



Open Science Grid

# OSG Production Support

Bo Jayatilaka  
*Fermilab*

OSG Staff Retreat  
May 19, 2015

# Mission of Production Support

---

1. Growing the opportunistic pool of the OSG
  - Add sites to opportunistic sphere
  - Improve on methods and technology to access opportunistic resources more efficiently
2. Help experiments (and VOs) access the opportunistic sphere
  - *e.g.* find the right VO for experiments (OSG for PHENIX, Fermilab for the FNAL IF experiments)
3. Be a catalyst for projects and technologies that help VOs run opportunistically
  - *e.g.* XRootD-based StashCache system for distributed storage access



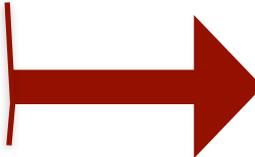
# Staffing

## Production Support

## User Support

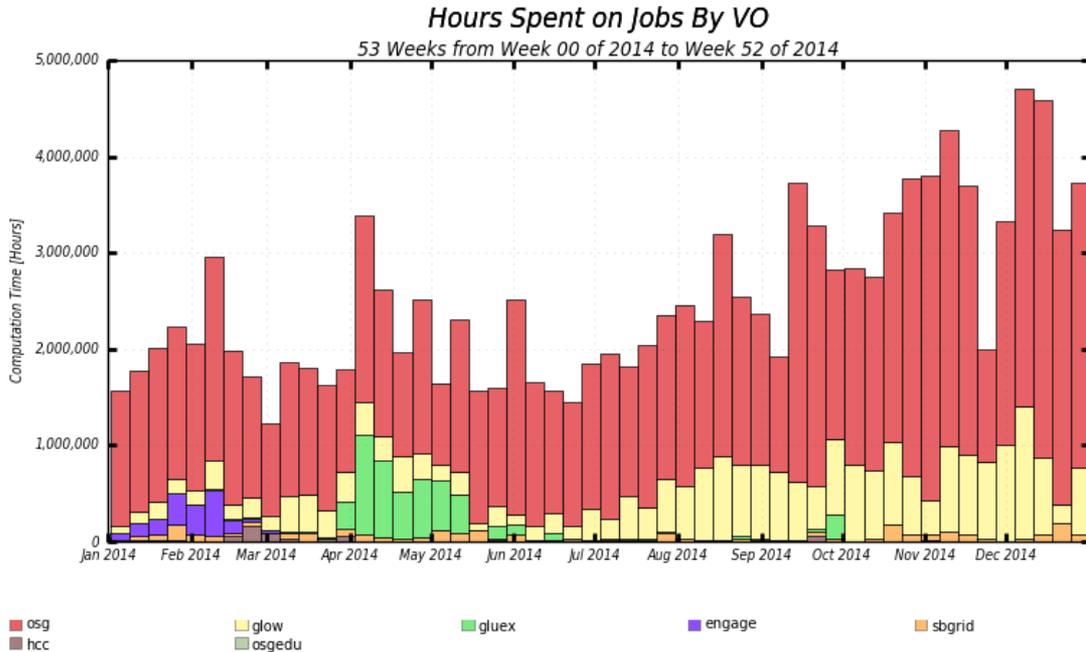
<i>Name</i>	<i>Institution</i>	<i>FTE</i>
Alex Zaytsev	BNL	0.10
Marko Slyz	FNAL	0.60
Tanya Levshina	FNAL	0.25
Bo Jayatilaka	FNAL	0.75
Chander Sehgal	FNAL	<i>ex officio</i>
Robert Illingworth	FNAL	0.50
Juan Morales	FNAL	1.00

<i>Name</i>	<i>Institution</i>
Alex Zaytsev	BNL
Marko Slyz	FNAL
Tanya Levshina	FNAL
Bo Jayatilaka	FNAL
Chander Sehgal	FNAL
Mats Rynge	ISI
Emelie Harstad	Nebraska

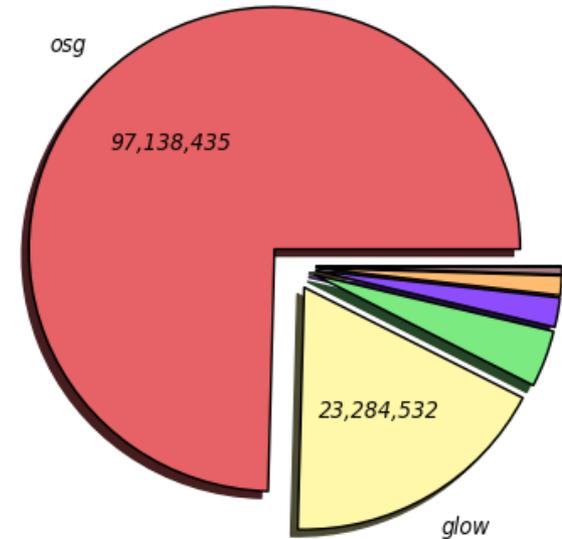


**Campus  
Grids**

# Opportunistic Computing in 2014



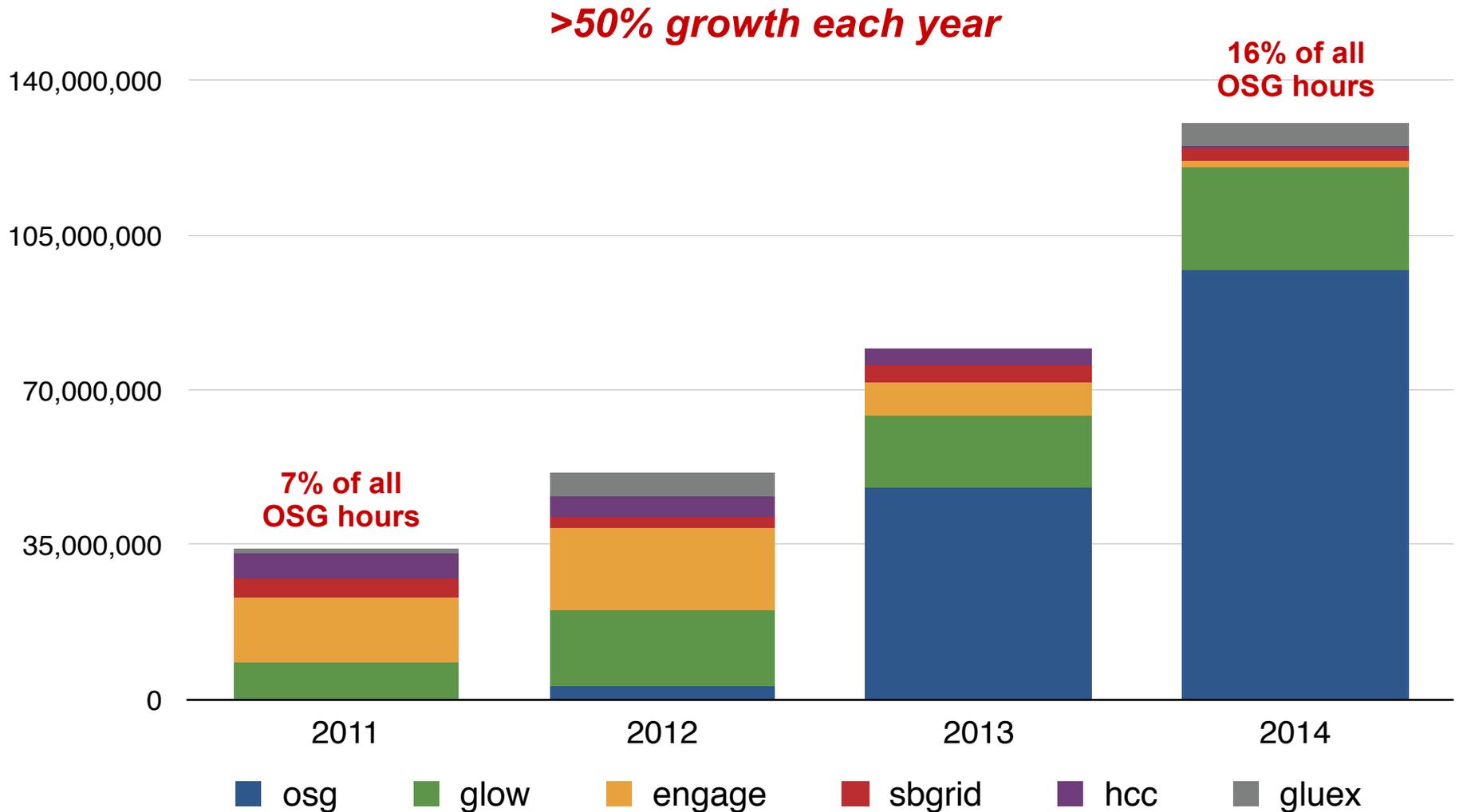
Maximum: 4,699,677 Hours, Minimum: 45,027 Hours, Average: 2,457,438 Hours, Current: 3,728,689 Hours



- Primarily opportunistic VOs (osg, glow, gluex, engage, sbgrid, hcc) received **130M wall hours** in 2014
  - 16% of all OSG hours

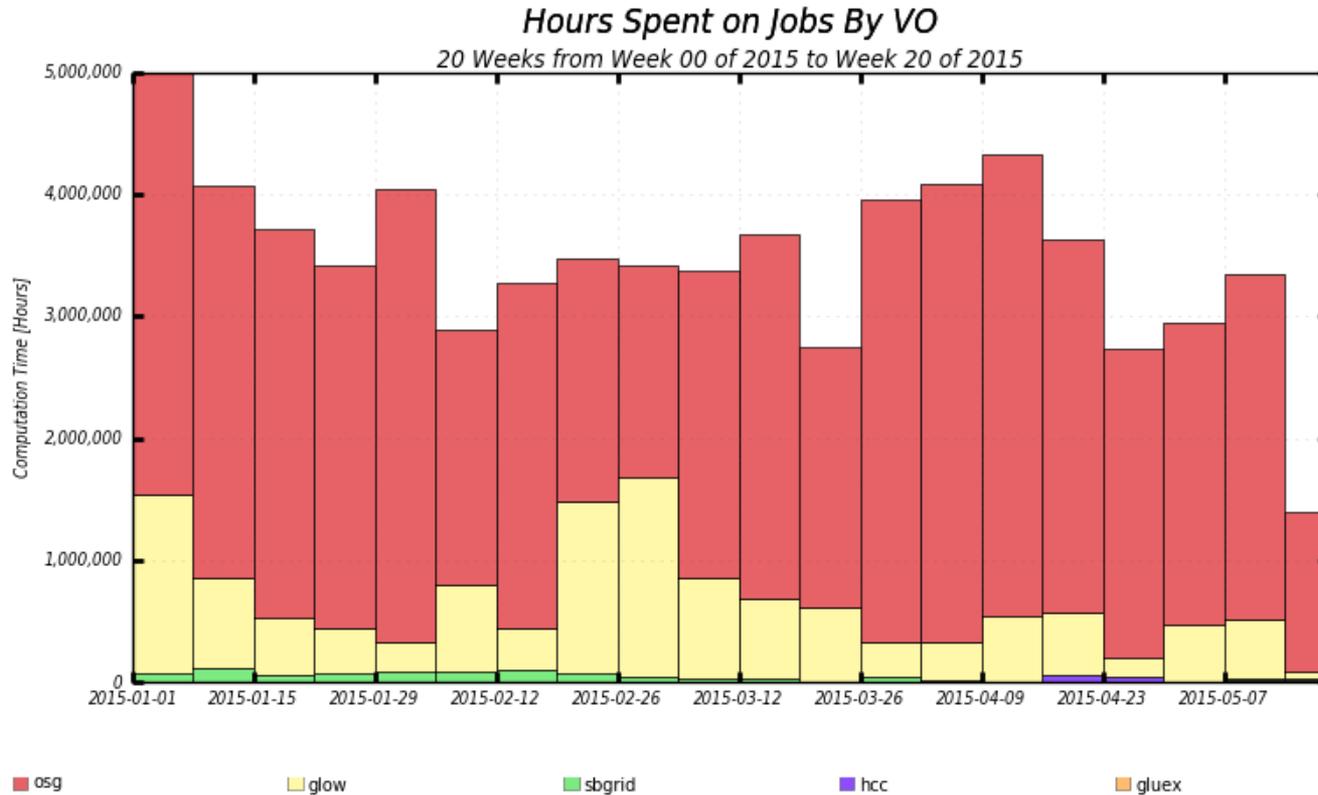
Year 3 goal for user support

# Growth of Opportunistic VOs





# 2015 so far



Maximum: 4,998,586 Hours, Minimum: 1,397,640 Hours, Average: 3,479,637 Hours, Current: 1,397,640 Hours

- 70M wall hours
  - 22% of all OSG hours



# Year 3 goal: Integrate Fermilab IF

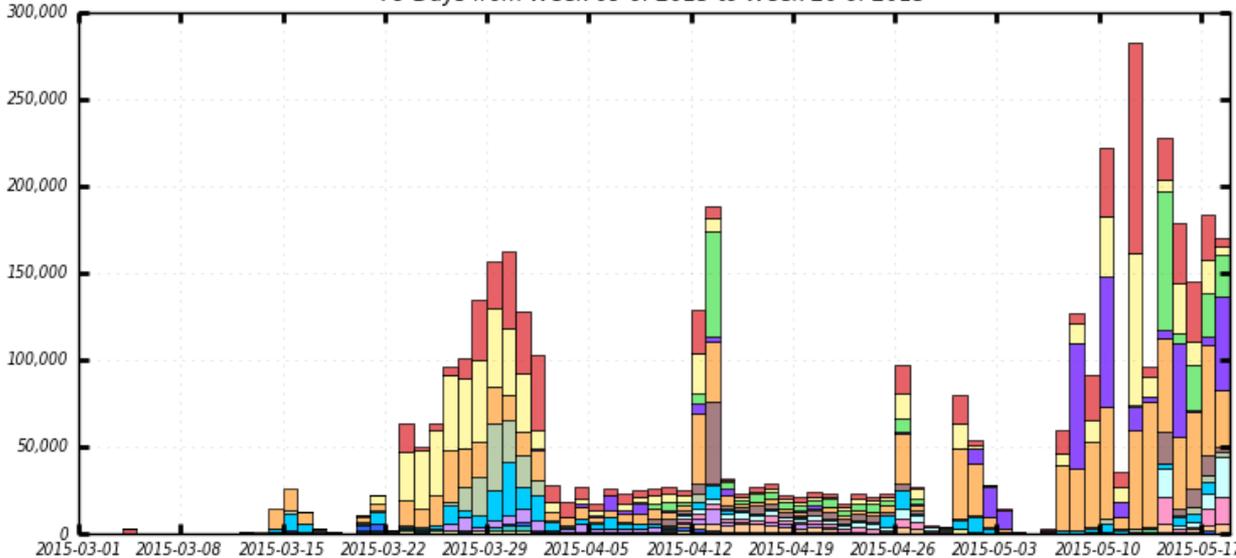
---

- Use the “Fermilab” VO rather than individual experiment VOs when practical
  - Sites only have to deal with enabling one VO
  - Group accounting within VO
  - Individual experiment VOs still available if a site wants to support subset of experiments
- Enabled Fermilab VO at most sites that support other opportunistic VOs
- Worked with sites to turn on access to Fermilab VO and FIFE support to test functionality for a range of experiments

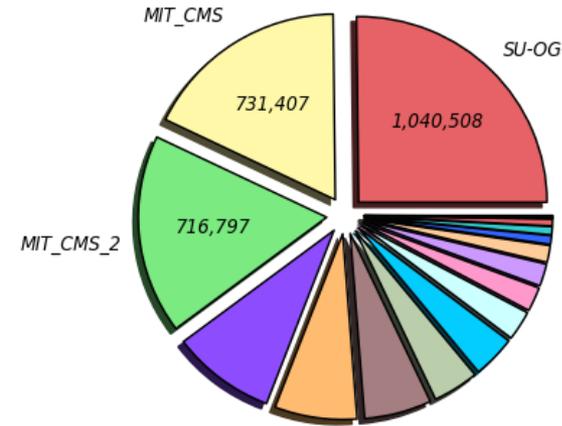


# Recent success: mu2e

Hours Spent on Jobs By Facility  
79 Days from Week 09 of 2015 to Week 20 of 2015



Wall Hours by Facility (Sum: 4,131,281 Hours)  
12 Weeks from Week 09 of 2015 to Week 20 of 2015



- MIT\_CMS\_2
- CIT\_CMS\_T2
- MWT2\_CE\_UIUC
- MWT2
- MIT\_CMS
- GLOW-OSG
- Nebraska
- UCSDT2-D
- CIT\_CMS\_T2B
- Tusker-CE1
- Crane-CE1
- UCSDT2-C
- USCMS-FNAL-WC1-CE3
- red-gateway2
- BNL\_ATLAS\_1
- OSC\_OSG\_CE
- SU-OG-CE
- red-gateway1
- FNAL\_CDFOSG\_2

Maximum: 282,500 , Minimum: 17.52 , Average: 52,294 , Current: 170,065

- Large production campaign through ~September
- Over 4M non-Fermigrid hours since March
- Some site-specific hiccups but support being utilized



# mu2e: (ongoing) lessons

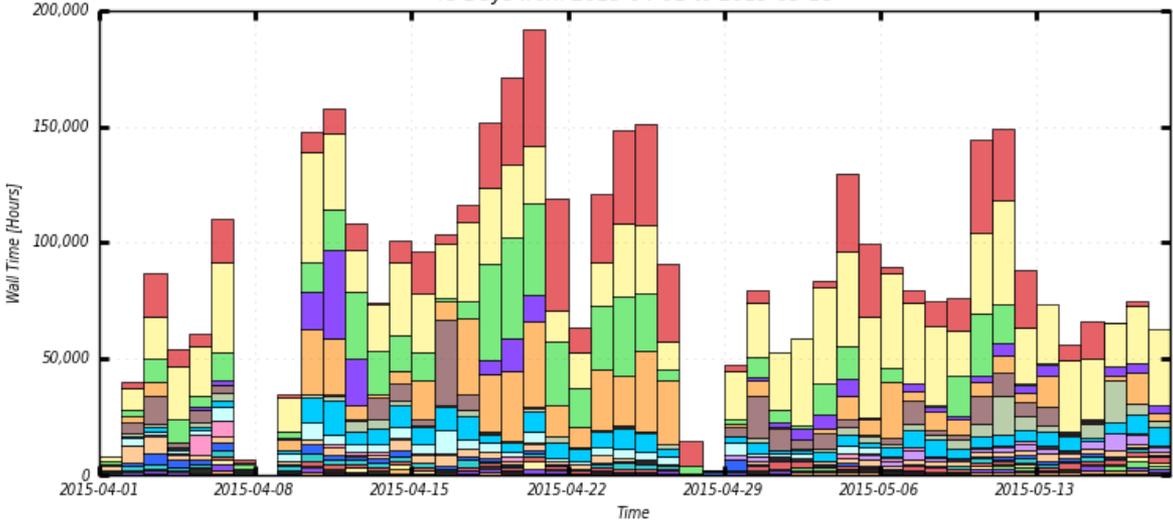
---

- A case study in the “mid sized” experiment using OSG resources
- Support came from GOC and FIFE (Fermilab)
- Issues dealt with as they came up
  - Site black holes
  - Libraries/products not being available at sites (ups problems in particular)
  - Sites simply not supporting the VO (or thinking they did but didn't in reality)
- Effort particularly reliant on one mu2e physicist willing to help diagnose problems on the user end
  - Will be harder with experiments/VOs that don't have that kind of initiative on their side
- **Should reduce effort for other experiments using FIFE tools to get on OSG**



# Recent success: sPHENIX

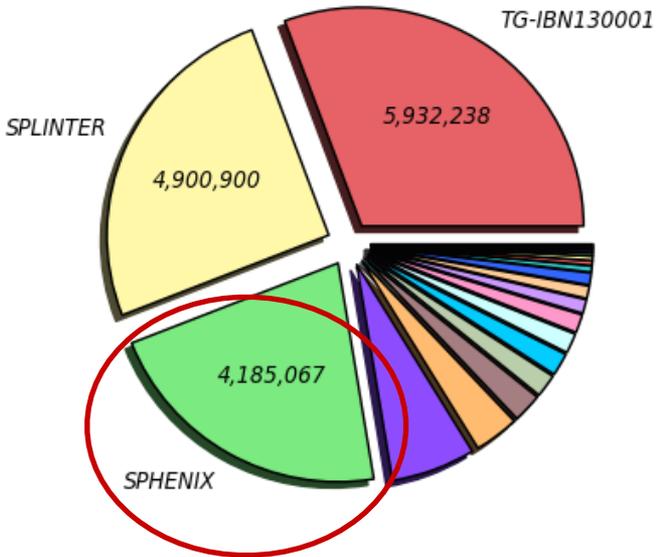
Daily Hours By Project and Site  
48 Days from 2015-04-01 to 2015-05-18



- SPHENIX , CIT\_CMS\_T2
- SPHENIX , SU-OG
- SPHENIX , Nebraska
- SPHENIX , USCMS-FNAL-WC1
- SPHENIX , MIT\_CMS
- SPHENIX , FNAL\_FERMIGRID
- SPHENIX , UConn-OSG
- SPHENIX , UCSDT2
- SPHENIX , BNL-ATLAS
- SPHENIX , Tusker
- SPHENIX , GridUNESP\_CENTRAL
- SPHENIX , MWT2
- Other
- SPHENIX , CIT\_HEP\_CE
- SPHENIX , TTU-ANTAEUS
- SPHENIX , GLOW
- SPHENIX , NUMEP-OSG
- SPHENIX , UCD
- SPHENIX , FLTECH
- SPHENIX , cinvestav

Maximum: 191,837 Hours, Minimum: 283.78 Hours, Average: 87,188 Hours, Current: 62,557 Hours

Wall Hours by VO (Sum: 19,358,709 Hours)  
48 Days from 2015-04-01 to 2015-05-18



- Studies for upgraded PHENIX detector at BNL (~5 trillion collisions)
- Expected to continue through the end of this month
  - Able to ramp up rapidly - #3 project on OSG in that time
  - Support primarily via BNL (Zaytsev)



# Visualization tools: treemaps



[http://fermicloud033.fnal.gov:8100/gratia/xml/osg\\_hours\\_efficiency\\_tree\\_map\\_by\\_vo\\_project\\_facility](http://fermicloud033.fnal.gov:8100/gratia/xml/osg_hours_efficiency_tree_map_by_vo_project_facility)

# Upcoming: SDSC Comet

---

- New HPC Installation at SDSC: Comet
  - “HPC for the 99%” (see M. Norman’s [talk at 1/14/15 Council meeting](#))
  - 45K cores, online now
  - Targeted user base similar in profile to OSG/DHTC users (jobs that don’t require massive MPI and tend to be shorter)
- Both ATLAS and CMS have had success using XSEDE sites like Comet recently
- Plan: implement OSG-CE in front of Comet HPC sites
  - Allows users with OSG experience who also have a Comet allocation to run on both easily using the same code/tools



# “OSG-HPC-CE”

- 
- Project defined as a CE interfacing to HPC installations (starting with Comet)
    - Plenty of discussion about this already
  - Also need to define the goals of the project
    - Allow users to use HPC installations using (familiar) OSG tools
  - Issues (many already discussed by Brian yesterday)
    - HTCondor-CE to support multiple PBS backends
    - Dealing with project accounting
      - Including at the factory level— gWMS is (not yet) project-aware
    - Resource provisioning for users with multiple allocations(+opportunistic access)
  - Whitepaper outlining goals and rough outline expected within the next two weeks



# New site integration

---

- No coherent procedure for new site integration
- User support used to be initial point of contact for interested sites and follow the process throughout
  - Was this by design or a result of the front door?
- Most of the heavy lifting was done by Software (and still is)
- Does there still need to be a designated facilitator role?
  - Should it come from Production Support or elsewhere?
  - I **have** done this for FIU have **not** for UWa or Clemson



# Current concerns/work

---

- Can we grow any more?
  - Current accounting of “opportunistic” is incomplete
    - New options for gratia to truly show opportunistic hours (run on sites not owned by VO) being tested and soon deployed
  - Rough estimate is ~17M/month for the past 3 months
  - Are there structural limitations? Test proposed at last council meeting to be carried out
- New site integration
  - FIU in testing now
- StashCache testing for use cases outside of OSG Connect
  - Timing tests for NOvA flux files comparing dCache and Stash done at Fermilab and at Nebraska
    - To be one of the initial customers of Stashcache



# Year 4 plans

---

- Continue to understand availability and limits of opportunistic pool
  - Maintain over 150M hours/year
- Continue to expand opportunistic pool where possible
  - May be at a point where ET's help is needed here
- Coordinate effort to implement OSG-HP-CE
  - Demonstrate functionality of OSG interface to Comet (CMS first?)
- Expand access to opportunistic resources across VOs
  - Demonstrate functionality of StashCache for NOvA and other Fermilab VO experiments



Open Science Grid

# Backup



# 10 Largest Sites in April

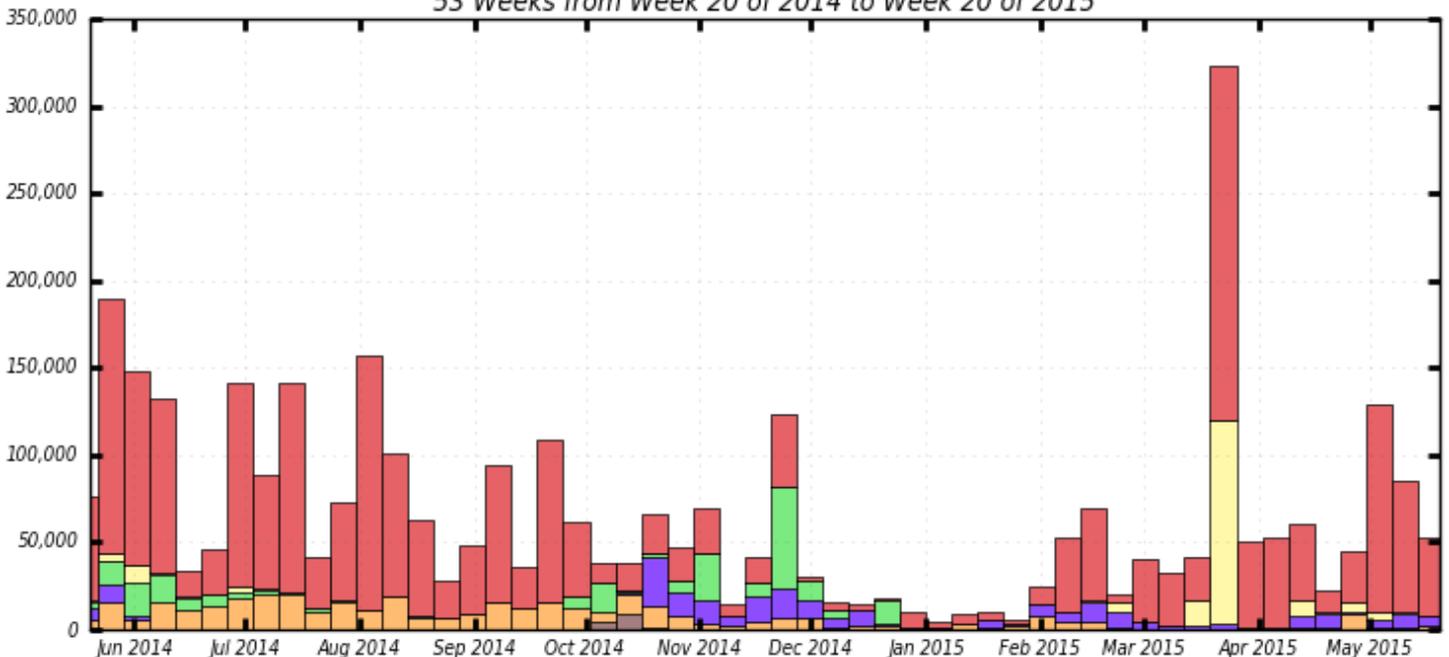
Site	Total	Percent		Opportunistic VOs			Total Opp
		Opportunistic	glow	hcc	osg	sbgrid	
<i>Total (all sites)</i>	71,055,170	22%	1,624,498	105,123	13,983,722	101	15,713,444
CIT_CMS_T2	3,740,081	65%	433,164	5,718	1,994,352	10	2,433,244
Nebraska	3,208,588	75%	327,814	2,690	2,091,254		2,421,758
SU-OG	2,280,215	89%	104,859	54,777	1,878,958	2	2,038,596
MIT_CMS	2,979,428	58%		1,945	1,720,257	26	1,722,228
Tusker	1,443,098	84%	57,527	461	1,147,567		1,205,555
Purdue-Hadoop	1,445,378	50%	65,955	904	655,657	14	722,530
FNAL_FERMIGRID	7,836,830	9%	90,873	227	592,046		683,146
USCMS-FNAL-WC1	7,462,100	9%	82,532	2,035	570,207		654,774
UCSDT2	1,452,613	41%	100,008	2,027	493,140	31	595,206
MWT2	5,205,889	10%	137,281	6,405	388,295		531,981



# CMS opportunistic usage

### Opportunistic Wall Hours by Site

53 Weeks from Week 20 of 2014 to Week 20 of 2015



- Tusker-CE1
- Crane-CE1
- Sandhills
- Baylor-Tier3-CE
- NYSGRID\_CORNELL\_NYS1
- NWICG\_Earth
- GridUNESP\_CENTRAL
- BNL\_ATLAS\_3
- BNL\_ATLAS\_5
- WT2
- FNAL\_GPGRID\_1
- OU\_OSCER\_ATLAS
- BNL\_ATLAS\_1
- gpce02.fnal.gov
- BNL\_ATLAS\_2
- BNL\_ATLAS\_4

Maximum: 323,380 , Minimum: 4,531 , Average: 65,477 , Current: 52,858

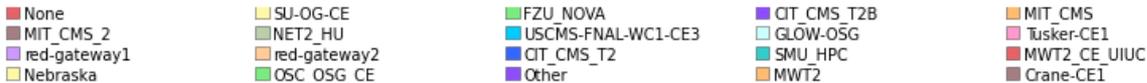
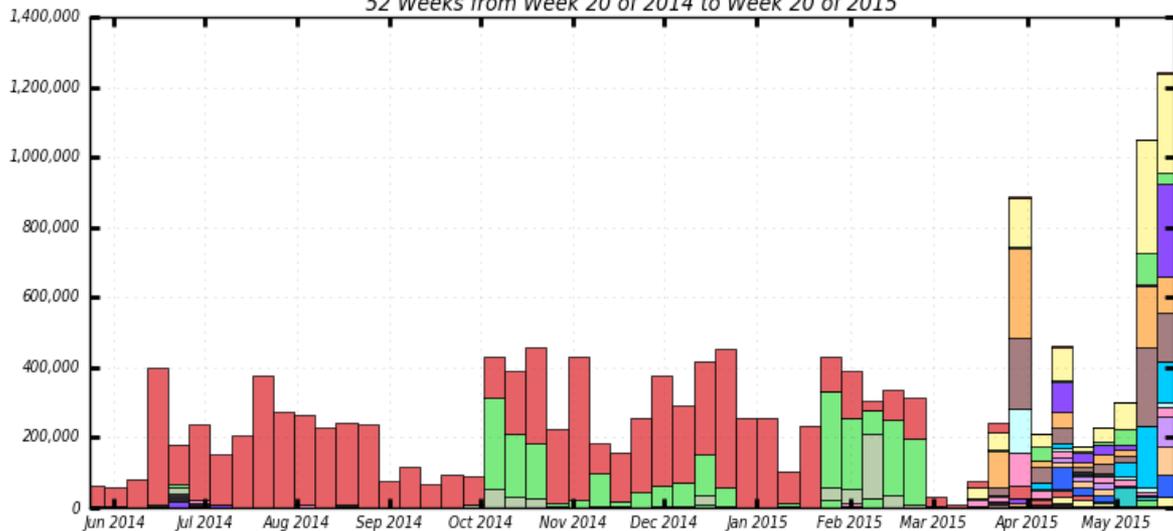
## 3.5M hours in the past year



# Fermilab\* opportunistic usage

### Hours Spent on Jobs By Facility

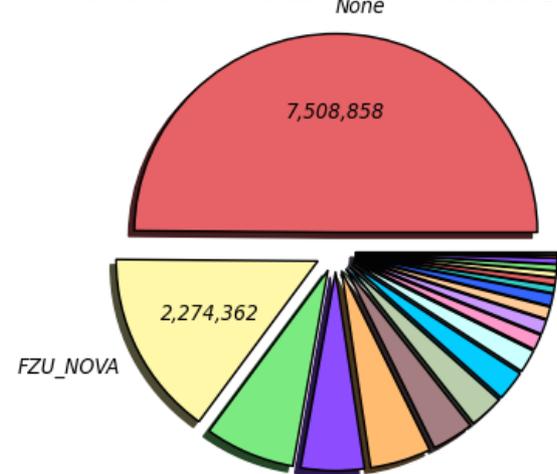
52 Weeks from Week 20 of 2014 to Week 20 of 2015



Maximum: 1,242,046 , Minimum: 7,089 , Average: 289,591 , Current: 1,242,046

### Wall Hours by Facility (Sum: 15,058,761 Hours)

52 Weeks from Week 20 of 2014 to Week 20 of 2015



- \*nova and mu2e