# The OSG Campus Infrastructures Community

## Background

OSG has had a number of activities related to HTC outreach, education and training that pertain in some fashion to the context of using or building campus infrastructures. These have either been in the form of summer school events [1], site administrator workshops [2], or in the case of Grid Colombia a full-scale “build-a-grid” curriculum which comprised two weeks from Unix how-to all the way to deploying GOC-like services [3]. The style of each was designed to fit the target audience: the OSG summer schools for example have focused on users and students, not infrastructure. The site administrator workshops on infrastructure and core OSG and network services, not use. In context of the Campus Grids activity the challenges are different as they involve **both** use (from the research domain engagement goal) and the necessary core infrastructure and services to effectively use resources on (and off) campus. To date the focus has been development of tools to access cycles, but over the course of OSG N5Y this will necessarily involve data (and other advanced services) in more integral ways.

There are examples of university community engagement of national cyberinfrastructure projects. In the HPC space, TeraGrid and now XSEDE have a very large Campus Champions [5] program which has effectively energized a community from more than 100 universities. Associated with this effort has been a regular set of training events and seminars leading to a body of documentation for new users, with a focus solely on use. In the networking space, Internet2 has fostered community building training events either through there long-standing Joint-Techs meetings, or other topical meetings such as campus network troubleshooting [6].

There are interesting approaches we can adopt from successful programs elsewhere. Part of the CIC program of work is to identify those methods which have found success and adapt/adopt as appropriate for the OSG context.

### Infrastructure versus Applications versus Access

Without question, our goal in OSG is to bring the advantages of HTC as efficiently as possible to domain end-users so as to optimize science. Clearly, our target is not solely IT personnel at campuses who may, depending on their local role and responsibilities, have limited contact (and therefore a sense of success-ownership) with research groups. Our mandate should not be to replicate what folks can get elsewhere, or to introduce a scope as fabric-centric as HEPiX [7] (though much of what is discussed there could find application in an HTC CIC), but to be driven by research needs from (O(100 universities x10 groups)) of research groups that can directly benefit from OSG inspired technologies and practices. To scale, it must have a strong community providing broad knowledge base of technical and research computing.

The CIC area therefore has been designed with these points in mind, and has the following main components:

* Development of a topical seminar series and forum highlighting concepts in the development and use of campus infrastructures
* Convening face-to-face meetings of the OSG CIC for both infrastructure providers and domain experts/leaders on campuses
* Development of a campus engagement program which programmatically develops ties between research domain experts, campus infrastructure providers and the CIC.
* Developing a program of engagement XSEDE.

## CIC Topical Seminar Series

These would be in the context of monthly campus grid teleconferences which would be re-formatted to include a top-of-the meeting “Campus Roundtable” where news from the campuses can be reported, followed by the seminar. In some cases we may facilitate in-person events at centers with strong OSG affiliation (following the XSEDE model). Part of the CIC program of work involves:

* Working with the campus grid team and others in OSG, identifying topical, compelling issues suitable for the campus infrastructures community
* Development of templates for practical “how-to” modules for domain experts and providers of campus infrastructure
* Development of materials in-house for applications and services with direct expertise from the campus grid team, as well as other core OSG staff
* Recruitment and development of materials from affiliated sources – many closely related to OSG (eg. Condor, Globus Online, Internet2 or ESnet, US LHC, etc.)
* Seminar materials preparation, logistics, advertising
* Facilitation of OSG CIC affiliated, in-person training events

An example format for advertised modules:

|  |  |  |  |
| --- | --- | --- | --- |
| OSG CAMPUS INFRASTRUCTURES SERIES | | | |
| **Enabling unified job management on your Campus using Bosco** | Intermediate | WEB  1 hour | Brookline Gore |
| ABSTRACT: In this tutorial, OSG campus grid infrastructure providers learn how to deliver easy-to-use job submission tools to their research communities using the Bosco toolkit. Based on Condor technology, Bosco provides a simple command interface that addresses the problem of coherent job management across multiple backend campus cluster resources. This tutorial will demonstrate how a campus can effectively use Bosco to simplify access for their end-users presenting in detail how to prepare multi-user submit hosts, with typical job submission examples.  WHO: campus infrastructure providers, intermediate HTC users  REQUIREMENTS: Submit host, target backend cluster which uses either of the PBS job scheduler (Torque or PBS-pro) or Condor. | | | |

Note a scheme for classifying the module according to skill level, time, web (versus a possible in-person event), presenter, intended audience and requirements. Web events could use Adobe Connect, which has the following capabilities:

* Simple web access for video and audio
* Built-in phone bridging
* Document sharing
* Desktop sharing

### Developing a CIC Topical Seminars list

We need broad input to develop a suitable program. Working with the campus grids team we will form a **CIC Topical Seminars Committee** to develop a first year “curriculum” the kick-start the CIC community. Associated with this effort will be deployment of an community organizational system incorporating tools for outreach, communication, content organization, and social networking. Here are some straw-man ideas for the first year:

1. HTC Fundamentals (XSEDE community would be interested)
2. Enabling unified job management with Bosco
3. Adapting Mathematica applications to your campus HTC environment
4. Network troubleshooting with PerfSONAR
5. Distributing user applications using CVMFS and Parrot
6. Adapting “R” applications to your campus HTC environment
7. Techniques for optimizing application IO performance
8. Adapting Matlab applications to your campus HTC environment
9. HTC and Genome sequencing
10. Using MapReduce in a campus grid
11. Bridging your campus with opportunistic resources of the OSG (infrastructure providers)
12. Accessing resources off-campus with GlideinWMS (users)
13. Unified DHTC computing with OSG and XSEDE
14. Enabling a sharing infrastructure on your campus with Globus Online Multi-User

An incentive for the series would be the visibility to those making contributions in terms of their professional development and as an employment networking opportunity.

## OSG CIC Face-to-Face Meetings

To boost attendance and limit travel for participants, we intend to co-locate these with meetings from affiliated communities. So in addition to the OSG All Hands, these could be US LHC related facilities workshops, OSG Summer Schools, perhaps joint XSEDE training events, Internet2 training events, tutorials associated with Supercomputing [8], etc. However we should not rule out, over the course of OSG N5Y, stand-alone workshops.

We have begun discussions with Jason Nielson to hold the first OSG CIC meeting at University of California, Santa Cruz, on November 14-15, co-located with the Fall US ATLAS Computing Facilities workshop.

To make workshops compelling and boost attendance, we have discussed having keynote speakers attend. For this we would need to budget travel and honorarium as appropriate.

We have discussed hosting dedicated “hackathon” or “data-challenge” type meetings in this context.

## OSG CIC Campus Engagement Program

We will work with US LHC Tier3 universities where in many cases there are natural ties to the OSG. In these locations there additionally is an obvious need for campus technologies such as Bosco to reach (in many cases substantial) non-program funded computing resources on campus. These are great locations to develop engagement conduits to researchers in other domains who have workloads suitable for HTC. A challenge here is the potential scale of the engagement activity (n.b. XSEDE Campus Champions has 159 members with committed responsibilities for XSEDE outreach across 100 campuses). Part of the CIC program of work in this context is to develop a regular process and system for maintaining engagement, development of engagement expectations, training materials, etc. Portfolios, for example, to chronicle issues, successes and progression of the engagement.

Another component of campus engagement is to provide summarize highlights of activities within the CIC for the OSG newsletter in a regular feature. This could include a sketch of the monthly topical seminar and other news reported in the Campus roundtable.

## OSG CIC and XSEDE Campus Champions

We need to understand how this activity works in two ways:

1. Our existing efforts at providing OSG resources as a virtual cluster to XSEDE users
2. Coupling to the effort/persons allocated by XSEDE to engage OSG

Campus researchers of course are driven to acquire resources anyway they can, and our goal is to facilitate that to every extent possible given our limited effort. Within OSG we will need to provide consistency among documentation and support efforts between the Users and CG areas with regards to presenting a consistent view to campuses. As part of CIC engagement of campuses, working with both the responsible within XSEDE for OSG liaison issues as well as directly with Campus Champions at universities where DHTC campus infrastructures are being developed will be key to helping them effectively use local campus, OSG and XSEDE resources.

## Supporting Infrastructure Ideas

Some random ideas for supporting infrastructure

* Creation of a campus infrastructures “geek centers”
* Google + page, hangouts
* Community forum and easy-to-use repository for seminar materials, recordings, postings for the CIC
* Adobe Connect for seminar meeting broadcast and recordings

## Milestones

Here are some initial milestones for the CIC sub-area of Campus Grids, first year:

1. Define the appropriate metrics for telling the campus story in OSG. We have discussed these in terms of:
   1. Making distributed high throughput computing easy, visible (awareness) and ubiquitous
   2. Finding the appropriate metric for measuring “presence” on campuses
   3. Capture science success stories, indicating the multiplicative effects of using campus and distributed HTC resources
   4. Classification of infrastructures with a maturity model
2. Establish the CIC Topic Seminar series as a staple for community building and knowledge sharing
3. Convene one face-to-face CIC meeting with a broad technical program compelling to the campus infrastructure providers and users
4. Promote community through use of a CIC resource center (social contacts, topical seminar materials, pointers to tools and guides)

## References

The 2012 OSG Summer School,  
 <https://opensciencegrid.org/bin/view/Education/OSGUserSchool2012>

OSG Site Administrator Workshops, <https://www.opensciencegrid.org/bin/view/SiteCoordination/WebHome>

Grid Colombia Workshop, <https://twiki.grid.iu.edu/bin/view/ReleaseDocumentation/GridColombiaWorkshop2010>

High Throughput Computing and Sequence Analysis, University of Arizona, <http://bcf.arl.arizona.edu/htc>

XSEDE Campus Champions, <https://www.xsede.org/web/xsede12/campus-champions>

Internet2 communities, <http://www.internet2.edu/communities/index.cfm>, and workshop series <http://events.internet2.edu/events-future.cfm?type=4>

See for example, the Spring 2012 HEPiX workshop, <https://indico.cern.ch/conferenceTimeTable.py?confId=160737#all.detailed>

Supercomputing 2012 tutorials, <http://sc12.supercomputing.org/content/call-tutorials>

CI days, <http://orci.research.umich.edu/news-events/cidays/ci-days-2011/plenary-presentations/> , <https://wiki.internet2.edu/confluence/display/cidays/Home?title=Main_Page> , planning guide: <http://www.clemson.edu/ccit/cidays/planning/index.html>

CI Community at University of Michigan, <http://orci.research.umich.edu/ci-community/> , Computational Discovery Days, and contacts on campus <http://orci.research.umich.edu/ci-community/ci-related-research-groups/>

OSG as an XSEDE high throughput resource: <https://www.opensciencegrid.org/bin/view/VirtualOrganizations/OSGasXsedeSp>; guide for XSEDE users of OSG: <http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=1081>

Campus Grid Maturity Model and deployed campus grids: <https://www.opensciencegrid.org/bin/view/CampusGrids/DeployedCampusInfrastructures>