

Accelerating Campus Research with Connective Services for Cyberinfrastructure

Rob Gardner
Steve Tuecke

A collection of approximately 15 red ovals of various sizes scattered around the 'ci connect' logo.

ci connect

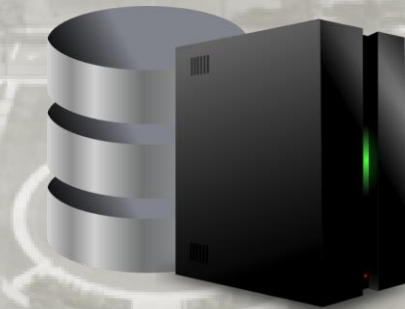


THE UNIVERSITY OF
CHICAGO

**A typical campus: distributed departments,
distributed researchers, distributed resources**

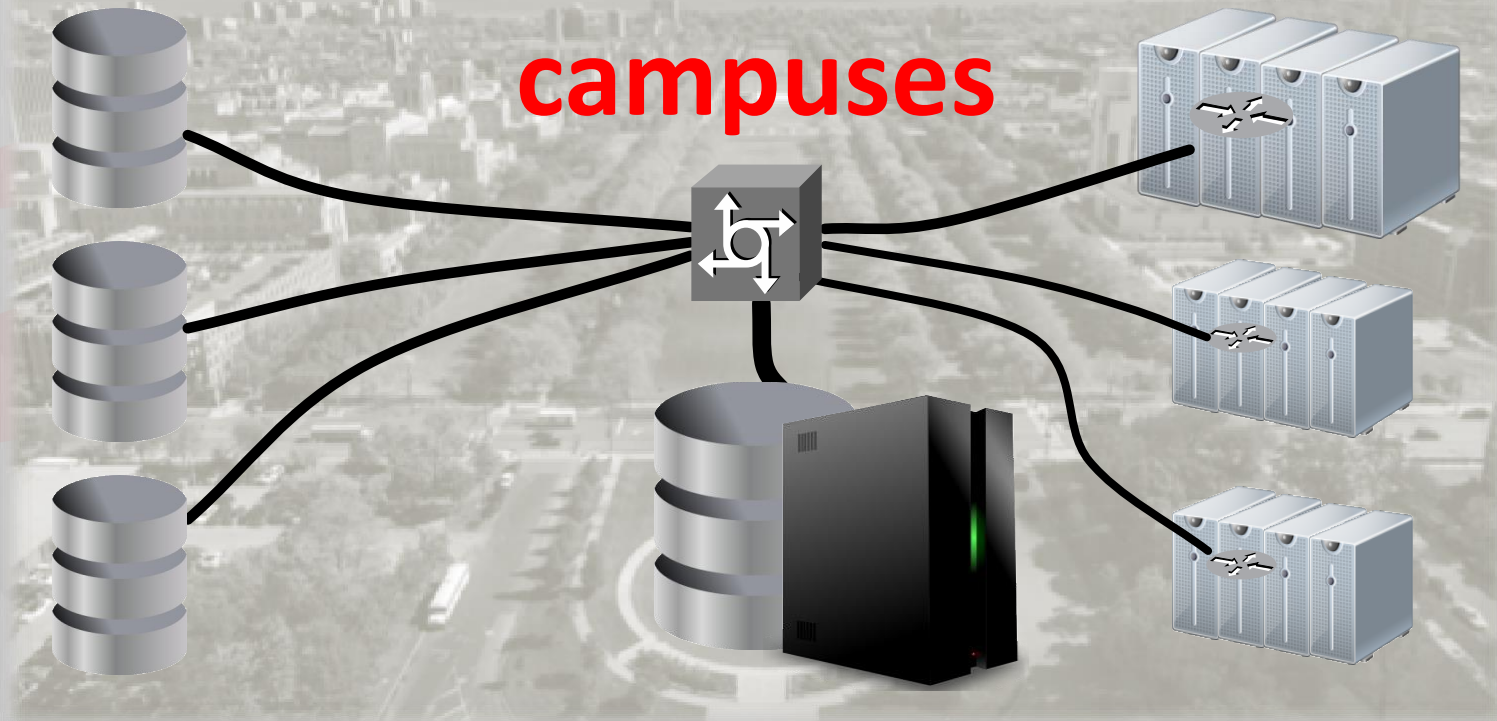


Commodification, cloud technologies, and practices achieving economies of scale drive centralization of resources on campuses, condo clusters, etc.

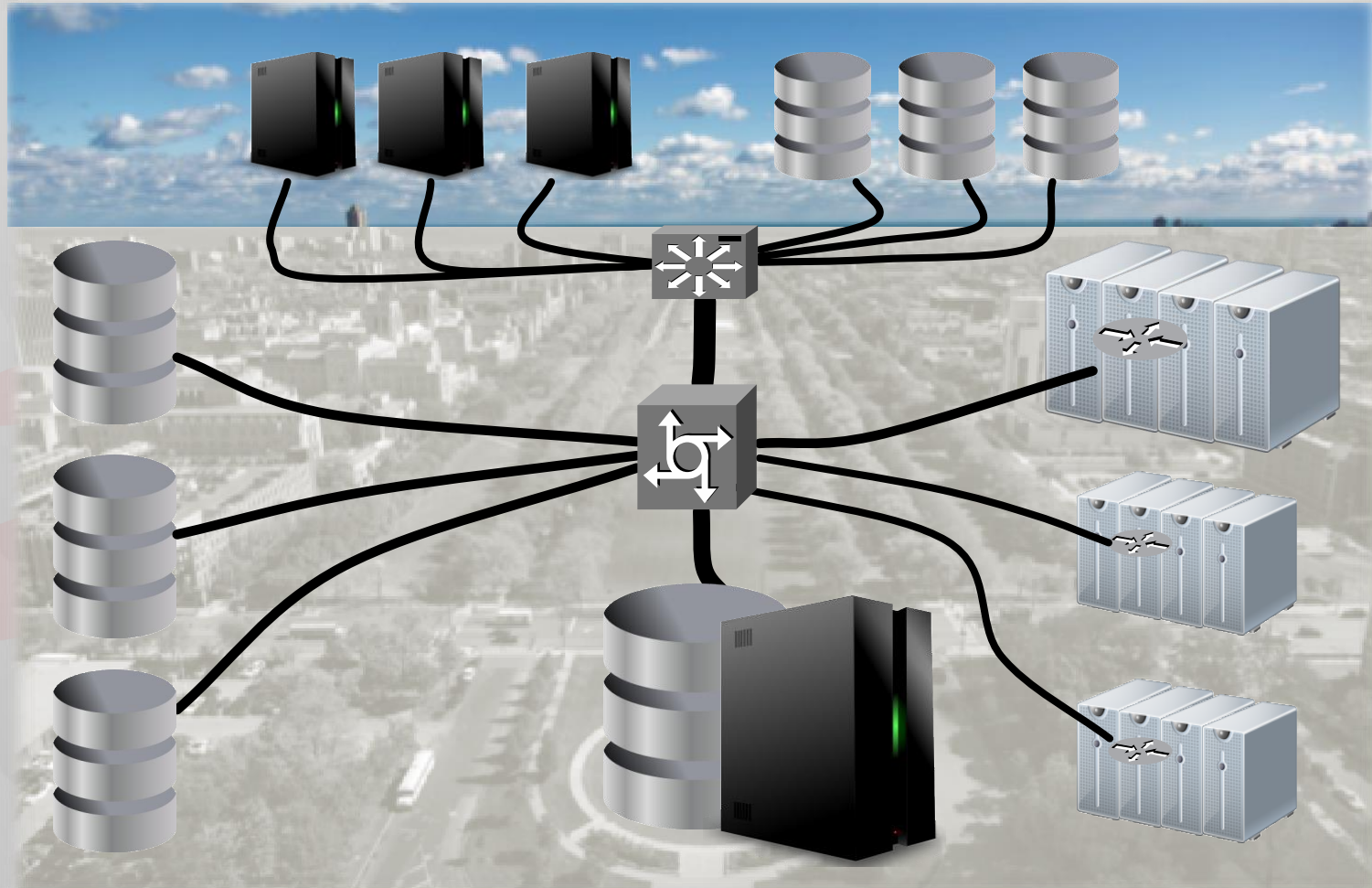


College-level funding discouraging “closet clusters”

Yet, for human and practical reasons there often remains a distribution of resources on campuses

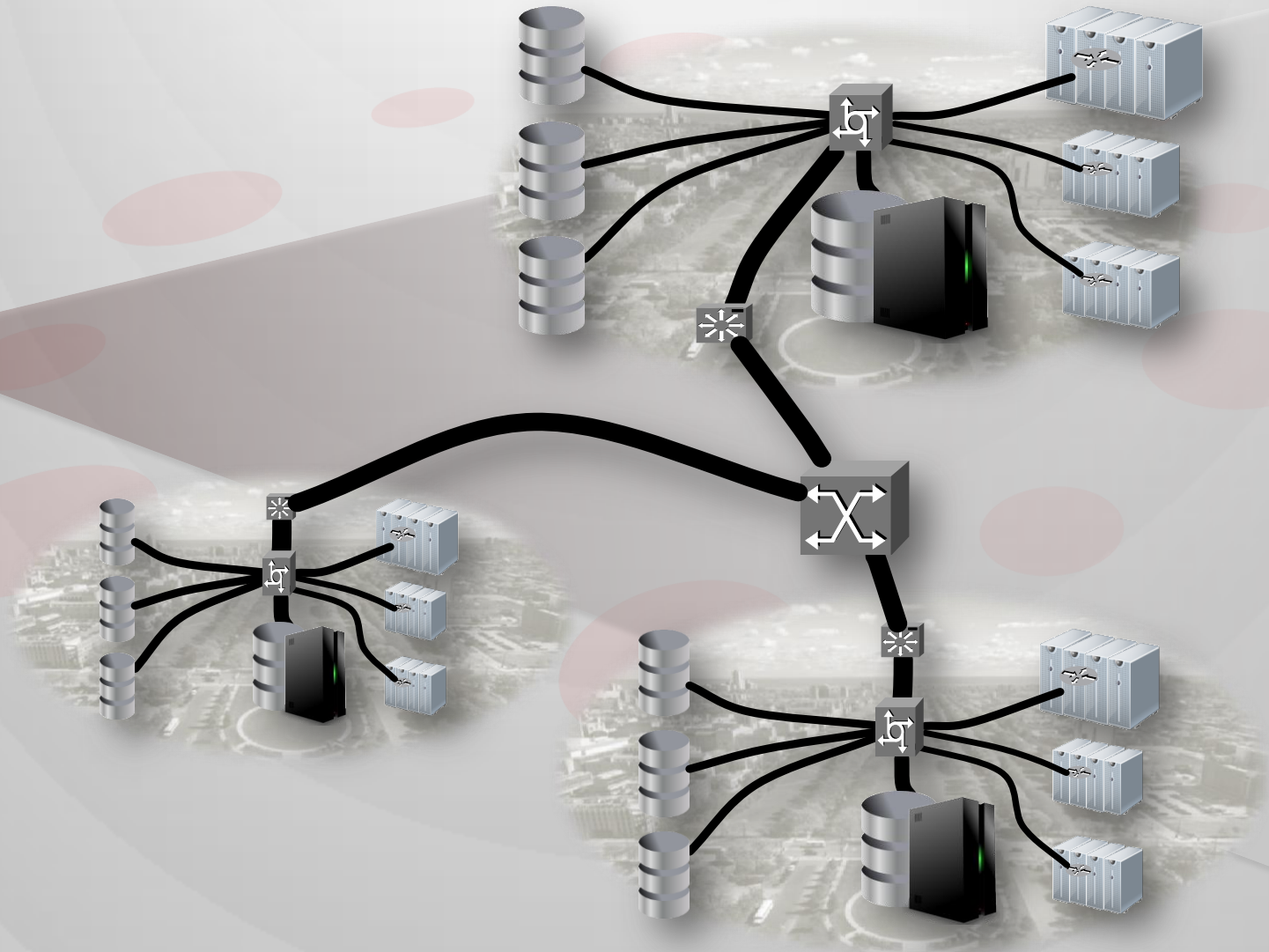


And off-campus too...

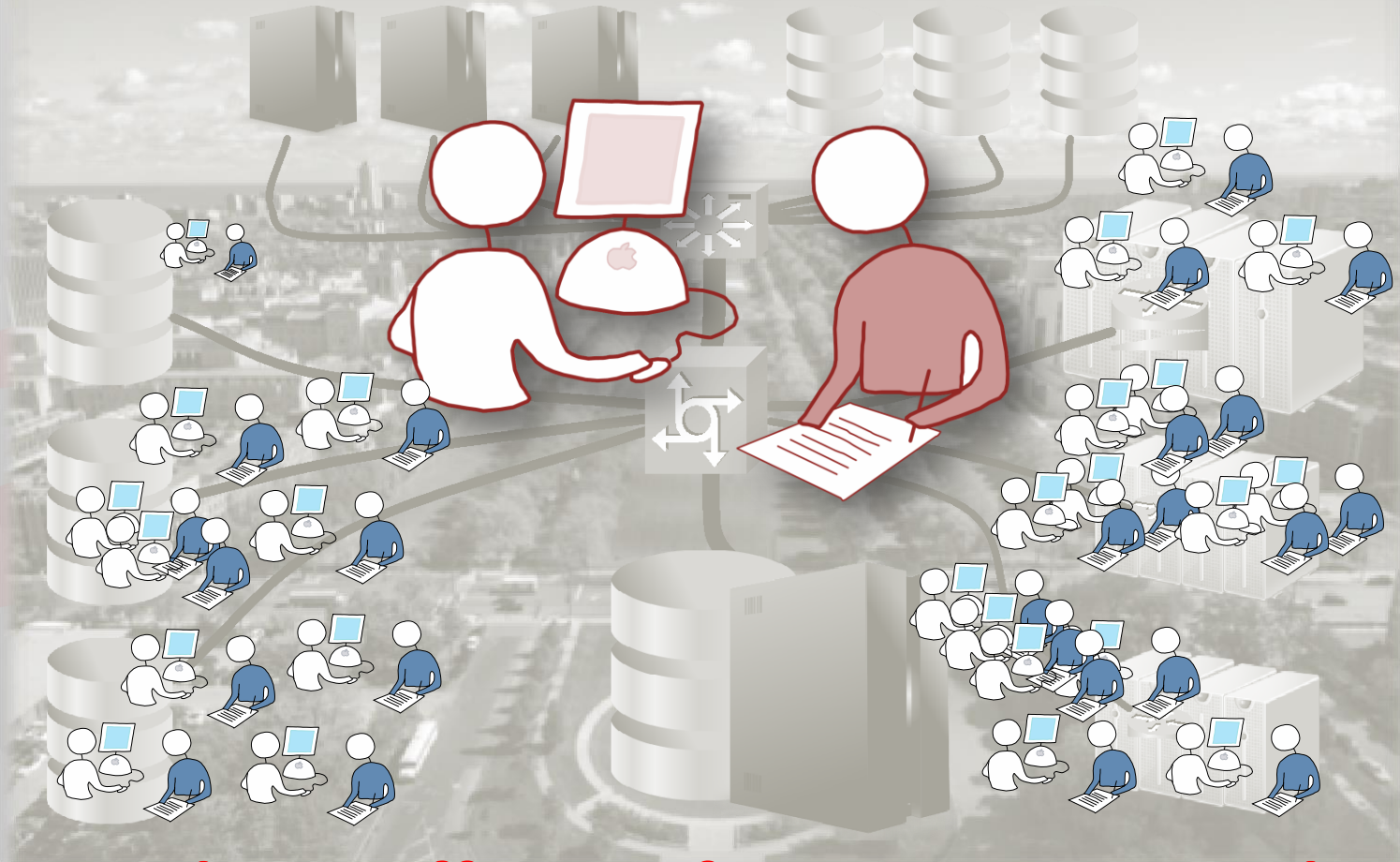


**E.g. National CI grids, HPC centers,
commercial clouds, science clouds ...**

Bridged Campus Cyberinfrastructures



Needed: services to make distributed resources transparent to campus users...



...and cost effective for resource providers, wherever those resources are provisioned.

CI Connect Vision

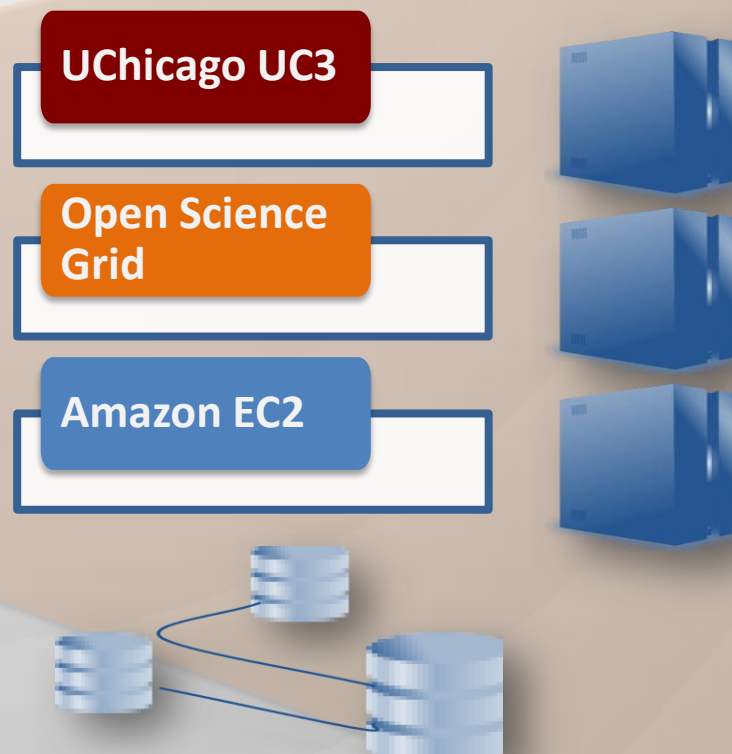
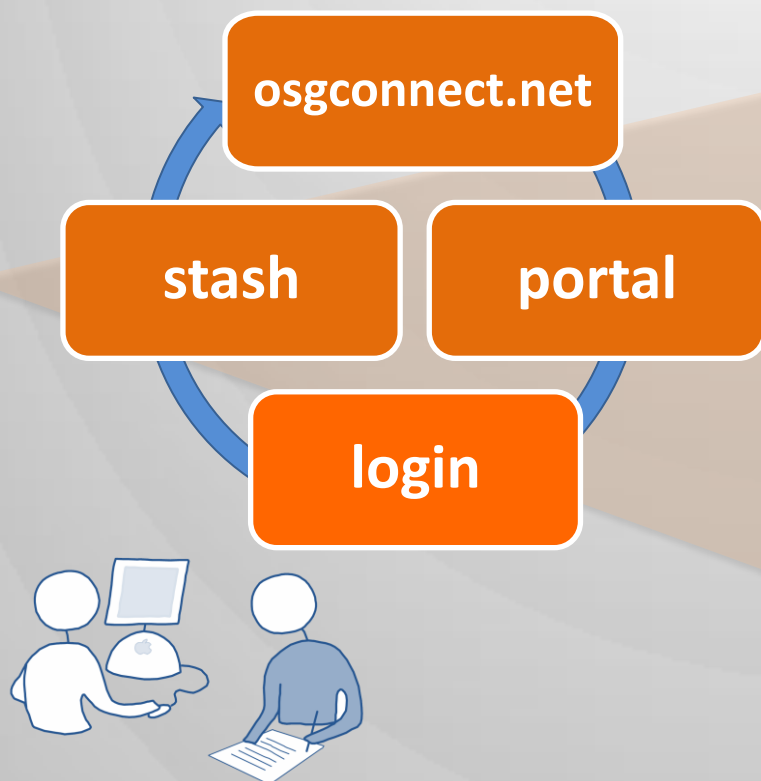
**Accelerate research on campus by
providing connective services for local,
cloud and national cyberinfrastructure**

- OSG Connect is a service for connecting users and project Groups to the national-scale Open Science Grid
- Provisioned as a service, using the CI Connect platform





osg connect



OSG Connect resources: campus grid, opportunistic OSG, & Amazon cloud in one place – integrated with identity and research data management services



osg connect

CycleServer

Welcome

Pool Summary

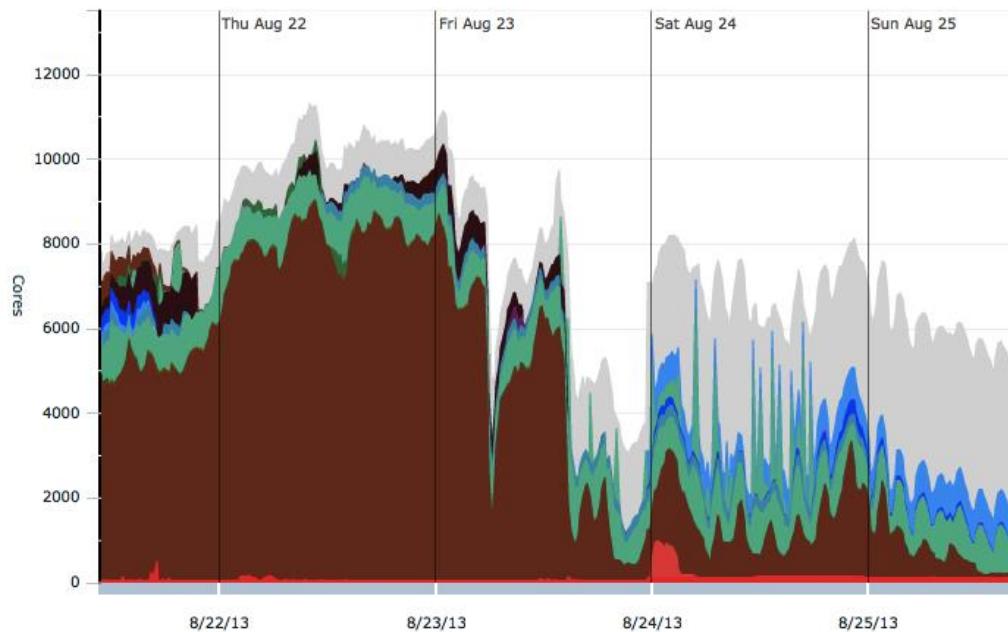
Pool	Total Slots
Open Science Grid	2581
OSG-Connect Cloud	120
University of Chicago Computing Cooperative	529
University of Chicago Computing Cooperative (ITS)	240
University of Chicago Computing Cooperative (MWT2)	3144
University of Chicago Computing Cooperative (UCT3)	258
Total	6872

Jobs by State

Jobs by Owner

Show: Historical grid usage in all pools

Time Frame: 3 Hours | Day | Week | Month



Scheduling to OSG via flocking to the GlideinWMS service

View as: Area | Line

Legend
Used by owner
rynge
donkri
high.psiders
jstupak
malik
nitish
pwolberg
spadhi
jdandoy
kkrizka
vo.engage.engage
vo.glow.glow
vo.osg.osg
vo.hcc.hcc
yadunand
high.hayashis
izheng
lincoln
johnda
jwebster
yx5
high.rynge
strolog
dweitzel
high.kimfwong
high.wdi1114

OSG Connect projects



osg connect

Support ▾

Resources ▾

Connect ▾

Transfer ▾

rwg ▾

My Groups

osg.divulsi
osg.CompChem
osg.CompNeuro
osg.ConnectTrain
osg.Duke-QGP
osg.EvoTheory
osg.GlassySystems
osg.KnowledgeSys
osg.NESCent
osg.NRELMatDB
osg.OSG-Staff
osg.PathSpaceHMC
osg.PlantBio
osg.RADICAL
osg.RDCEP
osg.SouthPoleTelescope
osg.Swift
osg.UChicago-RCC
UC3 Support

MEMBER OF

atlas.org.uchicago

Create New Group »

● admin action required

Search

osg.KnowledgeSys

Home

Members

SubGroups

Settings



connect > osg.KnowledgeSys

edit

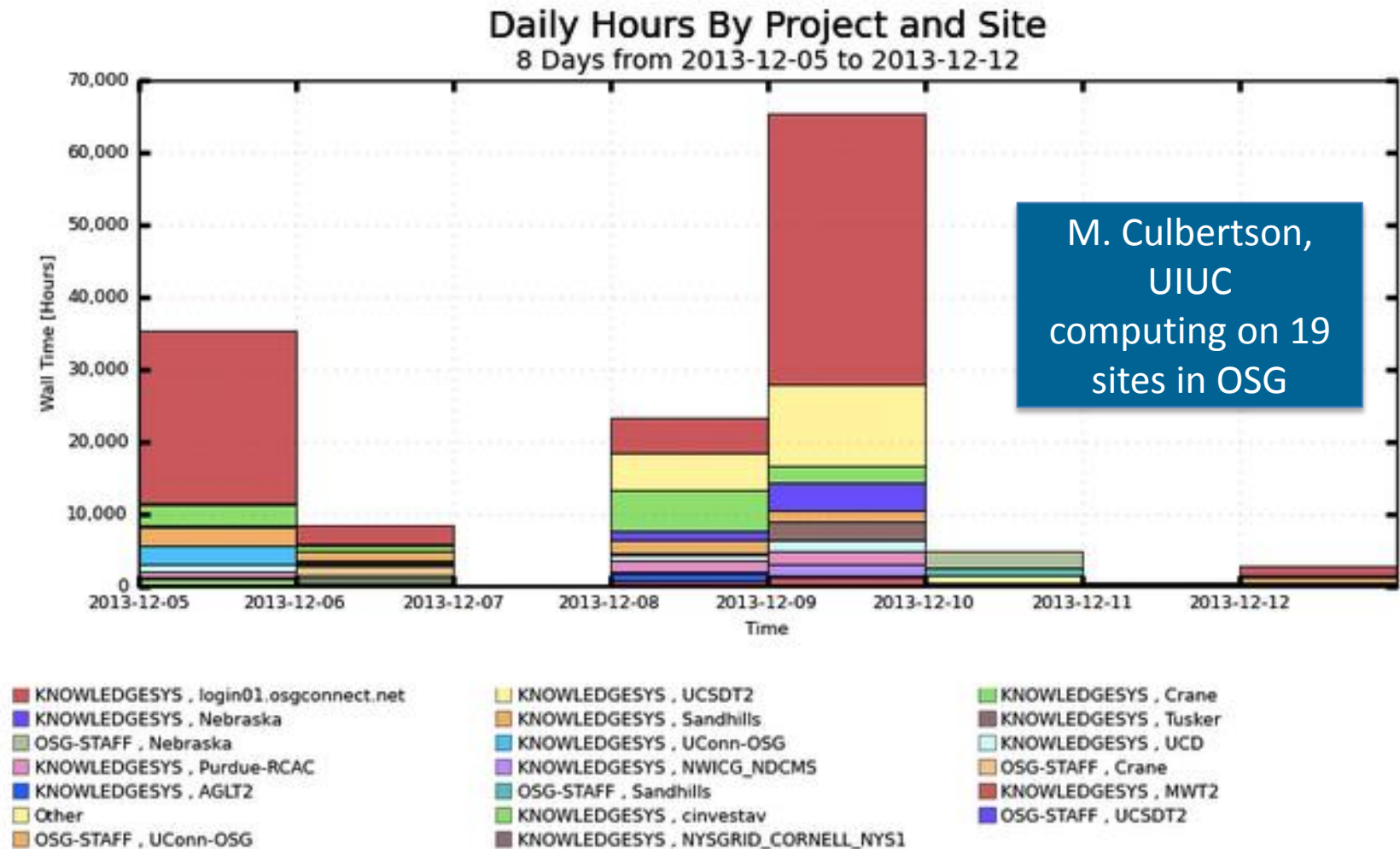
Project Name: Knowledge Systems
Field of Science: Knowledge Models and Cognitive Systems
PI Name: Michael J. Culbertson
PI Email: culbert1@illinois.edu
Organization: University of Illinois, Urbana-Champaign
Department: Educational Psychology
Join Date: September 16, 2013
Sponsor: OSG Connect
OSG Contact: Rob Gardner
Project Contact:
Telephone:

Enabling
individual
researchers

In educational assessment, several questions must be answered when constructing a test, such as "How many items are necessary for adequate knowledge measurement precision?", "How many field-test students are needed to adequately calibrate model parameters?", or "Which computerized adaptive testing (CAT) algorithm performs best?" For complex non-linear models, these questions are typically approached by simulation: Model parameters are calibrated (as if unknown) from simulated student item responses, or the emergent properties of particular CAT algorithms are investigated with a large number of simulated test takers. Since the design space grows quickly, many simulations are necessary to understand general trends.

Simulations throughout the test design space can be run independently, requiring little coordination between cores. Computations generally do not have high memory requirements or unusual library/code dependencies, and computations can be recovered from checkpoints easily. The large number of simulations suggests parallel computing, but the independence allows an asynchronous, distributed environment, such as OSG.

Knowledge Synthesis Studies enabled by d-HTC

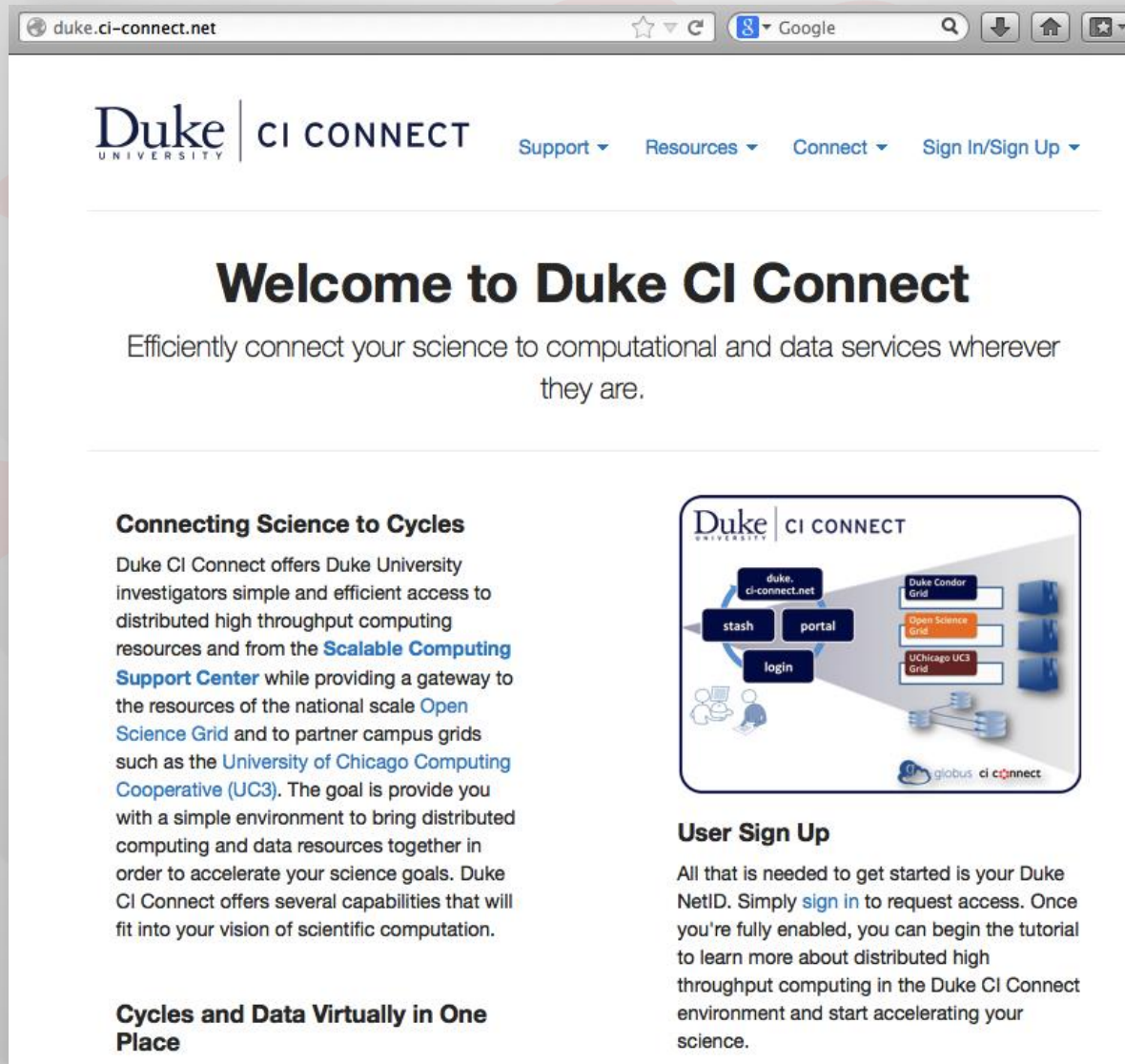


Maximum: 65,511 Hours, Minimum: 233.58 Hours, Average: 17,671 Hours, Current: 2,870 Hours

- An HTCondor-based job submission environment connecting local campus cluster resources, the OSG, bridged campuses, or optionally Amazon cloud resources
- Offered as a service (PaaS model)
- Deployed Duke University as the first branded CI Connect campus
- Deployments in progress for the University of Chicago
- Discussions with University of Michigan

Collaboration between Scalable Computing Center at Duke University and the Computation Institute at UChicago

<http://duke.ci-connect.net>



The screenshot shows a web browser window with the address bar displaying "duke.ci-connect.net". The page features the Duke University logo and "CI CONNECT" header. Navigation links include "Support", "Resources", "Connect", and "Sign In/Sign Up". The main heading is "Welcome to Duke CI Connect" with the tagline "Efficiently connect your science to computational and data services wherever they are." Below this, the "Connecting Science to Cycles" section describes the service's purpose and lists connected resources like Duke Condor Grid, Open Science Grid, and UChicago UC3 Grid. A diagram illustrates the workflow from user login to resource access. The "User Sign Up" section provides instructions for getting started with a Duke NetID.

duke.ci-connect.net

Duke UNIVERSITY | CI CONNECT

Support ▾ Resources ▾ Connect ▾ Sign In/Sign Up ▾

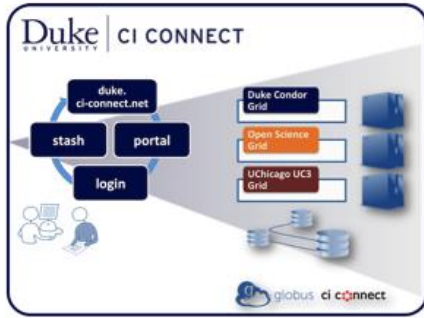
Welcome to Duke CI Connect

Efficiently connect your science to computational and data services wherever they are.

Connecting Science to Cycles

Duke CI Connect offers Duke University investigators simple and efficient access to distributed high throughput computing resources and from the [Scalable Computing Support Center](#) while providing a gateway to the resources of the national scale [Open Science Grid](#) and to partner campus grids such as the [University of Chicago Computing Cooperative \(UC3\)](#). The goal is provide you with a simple environment to bring distributed computing and data resources together in order to accelerate your science goals. Duke CI Connect offers several capabilities that will fit into your vision of scientific computation.

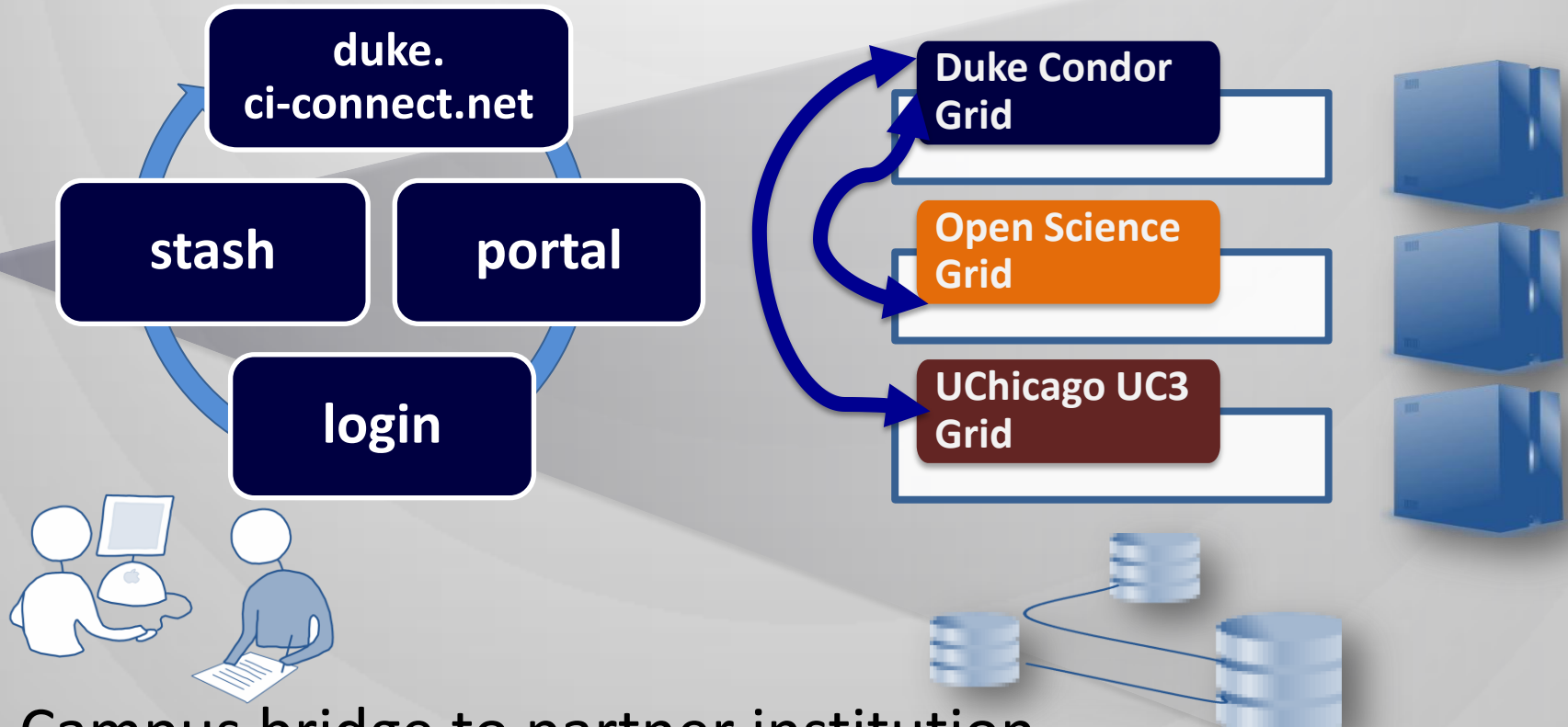
Cycles and Data Virtually in One Place



The diagram illustrates the workflow for connecting science to cycles. It shows a user interacting with the "duke.ci-connect.net" portal, which includes "stash", "portal", and "login" components. The portal connects to various grids: "Duke Condor Grid", "Open Science Grid", and "UChicago UC3 Grid". The workflow is supported by "globus" and "ci-connect" infrastructure.

User Sign Up

All that is needed to get started is your Duke NetID. Simply [sign in](#) to request access. Once you're fully enabled, you can begin the tutorial to learn more about distributed high throughput computing in the Duke CI Connect environment and start accelerating your science.



Campus bridge to partner institution

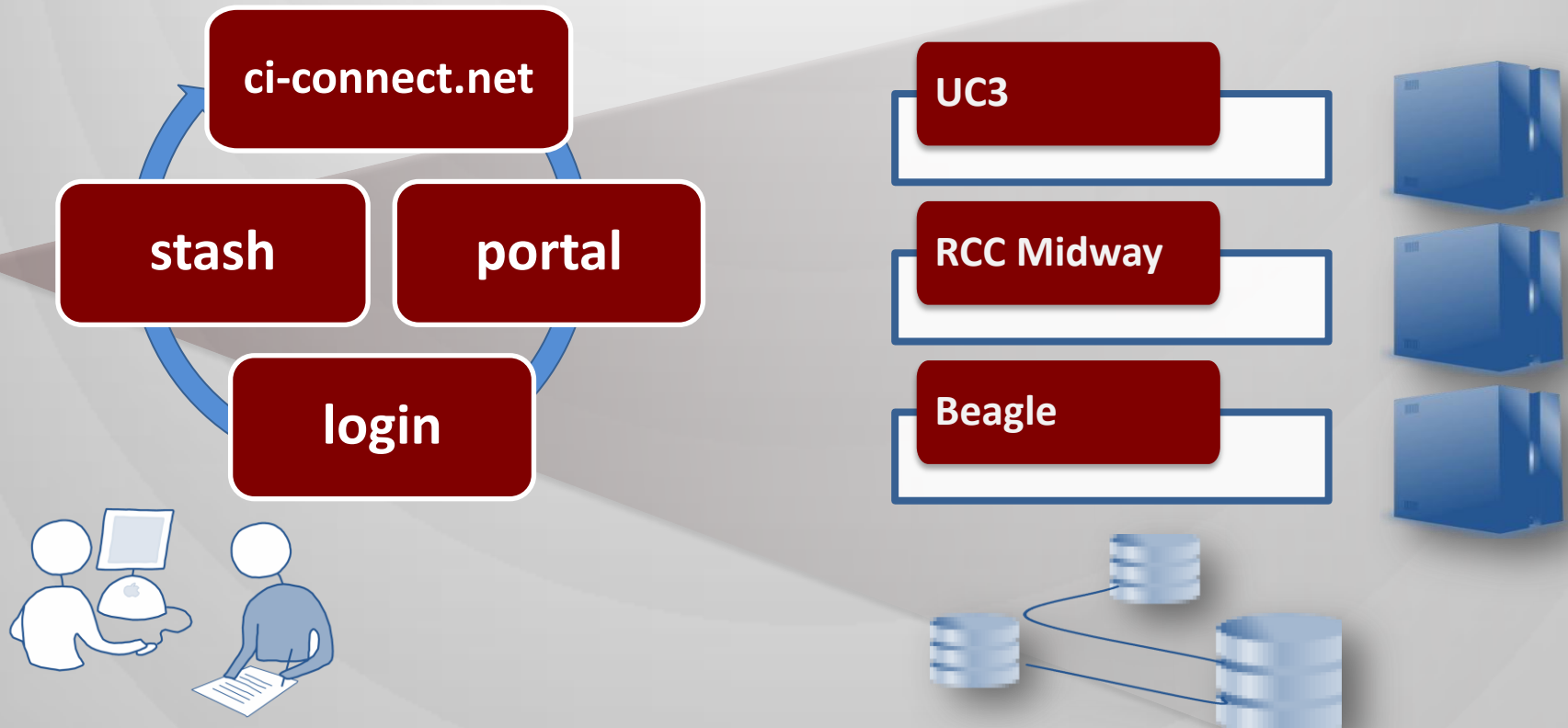
Connection to OSG

Access controlled by Duke SCC



THE UNIVERSITY OF
CHICAGO

CI CONNECT



Will set this up for UChicago Computation and Enrico Fermi Institutes
(Tier3 users will migrate to ATLAS Connect)



CI Connect Summary



- CI Connect offers a way to bring more CPU into users home environment without deploying IT
- Offers a means to share resources on and off campus with non-physics groups
- Integrated Globus for data management
- Instances deployed for:
 - OSG
 - Duke Scalable Support Center (Duke Grid)
 - ATLAS
- Started discussion with U Michigan
- Interest from XSEDE campus bridging