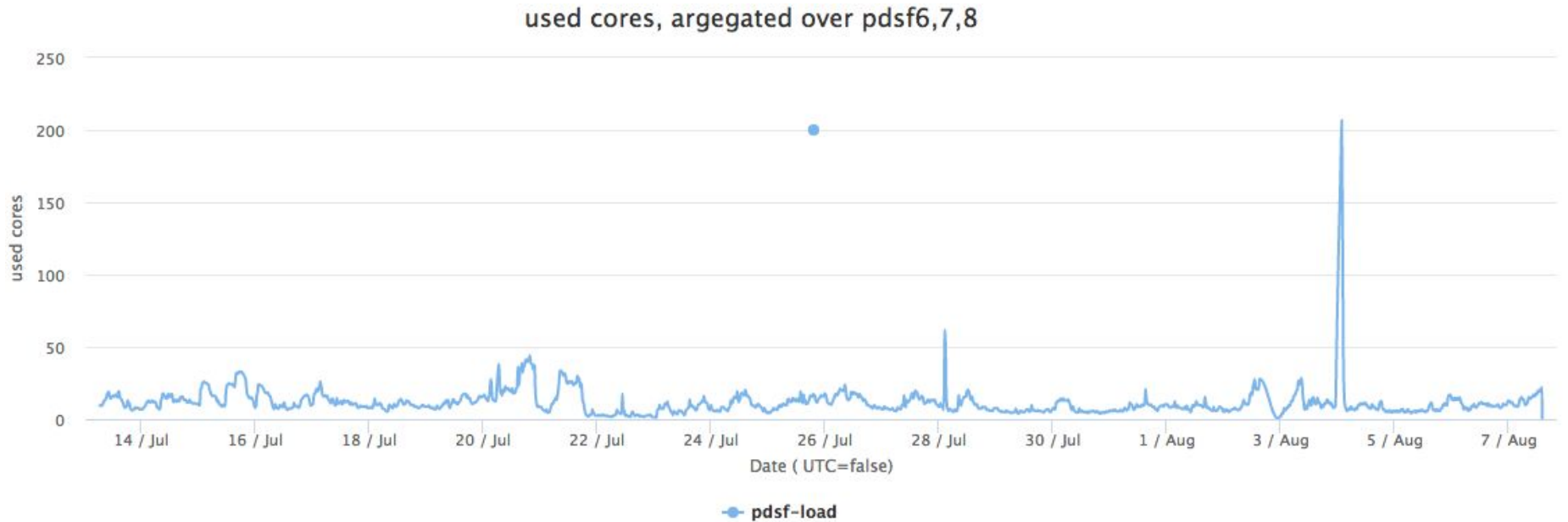


PDSF User Meeting

- PDSF performance
- announcements
- SLURM @ PDSF
- AOB

aggregated load on PDSF interactive nodes

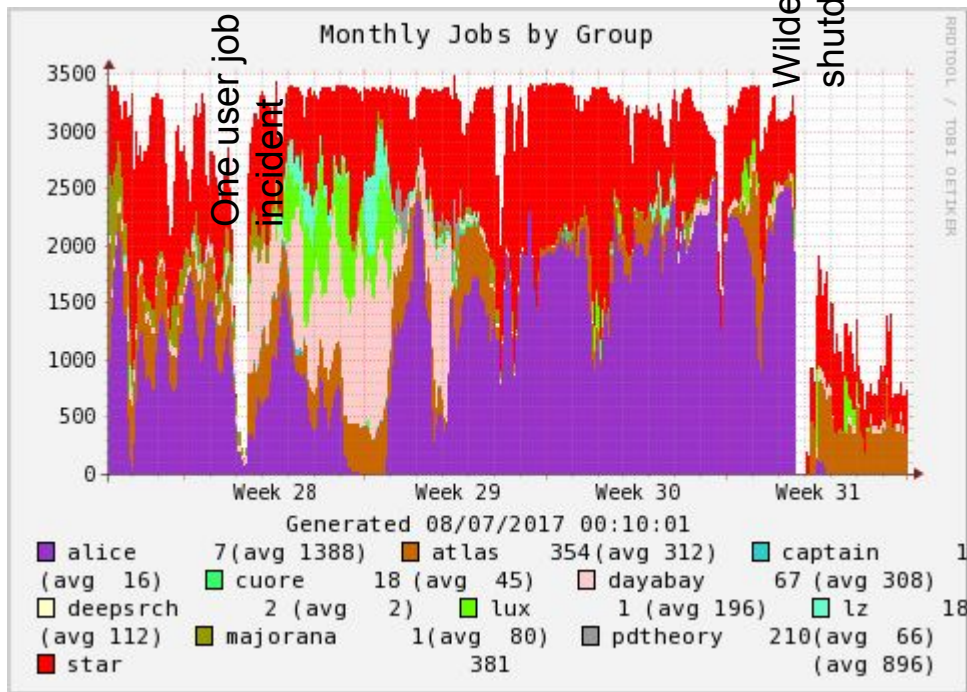
<https://portal-auth.nersc.gov/pdsf-mon/>



CPU aggregated over month

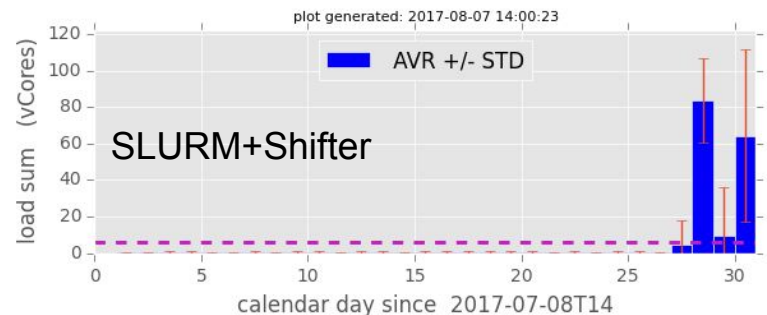
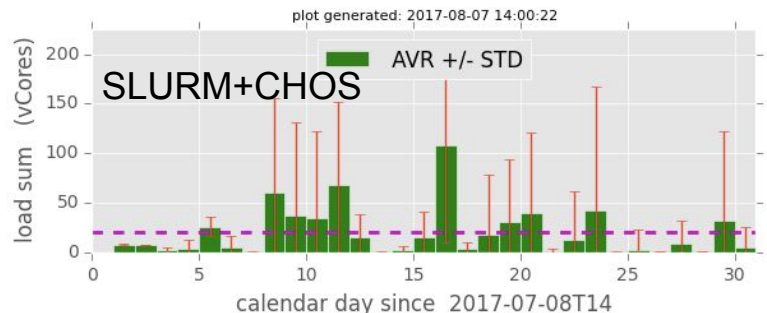
UGE+CHOS

UGE load $\sim 1/3$ after power dip, no Alice till today morning



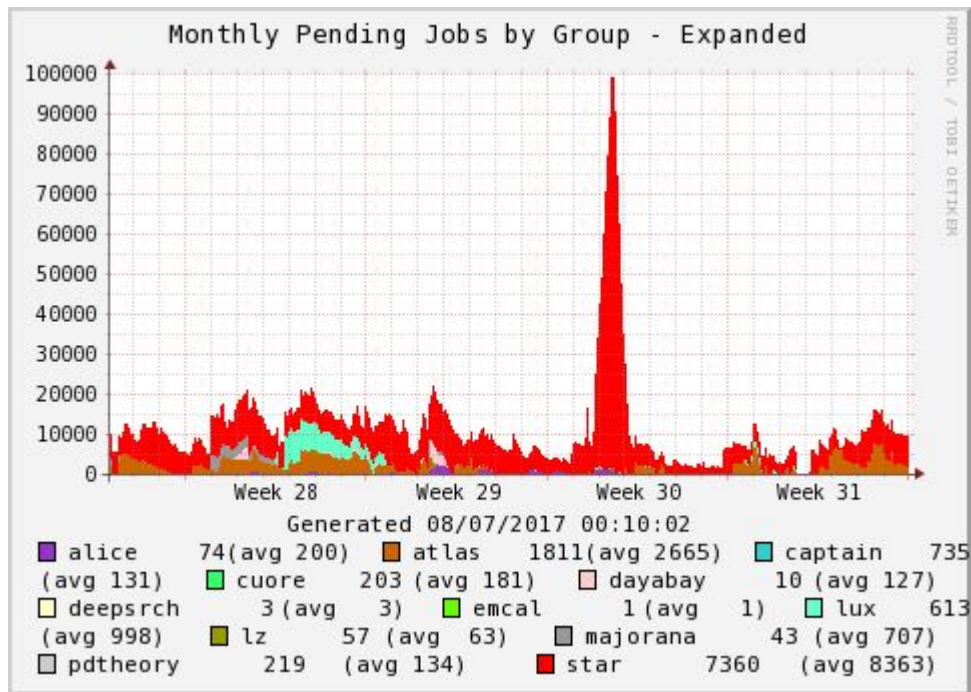
SLURM is used

<http://portal.nersc.gov/project/mpccc/balewski/pdsf3Load/latest/>



queue(s) load

UGE - looks fine



SLURM - no queue limits
(no monitor vs. time)

Partition	Nodes per Job	Physical Max Cores per Job	Max Walltime per Job
shared-chos	1	1-32	2 days
shared	1	1-32	2 days

Partition	Total vCores [nt1]	Total Nodes	Default Mem per vCore (MB)	Max Mem per vCore (MB)
shared-chos	420	7	1000	2000
shared	176	4	1000	2000

Aggregated utilization of UGE

user \ timeRange	sum last 30_days
0-all-users : owner	228.3 (cpu*year) , 1300453 jobs wallT fract=1.000, CPU/wallT=0.87
alice : project	96.6 (cpu*year) , 101676 jobs wallT fract=0.423, CPU/wallT=0.84
atlas : project	23.2 (cpu*year) , 494232 jobs wallT fract=0.102, CPU/wallT=1.00
dayabay : project	20.0 (cpu*year) , 51732 jobs wallT fract=0.087, CPU/wallT=0.52
dybprod : owner	0.0 (cpu*year) , 0 jobs wallT fract=0.000, CPU/wallT=0.00
hack : owner	1.9 (cpu*year) , 16028 jobs wallT fract=0.008, CPU/wallT=0.93
lux : project	8.0 (cpu*year) , 29285 jobs wallT fract=0.035, CPU/wallT=0.91
lz : project	3.8 (cpu*year) , 4793 jobs wallT fract=0.017, CPU/wallT=1.00
majorana : project	4.1 (cpu*year) , 60960 jobs wallT fract=0.018, CPU/wallT=0.48
star : project	70.9 (cpu*year) , 551245 jobs wallT fract=0.311, CPU/wallT=0.95
staremb : owner	0.0 (cpu*year) , 0 jobs wallT fract=0.000, CPU/wallT=0.00

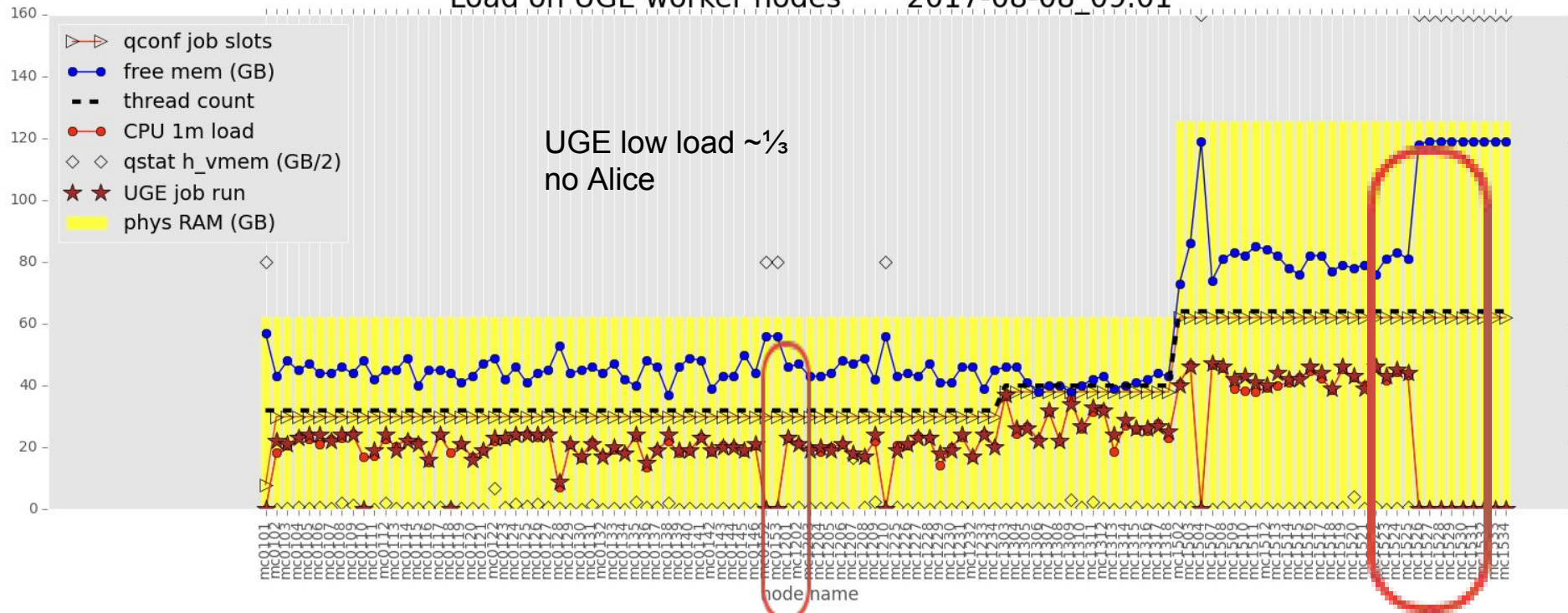
July average
Served 3.1k vCores
Capacity 2.8k vCores
CPU/WallT 0.87

June average
Served 3.1k vCores
Capacity 3.4k vCores
CPU/WallT 0.74

user \ timeRange	sum last 30_days
0-all-users : owner	251.1 (cpu*year) , 1342121 jobs wallT fract=1.000, CPU/wallT=0.74
alice : project	113.8 (cpu*year) , 99351 jobs wallT fract=0.453, CPU/wallT=0.70
atlas : project	16.6 (cpu*year) , 386661 jobs wallT fract=0.066, CPU/wallT=0.93
dayabay : project	35.2 (cpu*year) , 105573 jobs wallT fract=0.140, CPU/wallT=0.36
dybprod : owner	0.0 (cpu*year) , 0 jobs wallT fract=0.000, CPU/wallT=0.00
hack : owner	3.3 (cpu*year) , 5687 jobs wallT fract=0.013, CPU/wallT=0.97
lux : project	3.5 (cpu*year) , 4195 jobs wallT fract=0.014, CPU/wallT=0.95
lz : project	2.3 (cpu*year) , 181794 jobs wallT fract=0.009, CPU/wallT=0.92
majorana : project	9.4 (cpu*year) , 141214 jobs wallT fract=0.037, CPU/wallT=0.74
star : project	69.8 (cpu*year) , 415768 jobs wallT fract=0.278, CPU/wallT=0.92
staremb : owner	0.0 (cpu*year) , 0 jobs wallT fract=0.000, CPU/wallT=0.00

Mendel utilization - snapshot

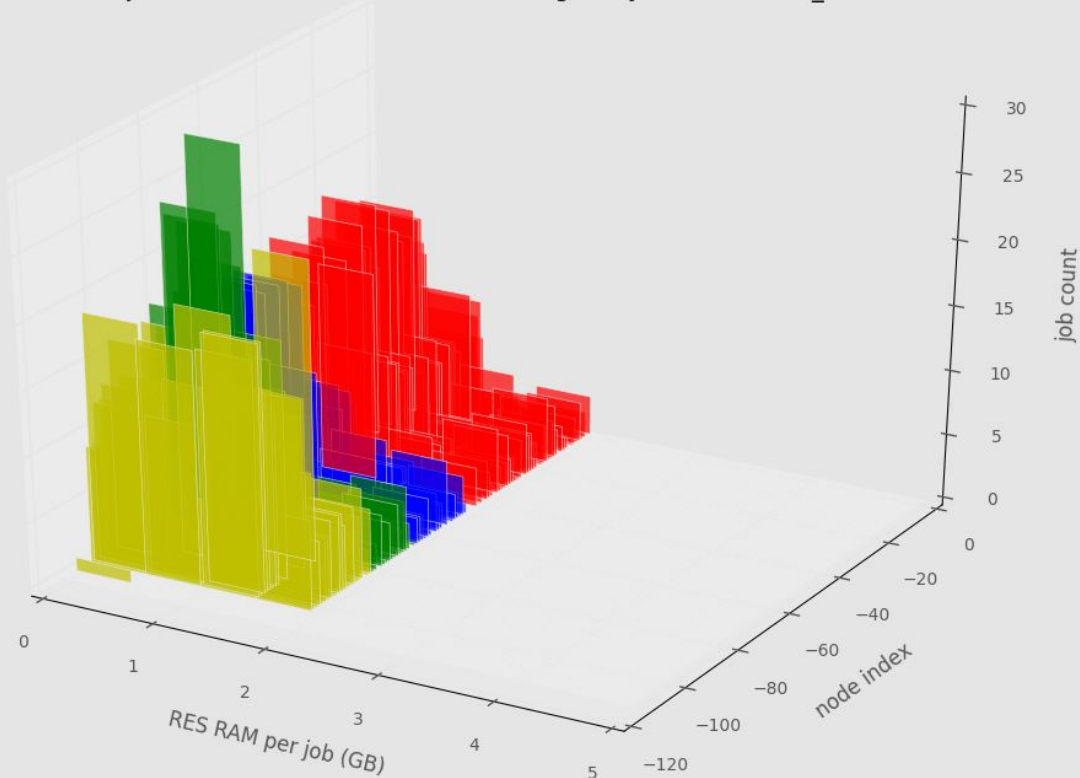
Load on UGE worker nodes 2017-08-08_09.01



SLURM nodes

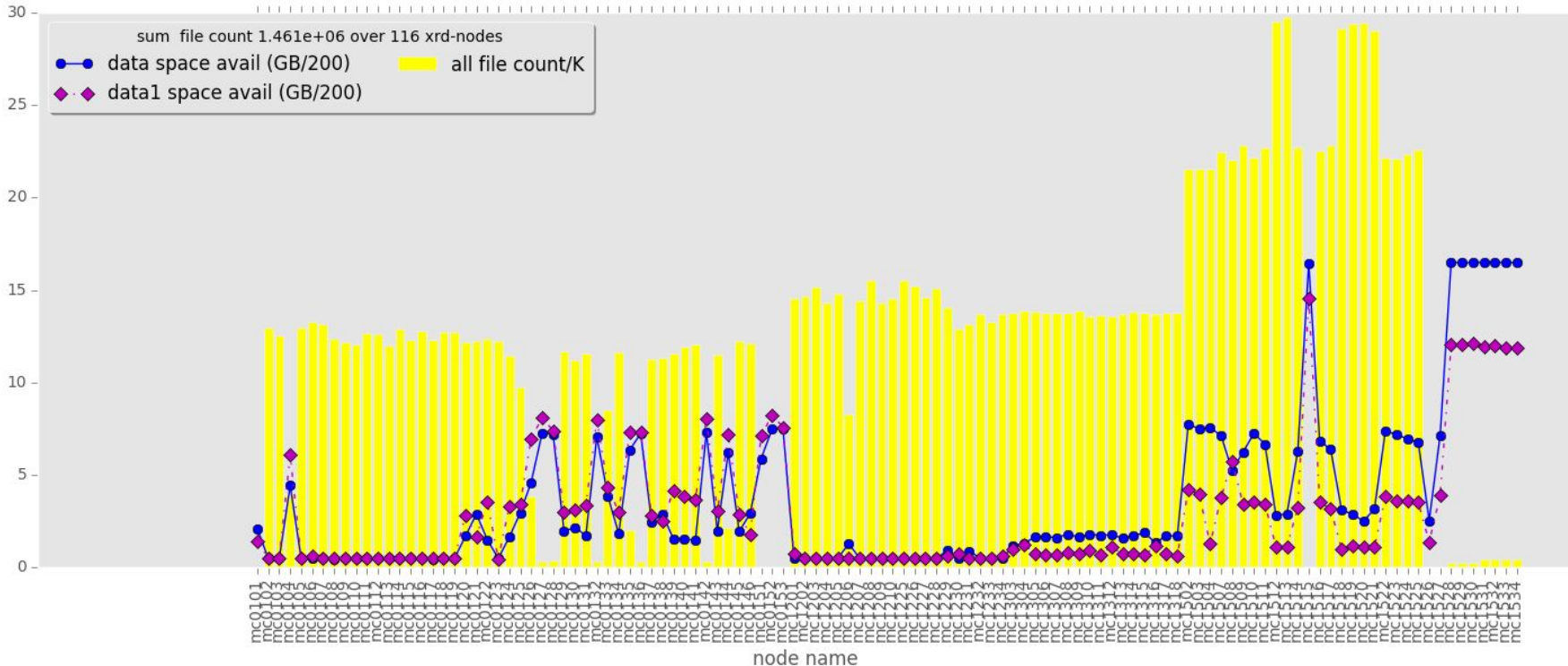
Mendel RAM/task usage - snapshot

RAM/job, rack colors: mc01=r,12=b,13=g,15=y 2017-08-08_09.01



r	qw	Project
1242	0	alice
159	1080	atlas
58	417	lux
2	0	majorana
1307	5357	star
2768	6854	Totals

XrootD availability - snapshot



/project(a) utilization - snapshot

<http://portal.nersc.gov/project/star/jthaeder/diskUsage/overview/indexExt.html>

<https://my.nersc.gov/data-mgt.php>

pdsf6 \$ prjqquota dayabay

Project	Usage	Quota	InDoubt	Usage	Quota	InDoubt
dayabay	794759	870400	69	135478754	150000000	558

pdsf6 \$ prjaquota dayabay

Project	Usage	Quota	InDoubt	Usage	Quota	InDoubt
dayabay	706679	716800	0	2292741	10000000	0

pdsf6 \$ prjqquota majorana

Project	Usage	Quota	InDoubt	Usage	Quota	InDoubt
majorana	38869	40960	0	2345664	4000000	0

pdsf6 \$ prjaquota majorana

Project	Usage	Quota	InDoubt	Usage	Quota	InDoubt
majorana	57901	61440	9	4445767	10000000	495

FillStatus (Quota): **PROJECT** (2017-08-07 13:08)

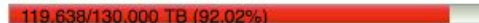
star - size



star - inodes



starprod - size



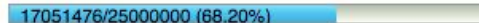
starprod - inodes



alice - size



alice - inodes



FillStatus (Quota): **PROJECTA** (2017-08-07 13:08)

starprod - size



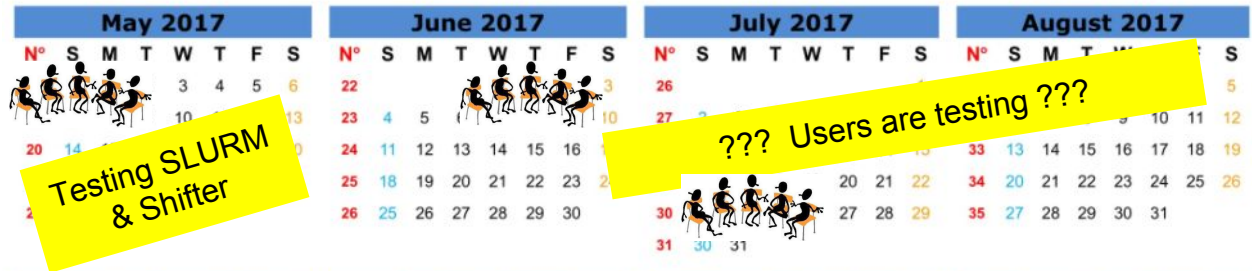
starprod - inodes



Set date for SLURM at PDSF

Please ask your users to test SLURM+CHOS at scale.

2017 Calendar



UNIVA vs. SLURM tutorials



univa vs. slurm



All

Images

Videos

Shopping

News

More

Settings

Tools

About 3,490 results (1.04 seconds)

SGE vs SLURM comparison - Uppsala Multidisciplinary Center for ...

www.uppmax.uu.se/support/user-guides/sge-vs-slurm-comparison/ ▼

SGE vs **SLURM** comparison. A guide comparing common commands in **slurm** and **sge**. Some common commands and flags in **slurm** and **sge**: ...

SGE to SLURM conversion | Stanford Research Computing Center

<https://srcf.stanford.edu/sge-slurm-conversion> ▼

Sun Grid Engine (SGE) and **SLURM** job scheduler concepts are quite similar. Below is a table of some common SGE commands and their **SLURM** equivalent. ... Also check out Getting started with **SLURM** on the Sherlock pages.

SGE vs Slurm

www.sdsc.edu/~hocks/FG/MSKCC.slurm.sge.html ▼

Univa Grid Engine development is led by CTO Fritz Ferstl, who founded the Grid Engine project and ran the business ... Comparison SGE vs **Slurm** comparison.

Few good tutorials:

<http://www.uppmax.uu.se/support/user-guides/sge-vs-slurm-comparison/>

<https://srcf.stanford.edu/sge-slurm-conversion>

<http://www.sdsc.edu/~hocks/FG/MSKCC.slurm.sge.html>

<https://www.sdsc.edu/~hocks/FG/PBS.slurm.html>

SLURM @ PDSF

<https://www.nersc.gov/users/computational-systems/pdsf/using-slurm-pdsf-batch-sbatch/>

Alternative: <https://bitbucket.org/balewski/tutornersc/src/master/2017-05-pdsf3.0/>

The screenshot shows the NERSC website interface. At the top left is the NERSC logo with the tagline 'Powering Scientific Discovery Since 1974'. A search bar is located at the top right. The navigation menu includes: HOME, ABOUT, SCIENCE AT NERSC, SYSTEMS, FOR USERS (highlighted), NEWS & PUBLICATIONS, R & D, EVENTS, LIVE STATUS, and STAFF. The 'FOR USERS' sidebar menu lists: Live Status, User Announcements, My NERSC, Getting Started, Connecting to NERSC, Accounts & Allocations, Computational Systems, Cori, Edison, PDSF, and About. The main content area shows a breadcrumb trail: Home » For Users » Computational Systems » PDSF » Using SLURM PDSF Batch - sbatch. The main heading is 'USING SLURM PDSF BATCH - SBATCH' in green. Below it, the text reads: 'PDSF-specific SLURM commands in a nut-shell'. A sub-section states: 'To execute in SLURM the equivalent of 'qsub jobscript.sh' do'. A code block shows the following commands and output:

```
$ ssh -X pdsf.nersc.gov
$ module load slurm
$ sbatch -p shared-chos -t 24:00:00 jobscript.sh
>>>Submitted batch job 102992
```

SLURM - outstanding high priority issues

Issue	Priority 1=high	Imparies	Experiment	status	Description
ganglia	2	SLURM	all	Known cause	Jobs running under SLURM are not reported on ganglia PDSF monitor page, ticket INC0103269
Ignored swap space	2	SLURM	all	discussed	SLURM is not seeing swap space on nodes, either partition (chos/shifter) INC0103391
OSG2	2	SLURM+grid jobs	ALICE, some STAR	Under eval.	deployment of 2nd OSG talking to SLURM on node mpdsfgrid02, will look like ...01. Jeff & James
AliEn VOBox2	2	SLURM	ALICE	In progress	grid jobs submission, monitoring. Need 2nd system for SLURM
Bad tcsh	3	Shifter	STAR	Under eval.	SLURM + Shifter is not working on PDSF for users who have the default NIM shell set to tcsh, ticket INC0101446
Smooth image selection	3	shifter	all	Under eval.	A user logs into PDSF and their shell is run inside the preselected Shifter image without any action on their part.

Multi-thread jobs on PDSF

When run code be aware if it is multi-thread

documentation for the package it looks like the default is
"maximum number of threads available if not set"

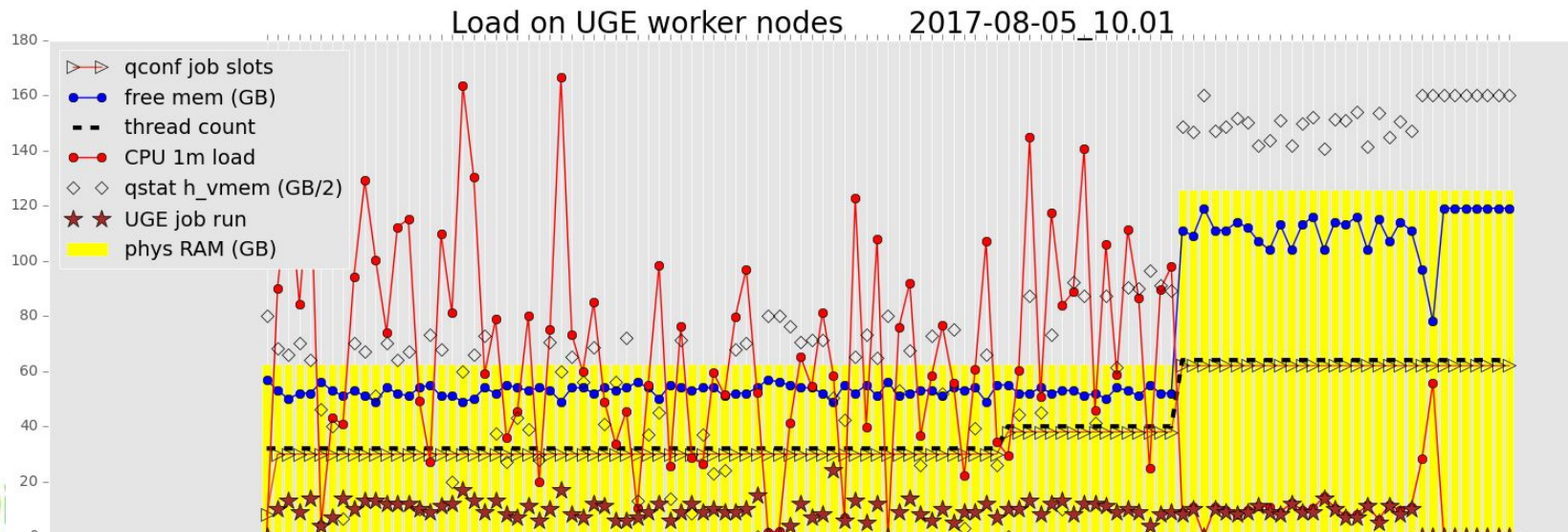
Not submit as 1 core UGE tasks

- MadGraph
- Nix
- Machine Learning: BDTs, tlearn

Solution in SLURM

#submit one task running on 32 vCores

```
sbatch -p shared-chos -t 24:00:00 -n32 jobscript.sh
```



Announcements

Bi-weekly office hours 12:30 -2:30pm
Thursday, August 17, 31, 59-4016-CR

PDSF user meeting

- Tuesday, September 12, **11:15am**, 59-3034-CR

Outages :

Cori maintenance: today

- There will be a training session on SLURM & Cori on 8/3, 1:00 - 3:00 pm in 149C. Zoom will be available for remote participants. [Please sign up if you plan to attend.](#)