Open Science Grid

Production Support Report

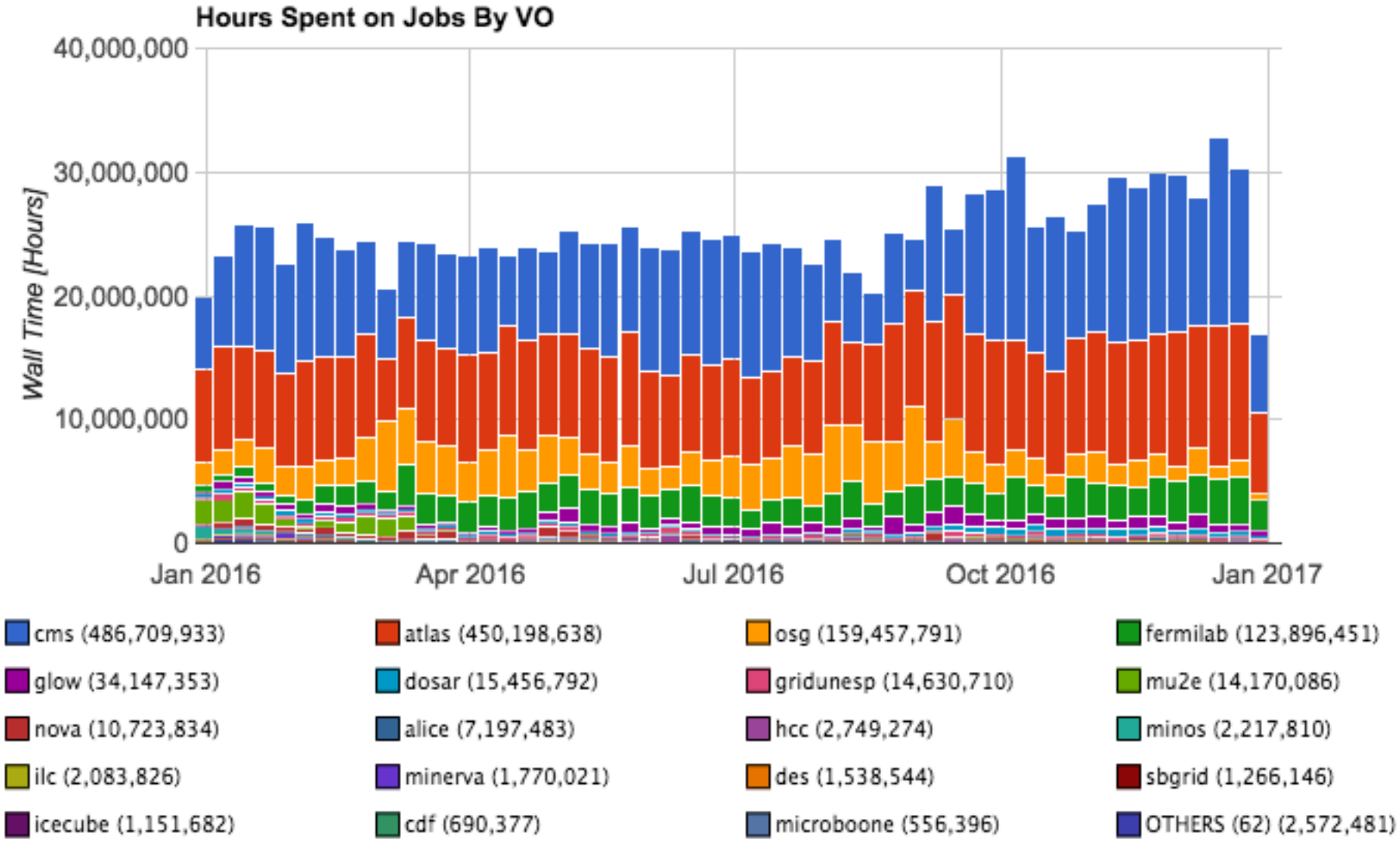
Ken Herner

18 Jan 2017

Open Science Grid

Part I: 2016 in Review

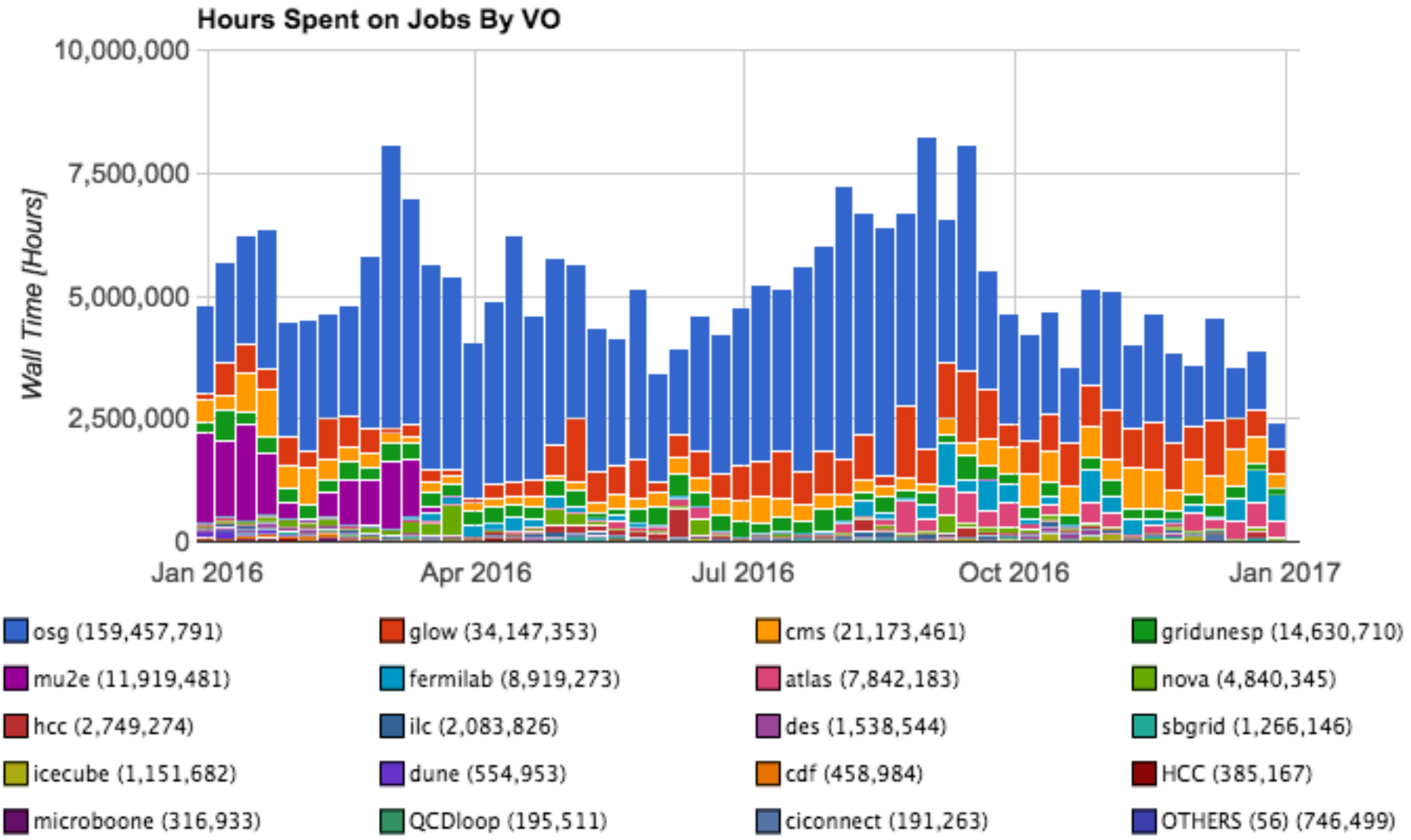
Total Hours: All sites, all VOs



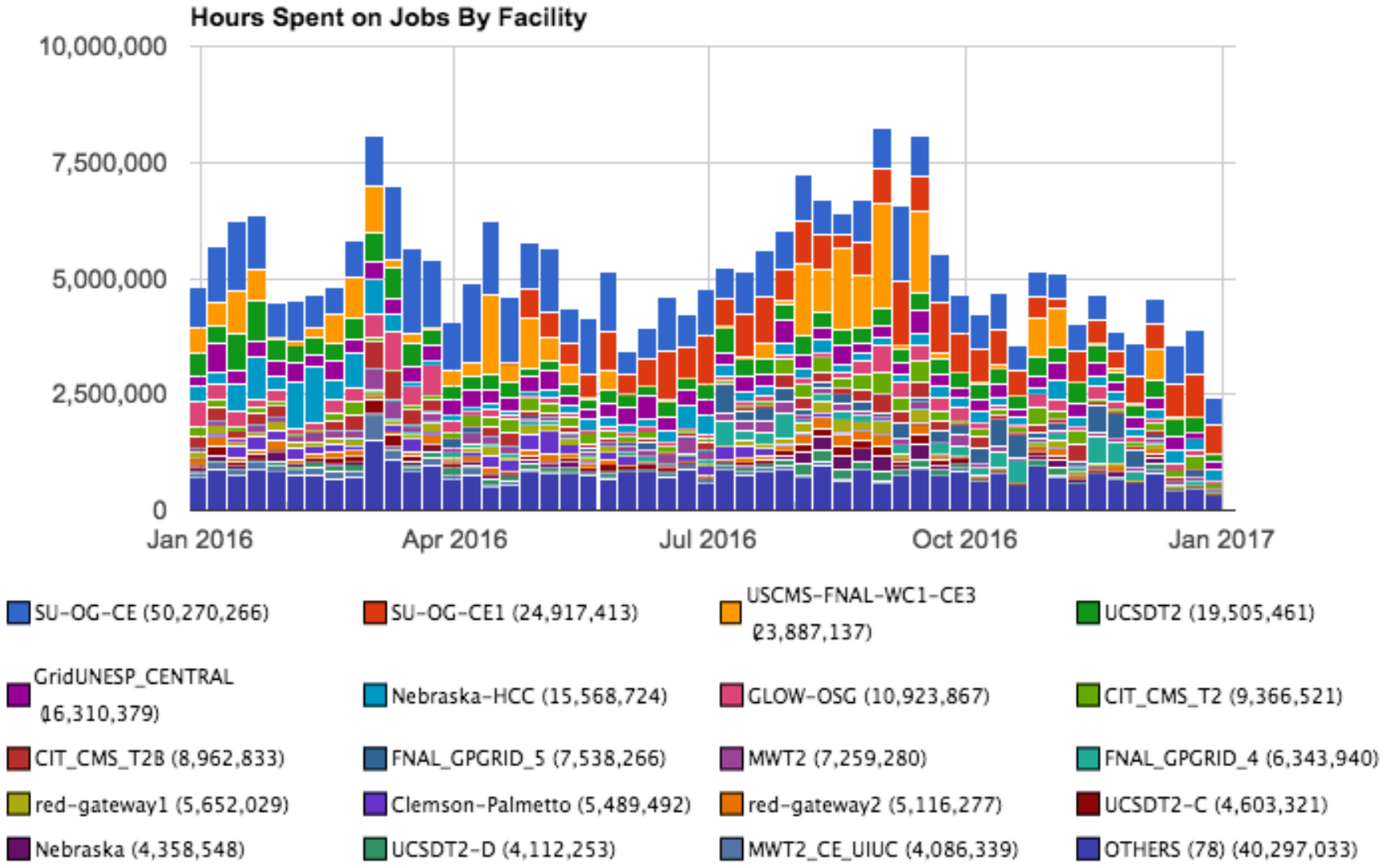
Total Hours: All sites, all VOs

- **1.3 billion hours total in 2016**
- ATLAS and CMS lead the way
- Some weeks exceeded 30M hours
- Monthly totals have improved at least 6 straight months
  - **130M in December 2016 alone**
- Numbers do not include Georgia Tech, so total is even higher

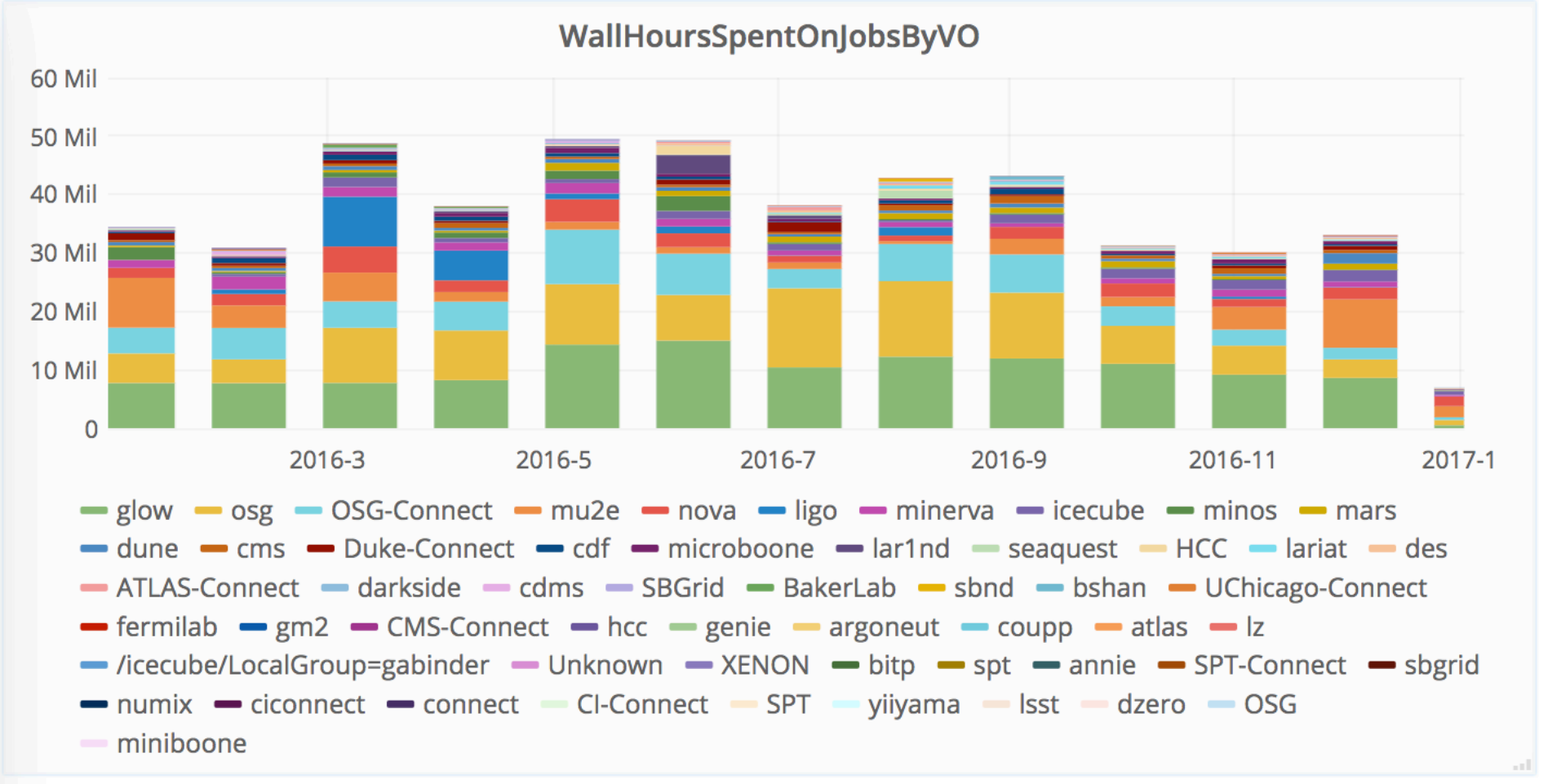
Opportunistic Totals



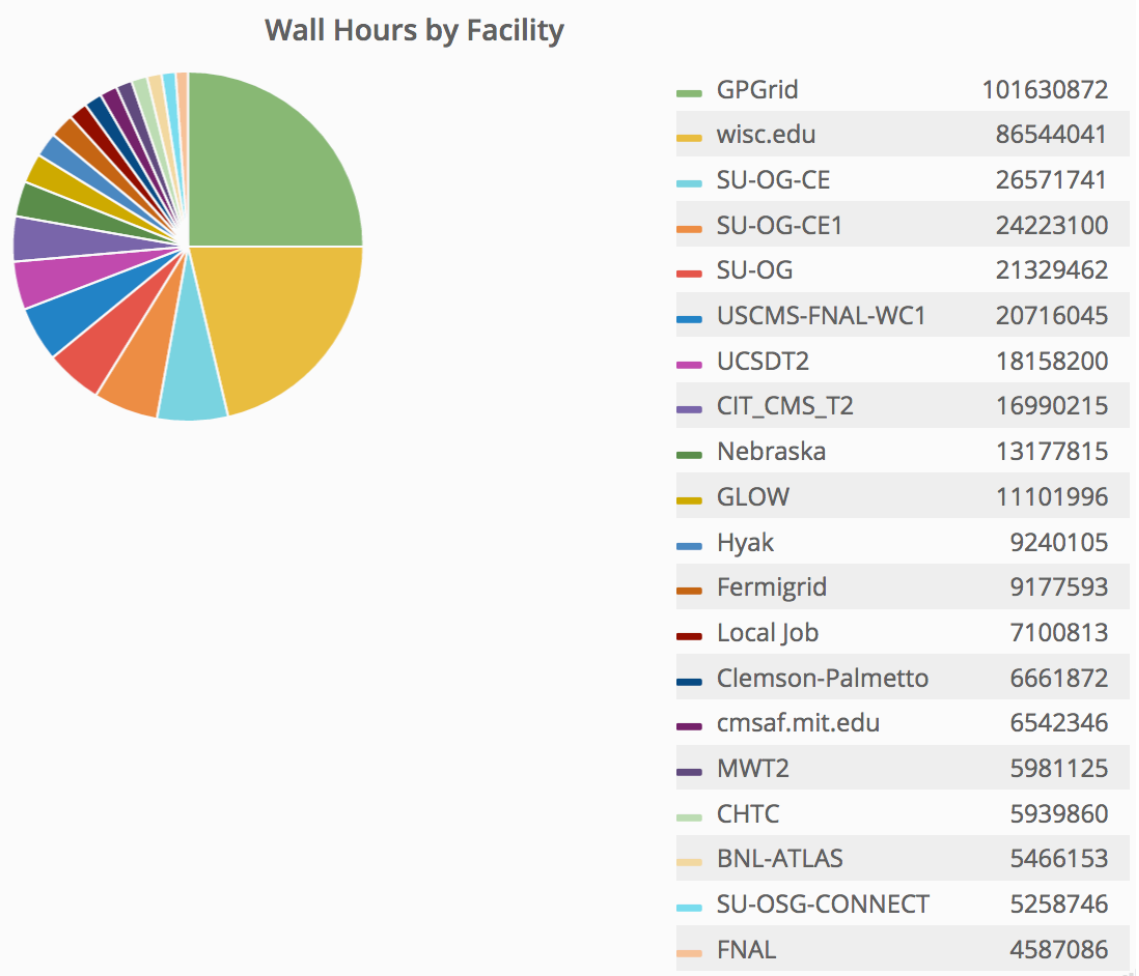
Opportunistic By Facility



Payload Totals

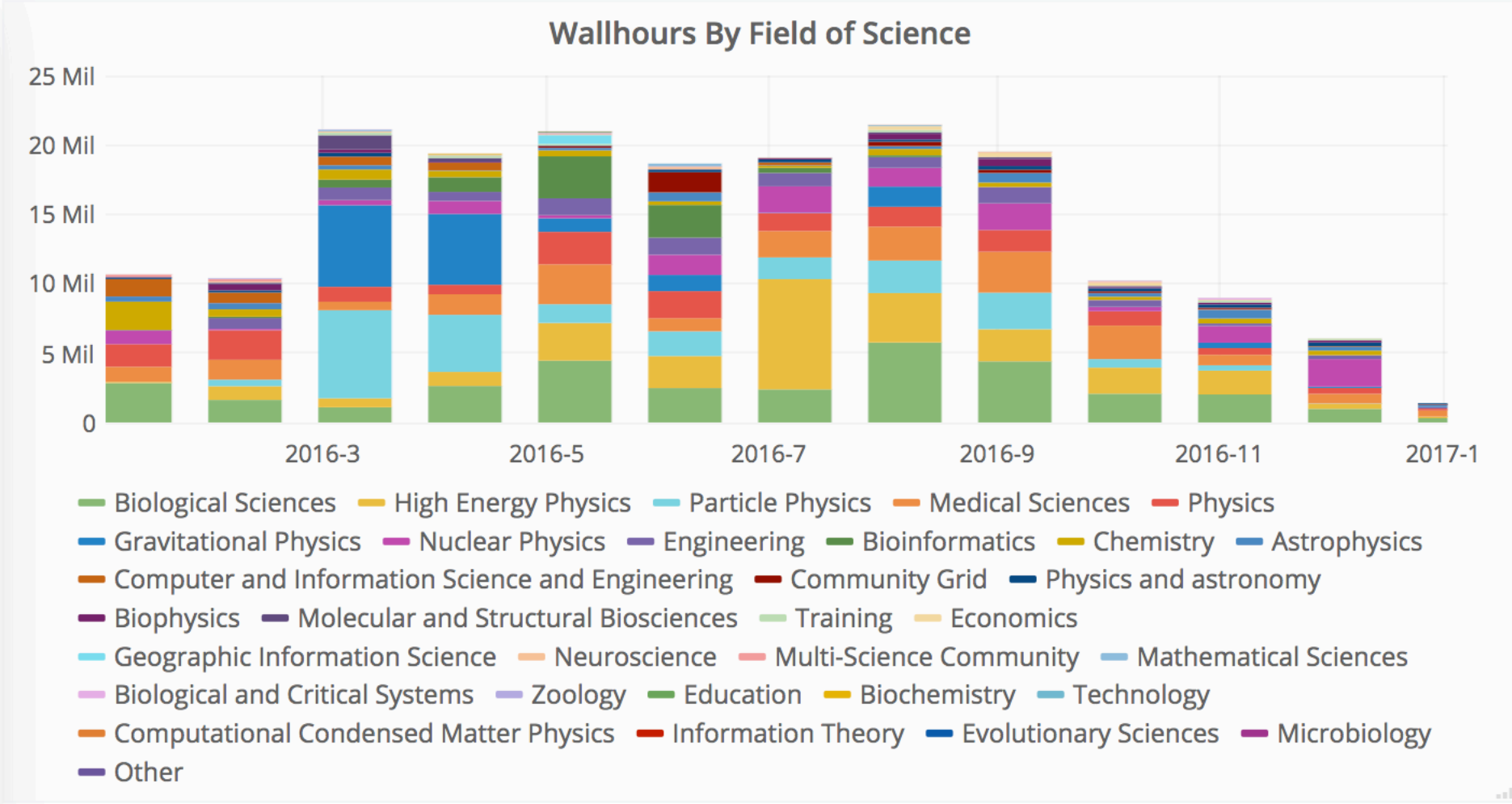


Payload By Facility



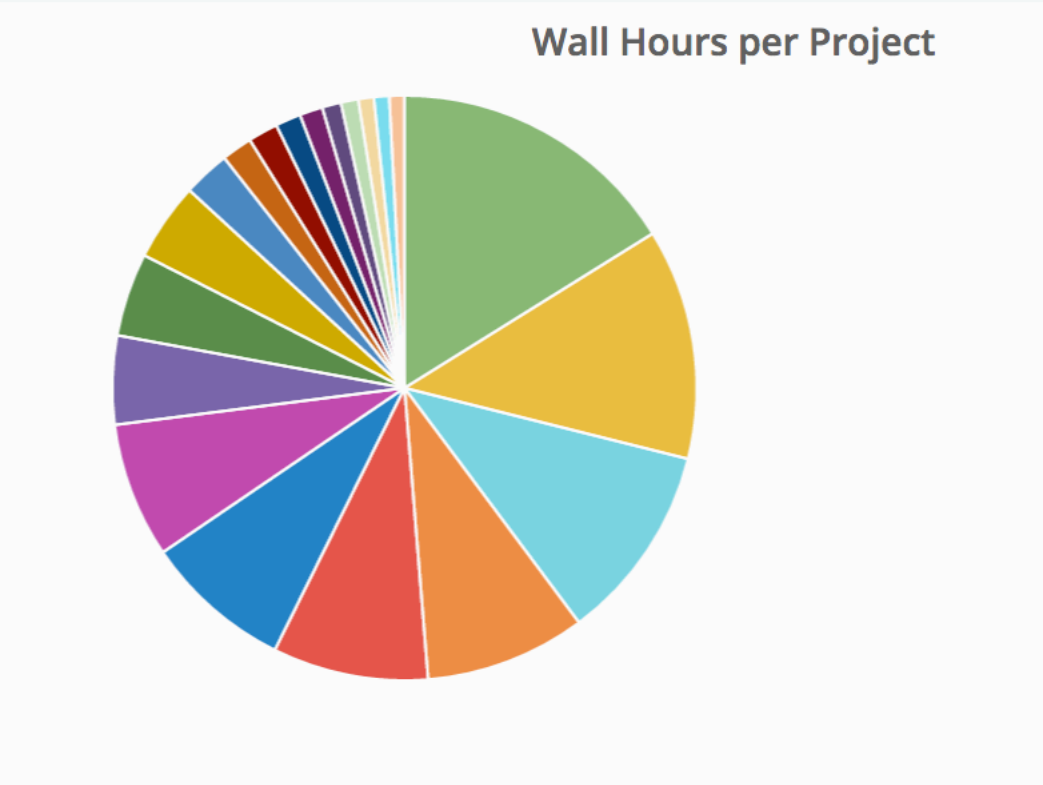
- Payload, not pilot jobs

Payload by Field



Usual suspects at work here...

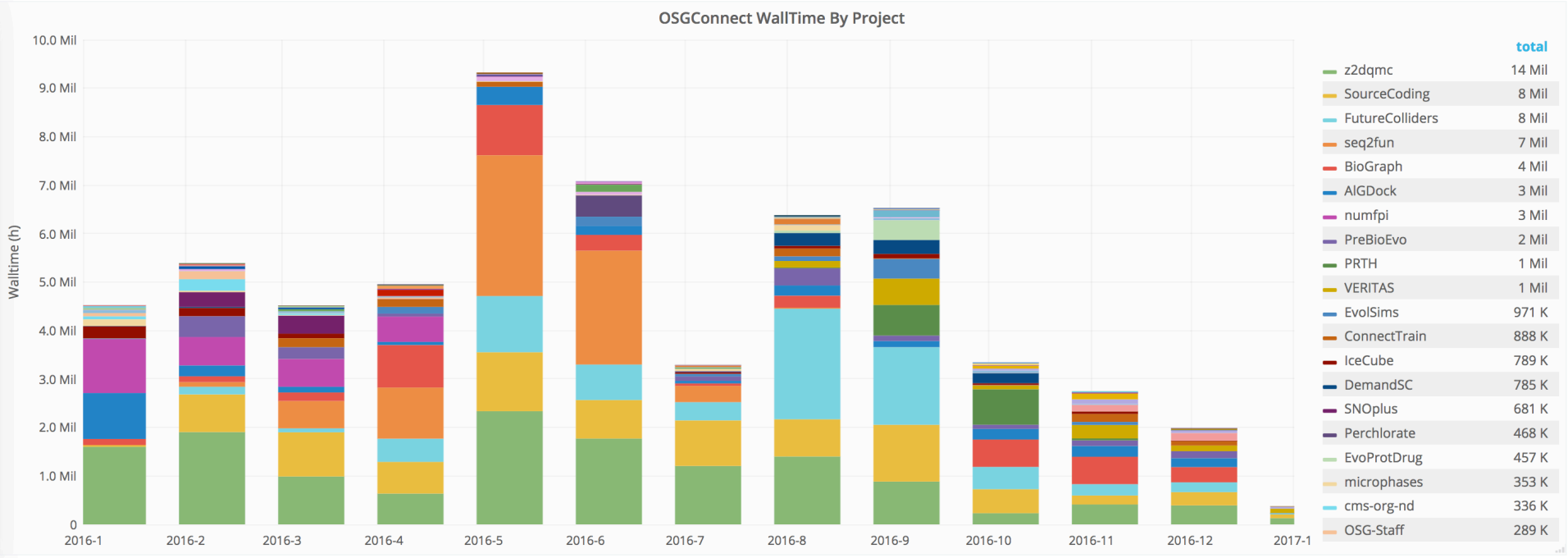
Project Totals



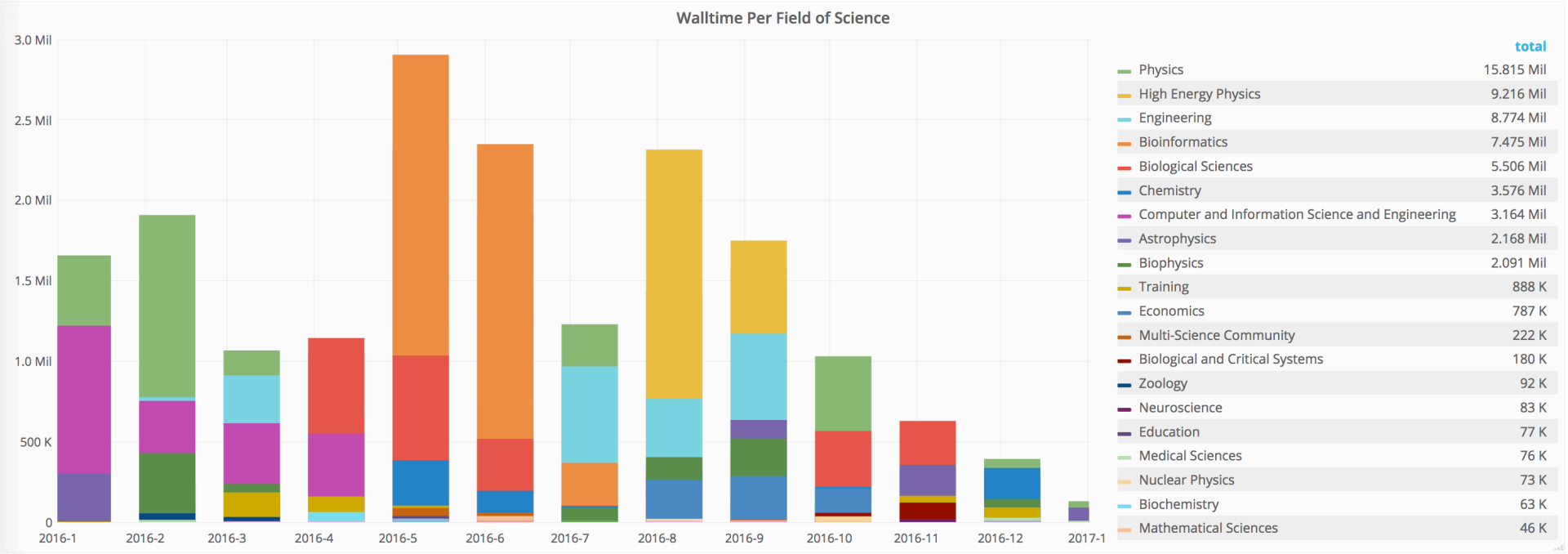
TG-IBN130001	27360795
AMS	21614274
SPLINTER	18499120
LIGO	15081818
CpDarkMatterSimulation	14593440
z2dqmc	13861248
Duke-QGP	12683005
SourceCoding	8309610
FutureColliders	7810953
seq2fun	7329788
BioGraph	4396233
AlGDock	2882869
numfpi	2799446
HCCLocalSubmit	2384529
TG-PHY150040	2165548
AmorphousOrder	1774525
PreBioEvo	1633669
DetectorDesign	1452756
atlas-org-uchicago	1431638
PRTH	1415732

- And the winner is... Donald Krieger!

OSG Connect by Project



Project time by Field



Open Science Grid

HPC

- MINOS and CMS ~~still attempting to~~ **running** at Stampede
  - MINOS still adapting workflow but test jobs do succeed
  - CMS
- Opportunistic running on Comet continues
  - still not yet with allocations
- Hopeful to get Stanford GPU cluster up and running in the semi-near future (get rest of site working first)

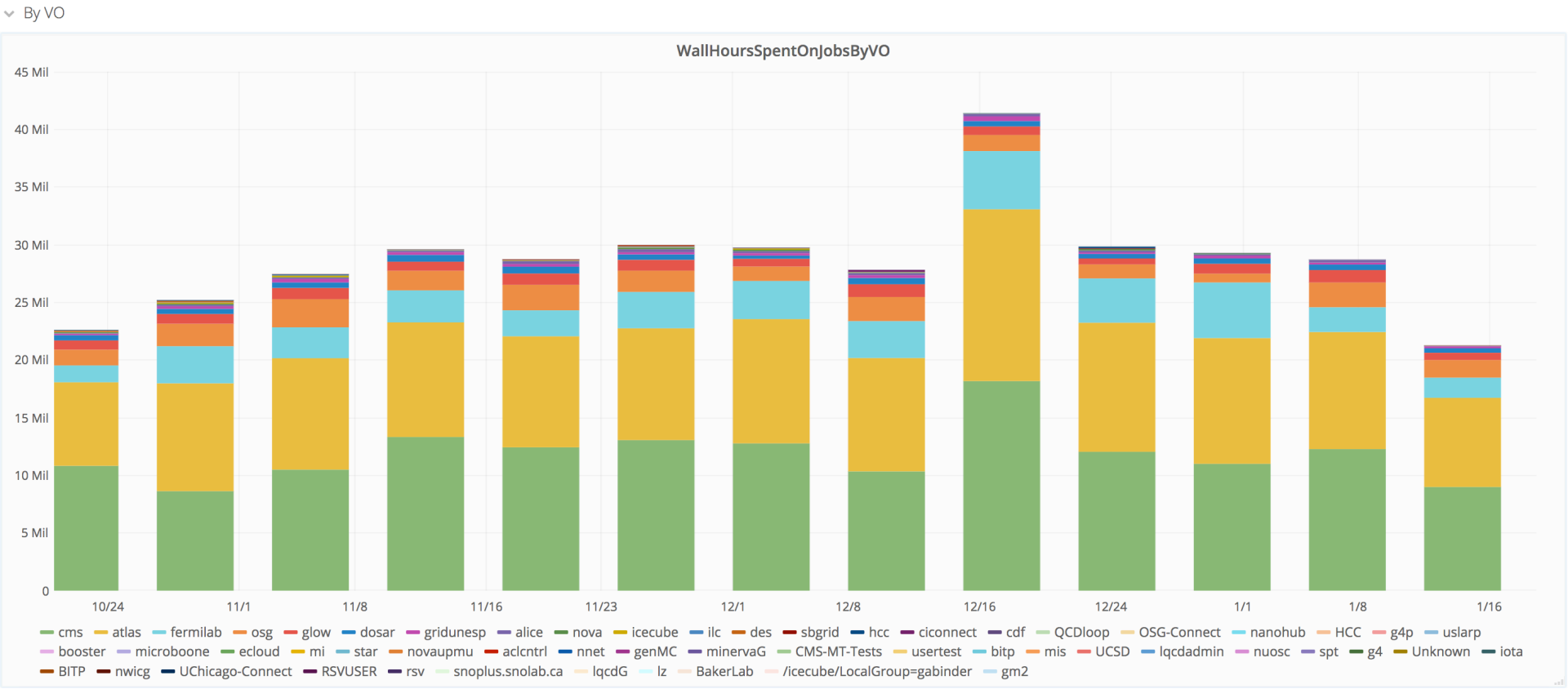
Trends and changes this year

- Steady growth throughout!
- Several new sites commissioned, including non-US sites
- Large push from IF experiments in spring for summer conferences, also from LHC for ICHEP
- Post-HEP conference lull in late summer. Huge boon for opportunistic running!!!
- Oct-Dec: Downturn in opportunistic resources for OSG, etc. VOs (-40%)
  - Academic year kicks in, affects campus sites
  - (The big factor) ATLAS and CMS putting greater emphasis on opportunistic running
  - **Very important to grow the pie and leverage non-LHC sites**
- Some sites are moving to SL7 (at least partially)
  - Some smoother than others
  - Could reduce overall availability for some workflows

Open Science Grid

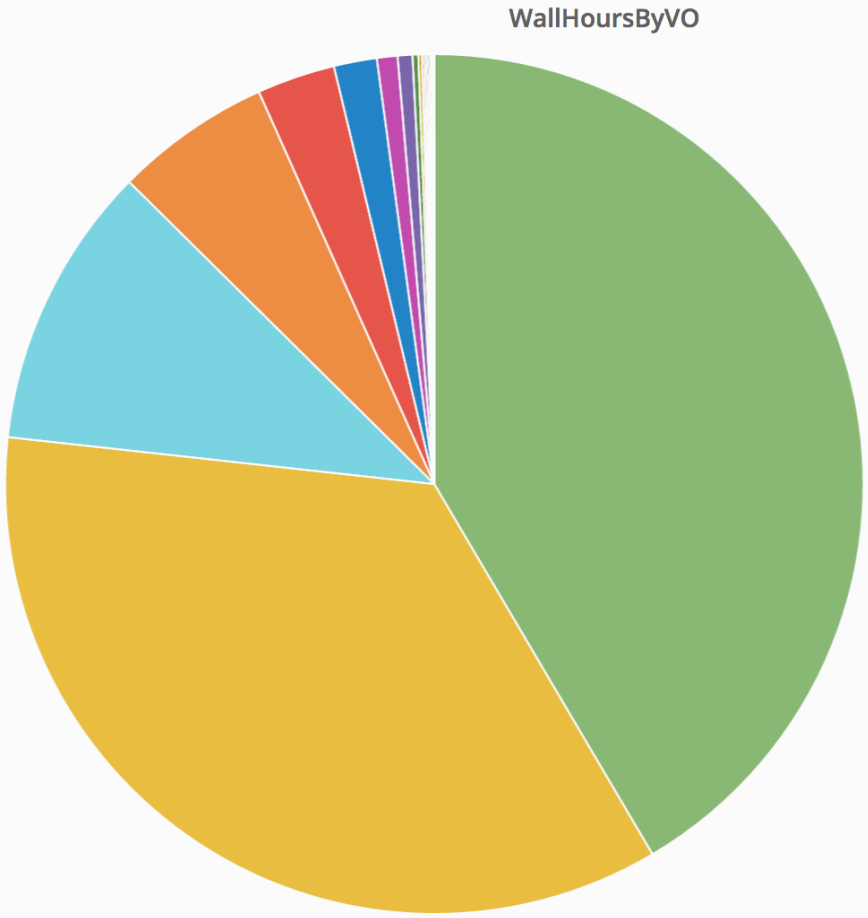
Part II: Recent status and plans

Pilot hours over 90 days



This is not opportunistic only
Excludes GaTech

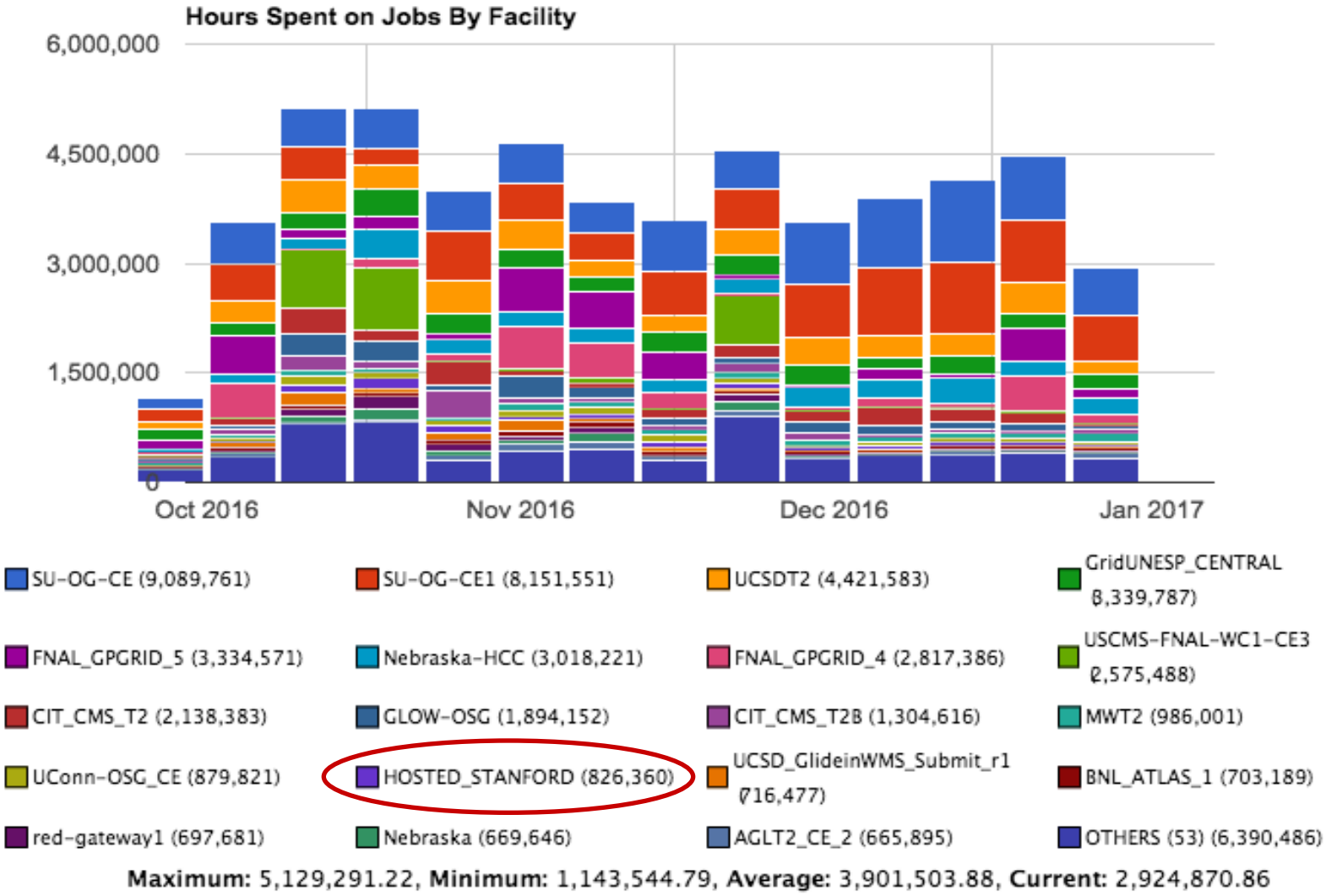
Pilot hours over 90 days (2)



cms	154348735
atlas	131010890
fermilab	39655547
osg	21886863
glow	10944336
dosar	6020063
gridunesp	2888526
alice	2061297
nova	805297
icecube	490234
ilc	330249
des	325963
sbgrid	303808
hcc	292076
ciconnect	226061
cdf	121892
QCDloop	48831
OSG-Connect	31630
nanohub	28933
HCC	18209

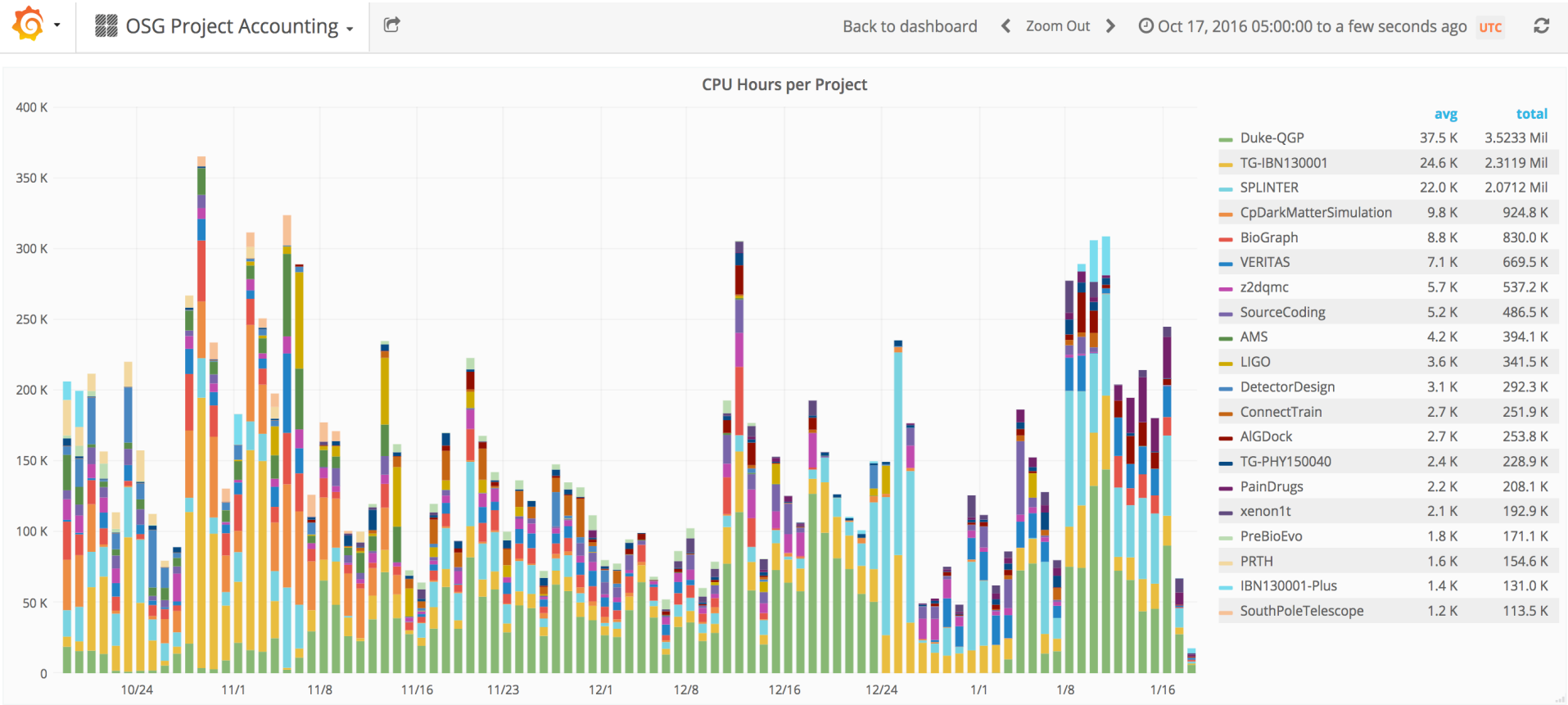
This is not opportunistic only
Excludes GaTech

Top opportunistic sites



- Syracuse again the top site

Projects: past 90 days



For reference: https://gracc.opensciencegrid.org/dashboard/db/projects-table

Recent progress with new sites

- JINR (NOvA + STAR)
  - CILogon Basic CA acceptance problem solved. A few final checks but hope to put in production in ~1 week
- Utah
  - Large increase in hours in the past month. Proxy lifetime issue fixed, apparently
- Ole Miss, Colorado, Puerto Rico
  - Seeing opportunistic glideins for OSG there now (credit to Mats)
  - Monitoring closely for additional VOs
- Will be putting out feelers to Utah, PR, Ole Miss soon (within next week)

Open Science Grid

GPUs

- Tusker (HCC) has nodes with GPUs on OSG
  - accessible via dedicated factory entry
  - Currently in use by IceCube (GLOW) and OSG VOs
  - Fermilab VO now testing. **Several experiments interested**.
- We should also express openness to sites that are willing to share GPUs
  - e.g., new BNL institutional cluster (non-ATLAS) with ~200 GPUs, Stanford
    - Spoke with Alex yesterday about BNL. Not yet ready to open up, but revisit again in a few weeks when they see how little usage there is going to be

Open Science Grid

Plans

- Continue efforts on new sites
  - Good success recently with LSJU, JINR (almost there!), Utah
  - Still work to do with CENF, Colorado, Ole Miss, Puerto Rico
- Continue to push HPC (e.g. Stampede) and GPU (Nebraska can be a good template
- Work out some outstanding accounting issues
  - Rebus doesn't like it when EGI sites are registered in OIM, "fixed" by disabling the site in OIM. Side effects on OSG reports?
  - Some sites report job start times in Epoch time (or with 0!!!), leading to wildly inflated hour totals
- **Let's take care not to get behind the curve with SL7**

Open Science Grid

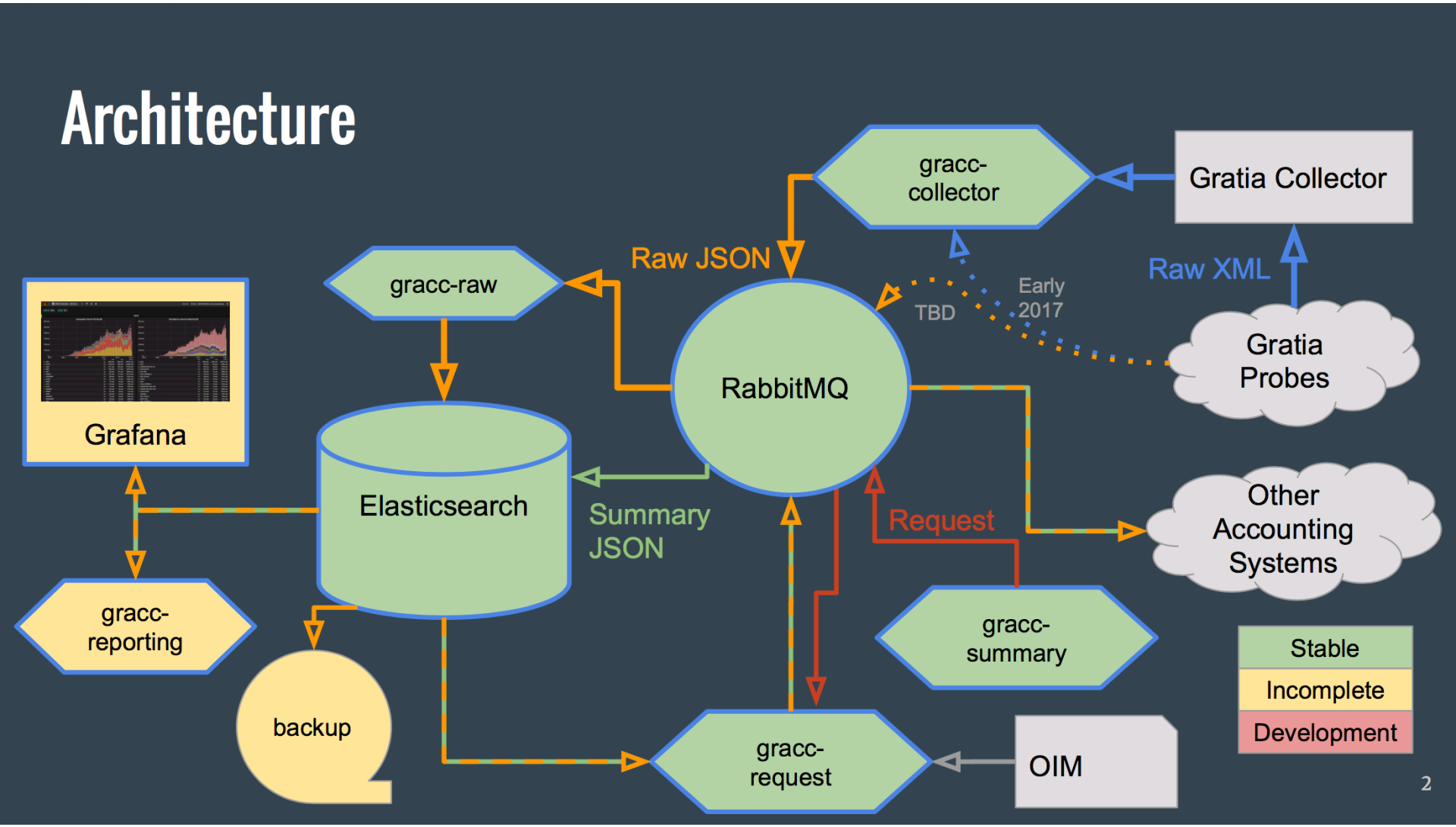
FIFE, etc. Updates

- NOvA using StashCache over CVMFS
  - Hven't been making big pushes lately but seems to be working well
  - Jobs at FNAL should be reading directly from FNAL now
- Expecting NOvA and Mu2e to be running more now (though not as much as last spring)
- KH working with DES to get some new workflows using StashCache (still in intro phase right now; stalled over holidays)

Part III: GRACC Update (Kevin)

GRACC Status

- Project meeting at Fermilab in mid-Dec. Detailed status report: https://docs.google.com/document/d/1MeyvPqii04mNEucyuvbiI5rPjD5ARzorFKiy Mc659z8/edit?usp=sharing
- Several critical items remaining for “GRACC Version 1” production release:
  - Njobs added to summary (GRACC-41)
  - Clean duplicates and “bad” records (GRACC-40)
  - Tape backups (GRACC-59)
  - Agent instability (GRACC-63)
  - Reports (GRACC-19, etc)
- V1 release by AHM; public rollout there?
- Many outstanding or in-progress features beyond critical functionality, prioritized in JIRA.



Courtesy K. Retzke

Open Science Grid

Summary

- Overall, a year of solid growth and great science
  - 1.3 billion hours!
  - Several new sites
  - Increased opportunistic running across all VOs
  - GPUs and HPC slowly coming along
- Some trends emerged
  - More opportunistic running from ATLAS and CMS (eats into OSG, GLOW, etc.)
  - Migrations to SL7 happening
  - Should work to accelerate HPC resources
  - Should continue working on adding new resources and becoming more efficient with what we have
- GRACC coming along nicely, initial release by AHM?