OSG User Support & Campus Grids

Rob Gardner • University of Chicago
Mats Rynge • ISI

Open Science Grid

OSG Area Coordinator's Meeting, **December 14, 2016**

In brief (updated)

- Overall things are humming along fine given the available manpower & scope
  - A steady stream of user support requests are handled daily (average ~15 issues/week)
  - Deep engagements with Xenon1T and soon South Pole Telescope **(both, going well!)**
  - New FreeSurfer execution service - all good - larger scale validations (**death march**)
  - New user software in OASIS modules
  - New campus “quick connection” steady production at Stanford
  - New StashCache (Xrootd) testing
  - OSG User Support Survey results
- Significant risk
  - Lack of compute resources for the OSG VO

`stashcp` testing (Bala)

- Tony Aburaad (UC undergrad) helped develop **stashcp,** with Ilija Vukotic
  - Systematic load & reliability testing over the summer
- Developments in the stashcp client:
  - Improved monitoring and error reporting (retries, code versions, etc.)
  - Correctly handling insufficient local space.
  - Correctly propagating and reporting xrdcp exit codes.
  - Fixed code doing timeouts of slow transfers.
  - Collecting data in elasticsearch for analysis, c.f. kibana.mwt2.org
  - For some reason this code was never "released"
- Derek cleaned up the code and released
- New testing from Bala

30 day analytics

Delivered from the origin site

304.9TB

Sum of link_out

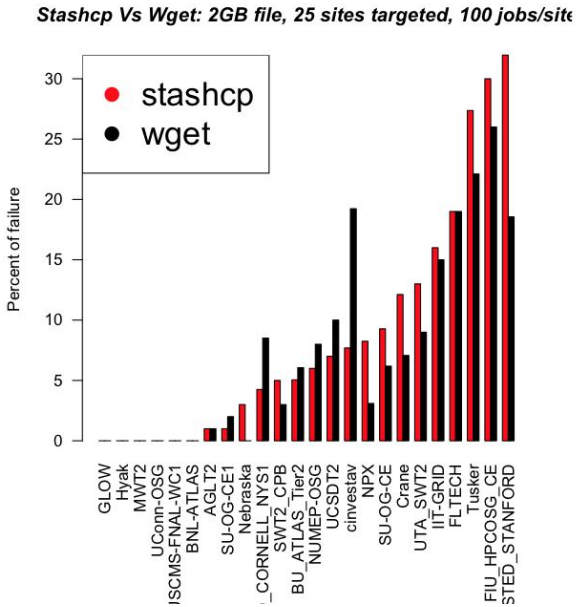
Delivered from caches

658.7TB

Sum of link_out

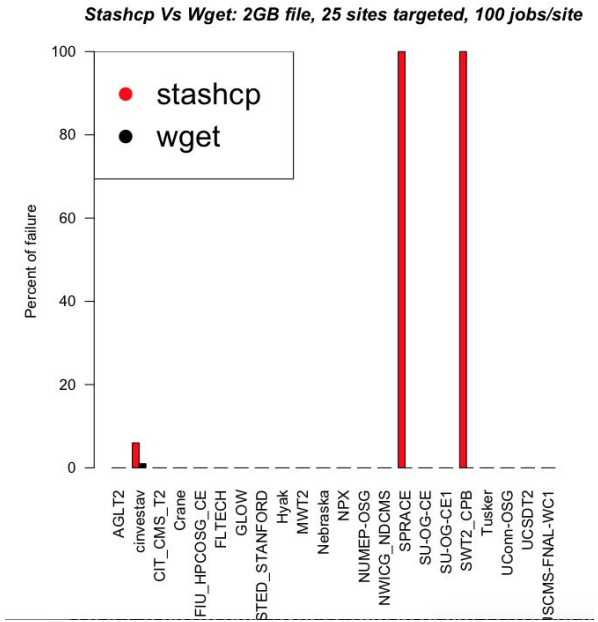
stashcp performance

Stashcp 3.4
Stashcp and wget failed on multiple sites about a month ago

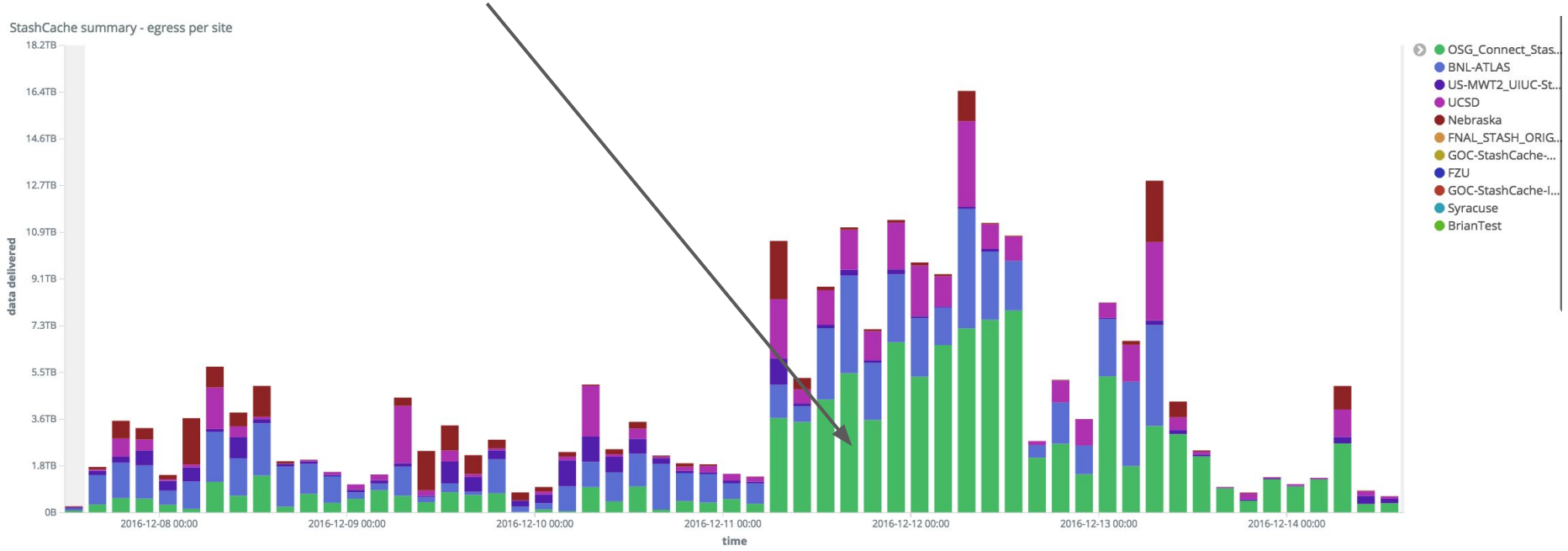


vs compute site

Stashcp 4.0 and HAS_SQUID == true
Performance is greatly optimized
(Test jobs ran during Dec 9-10)



Bala's testing



See more Kibana stashcp:

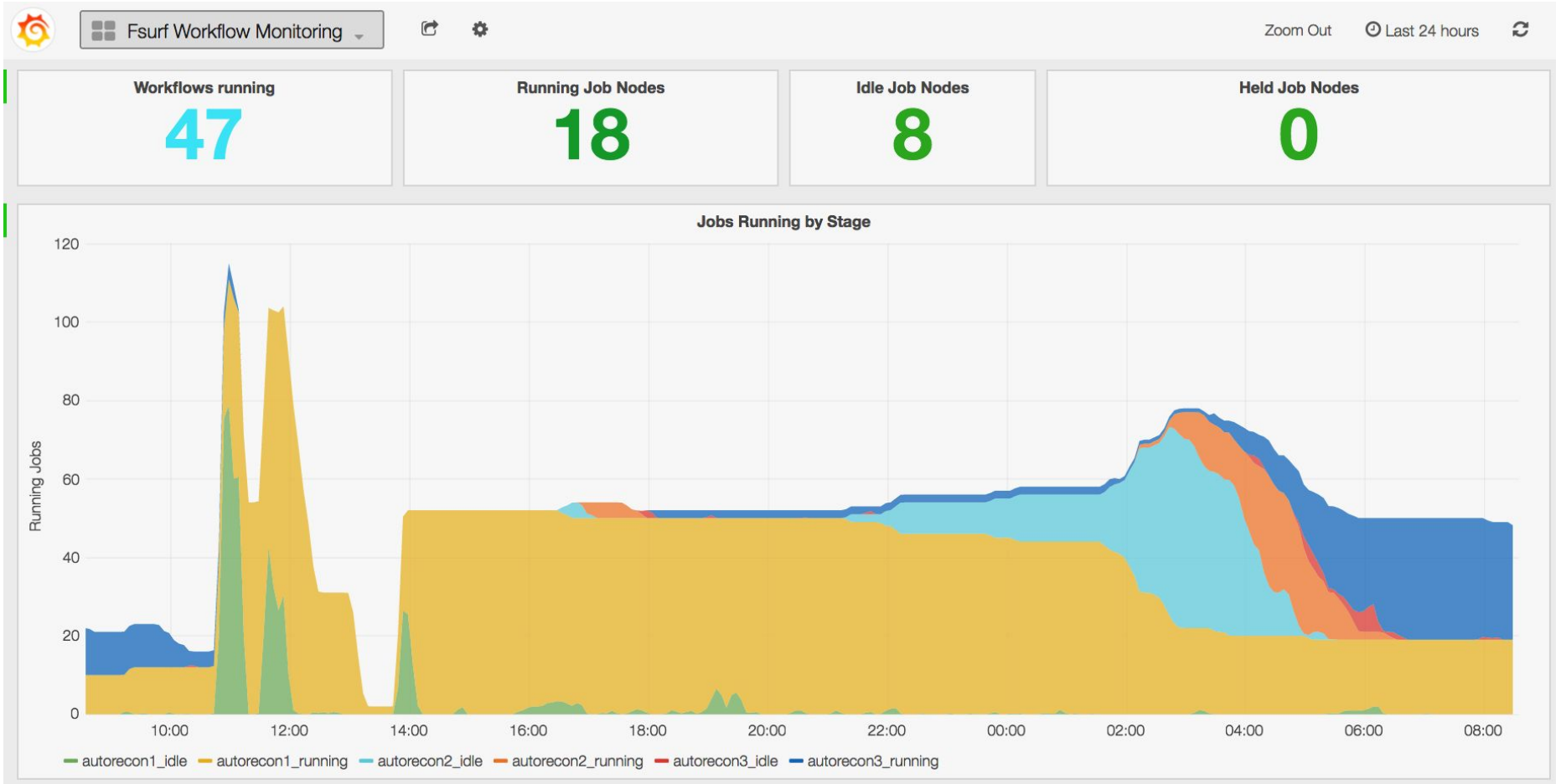
http://atlas-kibana.mwt2.org:5601/app/kibana#/dashboard/StashCache-summary

Fsurf Status (Suchandra)

- Ongoing Testing:
  - Don and Bala new batches of 150 and 50 test runs respectively
    - Most jobs ran successfully with about 13 jobs failing
      - UCR-HEP had compute nodes without a OSG_WN_TMP present
      - ND was accepting and preempting jobs
    - Updated scripts to handle missing OSG_WN_TMP
    - Updated HTCondor requirements to switch to another site after job failure
- Grafana dashboard created for monitoring
  - Linked on OSG Connect website
- Documentation all set and ready to go

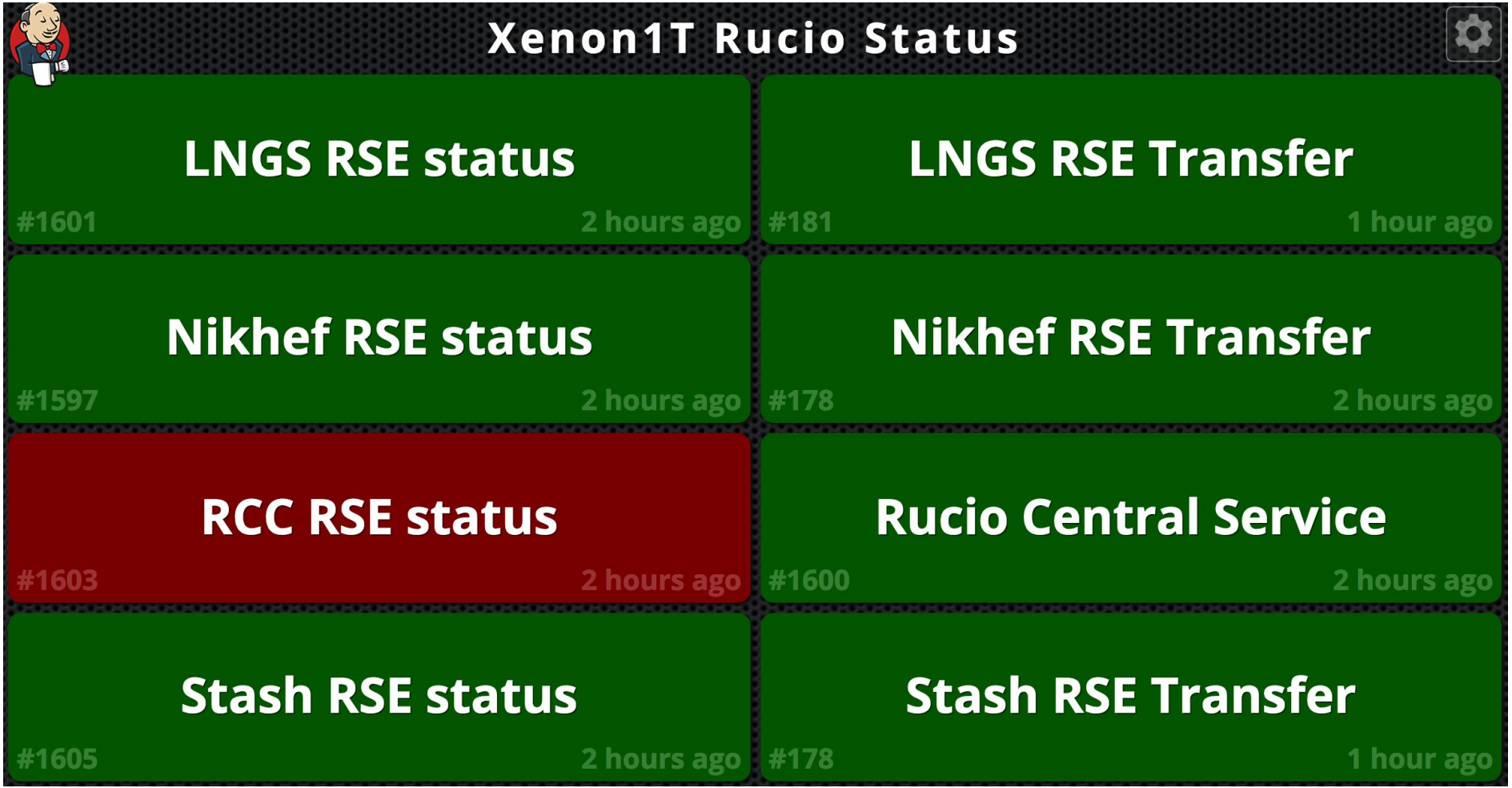
Fsurf next steps

- Run a final scale test with 450 MRI workflows to verify functionality at larger scales
- Final software and document touch ups and start planning release to Fsurf community with Don

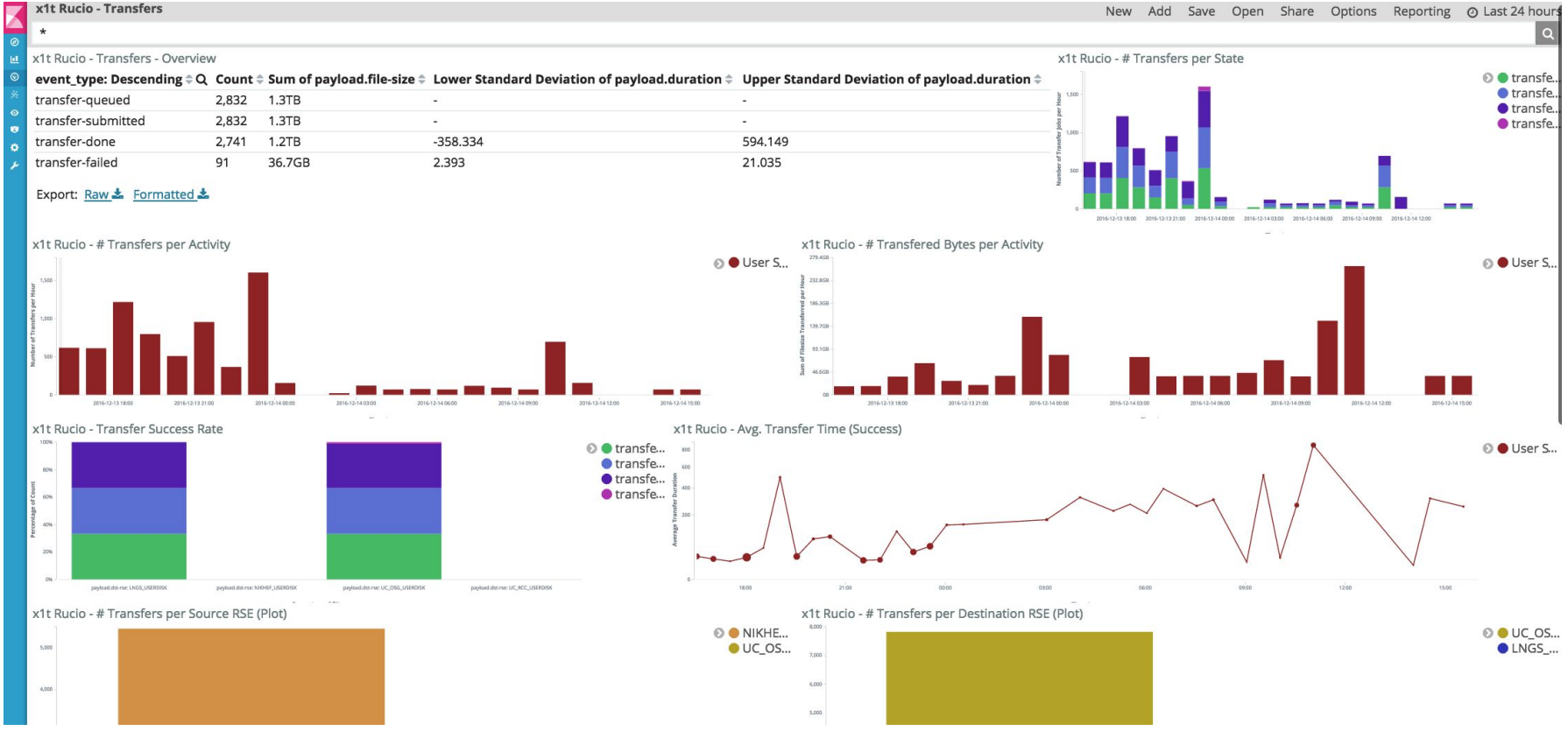

Fsurf Workflow Monitoring
Zoom Out
Last 24 hours
Workflows running
47
Running Job Nodes
18
Idle Job Nodes
8
Held Job Nodes
0
Jobs Running by Stage
Running Jobs
120
100
80
60
40
20
0
10:00
12:00
14:00
16:00
18:00
20:00
22:00
00:00
02:00
04:00
06:00
08:00
autorecon1_idle
autorecon1_running
autorecon2_idle
autorecon2_running
autorecon3_idle
autorecon3_running

Xenon1T Status (Benedikt, Suchandra)

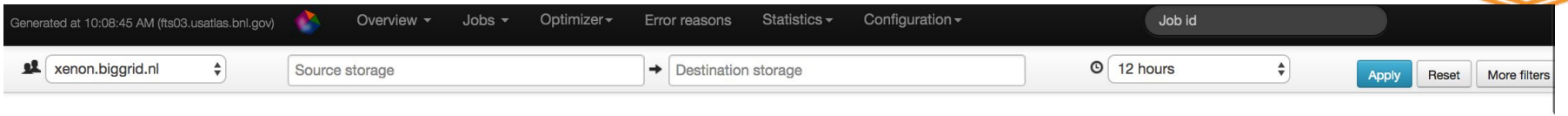
- OSG participated in week-long Xenon1T "tour de force" computing workshop (Nov 30 - Dec 7)
- Updated MC scripts to use pegasus in order to support submission to OSG and EGI sites
- Validating large scale (10M events) processing on OSG
- Validating Rucio distributed data management service
  - Now providing managed transfers between LNGS, Nikhef, Stash and UC-RCC gridftp endpoints
- Facilitated Xenon1T membership into LHCONE


Xenon1T Rucio Status
LNGS RSE status
#1601
2 hours ago
LNGS RSE Transfer
#181
1 hour ago
Nikhef RSE status
#1597
2 hours ago
Nikhef RSE Transfer
#178
2 hours ago
RCC RSE status
#1603
2 hours ago
Rucio Central Service
#1600
2 hours ago
Stash RSE status
#1605
2 hours ago
Stash RSE Transfer
#178
1 hour ago

Xenon1T Rucio analytics



Xenon1T FTS3 monitoring



Showing 1 to 50 out of 503

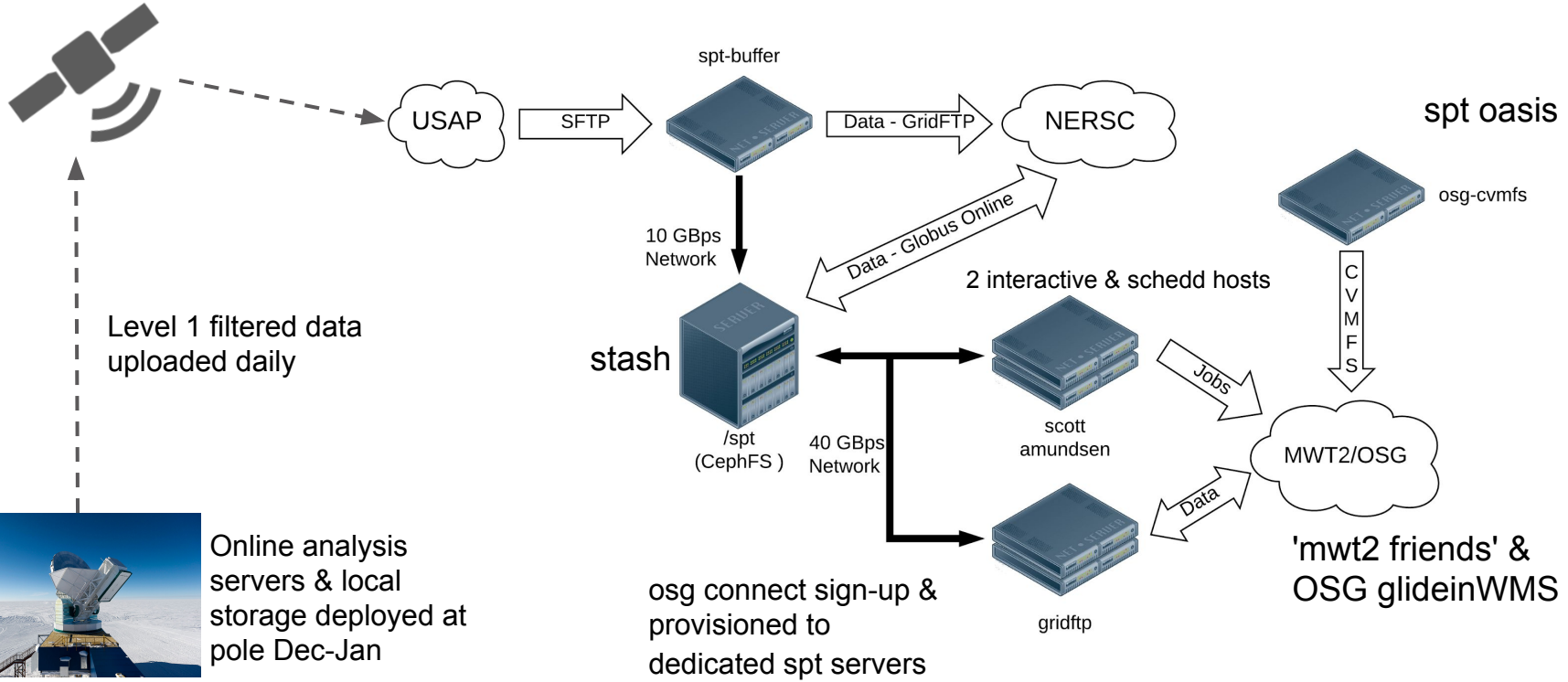
First | Previous | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | Next | Last

Job id	Submit time	Job state	VO	Source SE	Destination SE	Files	Priority	Destination space token
f131334f-ee3d-53ff-813d-4c0ba20ced34	2016-12-14T15:47:10	FINISHED	xenon.	srm://tbn18.nikhef.nl	gsiftp://gridftp.grid.uchicago.(	1	3	null
6f1b501c-9537-5b91-a2f7-db134bd2f1f5	2016-12-14T15:47:09	FINISHED	xenon.	srm://tbn18.nikhef.nl	gsiftp://gridftp.grid.uchicago.(	1	3	null
616522b5-dbe4-5215-b9f5-d07d0503d6ac	2016-12-14T15:47:09	FINISHED	xenon.	srm://tbn18.nikhef.nl	gsiftp://gridftp.grid.uchicago.(	1	3	null
7357d4e8-a918-5a6b-b5dc-df6b6d50de1d	2016-12-14T15:47:09	FINISHED	xenon.	srm://tbn18.nikhef.nl	gsiftp://gridftp.grid.uchicago.(	1	3	null
18df9fa7-e0fa-54f2-a632-dd56c6938520	2016-12-14T15:47:09	FINISHED	xenon.	srm://tbn18.nikhef.nl	gsiftp://gridftp.grid.uchicago.(	1	3	null

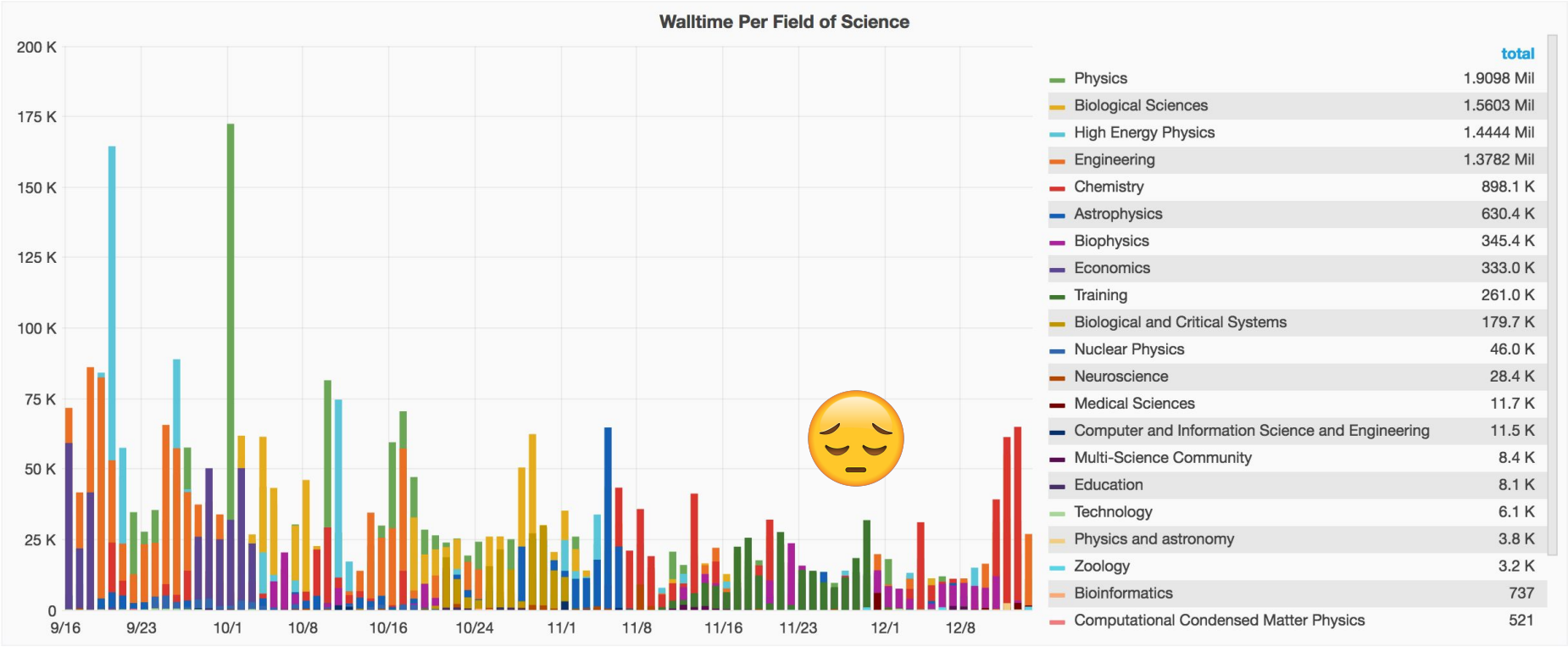
SPT Status (Benedikt)

- Software pipeline is based on IceCube framework
- Software repository for SPT in OASIS
- Analysis and archiving infrastructure procured, configured, and being delivered to the south pole (loaning a sysadmin for two months for on-prem commissioning)
- Setup analysis infrastructure co-located with OSG/MWT2

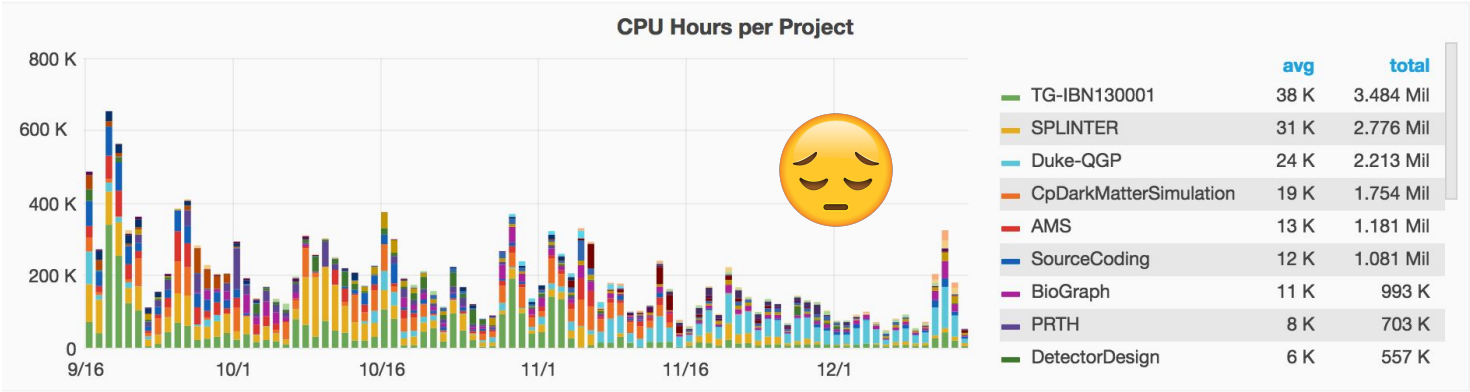
SPT Status, cont



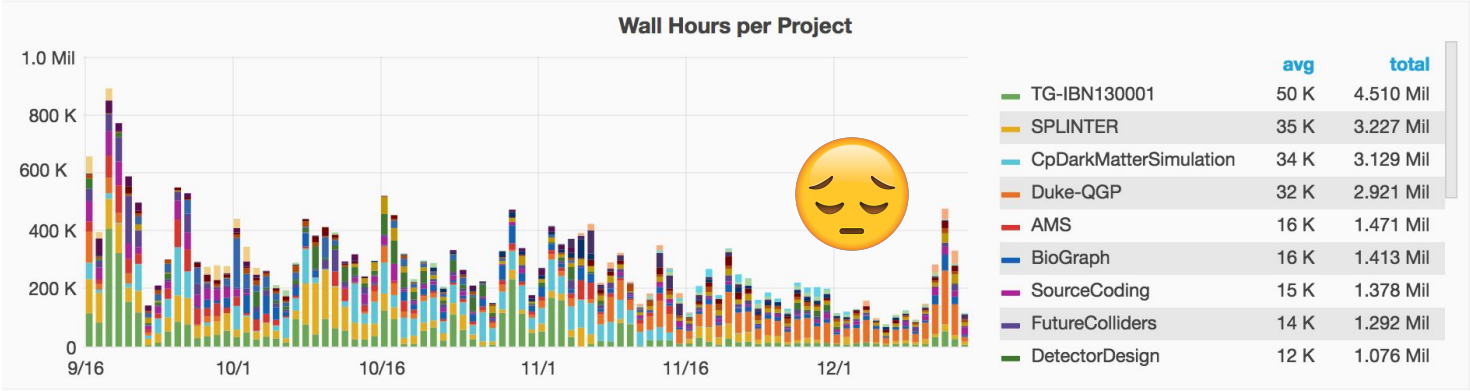
OSG Connect only



Combined Connect, XD & Direct

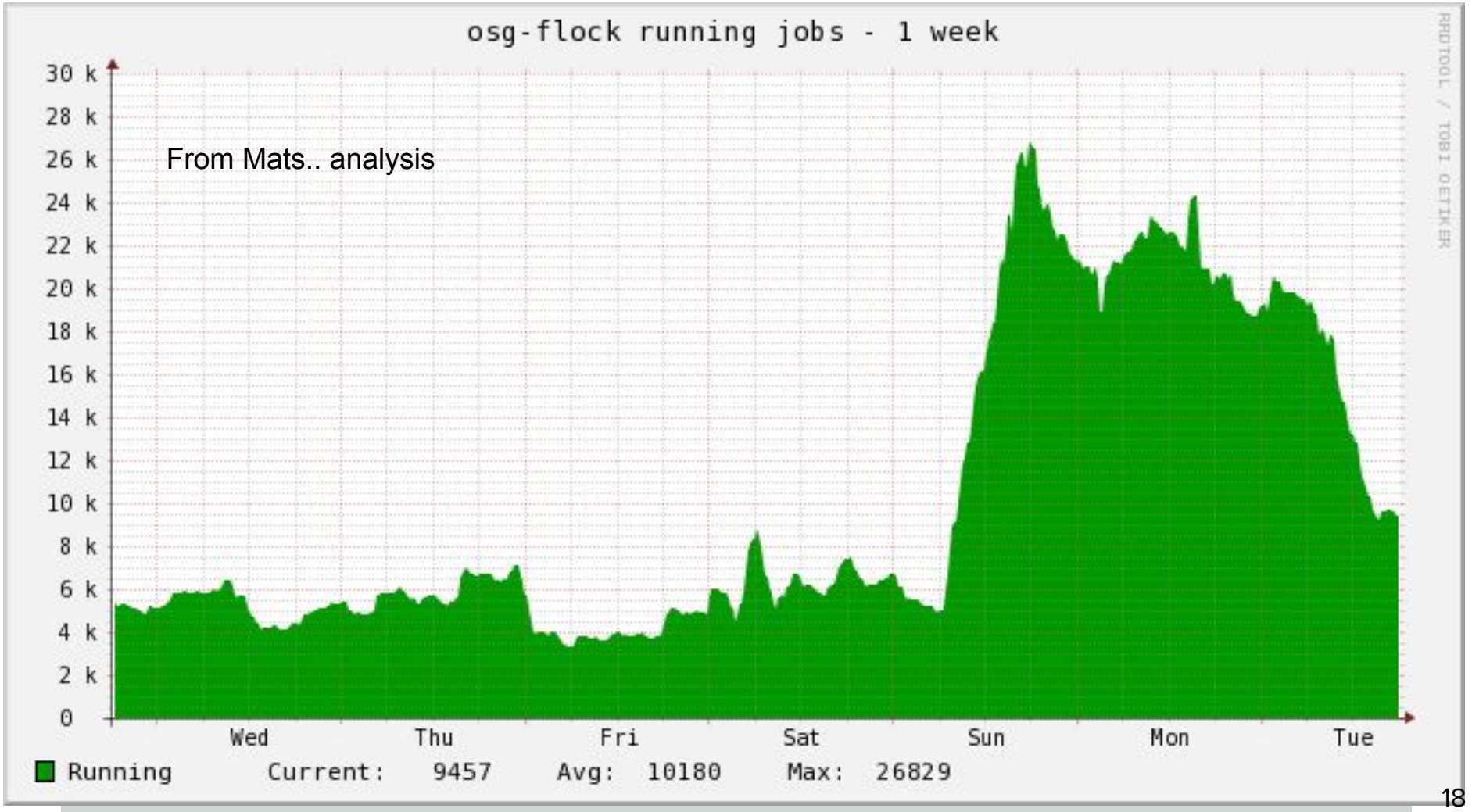


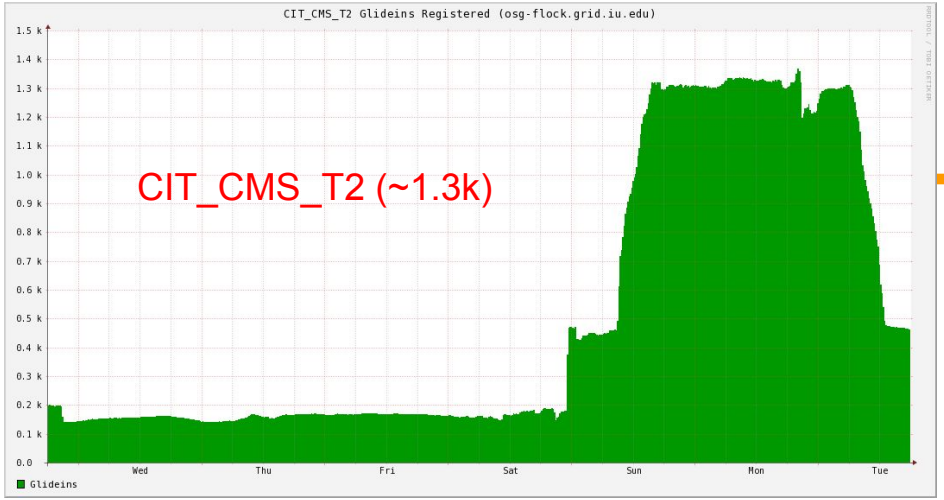
	avg	total
TG-IBN130001	38 K	3.484 Mil
SPLINTER	31 K	2.776 Mil
Duke-QGP	24 K	2.213 Mil
CpDarkMatterSimulation	19 K	1.754 Mil
AMS	13 K	1.181 Mil
SourceCoding	12 K	1.081 Mil
BioGraph	11 K	993 K
PRTH	8 K	703 K
DetectorDesign	6 K	557 K

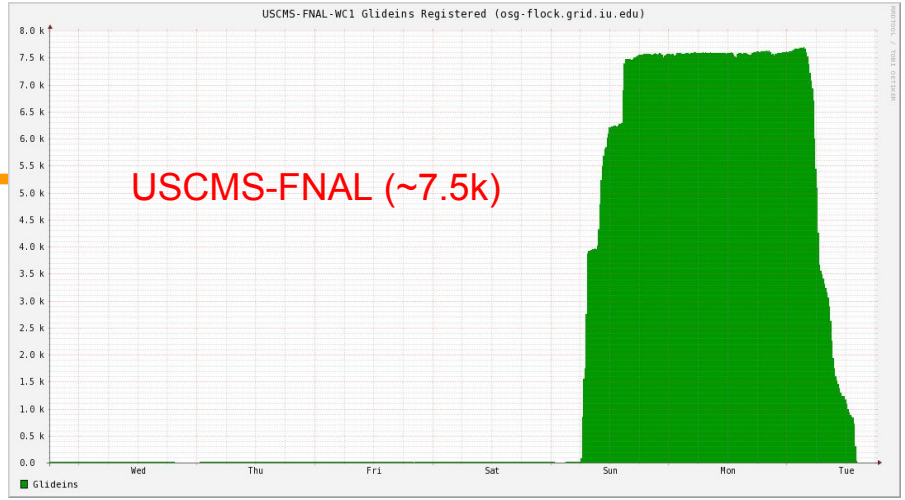


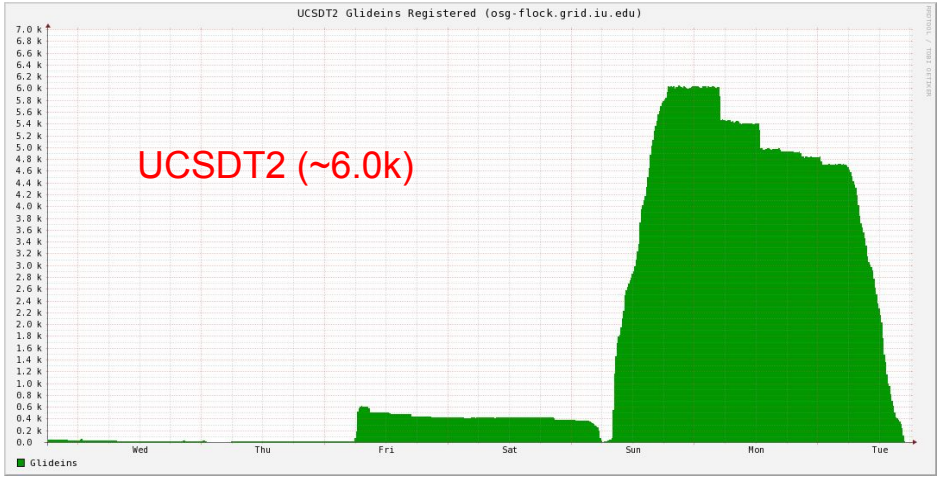
	avg	total
TG-IBN130001	50 K	4.510 Mil
SPLINTER	35 K	3.227 Mil
CpDarkMatterSimulation	34 K	3.129 Mil
Duke-QGP	32 K	2.921 Mil
AMS	16 K	1.471 Mil
BioGraph	16 K	1.413 Mil
SourceCoding	15 K	1.378 Mil
FutureColliders	14 K	1.292 Mil
DetectorDesign	12 K	1.076 Mil

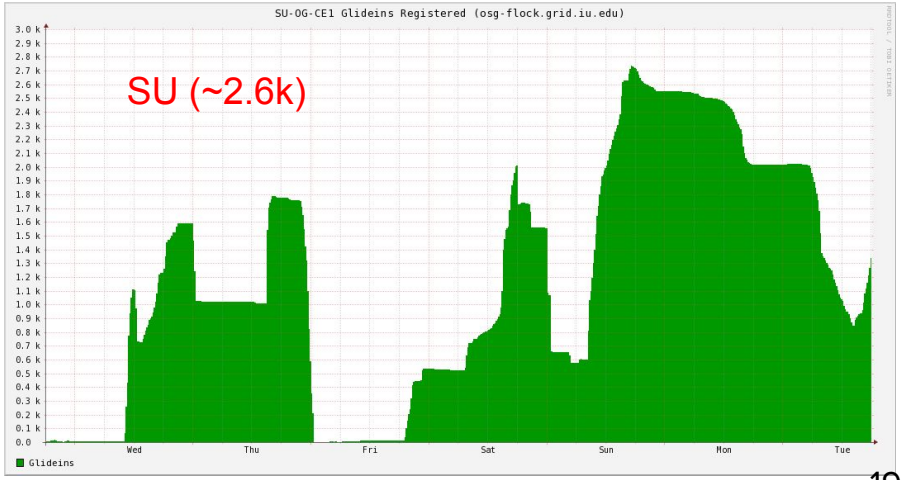
Top Projects

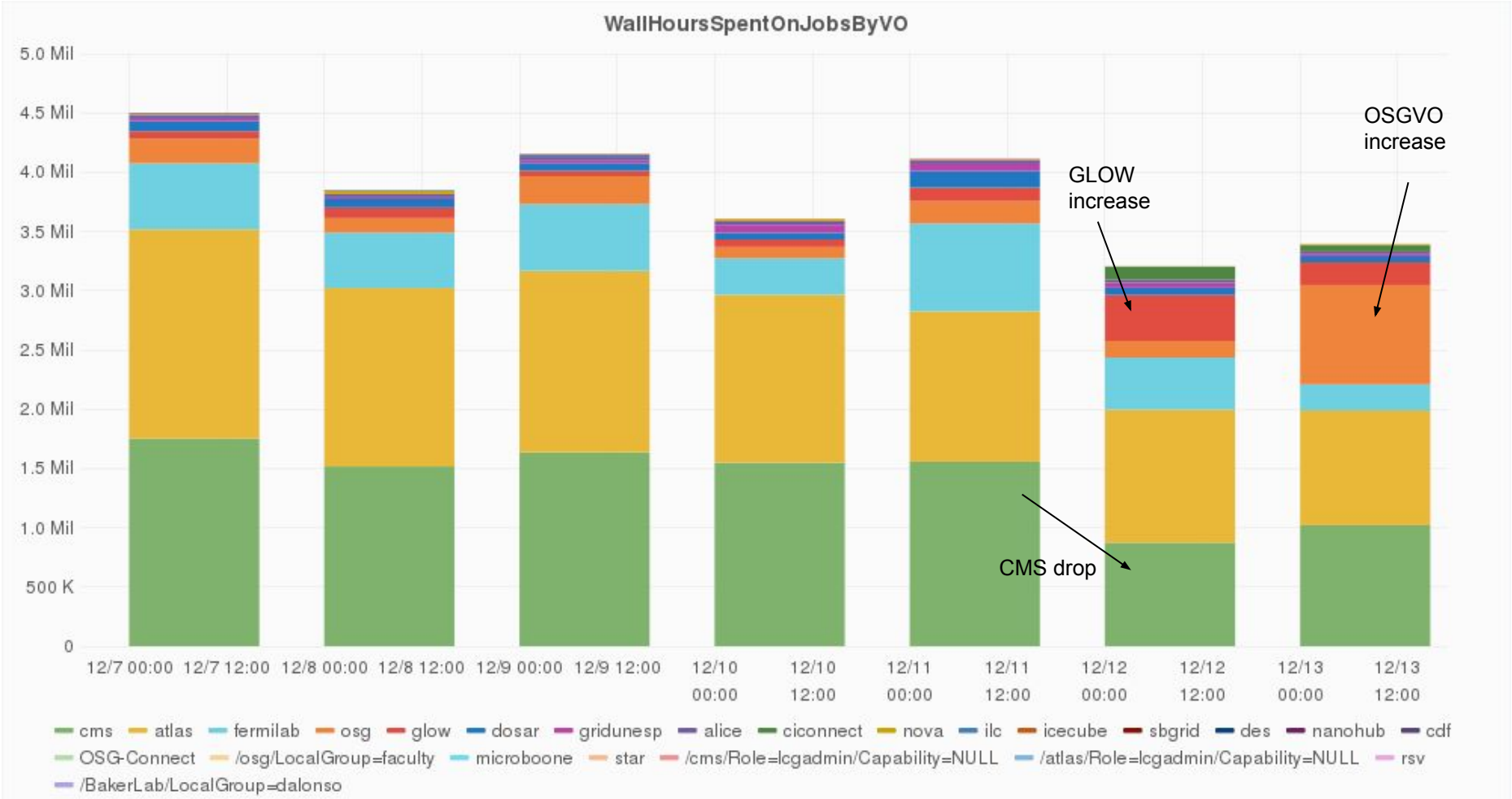

osg-flock running jobs - 1 week
From Mats.. analysis
30 k
28 k
26 k
24 k
22 k
20 k
18 k
16 k
14 k
12 k
10 k
8 k
6 k
4 k
2 k
0
Wed
Thu
Fri
Sat
Sun
Mon
Tue
Running
Current: 9457
Avg: 10180
Max: 26829
RRDTOOL / TOBI OETIKER


CIT_CMS_T2 Glideins Registered (osg-flock.grid.iu.edu)
CIT_CMS_T2 (~1.3k)
Wed
Thu
Fri
Sat
Sun
Mon
Tue
Glideins


USCMS-FNAL-WC1 Glideins Registered (osg-flock.grid.iu.edu)
USCMS-FNAL (~7.5k)
Wed
Thu
Fri
Sat
Sun
Mon
Tue
Glideins


UCSDT2 Glideins Registered (osg-flock.grid.iu.edu)
UCSDT2 (~6.0k)
Wed
Thu
Fri
Sat
Sun
Mon
Tue
Glideins


SU-OG-CE1 Glideins Registered (osg-flock.grid.iu.edu)
SU (~2.6k)
Wed
Thu
Fri
Sat
Sun
Mon
Tue
Glideins


WallHoursSpentOnJobsByVO
5.0 Mil
4.5 Mil
4.0 Mil
3.5 Mil
3.0 Mil
2.5 Mil
2.0 Mil
1.5 Mil
1.0 Mil
500 K
0
12/7 00:00
12/7 12:00
12/8 00:00
12/8 12:00
12/9 00:00
12/9 12:00
12/10 00:00
12/10 12:00
12/11 00:00
12/11 12:00
12/12 00:00
12/12 12:00
12/13 00:00
12/13 12:00
GLOW increase
OSGVO increase
CMS drop
cms
atlas
fermilab
osg
glow
dosar
gridunesp
alice
ciconnect
nova
ilc
icecube
sbgrid
des
nanohub
cdf
OSG-Connect
/osg/LocalGroup=faculty
microboone
star
/cms/Role=lcgadmin/Capability=NULL
/atlas/Role=lcgadmin/Capability=NULL
rsv
/BakerLab/LocalGroup=dalonso

OSG User Survey (Emelie)

- OSG User Support survey sent out to 814 users of the OSG VO (Connect, XD & Direct only)
- Questions to assess user satisfaction and identify common bottlenecks/roadblocks to using OSG

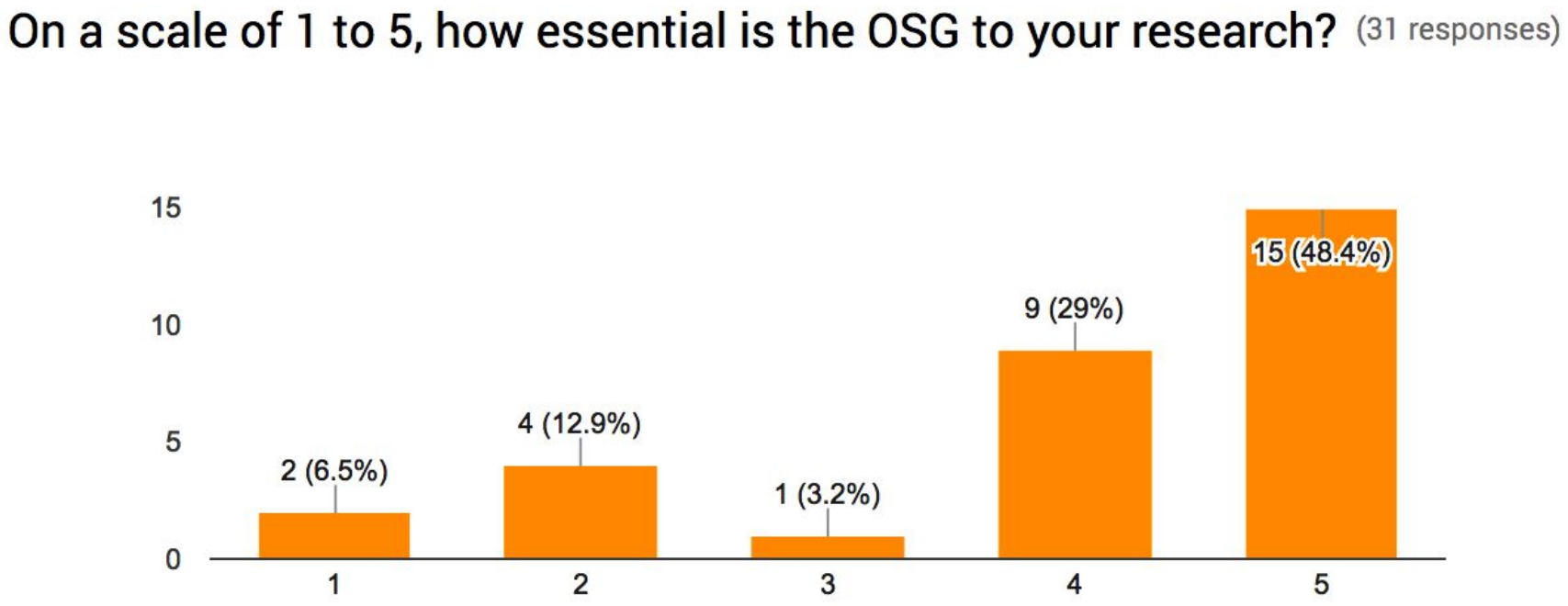
49 responses so far...

- Results indicate unanimous satisfaction with User Support team response to user issues
- OSG is very important or essential to 75% of respondents
- As expected, most respondents heard about OSG through a colleague, but 26% heard about OSG through a workshop or training event (and 24% heard from local campus resource center)*
- Many respondents were not aware of certain tools that might benefit their research, such as: *distributed software modules, StashCache, workflow management tools, Connect Client, and support for multicore, GPU, and large memory jobs*.
- Common factors that prevent/limit OSG usage among respondents were: *Jobs are too long-running, memory-intensive, require large input/output files, require multicore resources, or are waiting too long in the queue.* (47% responded that "Nothing" is limiting their use of OSG).
- None of the respondents replied "yes" when asked if they would like to be contacted by the User Support team to help resolve their issues, although 62% replied "maybe later". *Possible interpretation: they have already gotten as far as they can with User Support on these issues.*

*Users could select multiple responses to question

OSG User Survey

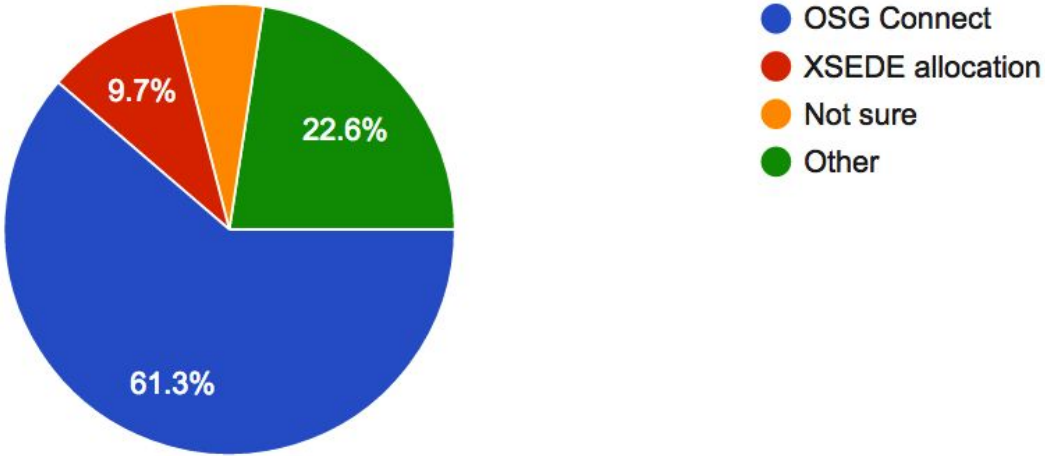
Preliminary (~ 50% returns)



OSG User Survey

Preliminary (~ 50% returns)

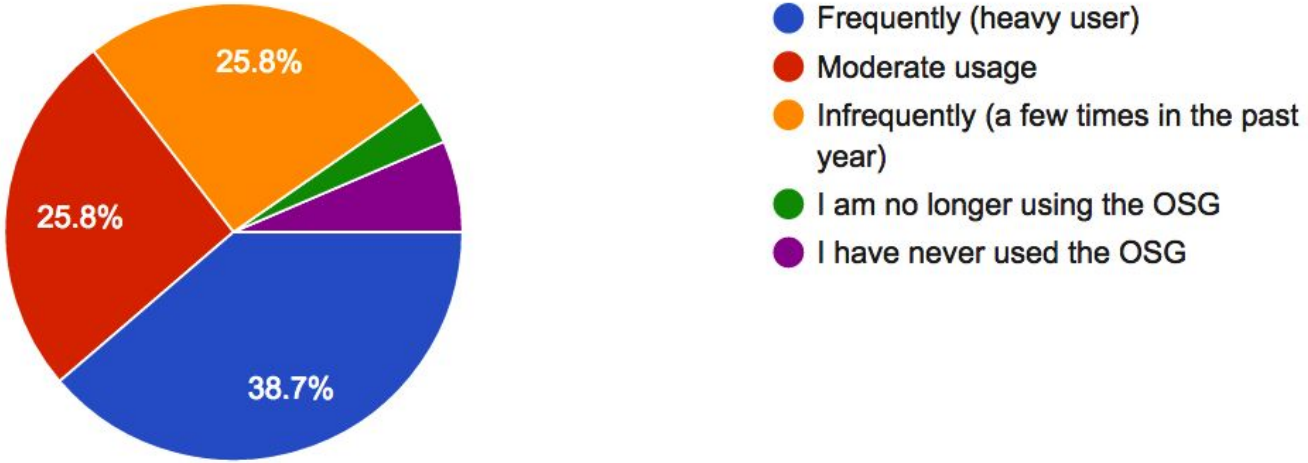
How do you primarily access the OSG? (31 responses)



OSG User Survey

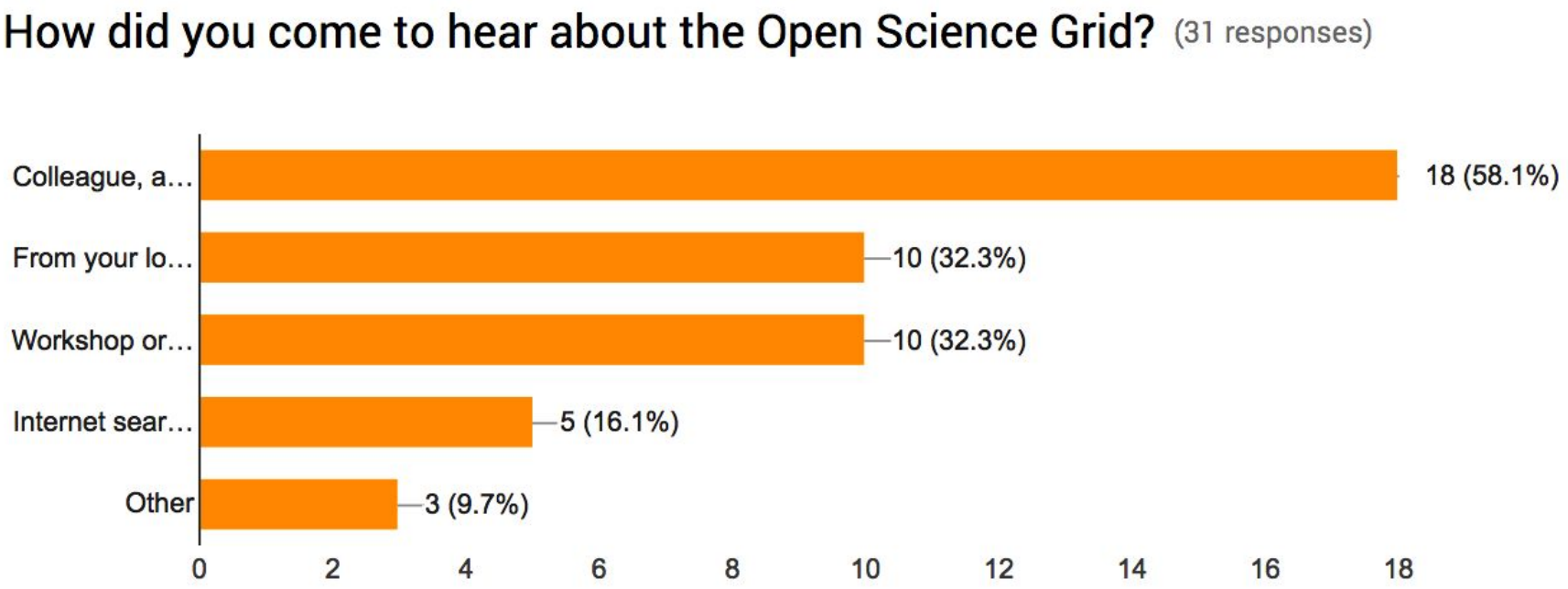
Preliminary (~ 50% returns)

How often do you use the OSG? (31 responses)



OSG User Survey

Preliminary (~ 50% returns)

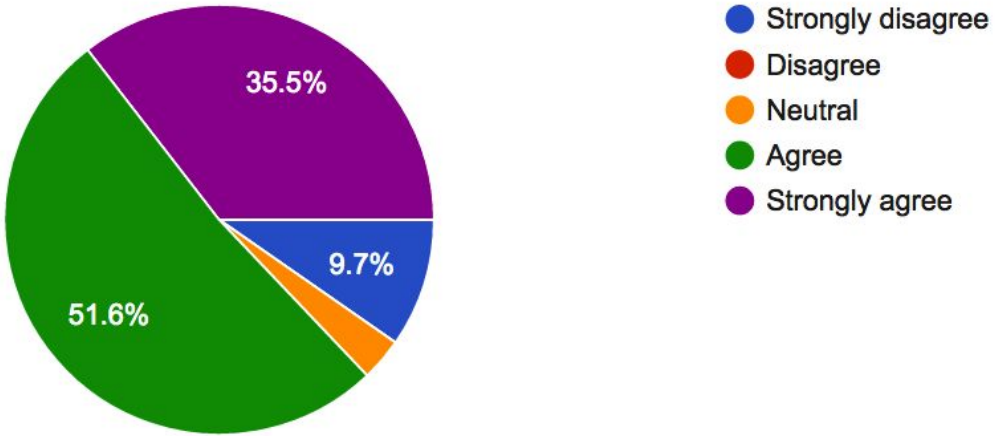


OSG User Survey

Preliminary (~ 50% returns)

How much do you agree with the following statement? "I feel that I have the tools, resources, and support I need to successfully utilize the OSG for my research."

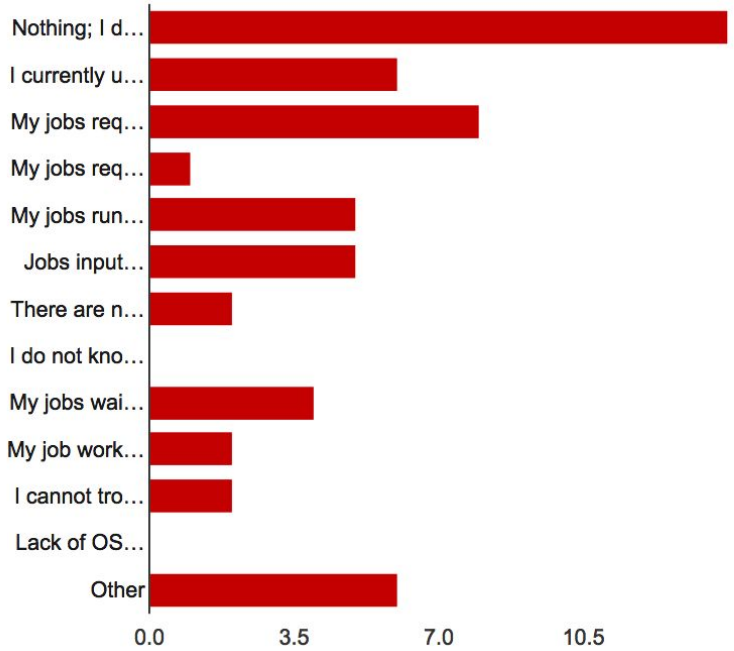
(31 responses)



OSG User Survey

Preliminary (~ 50% returns)

Which factors currently limit or prevent your use of OSG computing resources? (Check all that apply.)



Nothing; I don't feel like my OSG usage is being limited or prevented.	14	45.2%
I currently use other computing resources and do not need OSG resources	6	19.4%
My jobs require too much memory.	8	25.8%
My jobs require the use of proprietary software.	1	3.2%
My jobs run for too long and get pre-empted by resource owner's jobs.	5	16.1%
Jobs input or output data set is too large.	5	16.1%
There are not enough multicore resources for my jobs.	2	6.5%
I do not know how to submit my jobs to the OSG.	0	0%
My jobs wait too long in the queue (I am not getting enough resources).	4	12.9%
My job workflow is too complicated.	2	6.5%
I cannot troubleshoot job failures effectively.	2	6.5%
Lack of OSG user support	0	0%
Other	6	19.4%

Following are about tools & resources

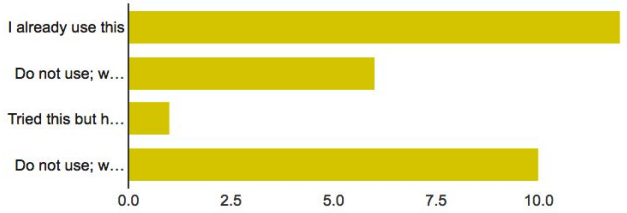
For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.

	I already use this	Do not use; would not be beneficial	Tried this but had problems	Do not use; would like to learn more
Accessing pre-installed packages on OSG through Distributed Environment Modules (OASIS)	○	○	◉	○
Transferring data to jobs using StashCache	○	○	○	◉
Managing workflows with DAGMAN/Pegasus/Makeflow	◉	○	○	○
Submitting jobs directly from your local campus cluster using Campus Connect Client	○	○	◉	○
Support for multicore jobs	○	○	○	◉
Support for GPU jobs	○	○	○	◉
Support for large memory jobs	○	○	○	◉
OSG User Support helpdesk (support.opensciencegrid.org)	◉	○	○	○

OSG User Survey

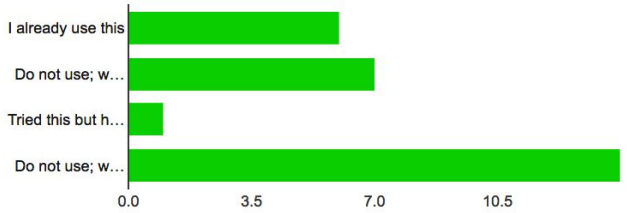
Preliminary (~ 50% returns)

Accessing pre-installed packages on OSG through Distributed Environment Modules (OASIS) [For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.]



I already use this	12	41.4%
Do not use; would not be beneficial	6	20.7%
Tried this but had problems	1	3.4%
Do not use; would like to learn more	10	34.5%

Transferring data to jobs using StashCache [For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.]

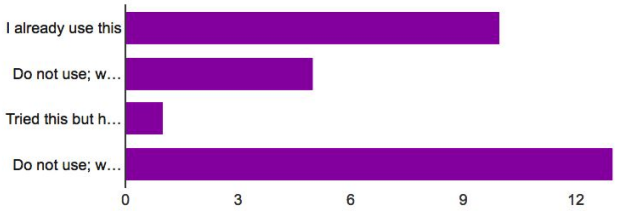


I already use this	6	21.4%
Do not use; would not be beneficial	7	25%
Tried this but had problems	1	3.6%
Do not use; would like to learn more	14	50%

OSG User Survey

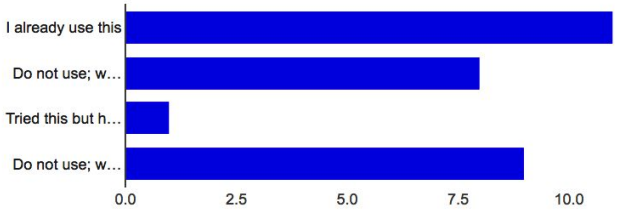
Preliminary (~ 50% returns)

Managing workflows with DAGMAN/Pegasus/Makeflow [For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.]



I already use this	**10**	34.5%
Do not use; would not be beneficial	**5**	17.2%
Tried this but had problems	**1**	3.4%
Do not use; would like to learn more	**13**	44.8%

Submitting jobs directly from your local campus cluster using Campus Connect Client [For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.]

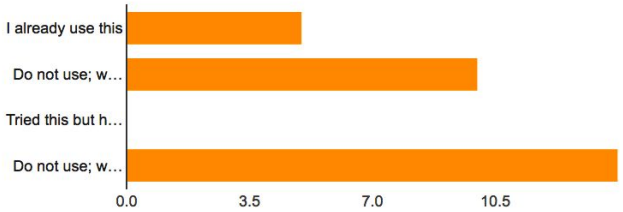


I already use this	**11**	37.9%
Do not use; would not be beneficial	**8**	27.6%
Tried this but had problems	**1**	3.4%
Do not use; would like to learn more	**9**	31%

OSG User Survey

Preliminary (~ 50% returns)

Support for multicore jobs [For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.]



I already use this	5	17.2%
Do not use; would not be beneficial	10	34.5%
Tried this but had problems	0	0%
Do not use; would like to learn more	14	48.3%

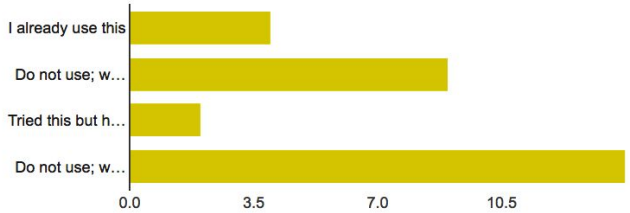
Support for GPU jobs [For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.]

I already use this	0	0%
Do not use; would not be beneficial	12	41.4%
Tried this but had problems	0	0%
Do not use; would like to learn more	17	58.6%

OSG User Survey

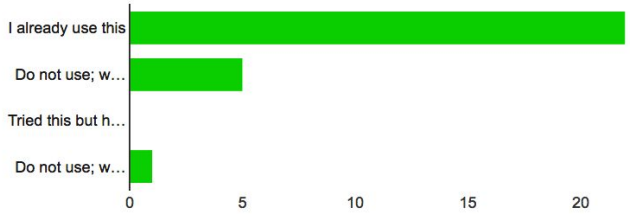
Preliminary (~ 50% returns)

Support for large memory jobs [For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.]



I already use this	4	13.8%
Do not use; would not be beneficial	9	31%
Tried this but had problems	2	6.9%
Do not use; would like to learn more	14	48.3%

OSG User Support helpdesk (support.opensciencegrid.org) [For each OSG tool or resource listed, please indicate whether you already use it for your research. If not, indicate whether you would like to learn more about it.]

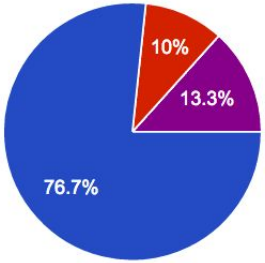


I already use this	22	78.6%
Do not use; would not be beneficial	5	17.9%
Tried this but had problems	0	0%
Do not use; would like to learn more	1	3.6%

OSG User Survey

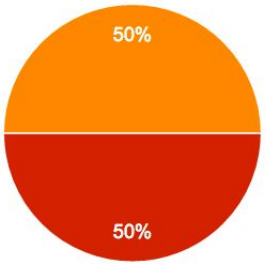
Preliminary (~ 50% returns)

How satisfied are you with the response of the User Support team to questions or support issues?



Very satisfied	**23**	76.7%
Reasonably satisfied	**3**	10%
Unsatisfied	**0**	0%
Very unsatisfied	**0**	0%
Not applicable	**4**	13.3%

Would you like the User Support team to contact you to discuss your OSG related computational needs?



Yes	**0**	0%
No	**15**	50%
Maybe later	**15**	50%

OSG User Survey - Comments

1. Bala is awesome. His help has been invaluable.
2. OSG computing and people (specifically Mats Rynge and Bala Desinghu) have been essential to my research. Thank you.
3. I would benefit for larger availability of OSG grid for large memory jobs (8 GB)
4. OSG is a great resource for scientists who don't have access for large-scale resources. However, once thing I notice is that the actual performance of a considerable number of applications running on OSG were not efficient. With reasonable efforts, it could be possible to run those applications by one or two orders faster.
5. User support staff has been very helpful, and gone above and beyond what I expected!
6. I think OSG is an amazing resource, and I'm disappointed that I'm not currently able to use it. It's possible that other projects in the future will have use of it, and if so, I'll definitely contact our local OSG/Condor people about using it.
7. I am getting significantly less OSG resources now than in past years (despite requesting the same amount). In the past I would typically have a few thousand jobs running at any given time; now I am lucky to get even 500.
8. The Globus file tranfer system is very useful even though it is not mentioned above. Also it would be very wonderful if it were possible to use condor_compile for job checkpointing.
9. my group is self-sufficient. We get from osg what we need, and when we need help we know how to get it.
10. The OSG school was quite beneficial and instructive. Although I may not be using it now because I have access to HPC resources, knowing about OSG and differences of HTC with HPC helps a lot to have stay open minded about the scale of the problem to tackle; HPC resources are much more efficient, but they have a hard limit (cores, memory, and even disk space) while HTC resources are dynamic and theoretically unlimited.
11. I'd like to take this opportunity to thank the support from OSG people. I am very grateful for their help. I'm also happy that OSG has a partnership with Duke.
12. Wonderful and important resource
13. OSG & its people are super awesome
14. The OSG support team responds more quickly and acts more effectively than any support unit I've used - thank you for all you do!!
15. OSG support service is excellent. But sometimes they took sometime to respond. Other than that it is very helpful

Additional comments

1. Please don't remove any of your educational material from your website. They are awesome.
2. OSG is a great tool. Thank you!
3. I am overall very satisfied with OSG, I sort of fell in love with pegasus too.
4. Fantastic computing resource ... absolutely essential to our research
5. I work at Florida State University HPC group and my main purpose to have this account is to show our users how to use OSG.
6. I am always answered extremely quickly when I submit a ticket to the help desk, and I really appreciate that. This is a great service with wonderful support staff.
7. thanks
8. OSG provides superior research computing facilitation. They go the extra mile to help the user. In my case, the phase transitioned from tech support to colleague and we are co-authors on a open source workflow (OSG-GEM) that could impact many genomics users. Thank you OSG!
9. I'm not currently involved in any research requiring OSG. My biggest hurdle was packaging my dependencies. I know OASIS was meant for this, but I frequently ran into problems with its third-party Python libraries, and it ended up being easier to do it myself. I feel like this could be a good approach if OSG have more direct support for user-packaged software rather than relying on support (who may be overwhelmed already). It's simple enough to environment-wrangle but most scientists wouldn't want to bother. Maybe some GitHub-like system where users can publish their packaged versions of commonly used dependencies?