



Open Science Grid

The Open Facility Access to DHTC for Science

August 19, 2014

Chander Sehgal
OSG User Support





Overview

Goal: Grow access to DHTC computing for US researchers (who are not already part of an existing OSG community)

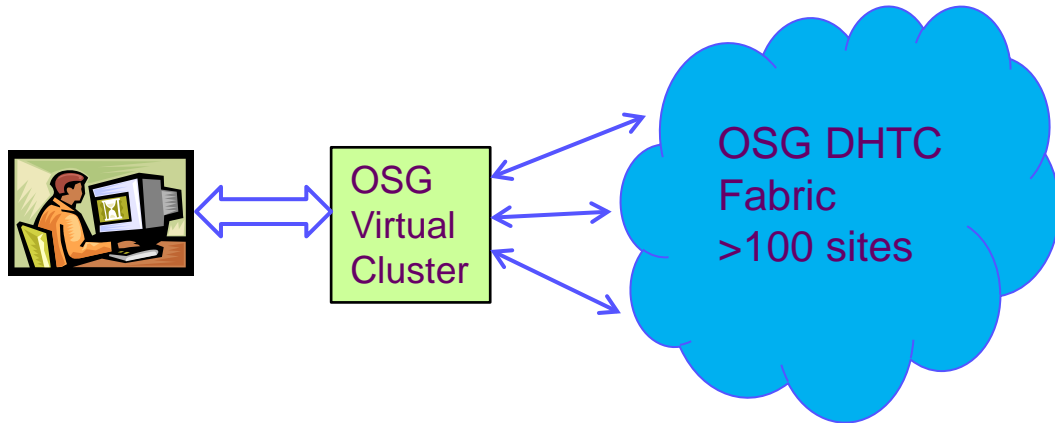
Strategy: Make it easier for users & communities to send scientific computing into OSG. Enable multiple entry paths that best fit the user's needs and environment.

Plan: Build on existing DHTC principles and technology of OSG; use the osg VO (which only exists to harvest opportunistic cycles) to connect researchers to OSG production fabric

1. OSG as level 2 SP in XD
2. Login Nodes for "OSG Direct" PIs/Users
3. Flocking path for community/campus access



OSG as XD Service Provider



In 12 months ending June 30, 2014

- 8 users: < 1K hours
 - 8 users: >1K and <100K hours
 - 4 users: >100K and <1M hours
 - 2 users: > 1M hours
- Total = ~32M wall hours

Service Introduced April 2012

- 2M hours allocated every quarter
- Established interfaces with allocations, security, operation and other teams in XSEDE
- As needed, interactions with ECSS and Campus Champions

OSG approach to XRAC allocations

OSG commits 2M hours per quarter to XRAC that we are “sure” we can deliver **but we strive to deliver more based on opportunistic availability**. We value the vetting of science performed by XRAC and for each request against OSG, we monitor:

1. SUs requested
2. SUs recommended
3. SUs awarded (often constrained by the 2M SUs committed by OSG every quarter)

OSG Treatment of XRAC Allocations:

- Provide SUs awarded (#3 above) in priority mode from the opportunistic pool
- When the SUs awarded are exhausted, continue to provide access to reach SUs recommended (#2 above) -- competing equally with other opportunistic users in OSG
- When end-date of award is reached, we terminate access **

** We have contemplated but not decided if we should terminate access when a PI reaches “SUs recommended”



OSG-Direct and OSG-Connect

OSG-Direct login node

- Provide DHTC access to researchers who know about and contact OSG directly
- Re-use OSG-XD setup
- Manual provisioning of logins, etc.

Available June 2012

OSG-Connect Service

- Provide DHTC access to campus researchers based on their existing Campus Identity
- Uses flocking method to access OSG DHTC
- Automated login creation

Available August 2013





Sending DHTC Jobs from Campus to OSG

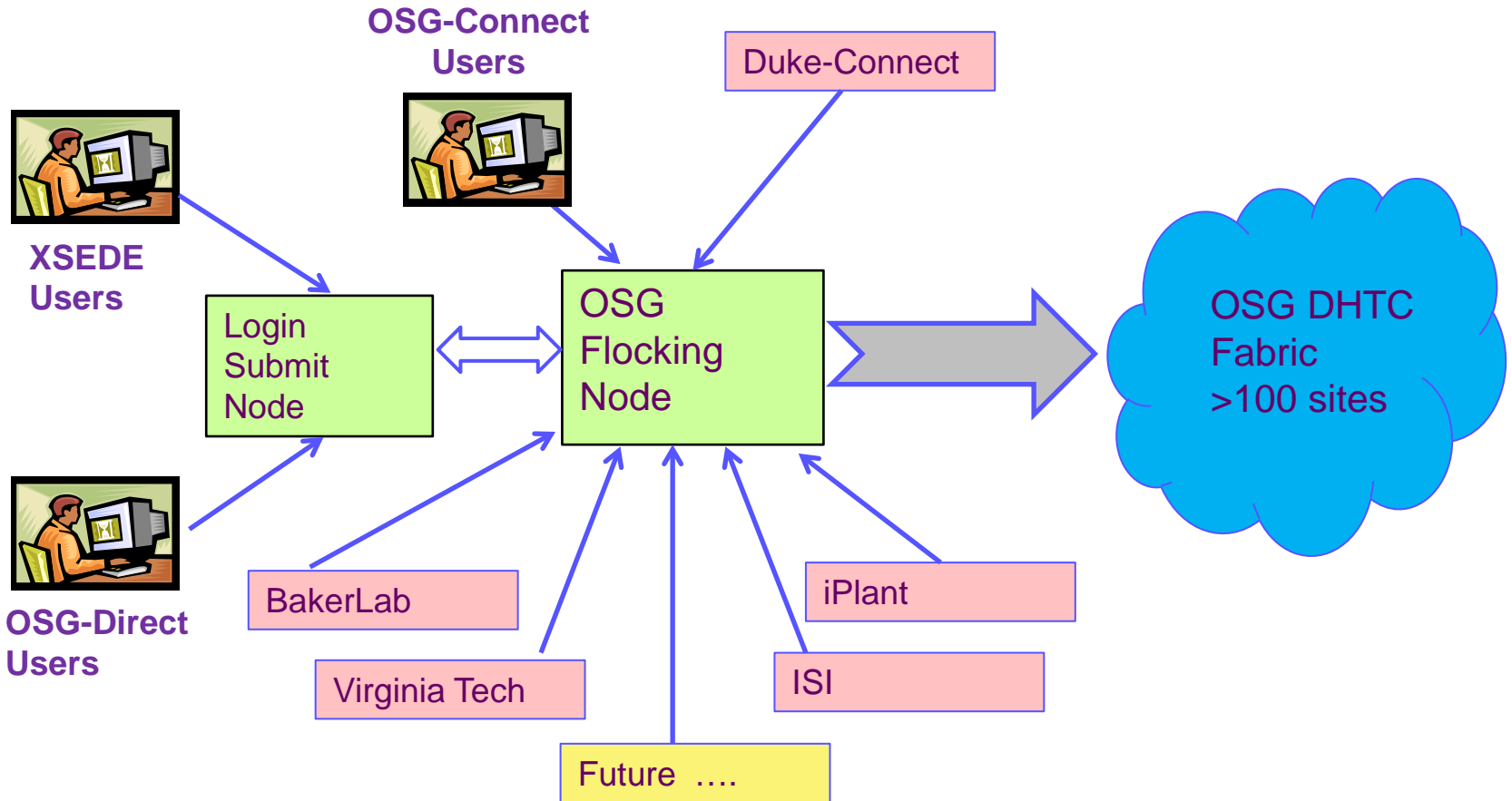
Improve ease of connecting to OSG production fabric for research groups who want to use DHTC; provide a “bulk” connection to the OSG fabric

HTCondor Flocking Connection

- Login submit node administered by local campus staff (they know and support their users)
- Simplified connection to OSG via a flocking node (removes much of the complexity of being a grid expert)
- OSG staff handles harvesting opportunistic resources from >100 sites and managing grid operations



Easier "On-Ramp" to the OSG DHTC Fabric

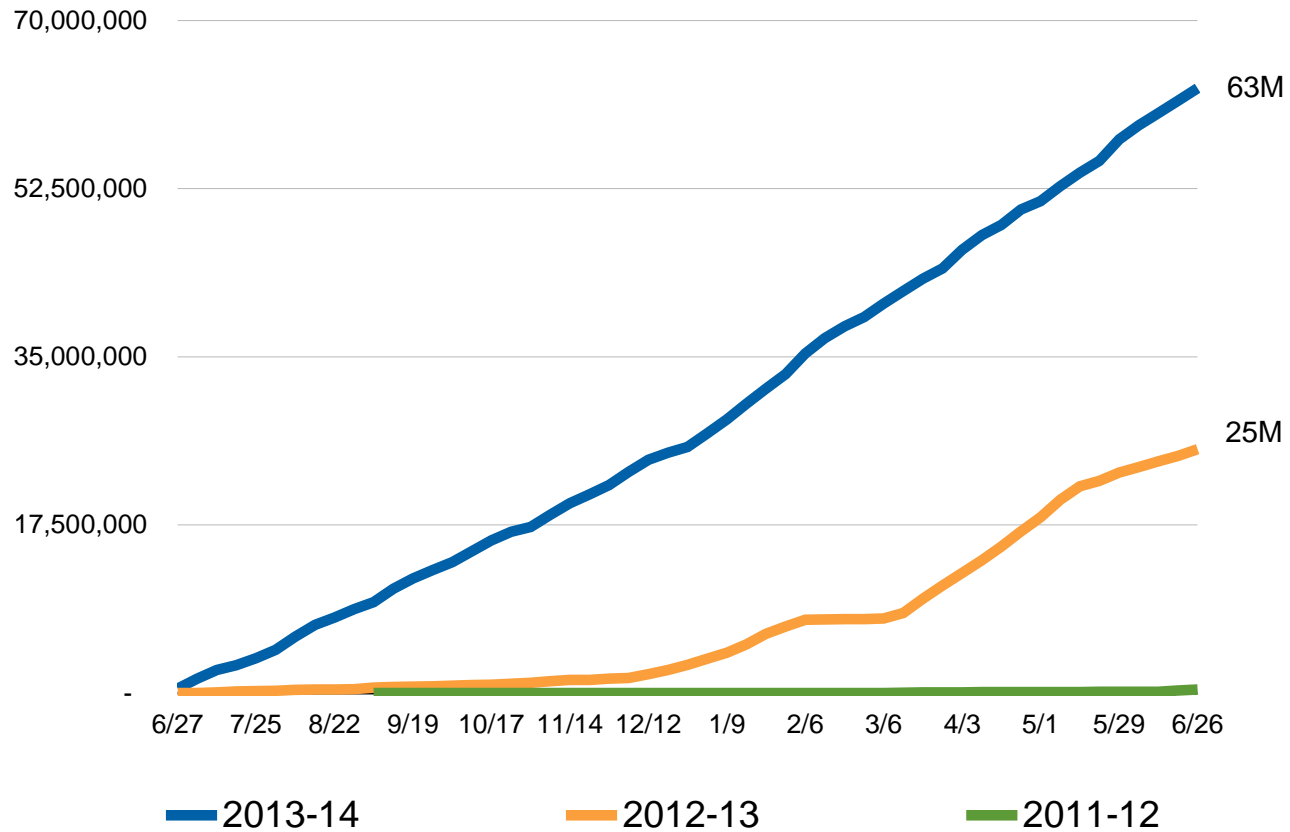


Access operates under the osg VO using glideinWMS pilot overlay system; currently engineered to access ~20K cores (on an average day total OSG usage is ~100K cores)

OSG VO Usage Trend

Last 3 Years (July – June)

OSG VO Wall Hours



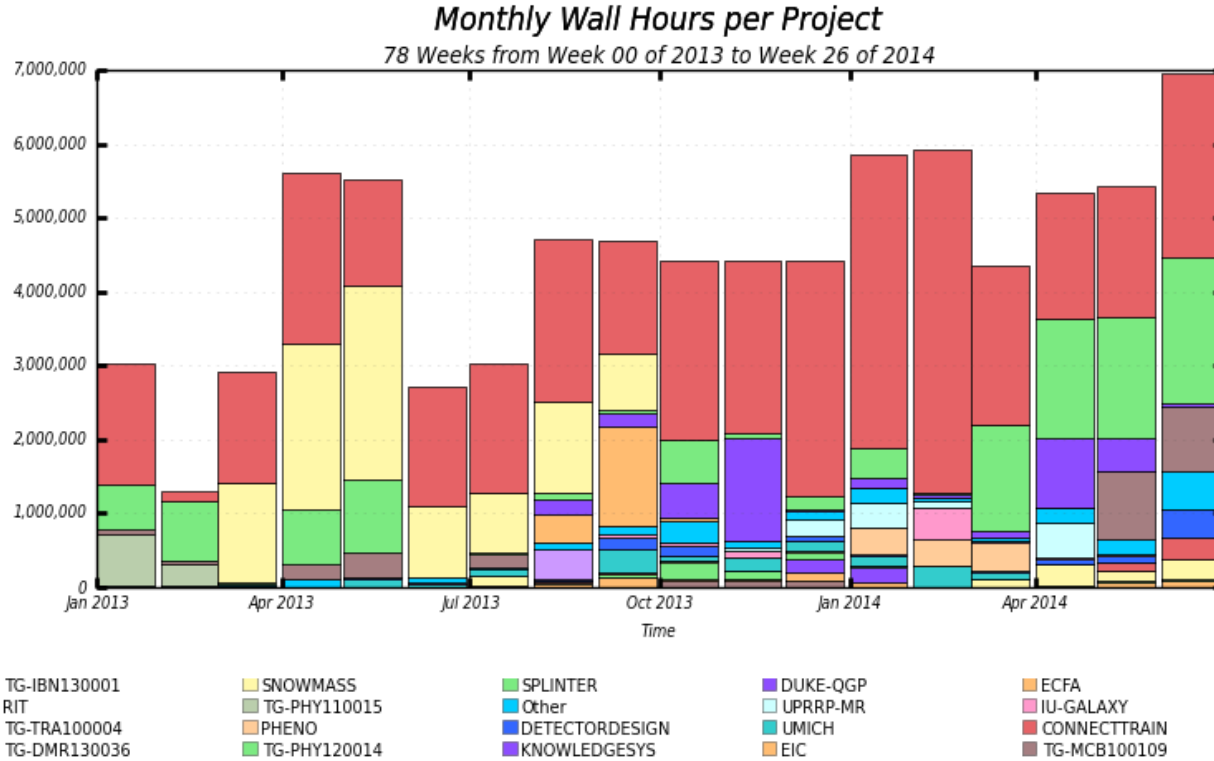
OSG DHTC Delivered by Access Method

July 2013 to June 2014

Access Method	Wall Hours
OSG-XD login node	33M
OSG-Direct login node	25M
OSG-Connect login node	2M
Flocking Connections	2M
Other (includes UCSD)	1M
Total	63M



DHTC Access via OSG login nodes



Maximum: 6,946,655 , Minimum: 0.00 , Average: 4,241,526 , Current: 6,946,655

	Q1 2013	Q2 2013	Q3 2013	Q4 2013	Q1 2014	Q2 2014
Users	10	13	30	31	35	43
Wall Hours	7,264,737	13,827,435	12,413,717	13,248,882	16,124,899	17,709,279



Science using OSG Open Facility

7/2013 – 6/2014

Field of Science	Wall Hours
Biological Sciences	30,160,878
Medicine	8,084,640
High Energy Physics	5,788,677
Nuclear Physics	3,973,122
Bioinformatics	2,078,561
Computer Science	2,016,424
Microbiology	1,280,064
Materials Research	1,002,362
Medical Imaging	990,332
Astro-Physics	780,080
Accelerator Physics	564,486
Educational Psychology	383,182
Molecular Biosciences	320,691
Chemistry	262,241
Cross-disciplinary Evolution	240,289
Information, Robotics, and Intelligent Systems	80,940
Neuroscience	67,048
Condensed Matter Physics	60,287
Plant Biology	36,976
Ocean Sciences	34,556
Neutrino Physics	18,823
Atmospheric Sciences	3,379
Other	820,509
Training	449,752
Total	59,498,299

Registering Projects in OSG

My Projects

[+ Add New Project](#)


You currently have no project that you are authorized to edit

Projects


AIGDock	AMFORA	aprime	atlas-org-duke	atlas-org-fresnostate
atlas-org-illinois	atlas-org-indiana	atlas-org-uchicago	AtlasConnect	BioMolMach
BioStat	BNLPET	CometCloud	CompChem	compcomb
CompNeuro	ConnectTrain			
duke-4fermion	duke-campus			
EIC	EvoTheory			
GRASP	gridsgenomes			
IU-GALAXY	KnowledgeLab			
NRELMatDB	NWChem			
OSGOpsTrain	P0-LBNE			
Proteomics	ProtFolding			
SBGrid	scicomp-analytics			
SouthPoleTelescope	SoyKB			
Swift	TG-ASC130043			
TG-CCR130001	TG-CDA080011			
TG-DEB140008	TG-DMR130036			
TG-MCB090183	TG-MCB090174			
TG-OCE130029	TG-PHY110015			
TG-TRA100004	TG-TRA120014			
UCSDPhysAstroExp	UCSDPhysAstroTheo			
UNC-RESOLVE-photometry	unlcpass			

Project / SouthPoleTelescope

SouthPoleTelescope

Name	SouthPoleTelescope
Description	The South Pole Telescope (or SPT) is a new telescope deployed at the South Pole that is designed to study the Cosmic Microwave background. Constructed between November 2006 and February 2007, the SPT is the largest telescope ever deployed at the South Pole. This telescope provides astronomers a powerful new tool to explore dark energy, the mysterious phenomena that may be causing the universe to accelerate.
Organization	University of Chicago
Department	Kavil Institute for Cosmological Physics
Sponsor Campus Grid	OSG Connect
Principal Investigator	 John Carlstrom
Field Of Science	Astrophysics

OIM Version 3.32 | [Report Bugs](#) | [Privacy Policy](#)

 Copyright 2011 The Trustees of Indiana University - Developed for Open Science Grid

<https://oim.grid.iu.edu/oim/project>

The Open Facility - Today

Accomplishment

Currently providing ~60M hours per year of DHTC to ~60 US researchers; we believe the Open Facility is currently demand limited

Our Challenges

1. Identify more users who can benefit from DHTC
2. Grow the opportunistic pool; how much can be harvested and delivered?
3. Make the system easier to use



Reach more Users

Identify more researchers who may benefit from access to OSG DHTC

- Provide more tutorials and documentation on how to structure jobs for use in OSG and how to submit jobs
- Grow number of XRAC requests against OSG
 - Contact each new startup/XRAC allocation to provide support
 - Assist expiring startup/XRAC allocations with writing new XRAC proposal
 - Use XSEDE Campus Champions to get the OSG DHTC message to more US campus researchers
- Establish partnership with ACI-REFs to provide 2nd layer support for them as they include OSG in their solutions for local campus researchers

>> This is a hard problem and we are open to additional suggestions

Grow the Opportunistic Pool

Goal: Grow opportunistic pool toward 20% of total OSG usage

- Modernize hardware at “on-ramp” to access ~30,000 cores
- Prototype methods to recognize sites that contribute opportunistic cycles; trial reports and recognition methods (e.g. show the contribution to science) that incentivize sites
- Identify sites that can contribute more efficiently and help them with configuration changes

Key Challenge: The opportunistic eco-system in a grid with >100 sites is complex and not well understood; we are working to:

- research the dynamics of available cores at sites;
- identify unused capacity and limits;
- develop recommendations to VOs on how to access opportunistic cycles more efficiently; and
- implement those recommendations for vo=OSG.




Improve the User Experience

- Improve data delivery to sites coordinated with jobs; OSG Public storage service (based on iRODS) is being transitioned to the GOC at Indiana (for vo=osg)
- Help users adapt commonly used analysis codes (e.g. NAMD) to run in the OSG DHTC
- Identify and install commonly used versions of software into a “library” via the OASIS (using CVMFS) – facilitate the delivery of software to worker nodes
- Address other user needs as they arise



Workplan Summary

- Initial Focus was to build DHTC Access capabilities for Users not already affiliated with an OSG community (1.8 FTE) 
- Now adding focus and staff to
 - Research and grow the opportunistic pool inside OSG (0.7 FTE)
 - Improve the User Experience and Identify new Researchers who can use OSG DHTC (0.5 FTE)



Backup Slides



OSG User Support Team

Name	Institution	FTE	
Mats Rynge	USC-ISI	0.50	
Emelie Harstad	Nebraska	0.50	Added April 2014 [2]
Marko Slyz	FNAL	0.60	
Tanya Levshina	FNAL	0.25	
Bo Jayatilaka	FNAL	0.75	Added March 2014 [1]
Alex Zaytsev	BNL	0.10	
Chander Sehgal	FNAL	0.30	Area Coordinator
TOTAL		3.00	

[1] Added to focus on growing the opportunistic pool

[2] Added to help improve the user experience and identify new users on Campuses



OSG Hosted DHTC Usage Details

		Q1 2013	Q2 2013	Q3 2013	Q4 2013	Q1 2014	Q2 2014
OSG-XD	Users	4	7	12	11	12	10
	Hours	4,329,339	5,559,332	6,233,763	8,826,166	10,997,878	6,776,372
OSG Direct	Users	6	6	11	10	11	14
	Hours	2,935,397	8,268,103	6,103,660	3,957,849	4,748,166	9,863,063
OSG-Connect	Users	--	--	7	12	13	20
	Hours	--	--	76,294	464,867	378,855	1,069,845
Total	Users	10	13	30	33	36	44
	Hours	7,264,736	13,827,435	12,413,717	13,248,882	16,124,899	17,709,280



XD Research Allocations

7/1/13-6/30/14

Project Name	PI	Institution	Field of Science	Wall Hours	Allocation	Used
TG-IBN130001	Donald Krieger	University of Pittsburgh	Biological Sciences	30,093,013	3,750,000	802%
TG-DMR130036	Emanuel Gull	University of Michigan	Materials Research	1,002,361	1,000,000	100%
TG-PHY120014	Qaisar Shafi	University of Delaware	Physics	584,567	1,200,000	49%
TG-MCB100109	Lillian Chong	University of Pittsburgh	Molecular Biosciences	264,362	500,000	53%
TG-CHE130103	Jeremy Moix	Massachusetts Institute of Technology	Chemistry	61,405	100,000	61%
TG-MCB090163	Michael Hagan	Brandeis University	Molecular Biosciences	56,314	154,000	37%
TG-ATM130015	Phillip Anderson	University of Texas at Dallas	Atmospheric Sciences	3,379	200,000	2%
TG-PHY110015	Pran Nath	Northeastern University	Physics	37	1,000,000	0%
Total				32,065,437	7,904,000	

>>Note: TG-DMR130036 just got another allocation for 2M hours starting July 1st.

XD Startup and Campus Champion Allocations 7/1/13-6/30/14

Project Name	PI	Institution	Field of Science	Wall Hours	Allocation	Used
TG-TRA100004	Andrew Ruether	Swarthmore College	Other	417,351	200,000	209%
TG-CHE130091	Paul Siders	University of Minnesota; Duluth	Chemistry	134,193	100,000	134%
TG-IRI130016	Joseph Cohen	University of Massachusetts; Boston	Information; Robotics; and Intelligent Systems	80,940	200,000	40%
TG-OCE130029	Yvonne Chan	University of Hawaii; Manoa	Ocean Sciences	34,547	33,120	104%
TG-TRA120014	Pol Llovet	Montana State University	Cross-Disciplinary Activities	19,479	200,000	10%
TG-IBN130008	Jorden Schossau	Michigan State University	Biological Sciences	16,857	200,000	8%
TG-DEB140008	Robert Toonen	University of Hawaii; Manoa	Environmental Biology	4,106	31,000	13%
TG-MCB120070	Joseph Hargitai	Albert Einstein College of Medicine	Molecular Biosciences	378	100,000	0%
TG-TRA120041	Hanning Chen	George Washington University	Computer and Information Science and Engineering	231	200,000	0%
TG-MCB140160	David Rhee	Albert Einstein College of Medicine	Molecular Biosciences	15	200,000	0%
TG-MCB130072	Robert Quick	Indiana University	Molecular Biosciences	11	200,000	0%
TG-CCR120041	Luca Clementi	San Diego Supercomputer Center	Computer and Computation Research	6	50,000	0%
TG-CDA100013	Mark Reed	University of North Carolina; Chapel Hill	Mathematical and Physical Sciences	6	100,000	0%
TG-CCR130001	Ruth Marinshaw	Stanford University	Other	2	50,000	0%
Total				708,121	1,864,120	

Startup Allocation
Campus
Champions



OSG-Direct users July 2013 to June 2014

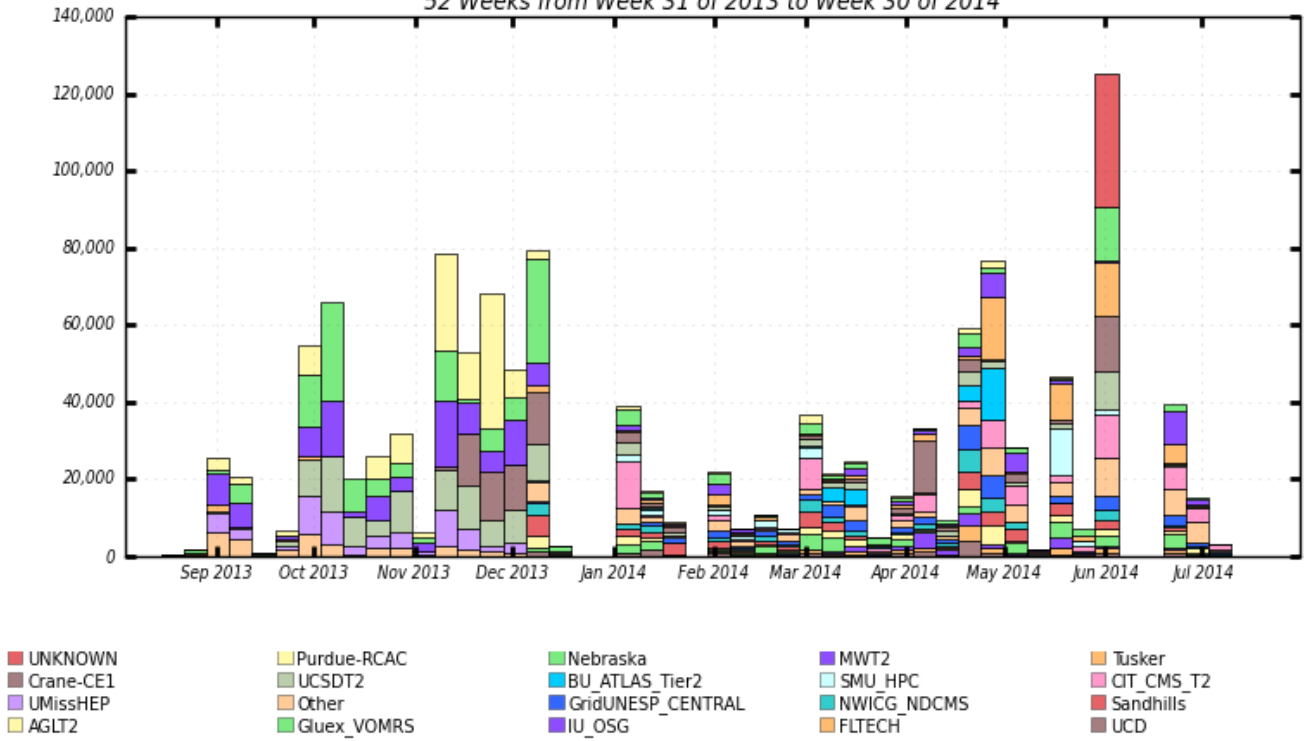
Project Name	PI	Institution	Field of Science	Wall Hours
SPLINTER	Robert Quick	Indiana University	Medicine	8,084,640
Duke-QGP	Steffen A. Bass	Duke University	Nuclear Physics	3,972,774
Snowmass	Meenakshi Narain	Brown University	High Energy Physics	2,812,541
RIT	P. Stanislaw Radziszowski	Rochester Institute of Technology	Computer Science	1,999,031
ECFA	Meenakshi Narain	Brown University	High Energy Physics	1,744,646
UMich	Paul Wolberg	University of Michigan	Microbiology	1,280,064
UPRRP-MR	Steven Massey	Universidad de Puerto Rico	Bioinformatics	1,192,943
Pheno	Stefan Hoeche	SLAC	High Energy Physics	1,108,623
DetectorDesign	John Strologas	University of New Mexico	Medical Imaging	989,043
IU-GALAXY	Robert Quick	Indiana University	Bioinformatics	665,197
EIC	Tobias Toll	Brookhaven National Laboratory	Accelerator Physics	564,486
HL-LHC-TP	Meenakshi Narain	Brown University	High Energy Physics	90,870
DeerDisease	Lene Jung Kjaer	Southern Illinois University	Biological Sciences	63,759
OSG-Staff	Chander Sehgal	Fermilab	Computer Science	40,514
SoyKB	Dong Xu	University of Missouri	Plant Biology	24,323
P0-LBNE	Maxim Potekhin	Brookhaven National Laboratory	Physics - Neutrino	18,334
BNLPET	Martin Purschke	Brookhaven National Laboratory	Medical Imaging	1,289
SNOplus	Joshua R Klein	University of Pennsylvania	Physics - Neutrino	489
Total		18 Users		24,653,567



The Baker Lab @ University of Washington Flocking Connection

Prediction and Design of Protein Structures and Protein-protein Interactions

WMS Hours Spent on Jobs By Facility (Glidein)
52 Weeks from Week 31 of 2013 to Week 30 of 2014



Maximum: 125,225 , Minimum: 0.00 , Average: 25,581 , Current: 3,224

1.4M Wall Hours

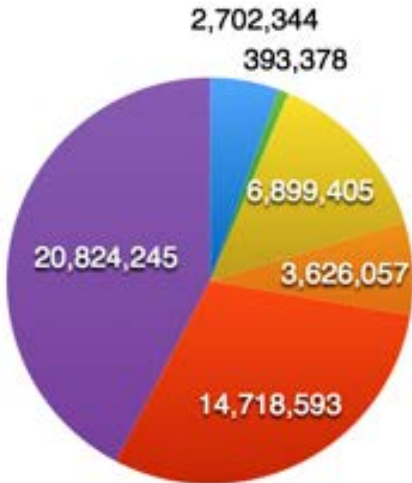


OSG Opportunistic Wall Hours

Last 3 years

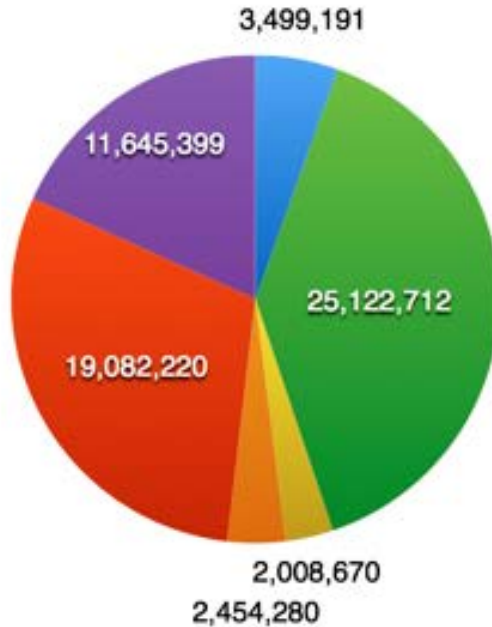
7/2011-6/2012

49,164,022



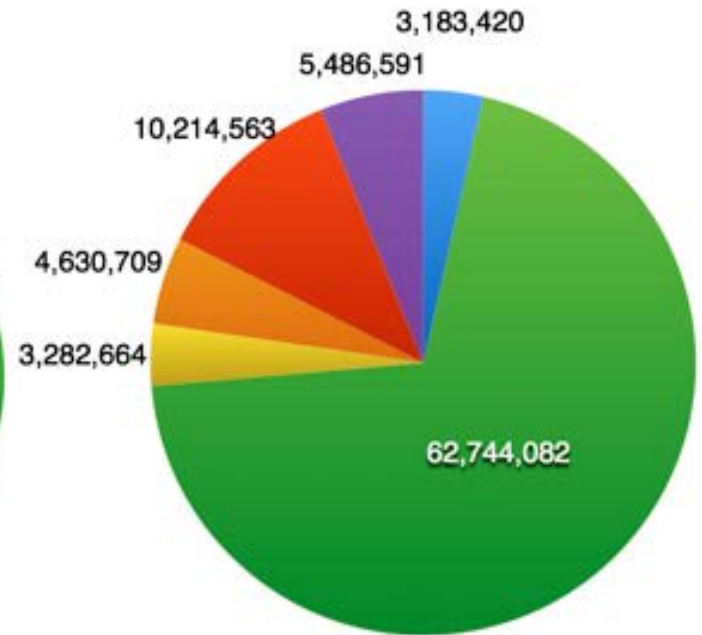
7/2012-6/2013

63,812,471



7/2013-6/2014

89,542,030

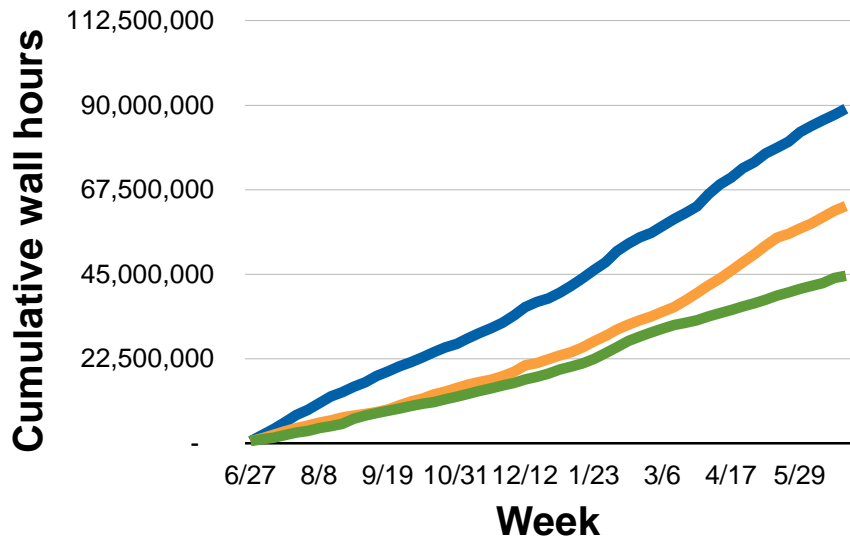


● sbgrid
 ● osg
 ● hcc
 ● gluex
 ● glow
 ● engage

Opportunistic Usage

Last 3 years

All opportunistic VO's



OSG VO

