FERMILAB-CONF-15-618-CD

The OSG open facility: A sharing ecosystem

B Jayatilaka[1], T Levshina[1], M Rynge[2], C Sehgal[1] and M Slyz[2]

[1]Fermi National Accelerator Laboratory, Batavia, IL 60510, USA

[2]University of Southern California, Information Sciences Institute, Marina del Rey, CA 90292, USA

E-mail: boj@fnal.gov

Abstract. The Open Science Grid (OSG) ties together individual experiments' computing power, connecting their resources to create a large, robust computing grid; this computing infrastructure started primarily as a collection of sites associated with large HEP experiments such as ATLAS, CDF, CMS, and DZero. In the years since, the OSG has broadened its focus to also address the needs of other US researchers and increased delivery of Distributed High Through-put Computing (DHTC) to users from a wide variety of disciplines via the OSG Open Facility. Presently, the Open Facility delivers about 100 million computing wall hours per year to researchers who are not already associated with the owners of the computing sites; this is primarily accomplished by harvesting and organizing the temporarily unused capacity (i.e. opportunistic cycles) from the sites in the OSG. Using these methods, OSG resource providers and scientists share computing hours with researchers in many other fields to enable their science, striving to make sure that these computing power used with maximal efficiency. We believe that expanded access to DHTC is an essential tool for scientific innovation and work continues in expanding this service.

1. Introduction

As computing needs of large high energy physics (HEP) experiments increased in the late 20th century, the collaborations behind these experiments increasingly turned towards distributed computing to meet these needs. The Large Hadron Collider (LHC) experiments were designed from the outset to depend on distributed high throughput computing (DHTC) across far-flung computing grids to meet their computational requirements. The Open Science Grid (OSG) [1] is a North American DHTC grid that is a key component of the wider LHC computing grid. Since its formation in 2006, the OSG has not only been a cornerstone in the computing efforts of the LHC experiments but has also provided an increasing volume of opportunistic resources to users from a variety of research disciplines via the OSG Open Facility.

2. The Open Science Grid

At present, the OSG consists of over 120 individual computing elements across over 50 institutions spanning North America. These institutions are a mix of universities and US national laboratories. While sites associated with LHC experiments—the US Tier-1 and Tier-2 sites for both ATLAS and CMS and a number of Tier-3 sites—make up a bulk of the resources on the OSG, an increasing number of university clusters with no direct association with LHC experiments are becoming constituents of the OSG. As with most grid computing infrastructures, the OSG uses the Virtual Organization (VO) trust model, where individual users are associated

Operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

with one or more VOs. Most VOs operating on the OSG reflect large physics experiments (*e.g.* ATLAS and CMS) as well as university communities. In the past twelve months, the OSG delivered over 800 million CPU-hours of computing to its users.

2.1. The evolving user base of the OSG

A summary of computational hours delivered across the OSG to all VOs over the the past five years can be seen in Fig. 1. The overall growth from delivering 30 million hours a month in 2011 to 70 million hours in 2015 is apparent, as is the change in VO-landscape. In 2011, 85% of the usage of the OSG was by large HEP experiments (ATLAS, CMS, CDF, and DZero). By 2015, the proportional share of OSG resources used by these large experiments had dropped to 67% despite ATLAS and CMS having record usage of the OSG in 2014. These shifts are due to an increasing delivery of resources that might otherwise remain idle when not used by their owners, or opportunistic resources, to a wider base of users. These opportunistic users come from VOs such as hcc [2] and glow [3] (research communities at the University of Nebraska and University of Wisconsin, respectively), sbgrid (software consortium for structural biology) [4], and other VOs which primarily consist of individual researchers. In 2011, a new such VO, the OSG VO, was formed to reflect the OSG Open Facility going forward.

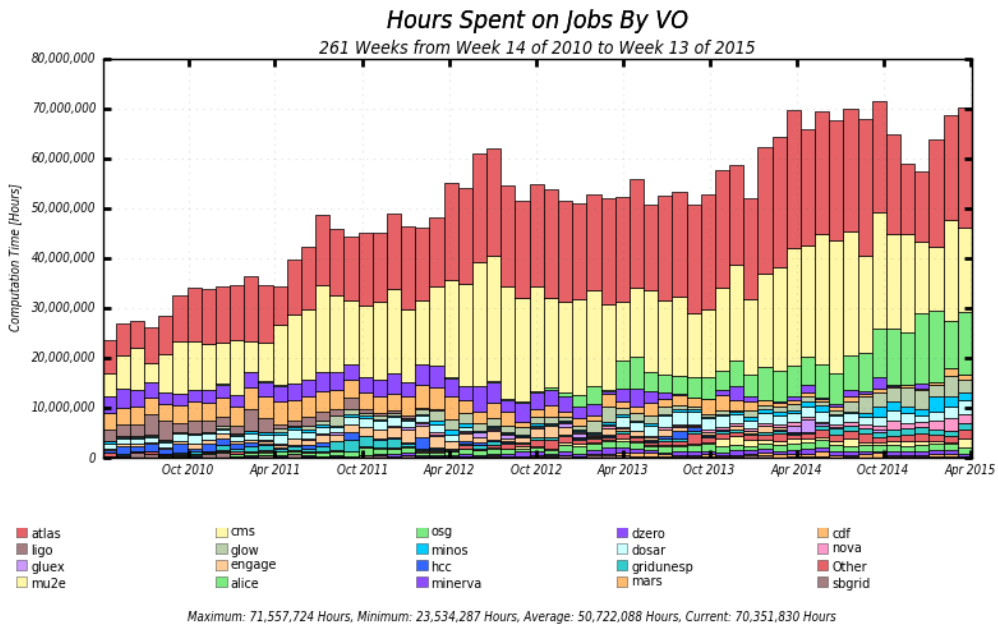

Figure 1. Computational hours on the OSG for the past five years by VO.

3. The OSG Open Facility

In 2011 an additional VO, simply called "OSG" (referred to as OSG VO to avoid ambiguity), was enabled on the OSG to increase access to opportunistic resources. Rather than being dedicated to a specific experiment or research community, the OSG VO was made available to any researcher at a US institution seeking access to DHTC resources. Conversely, researchers making use of the OSG VO did so fully understanding that the resources they were using were only opportunistic. Table 1 shows the growth in computational hours delivered to OSG VO users by calendar year since its inception in 2011; OSG VO users have utilized more than 100 million

hours in the past twelve months. Since 2011, the number of individual sites offering resources to users of the Open Facility has grown to over 30 (reflecting over 50 individual computing elements).

Table 1. Growth of the OSG VO in computational hours. For perspective, the total computational hours consumed by all VOs in the OSG are also presented.

Year	2011	2012	2013	2014
OSG VO (hours)	400,568	3,168,025	47,931,106	97,009,409
Total OSG (hours)	510,671,422	651,365,203	649,812,810	798,145,827

4. Resources available on the Open Facility

A majority of the sites accessible via the Open Facility are LHC computing sites, namely Tier-1 and Tier-2 sites for ATLAS and CMS. Large university-based distributed computing installations, such as the Holland Computing Center's at the University of Nebraska and Syracuse University's OrangeGrid also contribute a large volume of opportunistic resources to Open Facility users. The ten largest providers of resources to the Open Facility in 2014 are listed in Table 2. Sites primarily run Scientific Linux (or compatible flavors of Linux) and are Intel x86 architecture-compatible. User jobs tend to be single-threaded and utilize 2 GB of RAM or less, as the typical per-core RAM availability across the OSG is 2 GB. Most sites offer 30 GB of non-persistent storage per running job.

Table 2. The ten largest contributors to the OSG Open Facility by compute hours delivered in 2014.

Site name	Description	Compute hours
Fermigrid	Fermi National Accelerator Laboratory (Campus Grid gateway)	12,711,968
USCMS-FNAL-WC1	Fermi National Accelerator Laboratory (USCMS Tier 1)	11,073,515
UCSDT2	University of California San Diego (USCMS Tier 2)	10,713,436
CIT_CMS_T2	California Institute of Technology (USCMS Tier 2)	9,102,039
Tusker	The Holland Computing Center at the University of Nebraska-Omaha (HCC Campus Grid)	7,490,094
Nebraska	The Holland Computing Center at the University of Nebraska-Licoln (USCMS Tier2)	6,680,426
MWT2	University of Chicago, Indiana University, University of Illinois, Urbana-Campaign (USATLAS Midwest Tier 2)	6,219,981
MIT_CMS	Massachusetts Institute of Technology (USCMS Tier 2)	4,585,284
Purdue-Hadoop	CMS Hadoop cluster at Purdue University	2,385,959
NWICG_NDCMS	University of Notre Dame (USCMS Tier3)	2,281,689

5. Access to the Open Facility

Initially, access to the Open Facility was primarily obtained by users logging interactively into a submit host which runs the Condor batch system. From here, user jobs would be submitted to central flocking node which handles negotiation and submission to individual sites across the wider OSG. This mode of submission is referred to as "OSG-Direct" submission; a number of institutions (*e.g.* Duke, Virginia Tech) have their own submission nodes which also submit jobs to the central flocking node. While the OSG-Direct mode of submission is deprecated for new users, approximately 45 million compute hours per year are still delivered to users utilizing it. New users of the Open Facility are encouraged to use one of two access methods: via XSEDE allocation ("OSG-XD") or via the OSG Connect model for individual researchers ("OSG-Connect")

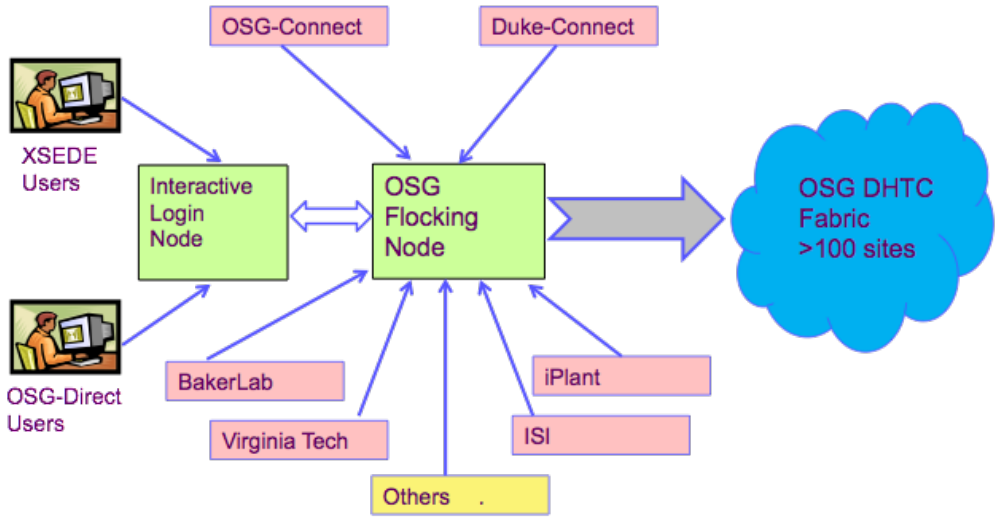

Figure 2. On-ramp to the OSG Open Facility.

5.1. OSG-XD

XSEDE [5] is a US National Science Foundation initiative that enables access and training for use of large computing resources. Since 2013, the OSG has been one of the facilities on which researchers can have allocation grants awarded. Users who obtain these allocation grants gain access to the same login node used by OSG-Direct users and are able to submit jobs via the Open Facility to the OSG. Unlike other XSEDE allocations (at HPC sites), users accessing the OSG via XSEDE allocations are not limited to only the computing resources awarded in their allocation. OSG-XD users who have exceeded their allocation are allowed to continue using OSG resources, albeit at a lower priority than users who have not. In the past year, approximately 50 million compute hours were utilized by OSG-XD users despite the sum of their XSEDE allocations being only 5 million hours.

5.2. OSG-Connect

Similar to OSG-XD, OSG-Connect provides an easy to use, hosted virtual environment. User and group management is provided by Globus Nexus [6], and new users can sign up using their own campus identities using InCommon [7] or CILogon [8], or create a new identity if neither

of those are available. The user can then access submit hosts via SSH, Globus Online and the local Stash [9] service for data management, and other services using the same credentials. The OSG-Connect environment is designed to be similar to a campus or HPC cluster. For example, a large set of software is preinstalled on the read-only distributed CVMFS filesystem, and the software is managed by the module tool so that users can load and unload modules and get the environment required by their codes. Compared to OSG-XD, OSG-Connect does not require allocations. Instead it is required that a user belongs to a project which has been set up and approved by OSG-Connect. Access to resources are opportunistic and scheduled using HTCondor's fair-share algorithm. OSG-Connect is a new offering in the OSG Open Facility ecosystem, but has already provided access to 135 users belonging to 42 projects, who used 9.7 million core hours during the last year.

5.3. Authentication and Authorization

One notable improvement which the OSG Open Facility has made is the discontinuation of requiring GSI [10] X.509 based proxies for user authentication and authorization. Users can now come in to the facility using an authentication method which makes sense to the individual user, such as using their own campus username and password, and will not be asked to request and maintain additional credentials such as X.509 certificates and proxies. This change has lowered the barrier of entry substantially, reducing both the time to get set up and the amount of necessary support interaction. Users who wants to access data on GSI authenticated systems can still obtain an X.509 user certificate, but it is entirely optional.

6. Usage of the Open Facility

Usage at the Open Facility is organized into projects. In the case of OSG-XD usage, each project corresponds to a specific XSEDE project, and thus, allocation. In other cases, each project typically corresponds to an individual PI's research, often with research staff of that PI also having access to the same project. Currently there are over 150 projects registered to use the OSG Open Facility, of which 39 are OSG-XD projects. Over the past year, 30-40 projects are active each month. Nearly 100 projects utilized the open facility at some point in 2014, and, as seen in Fig. 3, reflected a wide variety of applications. Furthermore, 23 peer-reviewed publications in 2014 resulted from work that utilized the OSG Open Facility.

7. Future prospects and challenges

While the overall OSG and the Open Facility continue to grow, the growth rate of the Open Facility is expected to shrink in 2015. This is mainly due to an expected decrease in opportunistic availability at ATLAS and CMS computing sites, which continue to be the largest source of resources on the OSG, as collisions on the LHC resume later this year. Providing support for users as well as identifying suitable new users continues to be a challenge. While VOs that reflect larger experiments or research communities often either have institutional experience using DHTC or internalized support for computing, many Open Facility users have neither. Another ongoing challenge is data movement. While large VOs such as ATLAS and CMS have federated data storage schemes that minimize the need to move data from site to site, the OSG VO currently has no such scheme. Consequently, users of the Open Facility can only practically move O(10 GB) of data to and from their job. An XRootD-based caching solution, Stash [9], is currently being developed which, when deployed, should allow for effective movement of O(1 TB) per job for Open Facility users.

8. Conclusions

The OSG Open Facility offers researchers from a variety of scientific disciplines a simple and efficient access to distributed high throughput computing resources. Not only does the

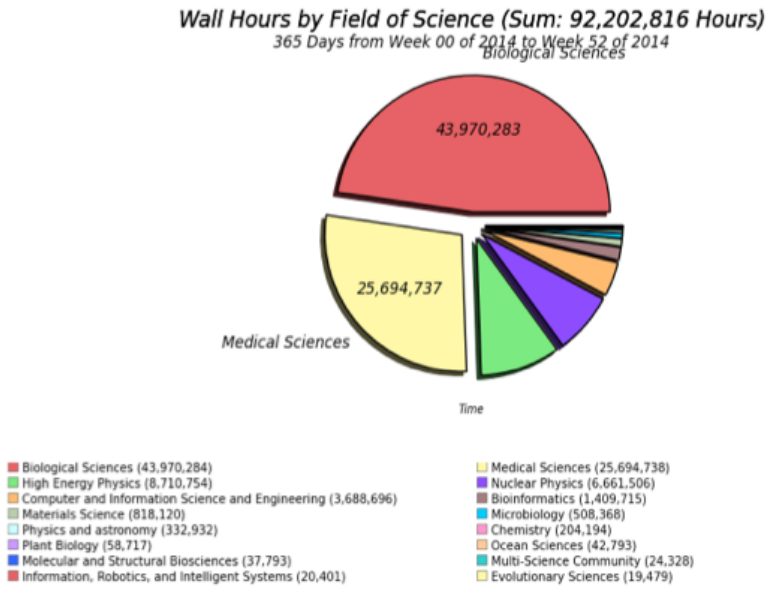

Figure 3. OSG Open Facility use in 2014 broken down by field of science.

facility provide user documentation, examples, and support staff, but the facility provides a common point for those services, which is invaluable when it comes to present a distributed grid infrastructure such as OSG as a coherent platform to user with no or little distributed computing knowledge. Even though the OSG Open Facility has proven to be a great step forward in usability, we know that the work is not complete. To increase, or at least maintain, the current level of cycle availability, it will be important to keep incorporating new non-LHC sites into the OSG. This is especially important now when the LHC is starting to produce data again, and opportunistic cycle availability at LHC sites will probably decrease. There is also always more work to be done in introducing new user communities to the OSG, and support existing communities and relationships, and the facility team will continue to work supporting science advancement via distributed high throughput computing and the power of resource sharing.

References

[1] Pordes R *et al. Journal of Physics: Conference Series* **78** 012057

[2] 2015 Holland computing center [Online; accessed 15-May-2015] URL `http://hcc.unl.edu/`

[3] 2015 Glow :: Grid laboratory of wisconsin [Online; accessed 15-May-2015] URL `http://research.cs.wisc.edu/htcondor/glow/`

[4] Morin A, Eisenbraun B, Key J, Sanschagrin P C, Timony M A, Ottaviano M and Sliz P 2013 *eLife* **2**

[5] Towns J *et al.* 2014 *Computing in Science and Engineering* **16** 62–74 ISSN 1521-9615

[6] Ananthakrishnan R, Bryan J, Chard K, Foster I, Howe T, Lidman M and Tuecke S 2013 *Cluster Computing (CLUSTER), 2013 IEEE International Conference on* pp 1–3

[7] 2015 Incommon: Security, privacy and trust for the research and education community [Online; accessed 15-May-2015] URL `https://www.incommon.org/`

[8] Basney J, Fleury T and Gaynor J 2014 *Concurrency and Computation: Practice and Experience* **26** 2225–2239 ISSN 1532-0634 URL `http://dx.doi.org/10.1002/cpe.3265`

[9] 2015 Osg connect stash [Online; accessed 15-May-2015] URL `http://stash.osgconnect.net/`

[10] Foster I, Kesselman C, Tsudik G and Tuecke S 1998 *Proceedings of the 5th ACM Conference on Computer and Communications Security* CCS '98 (New York, NY, USA: ACM) pp 83–92 ISBN 1-58113-007-4 URL `http://doi.acm.org/10.1145/288090.288111`