

The (Not-so) Virtual Reality of OSG on Blue Waters, Comet, and Jetstream

Open Science Grid All Hands Meeting 2017

7 Mar 2016

Edgar Fajardo

On behalf of OSG Software and Technology



Working in Blue Waters



What my friends
think I do



What Instagram
thinks I do



What I think I do



What my boss thinks I do



Blue Waters by the numbers

System Component	Specs
Number of CPU Cabinets	237
Computes nodes per rack	96
Cores per Node	16 x AMD 6276 "Interlagos" processors 16 core 2.3GHz
Ram per Node	64 GB
Total number of Cores	362400



How to submit to Blue Waters?

Glidelns by Hand:



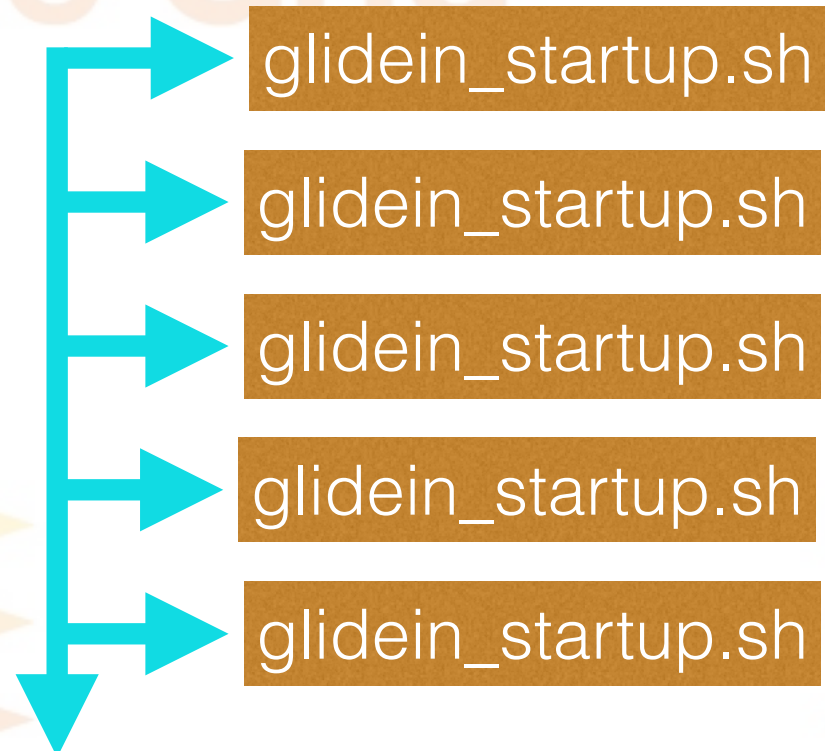
`glidein_startup.sh`

User



Login
Node

n times (where n is usually 10)



Because of the two factor authentication



How to submit to Blue Waters

- Still a “fake entry” is needed on the factory side.
- Then a “well configured” glidein_startup.sh is placed on the login nodes like:

```
exec $PBS_O_WORKDIR/glidein_startup.sh \  
-web http://glidein-1.t2.ucsd.edu/factory/stage \  
-sign a191bba36bd9ddb8e4eb4b5aeef1648e2d14200f \  
-signentry f8b022a148f33cf8ff00aac03582bd28475f479f \  
-signtype sha1 \  
-descript description.gbsehC.cfg \  
-descriptentry description.gbsehC.cfg \  
-dir OSG \  
-param_GLIDEIN_Client osg-ligo-1-t2-ucsd-edu_OSG_gWMSFrontend.blueWaters \  
-submitcredid 289405 \  
-slotslayout fixed \  
-clientweb http://osg-ligo-1.t2.ucsd.edu/vofrontend/stage \  
-clientsign 40d0c7dd61e2e4f605afcd02b00a535c38c9ac57 \  
-clientsigntype sha1 \  
-clientdescript description.gbsd47.cfg \  
-clientgroup blueWaters \  
-clientsigngroup dd0972166f1d07040589445da8cf93b28f8abb62 \  
-clientdescriptgroup description.gbsd47.cfg \  
-clientwebgroup http://osg-ligo-1.t2.ucsd.edu/vofrontend/stage/group_blueWaters
```



But the OS is SUSE: Solution: Shifter (aka Docker)

```
#!/bin/bash
#PBS -N testjob-shifter.Edgar.ligo
#PBS -v UDI=efajardo/centos6:osg-wn-client-v1
#PBS -l nodes=1:ppn=1
#PBS -l gres=ccm%shifter
##PBS -l walltime=06:00:00

module load shifter
mount | grep /var/udi
export CRAY_ROOTFS=UDI

cd $PBS_O_WORKDIR
mkdir -p /scratch/sciteam/$USER/$PBS_JOBID
export SCRATCH=/scratch/sciteam/$USER/$PBS_JOBID
aprun -n 1 -N 1 ~/edgar_tests/test_script.sh < input.data > output-shifter.$PBS_JOBID 2>outerr-shifter.
$PBS_JOBID
```



Achievements

- Run simple jobs inside the container, inside the pilot from a LIGO submit host.
- Access CVMFS through Parrot



Pending Problems:

- Pegasus seems to get stuck with Parrot. Possible solution: try David Lesny container with CVMFS without Parrot
- Automate the submission. Possible solution: Bosco may offer some hope with gsissh and a long lived proxy.



From Blue Waters to Comet



Update from last year's AHM presentation: OSG rides a Comet.

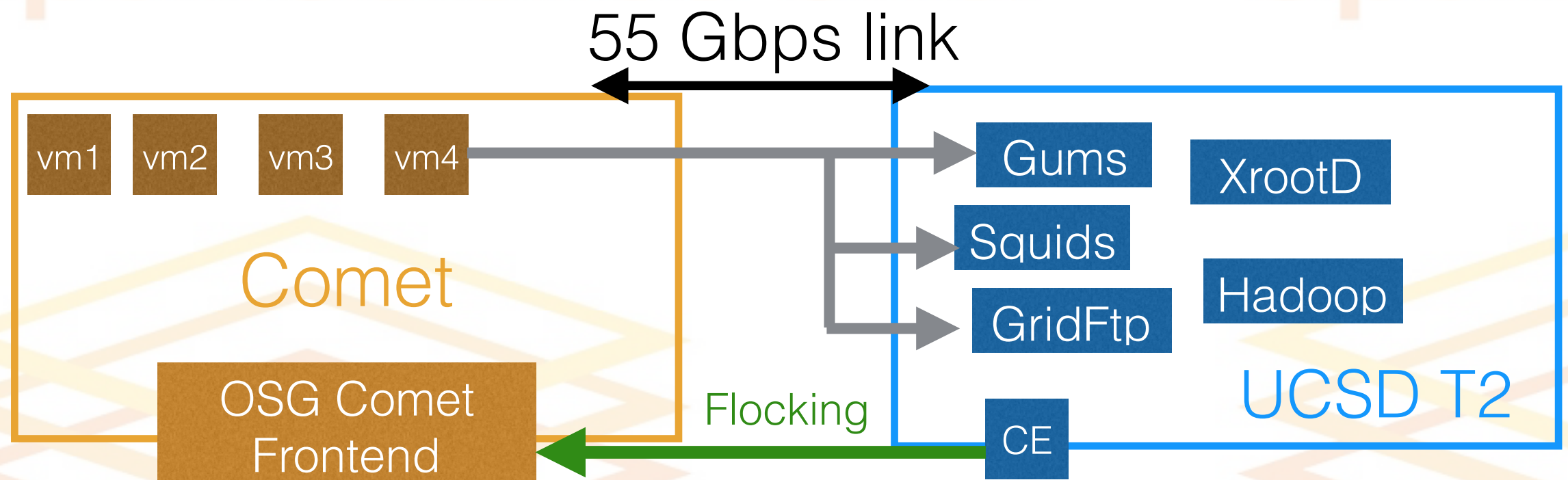
Last Year on Comet

- Running behind a NAT (limited to 1 Gbps)
- Using Comet rack dev opportunistic resources
- Only LIGO and OSG tested
- Not able to consume an allocation.



Where does OSG kick in?

Glideins can get into Comet using the already existing
UCSD T2 grid infrastructure

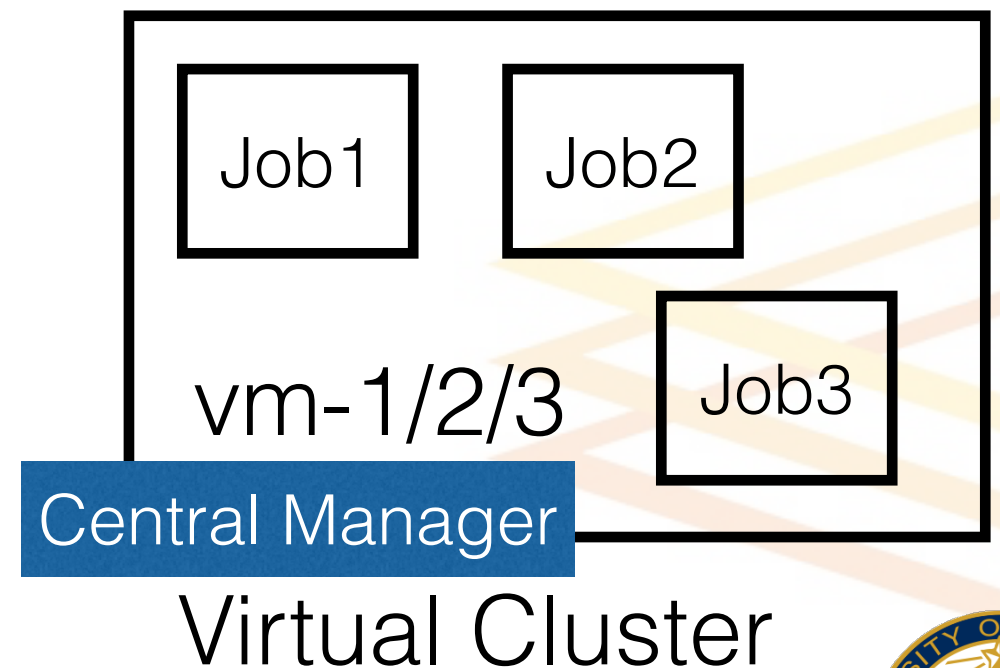


How Comet/OSG integration works



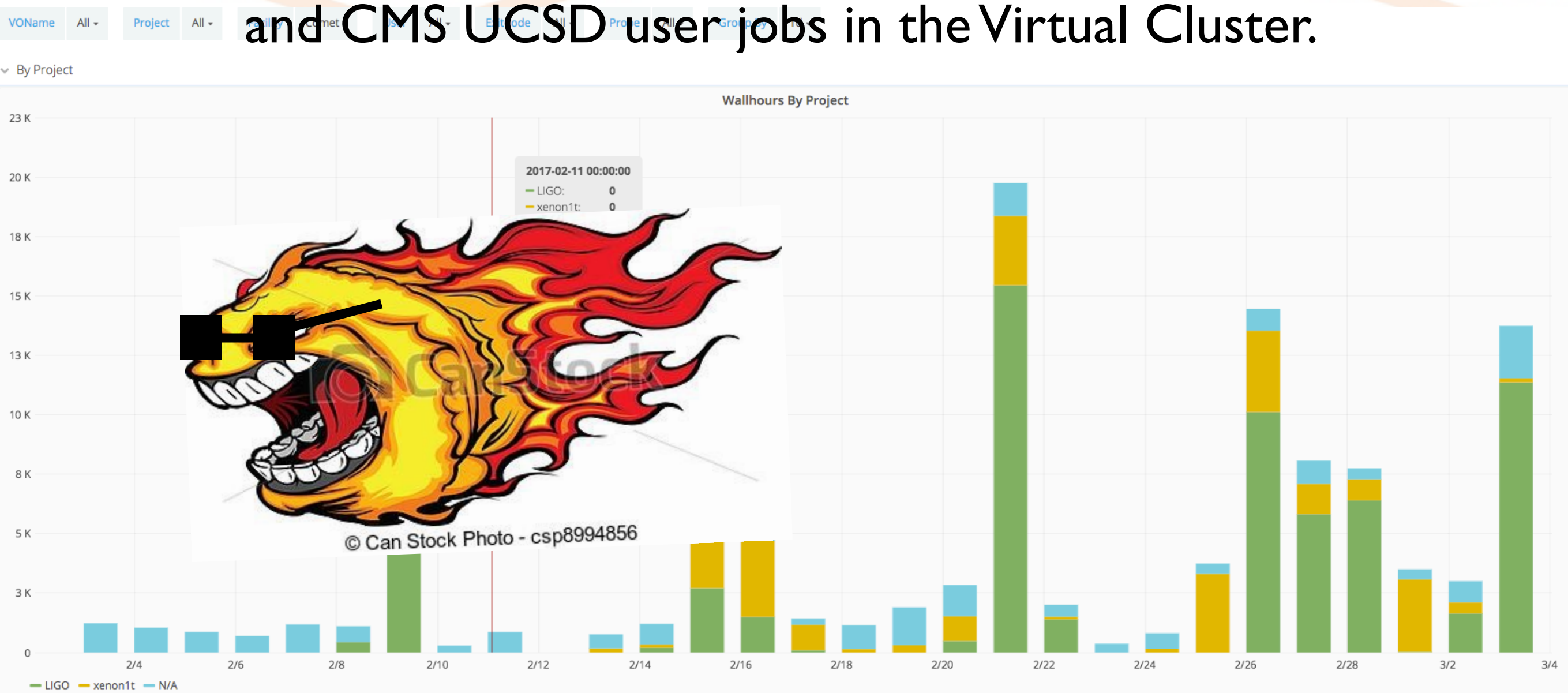
HTCondor -CE

- job1: +project_Name="allocation1"
+CometOnly=True
- job2:
+project_Name="allocation1"+CometOnly=True
- job3:
+project_Name="allocation1"+CometOnly=True



Achievements

- Successfully ran LIGO, Xenon1T, CMS Production and CMS UCSD user jobs in the Virtual Cluster.



Action items from last AHM

See slide 13 on last year's talk.

Short Term:

- ~~Spin up VM's given an allocation. Making sure only glide ins with that allocation run there.~~
- ~~Move to the production infrastructure (no longer behind a NAT).~~
- ~~Try to backfill flock CMS glideins to Comet.~~
- ~~Mount some lustre filesystem based on the allocation.~~



Action items from last AHM

See slides 14 on last years talk.

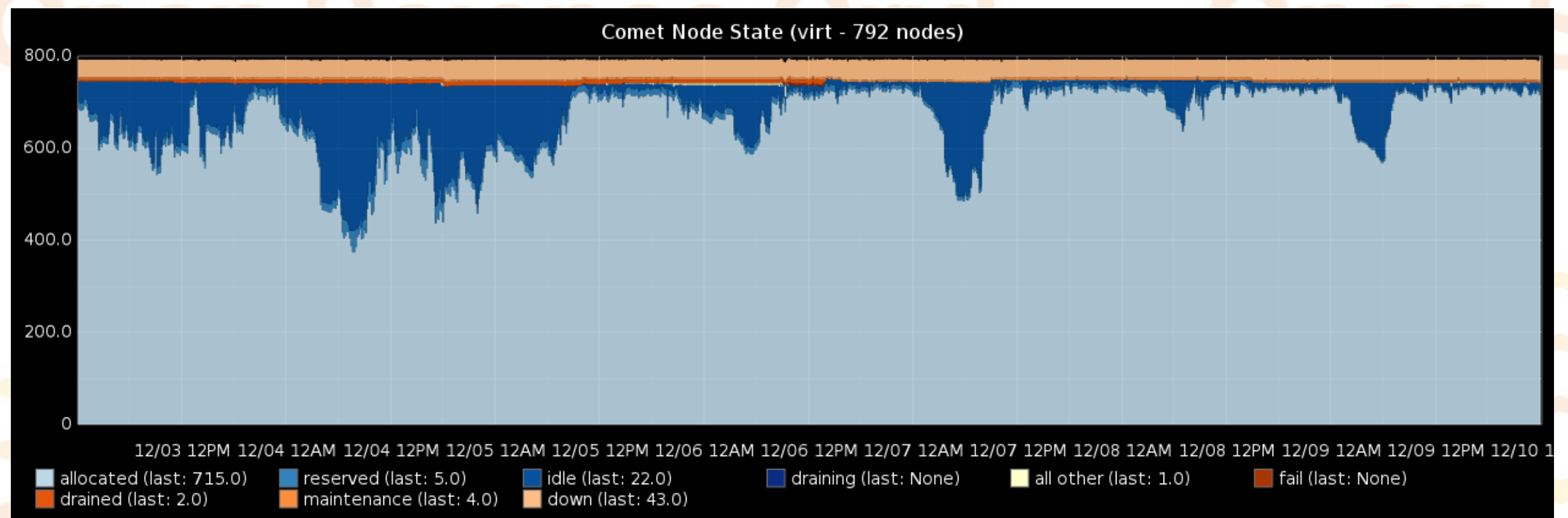
Long Term

- ~~Move to MultiCore~~
- Offer the possibility of a glidein taking over a whole virtualized rack. Multinode pilot (like Blue Waters).
- GPU access via the virtual interface. **Not gonna happen in Comet lifetime.**
- Backfill opportunistically
- Move beyond the 72 nodes limit right now for the Virtual Cluster.
- Figure out some other details when snapshotting.



Scavenged Used Cycles

OSG Comet Virtual Cluster would like to make use of *unused cycles*...
free science

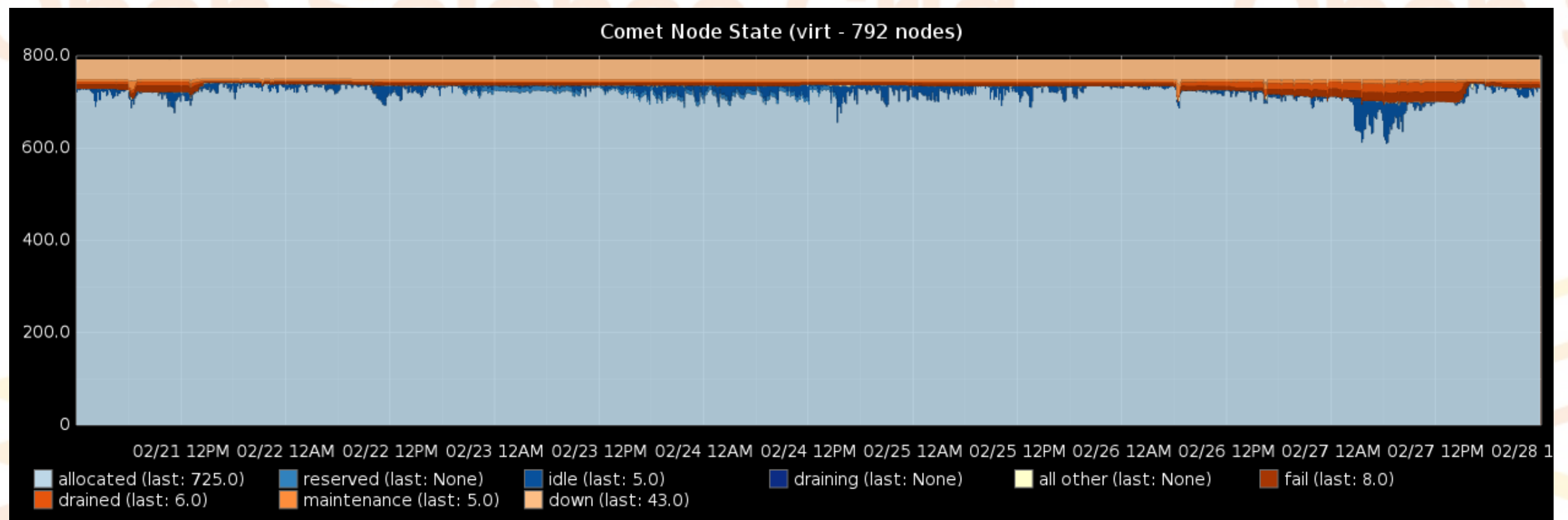


Comet available nodes shown in dark blue... 7 days in December 2016



Scavenged Used Cycles

OSG Comet Virtual Cluster would like to make use of *unused cycles*...



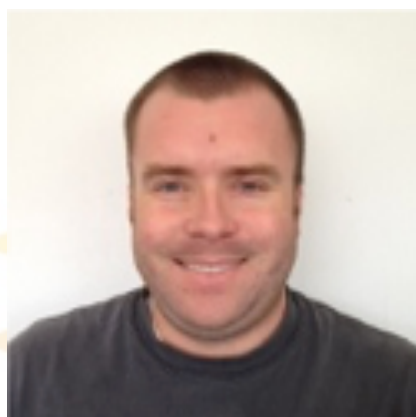
Comet available nodes shown in dark blue... 7 days in February 2017... where did they all go?



One More thing:

JET STREAM Integration:

Thanks to Marty Kandes (UCSD) for the slides:

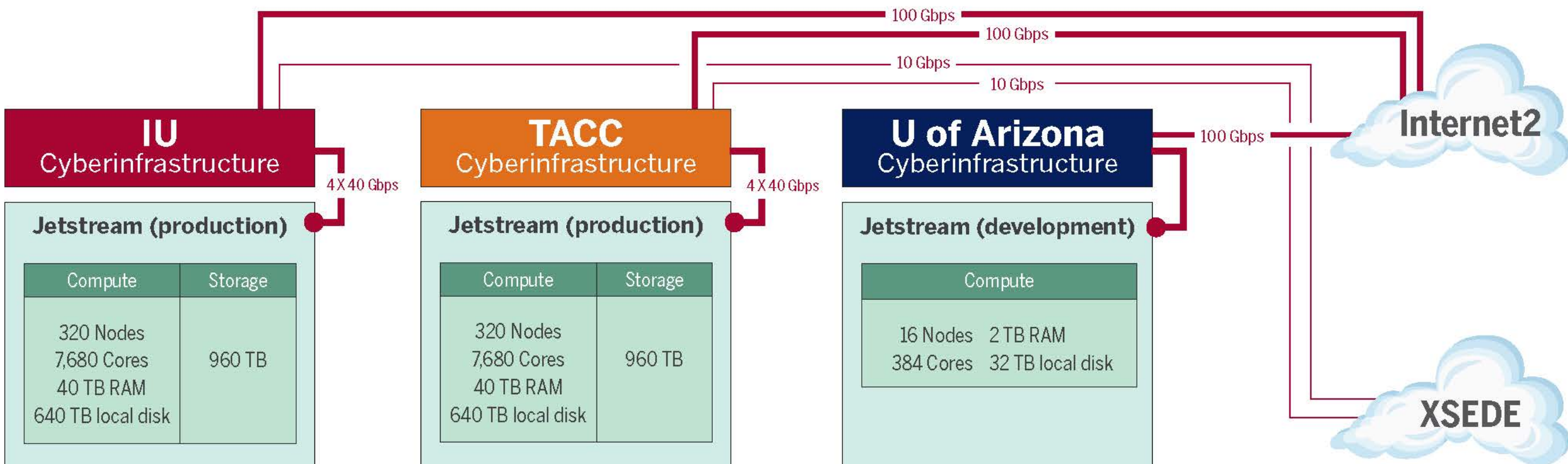


Jetstream





- First NSF-funded **cloud environment** designed to give researchers access to interactive computing and data analysis resources on demand.
- Distributed **Openstack-based infrastructure**; 0.5 PetaFLOPS
- **Jetstream team has offered to provide OSG with opportunistic usage when system load is low.**



OSG on *Jetstream*

Initial configuration attempts to follow **standard OSG model**.

- Glidein submission to an HTCondor-CE
- Local HTCondor Pool
 - Schedd + Central Manager running on same VM as CE
- Other supporting services: Squid, etc.

Developing **bootstrapping script(s)** to automate image builds and configuration, which should help facilitate long-term/shared management of site.

Some **cloud-related configuration issues**:

- Public/private network interfaces.
- Multiple public/private hostnames per network interface; e.g., Openstack's Nova (compute) and Neutron (networking) services do not share consistent hostnames by default.

Unknown: How to advertise size of available pool?



Acknowledgements

- Eliu Huerta (LIGO) and the whole team at Blue Waters.
- Trevor Cooper, Dmitry Mishin (SDSC) and the whole Comet team.
- Fugang Wang and Gregor von Laszewski (Indiana University) for the troubleshooting in the Comet Cloudmesh.
- Terrence Martin (UCSD) for the full integration setup and help debugging the network infrastructure at Comet Virtual Cluster.
- Mats Rynge, Rob Quick and Jeremy Fischer (Indiana University), Marty Kandes (UCSD).



Questions?

Contact us at:

1-900-OSG-HPC-Masters



Just Kidding

Contact us:

osg-software@opensciencegrid.org

Thank You

