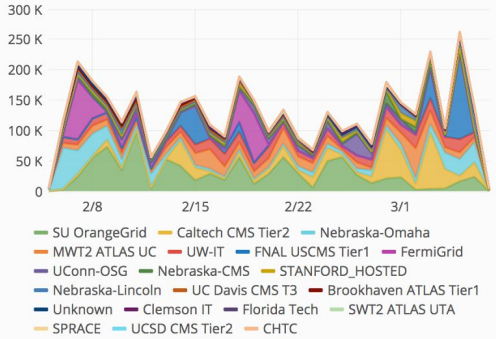
Total Jobs

495 K

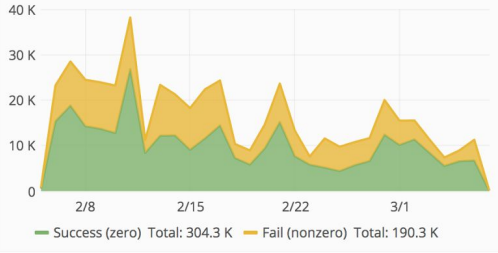
Wall Hours by Usage Model by 1d



Wall Hours by Site per 1d



Jobs by Exit Code per 1d



Sites

Site	Total Wall Hours
SU OrangeGrid	883.50 K
Caltech CMS Tier2	497.12 K
Nebraska-Omaha	413.63 K
MWT2 ATLAS UC	403.57 K
UW-IT	377.31 K
FNAL USCMS Tier1	311.97 K
FermiGrid	302.38 K
UConn-OSG	149.53 K
Nebraska-CMS	122.84 K

Opportunistic Sites

Site	Total Wall Hours
SU OrangeGrid	883.50 K
Caltech CMS Tier2	497.12 K
Nebraska-Omaha	413.63 K
MWT2 ATLAS UC	403.57 K
UW-IT	377.31 K
FNAL USCMS Tier1	311.97 K
FermiGrid	302.38 K
UConn-OSG	149.53 K
Nebraska-CMS	122.84 K

Payload Submission Nodes

Submission Node	Total Wall Hours
condor:submit-5.chtc.wisc.edu	1.89 Mil
condor:submit-3.chtc.wisc.edu	344.51 K
condor:submit-4.chtc.wisc.edu	12.82 K
condor:deepdivesubmit.chtc.wisc.edu	8.73 K
condor:cecc7test.hep.wisc.edu	0.00

OSG Connect Summary

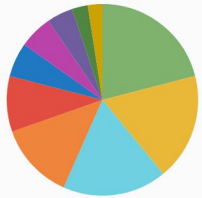
Zoom Out | Last 1 year UTC

Bin size: 7d | Access Point: osgconnect.net | Project: All | User: All

Total Wall Hours

52.2 Mil

Total Wall Hours for Top 10 Projects

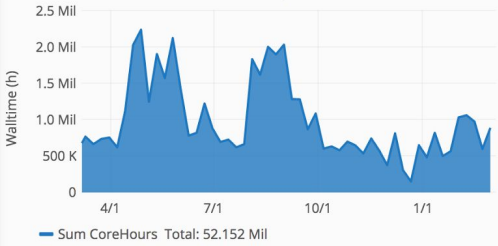


	values
z2dqmc	9128021
FutureColliders	7865270
SourceCoding	7654148
seq2fun	5654391
BioGraph	4088582
AlGDock	2503503
VERITAS	2476279
PRTH	1910856
PainDrugs	1166063
PreBioEvo	1073384

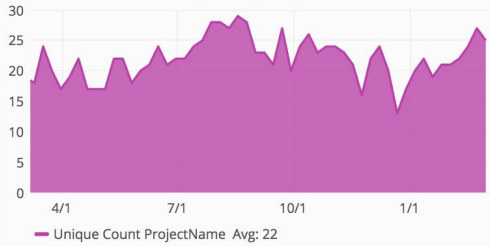
Total Jobs

41.3 Mil

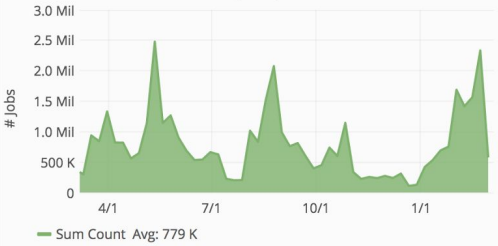
Wall Hours per 7d



Unique Projects per 7d



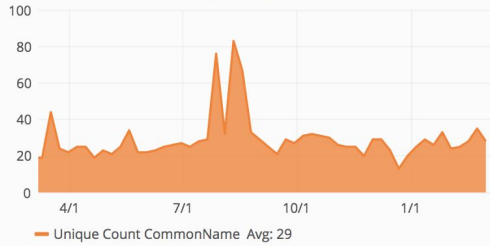
Total Jobs per 7d



Unique Users

283

Unique Users per 7d



Unique Projects

81

OSG Project Accounting

Zoom Out | Last 1 year UTC

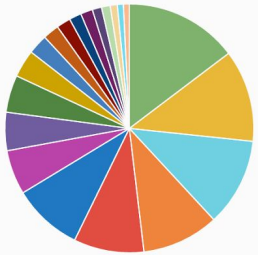
Bin auto

Projects

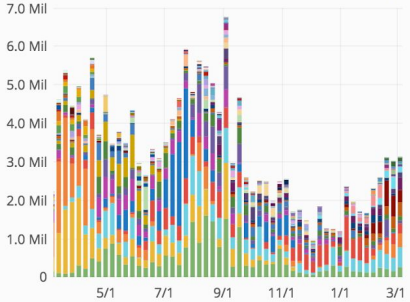
Project Name	PI Name	Institution	Field Of Science	WallHours
TG-IBN130001	Donald Krieger	University of Pittsburgh	Biological Sciences	23.22 Mil
AMS	Baosong Shan	MIT	Particle Physics	18.92 Mil
SPLINTER	Samy Meroueh	Indiana University	Medical Sciences	18.09 Mil
LIGO	Peter F. Couvares	Syracuse University	Gravitational Physics	15.81 Mil
Duke-QGP	Steffen A. Bass	Duke University	Nuclear Physics	14.42 Mil
CpDarkMatterSimulation	Christoph Paus	MIT	High Energy Physics	14.36 Mil
z2dqmc	Snir Gazit	University of California Berkeley	Physics	9.13 Mil
FutureColliders	Sergei Chekanov	Argonne National Lab	High Energy Physics	7.87 Mil
SourceCoding	David Mitchell	New Mexico State University	Engineering	7.65 Mil

1 2 3 4 5 6 7

Wall Hours per Project



Wall Hours per Project



	avg	total
TG-IBN130001	438 K	23.219 Mil
AMS	357 K	18.917 Mil
SPLINTER	341 K	18.092 Mil
LIGO	298 K	15.812 Mil
Duke-QGP	272 K	14.423 Mil
CpDarkMatterSimulation	271 K	14.361 Mil
z2dqmc	172 K	9.128 Mil
FutureColliders	148 K	7.865 Mil
SourceCoding	144 K	7.654 Mil
seq2fun	107 K	5.654 Mil
BioGraph	77 K	4.082 Mil
TG-PHY150040	63 K	3.328 Mil

CPU Hours per Project

5.0 Mil

avg total

Job Count per Project

7.0 Mil

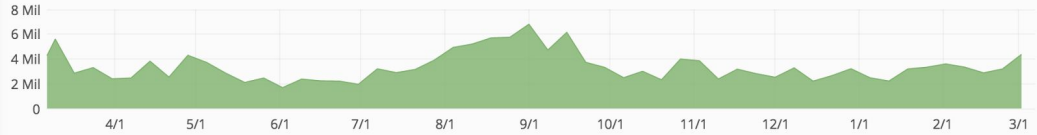
avg total

Opportunistic Usage

Zoom Out | Last 1 year UTC

Bin auto | Filter +

Opportunistic Wall Hours per 7d



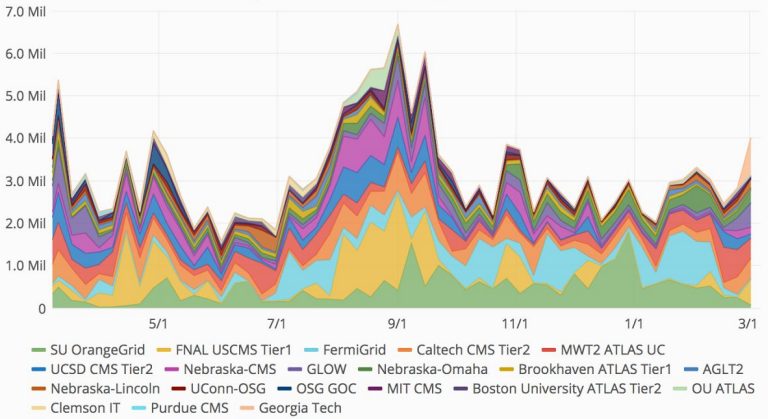
Total Opportunistic Wall Hours

177 Mil

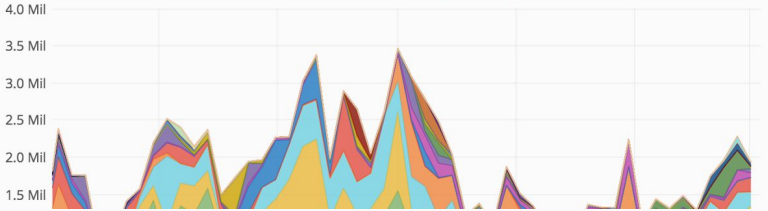
Top 20 Sites Providing Opportunistic Usage

OIM_Site	OIM_Facility	Wall Hours
SU OrangeGrid	Syracuse University	24 Mil
FNAL USCMS Tier1	Fermi National Accelerator Laboratory	20 Mil
FermiGrid	Fermi National Accelerator Laboratory	19 Mil
Caltech CMS Tier2	California Institute of Technology	17 Mil
MWT2 ATLAS UC	University of Chicago	15 Mil
UCSD CMS Tier2	University of California San Diego	13 Mil
Nebraska-CMS	University of Nebraska	13 Mil
GLOW	University of Wisconsin	10 Mil
Nebraska-Omaha	University of Nebraska	9 Mil
Brookhaven ATLAS Tier1	Brookhaven National Laboratory	5 Mil
AGLT2	University of Michigan	4 Mil
Nebraska-Lincoln	University of Nebraska	4 Mil
UConn-OSG	University of Connecticut	3 Mil
OSG GOC	Indiana University	3 Mil
MIT CMS	Massachusetts Institute of Technology	3 Mil
Boston University ATLAS Tier2	Boston University	2 Mil

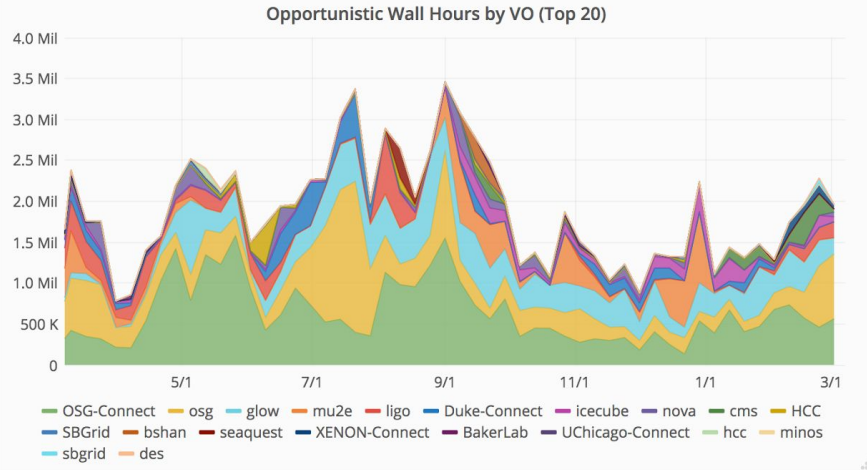
Opportunistic Wall Hours by Site (Top 20)



Opportunistic Wall Hours by VO (Top 20)



AGLT2	University of Michigan	4 Mil
Nebraska-Lincoln	University of Nebraska	4 Mil
UConn-OSG	University of Connecticut	3 Mil
OSG GOC	Indiana University	3 Mil
MIT CMS	Massachusetts Institute of Technology	3 Mil
Boston University ATLAS Tier2	Boston University	2 Mil
OU ATLAS	University of Oklahoma	1 Mil
Clemson IT	Clemson University	1 Mil
Purdue CMS	Purdue University	1 Mil
Georgia Tech	Georgia Institute of Technology	1 Mil



Top 10 Facilities

OIM_Facility	Wall Hours
Fermi National Accelerator Laboratory	40 Mil
University of Nebraska	25 Mil
Syracuse University	24 Mil
California Institute of Technology	17 Mil
University of Chicago	15 Mil
University of California San Diego	13 Mil
University of Wisconsin	10 Mil
Brookhaven National Laboratory	5 Mil
University of Michigan	4 Mil
University of Connecticut	3 Mil

Top 10 Resource Groups

ResourceGroup	Wall Hours
SU-OG	24 Mil
USCMS-FNAL-WC1	20 Mil
FNAL_FERMIGRID	19 Mil
CIT_CMS_T2	17 Mil
MWT2	15 Mil
UCSDT2	13 Mil
Nebraska	13 Mil
GLOW	10 Mil
BNL-ATLAS	5 Mil
Crane	5 Mil

Top 10 Resources

Resource	Wall Hours
USCMS-FNAL-WC1-CE3	20 Mil
SU-OG-CE	15 Mil
FNAL_GPGRID_5	10 Mil
SU-OG-CE1	10 Mil
GLOW-OSG	10 Mil
CIT_CMS_T2	9 Mil
FNAL_GPGRID_4	9 Mil
MWT2	8 Mil
CIT_CMS_T2B	8 Mil
Crane-CE1	5 Mil

What’s Ahead

Are we there yet?

Couple remaining items to be production-ready:

- Raw record archival to tape
- Transitioning reporting/interfaces to other accounting systems
- Operations

Future Work

More/Better Dashboards. What do you want to see?

More site- and user-specific fields. Customizable fields in OIM? Easily-extensible probes? User-defined ClassAds?

Integrate other sources of data. Grafana makes it easy to display diverse data on a single page. Elasticsearch can store *anything*. What else can we monitor?

Credits

GRACC Design & Development Team

- Derek Weitzel (UNL)
- Shreyas Bhat (FNAL)
- Carl Edquist (UW)
- Tanya Levshina (FNAL)
- Brian Bockelman (UNL)
- Bo Jayatilaka (FNAL)

Special thanks:

- OSG Area Coordinators
- GOC Staff
- “Voluntary” Beta testers

Links

Grafana: https://gracc.opensciencegrid.org

Kibana: https://gracc.opensciencegrid.org/kibana

Elasticsearch: https://gracc.opensciencegrid.org/q

Project docs: https://opensciencegrid.github.io/gracc

Source code: https://github.com/opensciencegrid

AMA

Live: Office hours Thursday morning

Chat: #gracc on OSG Slack

Email: gracc-project@opensciencegrid.org

Ticket: https://jira.opensciencegrid.org/browse/GRACC or GOC

There’s two hard problems in CS:

0. Cache invalidation
1. Naming things
2. Off-by-one errors