

# Tukey: The Open Science Data Cloud Graphical Interface

## Matthew Greenway, Allison Heath, Robert Grossman

### Tukey

"The graph paper - and transparencies - are there, not as a technique, but rather as recognition that the picture-examining eye is the best finder we have of the wholly unanticipated." John Tukey

We have designed and implemented Tukey, a graphical interface for accessing public and private cloud services. We run a central version of Tukey for the Open Science Data Cloud (OSDC) and it can also be deployed as a general interface for utility clouds.

The various components of Tukey provide users with these core services:

- Virtual server provisioning
- Usage and billing information
- Availability of resources
- Access to public data sets
- A Key service
- File sharing management

### Console

The Tukey Console provides a graphical interface for OSDC services. Our approach is similar to Amazon's AWS Management Console. The console gives OSDC users a way to launch, view and terminate virtual servers, and view monthly usage. This means that whenever an OSDC user has access to a web browser they can easily manage cloud resources.

After we approve a user's application they can sign in to the OSDC console to start using cloud services. To access features of Tukey that require authentication, the user must log in using a Shibboleth provider through the InCommon federation or an OpenID account such as University of Chicago and Google respectively.

Once authenticated, the user selects an image from Adler (a Eucalyptus cloud), or Sullivan (an OpenStack cloud) to launch an instance. They can then terminate them from the Instances view.

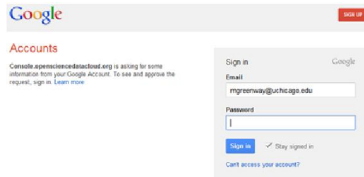


Figure 2-A. OpenID log in using a Google account.

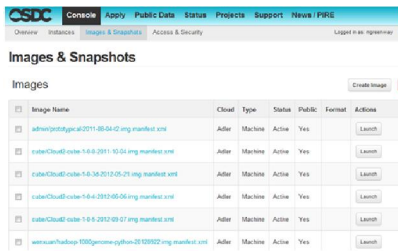


Figure 2-B. Lists virtual machine images. From this page, the user selects images and launches virtual machines.

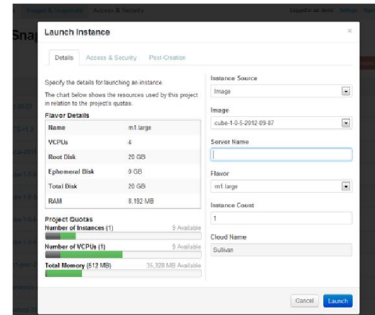


Figure 2-C. Demonstrates the launch instance workflow. The user sets the details of the instance and can see quota information.

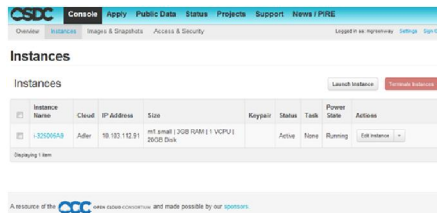


Figure 2-D. Provides a list of running instances. The user can view and terminate instances.

### File Permissions

The user can share files with each other using the file permissions interface. Users can manage groups of other users and collections of files to grant bulk permissions.

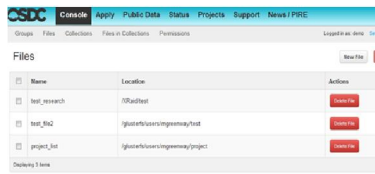


Figure 3. Displays the users file that they can share with other users and groups.

### Status

The OSDC Status page shows current statistics of compute and storage services indicating outages and availability of resources.

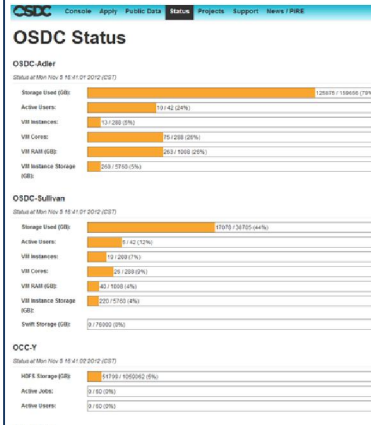


Figure 4. OSDC Status page.

### Key Service

Archival Resource Keys (ARK) are provided for data hosted on the OSDC by using the Open Cloud Consortium's Name Assigning Authority Number (NAAN). An ARK provides a persistent identifier for locating data and associated metadata. For example, the ARK for the Earth Observing-1 Mission is ark:/31807/opd2.



Figure 5. OSDC key service. Resolver.

### Data Sets

The Public Data Sets page displays a list of data sets mirrored on the OSDC-Root petabyte storage server. Each data set has its own page on how to download it or access it from OSDC compute resources.

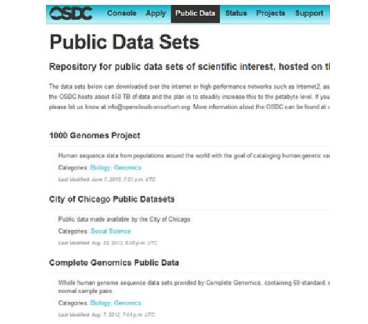


Figure 6. The Public Data Sets page.

### Create a Support Ticket

If users have any questions or problems, they can submit a support ticket directly from the OSDC site. The Open Science Data Cloud team is excited to know how we can improve usability and enjoy the technical challenges presented by big data and cloud computing.

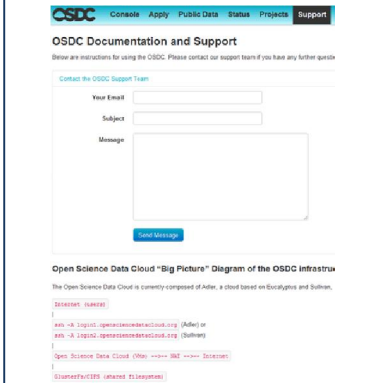


Figure 7. The support page.

### Apply for an Account

The first step to using the Open Science Data Cloud is to fill out an application for an OSDC account. The application page asks for information about potential users' projects and how much of the resource they think they will need.

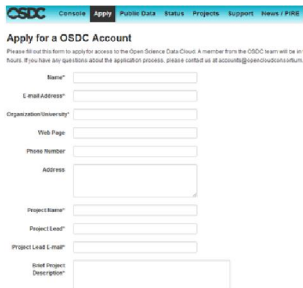


Figure 1. The OSDC application page.