Using OSG Connect

Lincoln Bryant

OSG User School

July 2014

What's OSG Connect?

- A "virtual cluster" for submitting directly into OSG

- Open to anyone at a US institution

- Free – no allocation required.

- HTCondor-centric, the tools and techniques you learn here transfer over to OSG Connect

Signing up for OSG Connect

- For the school, we've posted some materials here:

  - https://twiki.opensciencegrid.org/bin/view/Education/UserSchool14Connect

- Alternatively, you can visit the signup page here:

  - http://osgconnect.net/signup

What does OSG Connect offer researchers?

Resources

- Anywhere from 5000 to 20000 cores at any given time.

- 200TB shared storage system available for transient job data

- Access to the OASIS (OSG Application and Software Installation Service) software library

What else do I get?

- Share an identity service with Globus, the leading data transfer platform for research

- No grid certificates necessary, easy to sign up colleagues and collaborators.

- We're also working on building a web portal with custom, per-user analytics and job histories

But wait, there's more!

- Cool new stuff coming down the pipe:
  - Multicore jobs
  - Distributed software modules
  - GPUs and GPU-accelerated libraries

What are some services provided by OSG Connect?

Project accounting

- The fundamental organizational unit in OSG Connect is the project.
  - All accounting data is tracked via projects
  - Projects may have any number of users
  - Users can join as many projects as they like, but must be a part of at least one.
- For the school, we'll use the project "UserSchool2014"
  - If you continue to use OSG Connect after the school you should request a project:
  - https://osgconnect.net/newproject

Project accounting

Daily Hours By Project
15 Days from 2014-06-19 to 2014-07-03
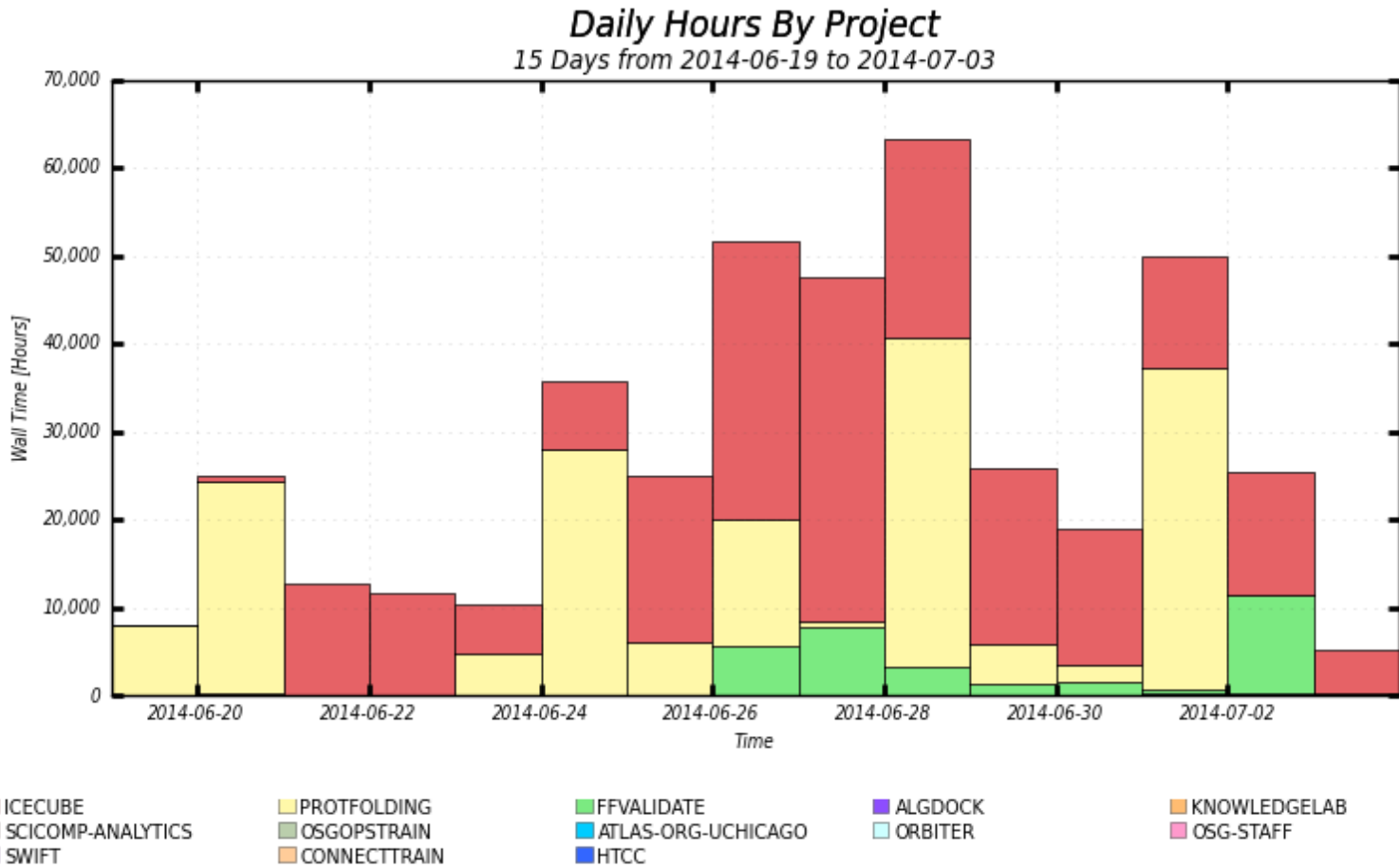
Maximum: 63,241 Hours, Minimum: 5,257 Hours, Average: 27,749 Hours, Current: 5,257 Hours

Viewing projects

- It's important to make sure you're using the correct project name so we can account your usage properly.

- We've created a small utility to let you see what project(s) you are a member of:

```
[lincolnb@login01 ~]$ connect show-projects
Based on your username (lincolnb), here is a list of projects you have access to:
  * atlas-org-uchicago
  * ConnectTrain
  * uchicago
  * UserSchool2014
```

Changing projects

- The "connect project" command gives you an interactive interface for changing project

```
[lincolnb@login01 ~]$ connect project
OSG Connect Project Selector - lincolnb


Select a project to be your default OSG Connect project.
Press ESCAPE or "Q" to quit without changes.



    1. atlas-org-uchicago
    2. ConnectTrain
    3. uchicago
    4. UserSchool2014
```

Software Access in OSG Connect

- Distributing software is hard. How can we make it easier?

- We want something that is...
  - Familiar (not reinventing the wheel)
  - Easy for admins to maintain
  - Easy for users to use

OSG Distributed Modules

- Most OSG sites have access to a software repository called "OASIS"

- We took the "module" tool that many HPC centers use and put it into OASIS

- We're now building a software library around this tool, which will hopefully make users' lives easier!

Using distributed modules

- The module tool has to be loaded into your environment via 'source'

```
[lincolnb@login01 ~]$ source /cvmfs/oasis.opensciencegrid.org/osg/modules/lmod/5.6.2/init/bash
```

- Once loaded, you can see all available modules

```
[lincolnb@login01 ~]$ module avail

-------- /cvmfs/oasis.opensciencegrid.org/osg/modules/modulefiles/Core ---------
   atlas        lapack             lmod/5.6.2 (D)    python/3.4
   blast        lmod/SiteHook      namd/2.9          settarg/5.6.2
   blender      lmod/SitePackage   python/2.7 (D)

  Where:
   (D):  Default Module

Use "module spider" to find all possible modules.
Use "module keyword key1 key2 ..." to search for all possible modules matching
any of the "keys".
```

Potential pitfalls

- Not all OSG sites have OASIS! You'll need to add a specific requirement to your HTCondor submit file:

```
Requirements = (HAS_CVMFS_oasis_opensciencegrid_org =?= TRUE)
```

- Your application wrapper will also need to setup the module tool and load the module!

Data transfer with Globus

- What is Globus?
  - A high-performance data transfer service for research
  - Allows users to easily move data between Globus "endpoints" at computing facilities.
  - Monitors performance, corrects errors, and retries failed transfers automatically
- OSG Connect shares an identity service with Globus.
  - All OSG Connect users are automatically Globus users!

Data transfer with Globus

- OSG Connect has an Globus endpoint named Stash

- With Globus Connect Personal, you can transfer data from Stash to your laptop and vice versa.
  - https://www.globus.org/globus-connect-personal

- Once installed, you can go here to start a transfer:
  - https://portal.osgconnect.net/xfer/StartTransfer

Finding support

- The OSG ConnectBook is a great place to start

    - http://osgconnect.net/book

- There's also the OSG Connect support mailing list

    - connect-support@opensciencegrid.org

Questions?

Next: OSG Connect hands-on session

Feel free to email me with any questions:
   lincolnb@uchicago.edu