A Lightweight Door into Non-Grid Sites

M Mascheroni[1], J M Dost[1], F Wuerthwein[1], I Sfiligoi[1], J Letts[1], E F Hernandez[1], R Gardner[2], L Bryant[2], J L Stephen[2], T Cartwright[3], B Lin[3], M Selmeci[3], D J Weitzel[4], B P Bockelman[5]

[1]UC San Diego, [2]University of Chicago, [3]University of Wisconsin-Madison, [4]University of Nebraska-Lincoln, [5]Morgridge Institute for Research

Motivation

Exponential growth in the amount of data collected in high energy physics and other sciences comes with corresponding growth in the need for computing resources. Allowing providers an easy way to share their resources is paramount to ensure the growth of resources is available to scientists. The Hosted CE initiative by the Open Science Grid provides the site administrator a way to reduce the effort needed to install and maintain a Compute Element (CE), the Grid portal to a compute cluster.

Traditional CEs

Batch-based cluster administrators typically deploy and operate the OSG Compute Element themselves on locally provisioned servers. This adds additional burdens due to:

- Knowledge of grid software
- An understanding of pilot authentication
- Human effort to configure and maintain the software

Hosted CEs

The Hosted CE is installed and maintained by OSG staff. This offloads the complexity and responsibility of maintenance from the site administrator to grid experts.

The site must simply provide an SSH login. From their point of view the Hosted CE users are like any other user login, just with special generic names, osg1, osg2…osg20 that map to different VOs.

OSG Staff takes care of the rest!

How to Participate[1]

Fill in a form to provide detailed contact and cluster information. Then OSG will contact you and ask you to grant access to a specific SSH key to your campus cluster head node. OSG will host and operate on your behalf a VM that runs the HTCondor-CE software! Also, WN client, managements of CA certs and CRL is automatically handled and nothing needs to be installed on the WN for the simplest use case (OSG jobs).

1: https://opensciencegrid.org/docs/compute-element/hosted-ce/

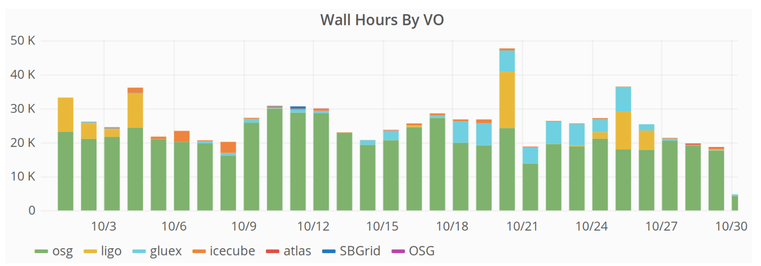
Support

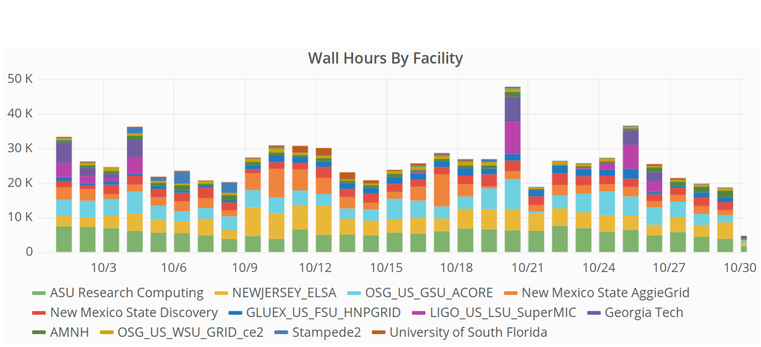
STATE	HOST	ICONS	OK	WA	UN	CR	PD
UP	hosted-ce07.grid.uchicago.edu		22	0	0	0	0
UP	hosted-ce13.grid.uchicago.edu		21	0	0	1	0
UP	hosted-ce16.grid.uchicago.edu		22	0	0	0	0
UP	hosted-ce20.grid.uchicago.edu		22	0	0	0	0
UP	hosted-ce23.grid.uchicago.edu		20	0	0	0	0
UP	hosted-ce28.grid.uchicago.edu		20	0	0	0	0

Tools to automatically detect issues with CE have been developed. OSG staff is in charge of communications with site administrators in case of problems.

Operations uses Checkmk based tools to:

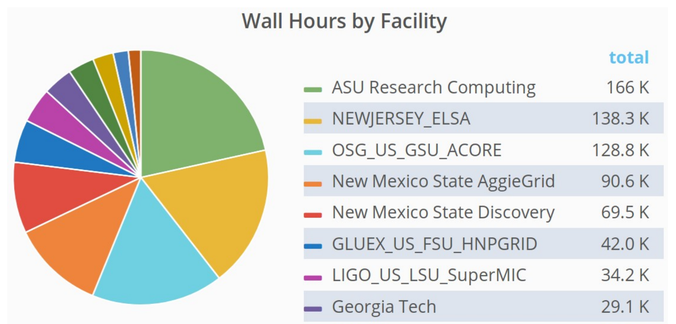
- Monitor health of the CEs
- Receive alerts when things go wrong
- Generate monthly availability email reports

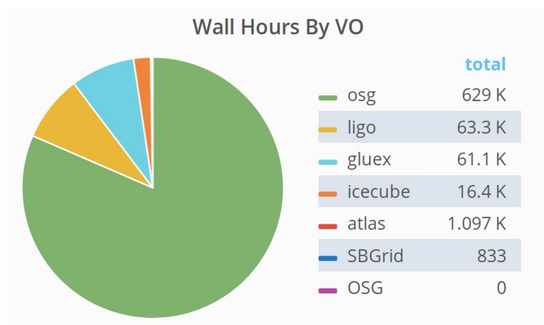


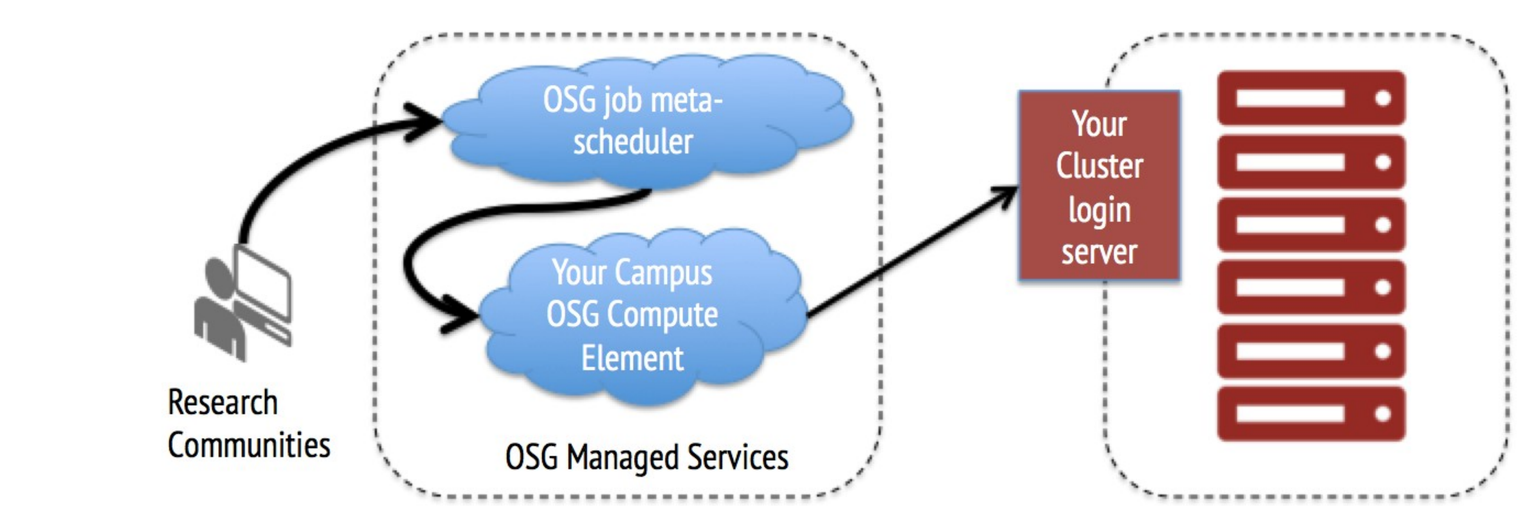


Site Requirements

- Existing compute cluster with a supported batch system running on a supported operating system (HTCondor, LSF, PBS / TORQUE, SGE, and Slurm)
- Red Hat Enterprise 6 and 7 and compatible platforms, for 64-bit Intel architectures
- Outbound network connectivity from the compute nodes (they can be behind NAT)
- Shared file system between the cluster head node and the compute nodes
- Temporary scratch space on each worker node







Scalability Improvements

Joint effort between the OSG Software and Operations teams has improved the management of the Hosted CEs by implementing tools to:

- Track Hosted CE configuration in version control
- Manage site-specific configuration and hot fixes
- Automate worker node data and software updates

Conclusions and future directions

With the Hosted CE initiative, the hardware/software stack needed to operate a CE is maintained by OSG Operations staff in a homogeneous and automated way. This provides a reduction in the overall operational effort needed to maintain the CEs: one single organization does it in an uniform way, instead of each single resource provider doing it in their own way. Solutions that use containers to further reduce the operational footprint needed to maintain a set of CEs are being explored.

This material is based upon work supported by the National Science Foundation under Grant No. NSF MPS-1148698