



THE UNIVERSITY OF
CHICAGO

Accelerating Science

with connective services for cyber
infrastructure

Rob Gardner
Computation Institute
University of Chicago

Science is Collaborative

- On campus
- Inter-campus
- National
- International



National Cyber Infrastructures

- XSEDE
- Open Science Grid
- Worldwide LHC grid
- Services ... e.g. Globus
- How do campuses connect to these, and to each other?

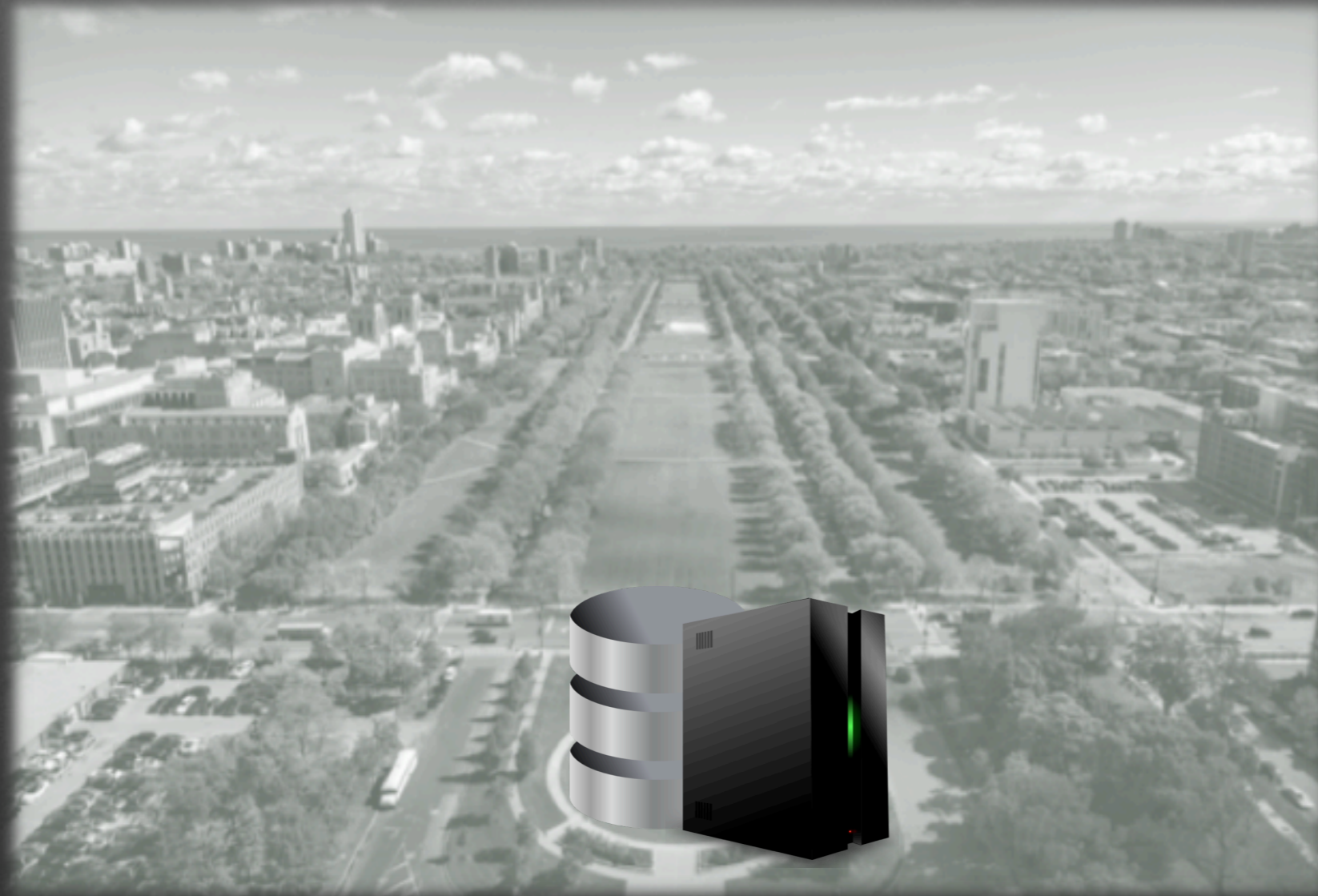
Campus Ecosystem



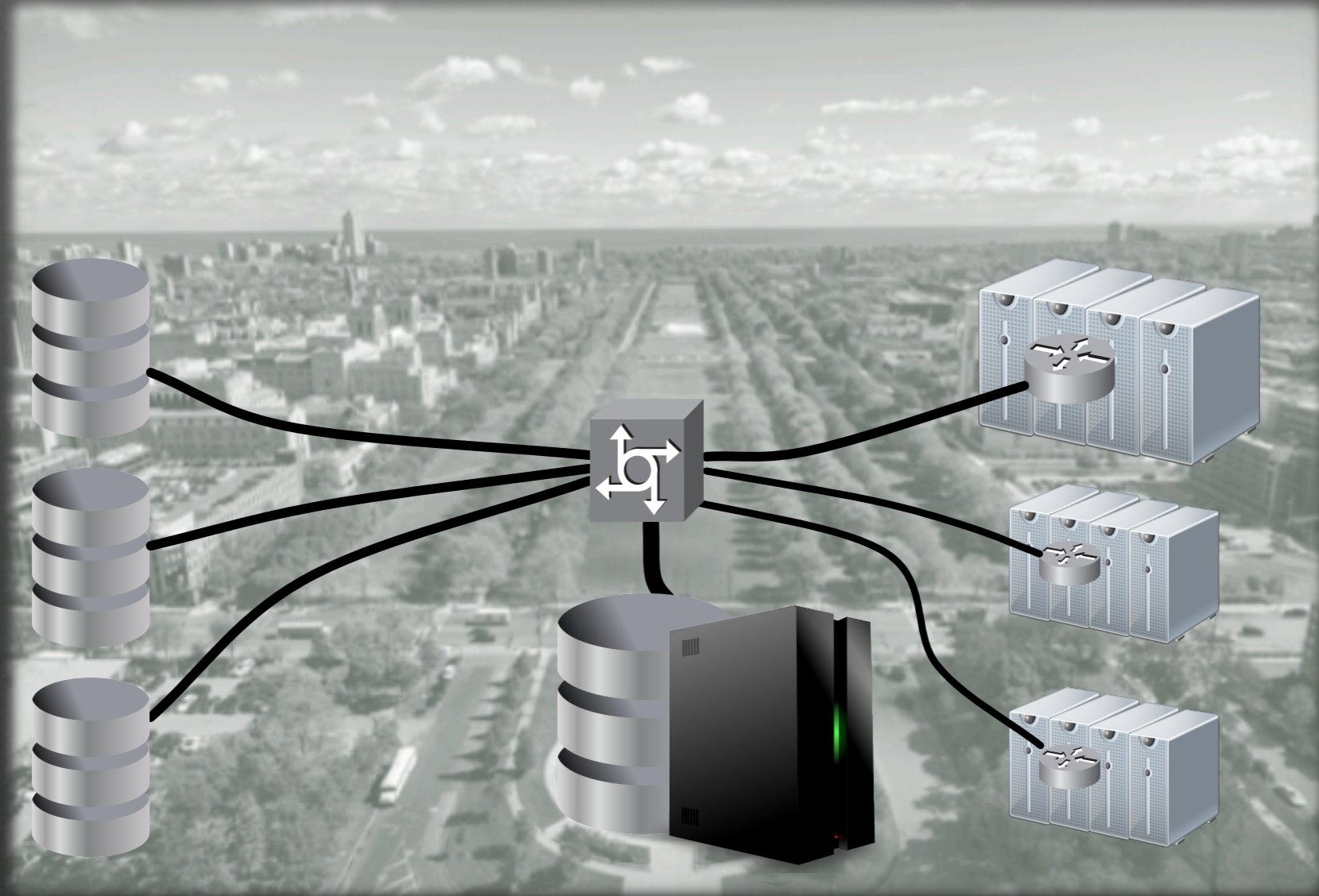
Challenges: limited budgets, operational costs, scale of science



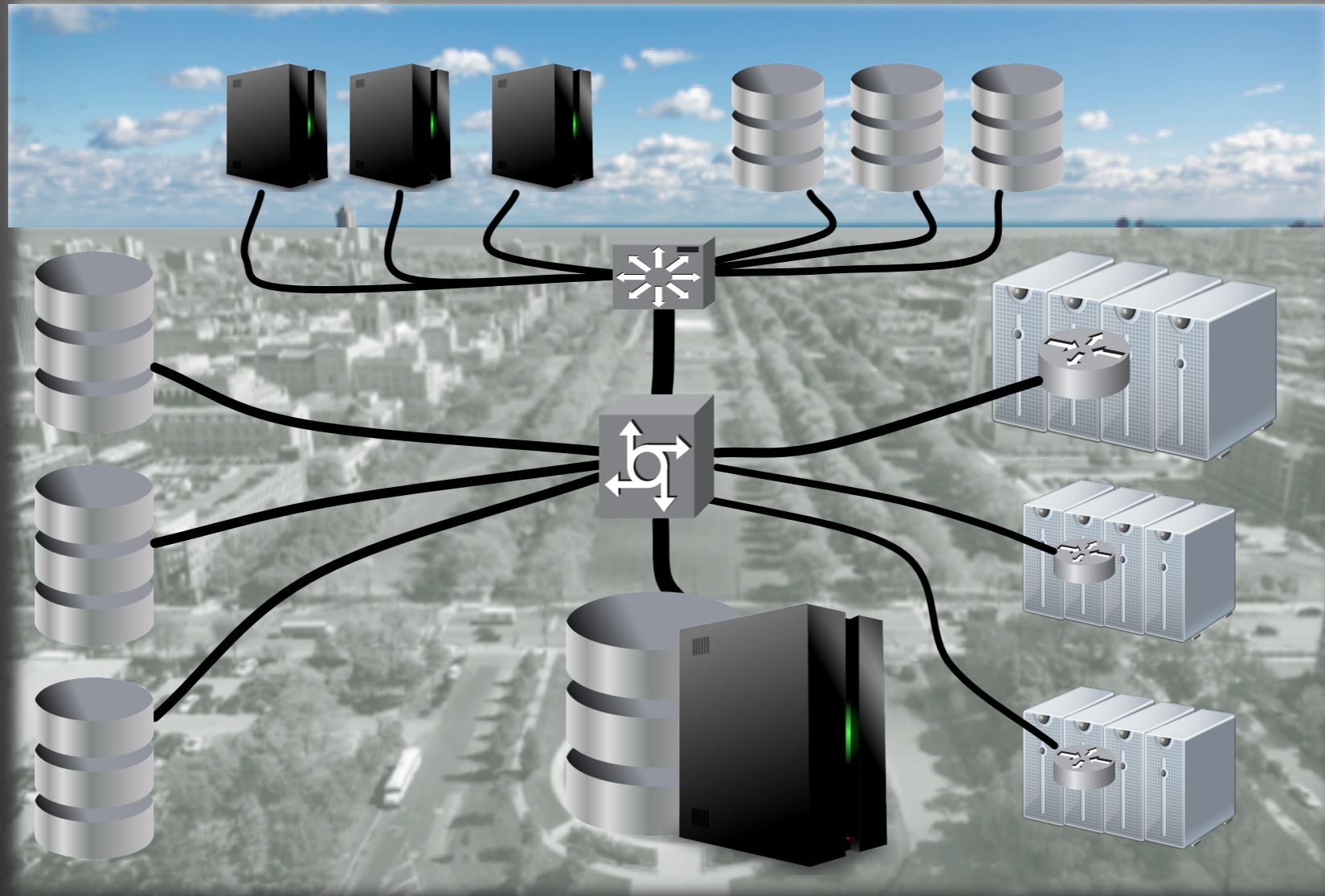
Commodification, Centralization



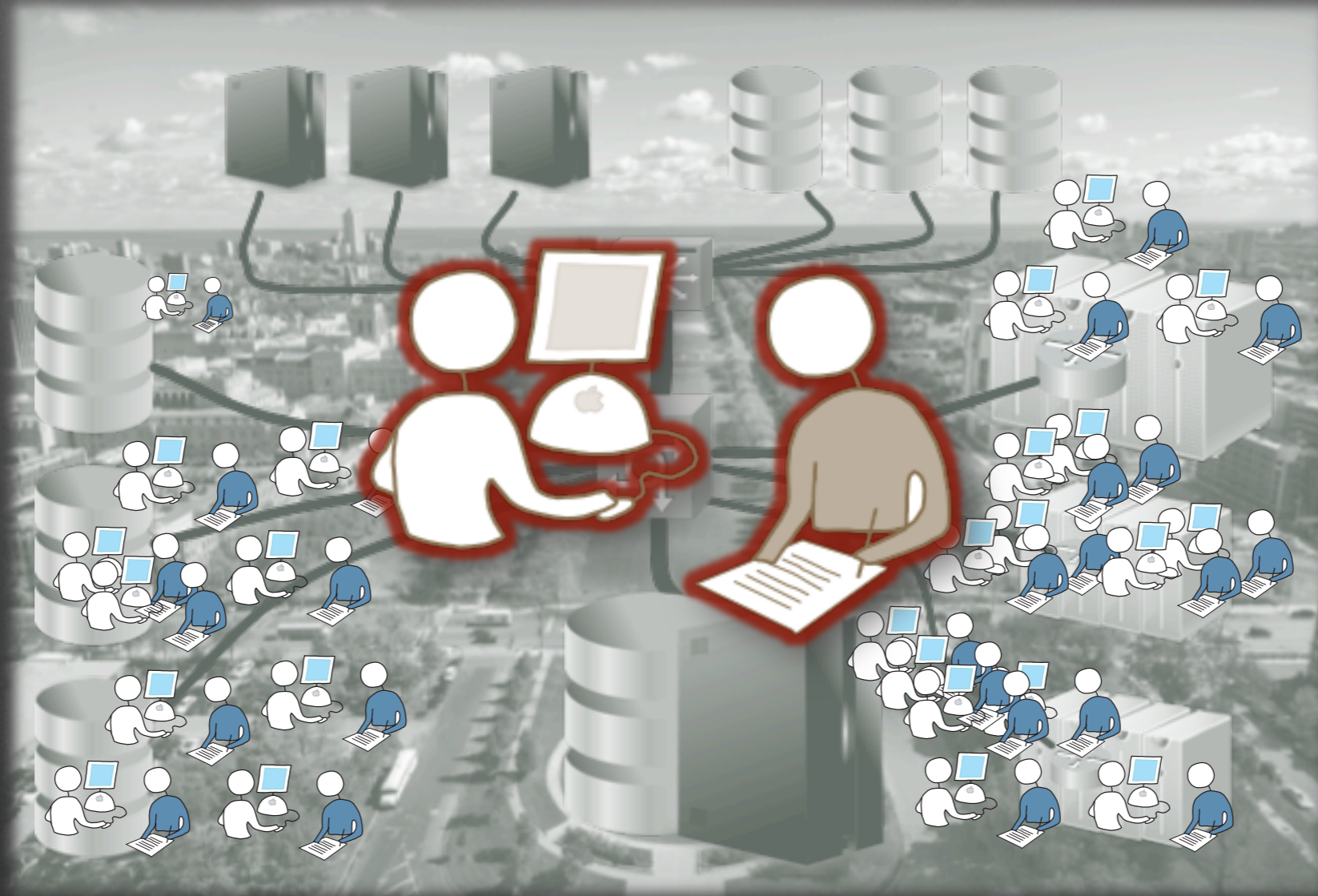
Even so, distributed resources remain



Off-campus too (the national ecosystem)



Need: transparency for users, cost effective for providers

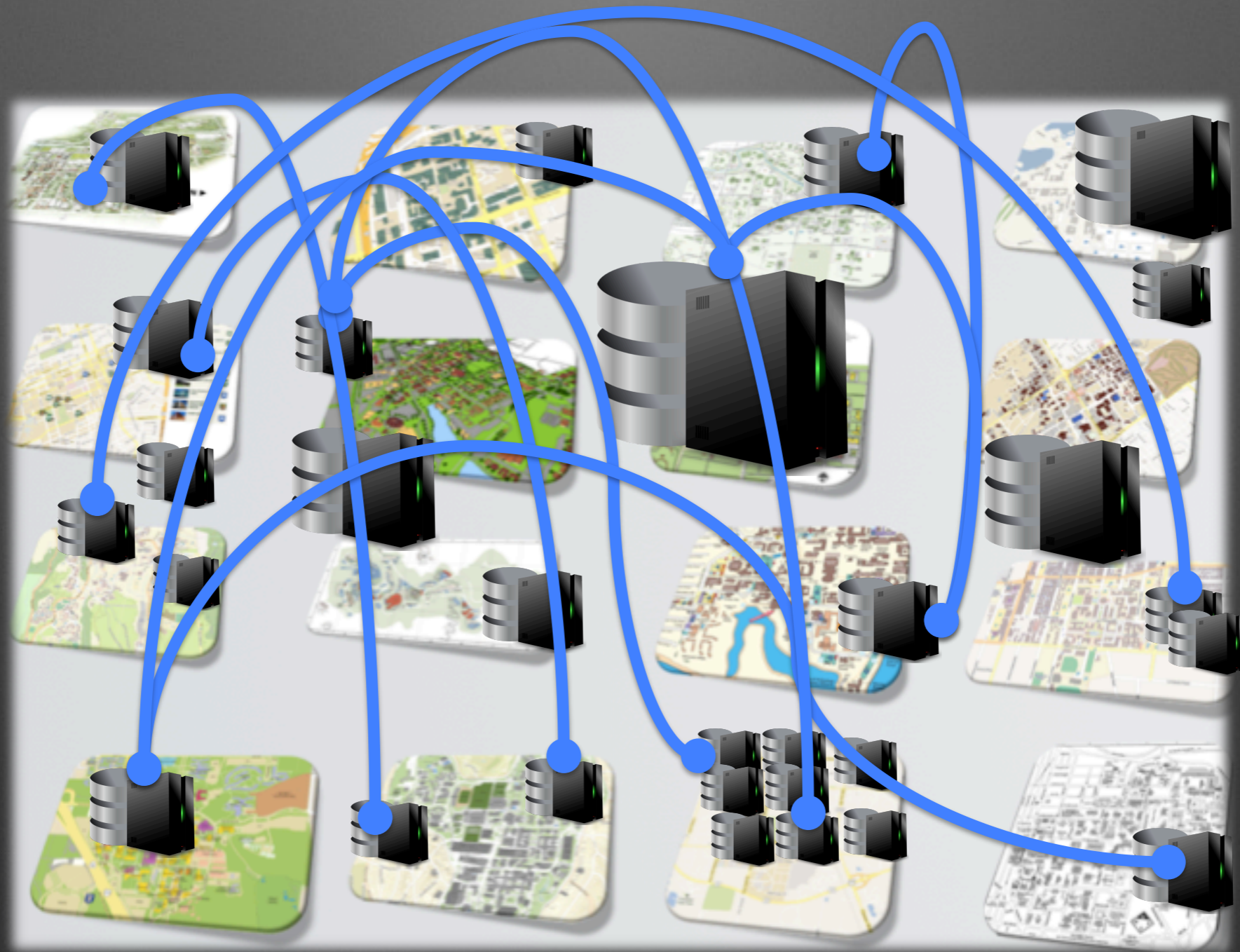


Campus Data Centers

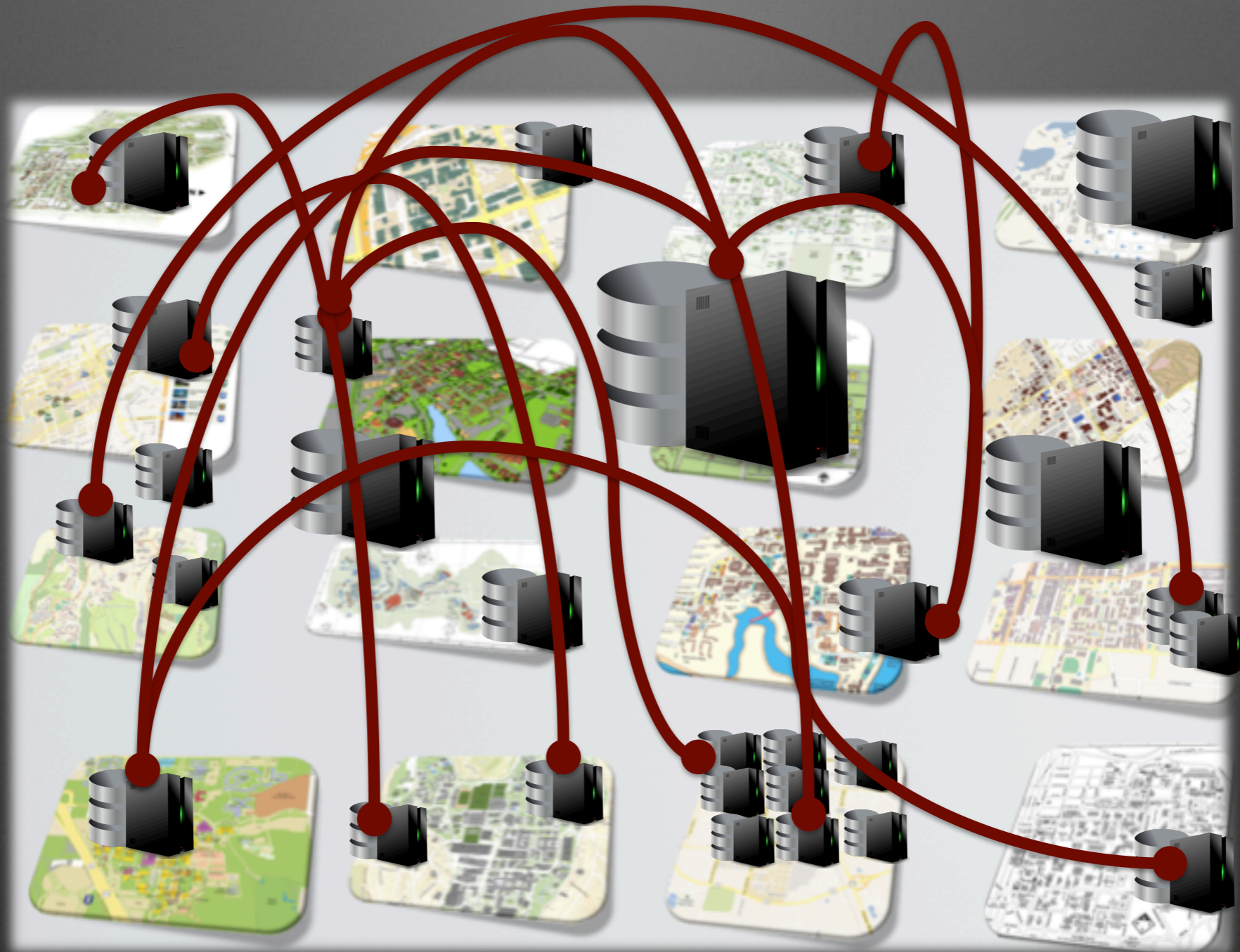
and the national cyber infrastructure



Globus data transfer, sharing (and now publication)



Cycle sharing ecosystem?



Distributed HT Computing

The OSG fabric of services

- The leading **distributed** high throughput computing service in the US
 - 104k cores, 75.6 PB, 123 compute endpoints
- Well connected
 - Most sites on OSG have 10 Gbps or greater to I2 or ESnet, many upgrading, plus SciDMZs
 - (at least) 2M transfers/day, 1 PB/day
- Friction free
 - The OSG VO with **GlideinWMS** offers transparent access to these resources for small groups

3

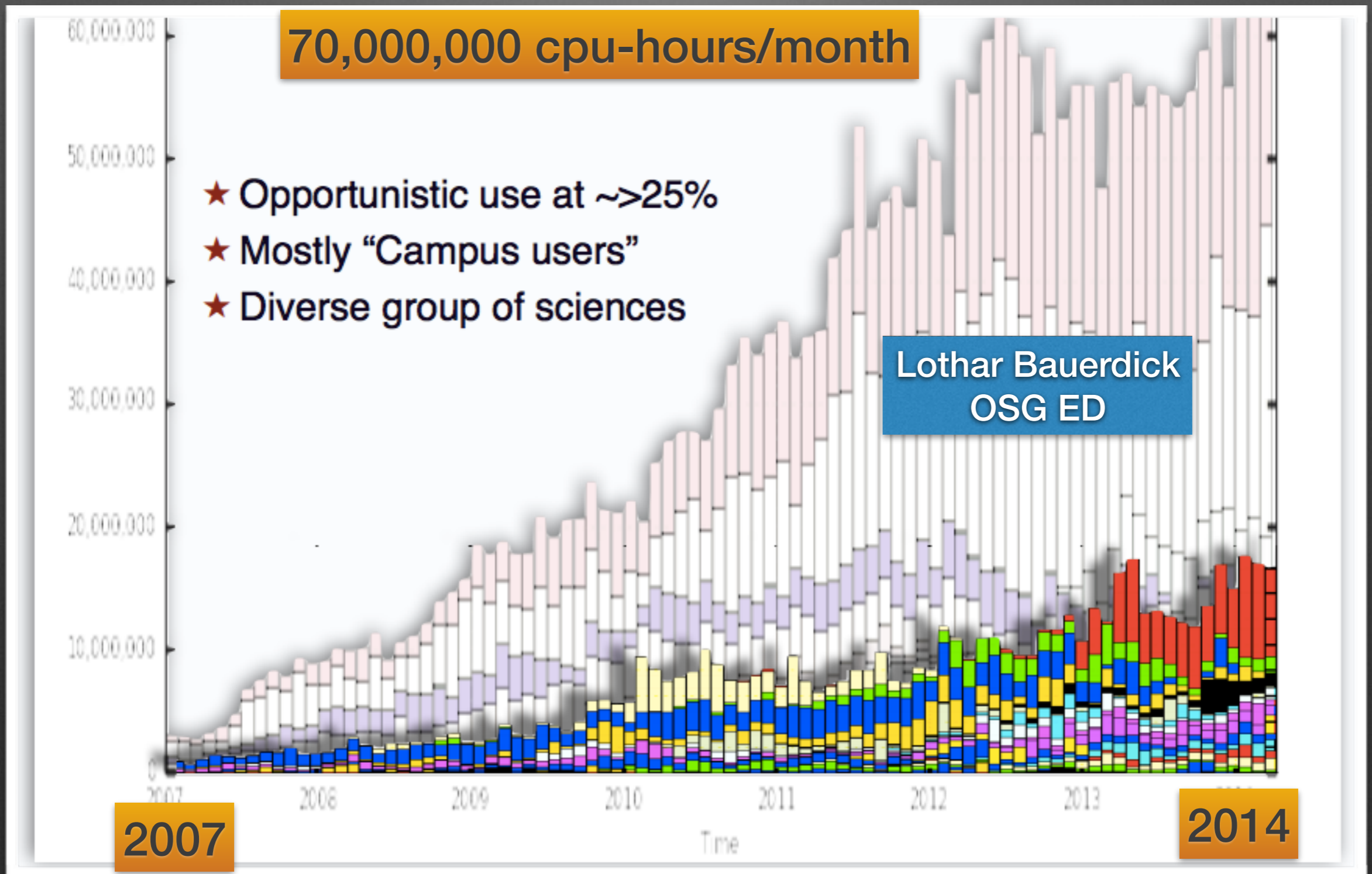


Open Science Grid



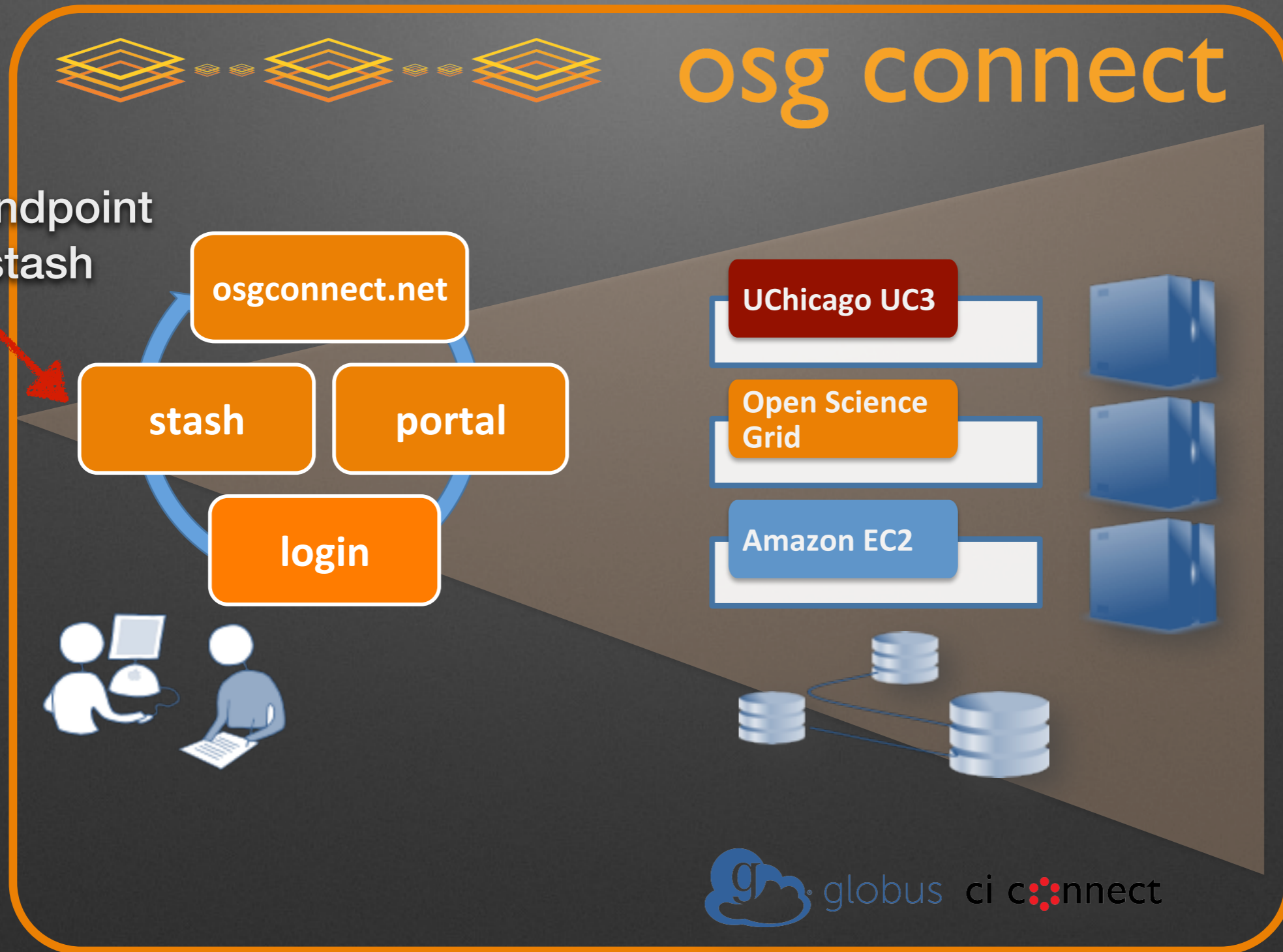
A great way to share empty job slots

Growing number of campus users



Login to the OSG

(direct access to those cycles)



Bringing resources together: Campus, OSG, Cloud

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

osg connect

Scheduling to OSG via flocking to the GlideinWMS service

Show: Historical grid usage in all pools

Time Frame: 3 Hours | Day | Week | Month

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
- lzheng
- lincoln
- johnnda
- jwebster
- yx5
- high.rynge
- strolog
- dweitzel
- high.kimfwong
- high.wdi1114

Version 4.0.15 licensed to University of Chicago, Computation Institute (OSG) Terms & Conditions © 2011 Cycle Computing, LLC. All Rights Reserved My Requests ▲

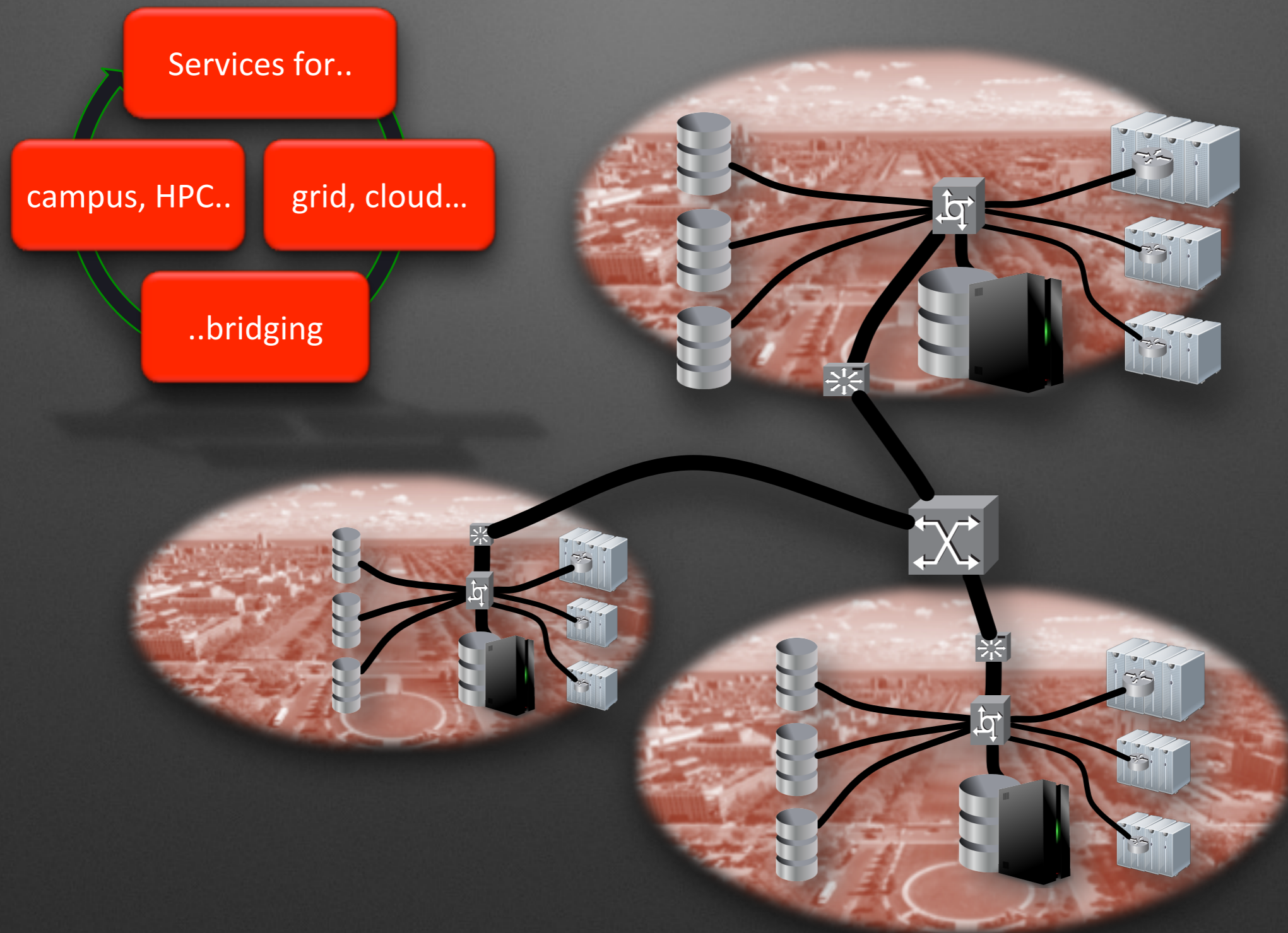
OSG Connect Suggested Campus Grids as Service

The logo for 'ci connect' is centered on the slide. It features the text 'ci connect' in a lowercase, sans-serif font. The 'c' in 'connect' is replaced by a cluster of six red dots arranged in a hexagonal pattern. The logo is surrounded by a circular arrangement of approximately 15 red ovals of varying sizes, creating a ring-like effect.

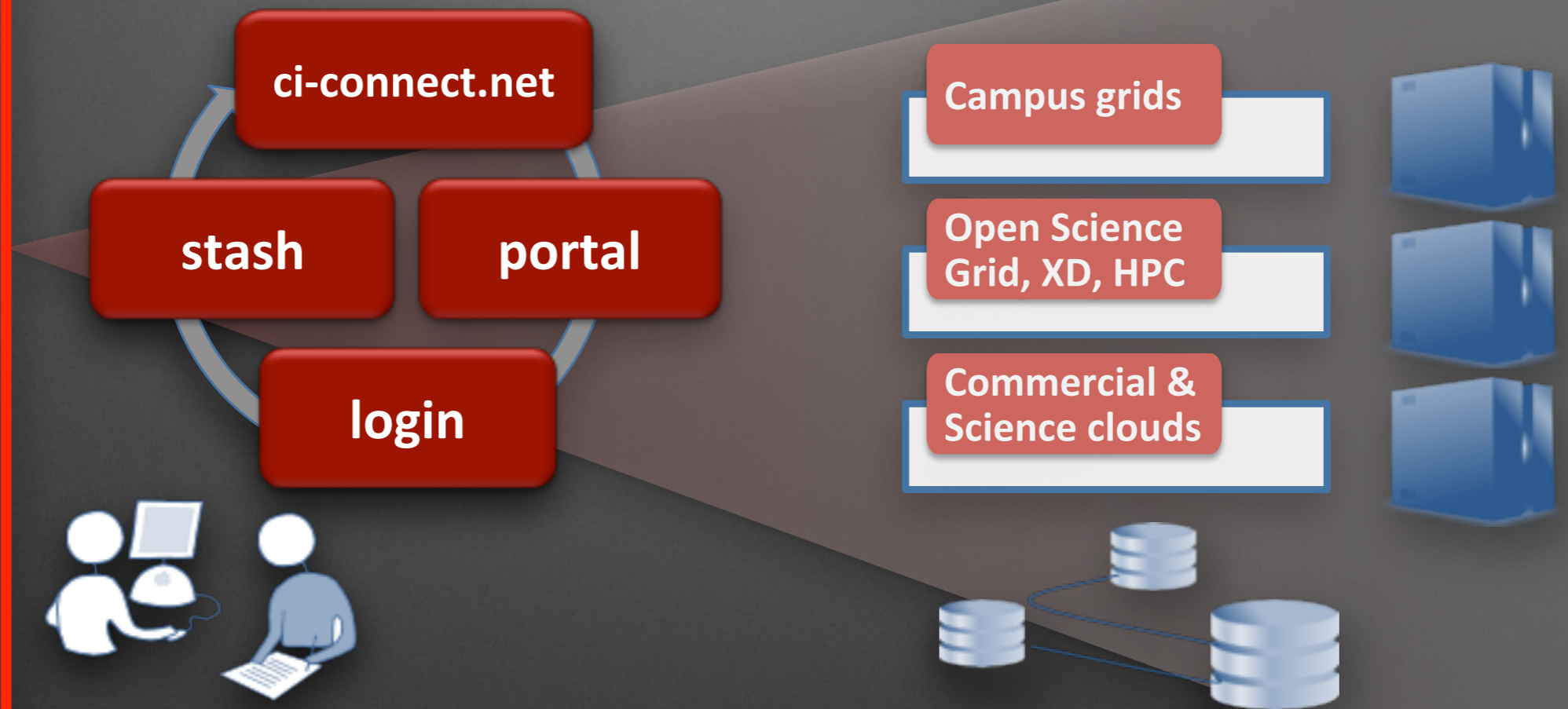
ci connect

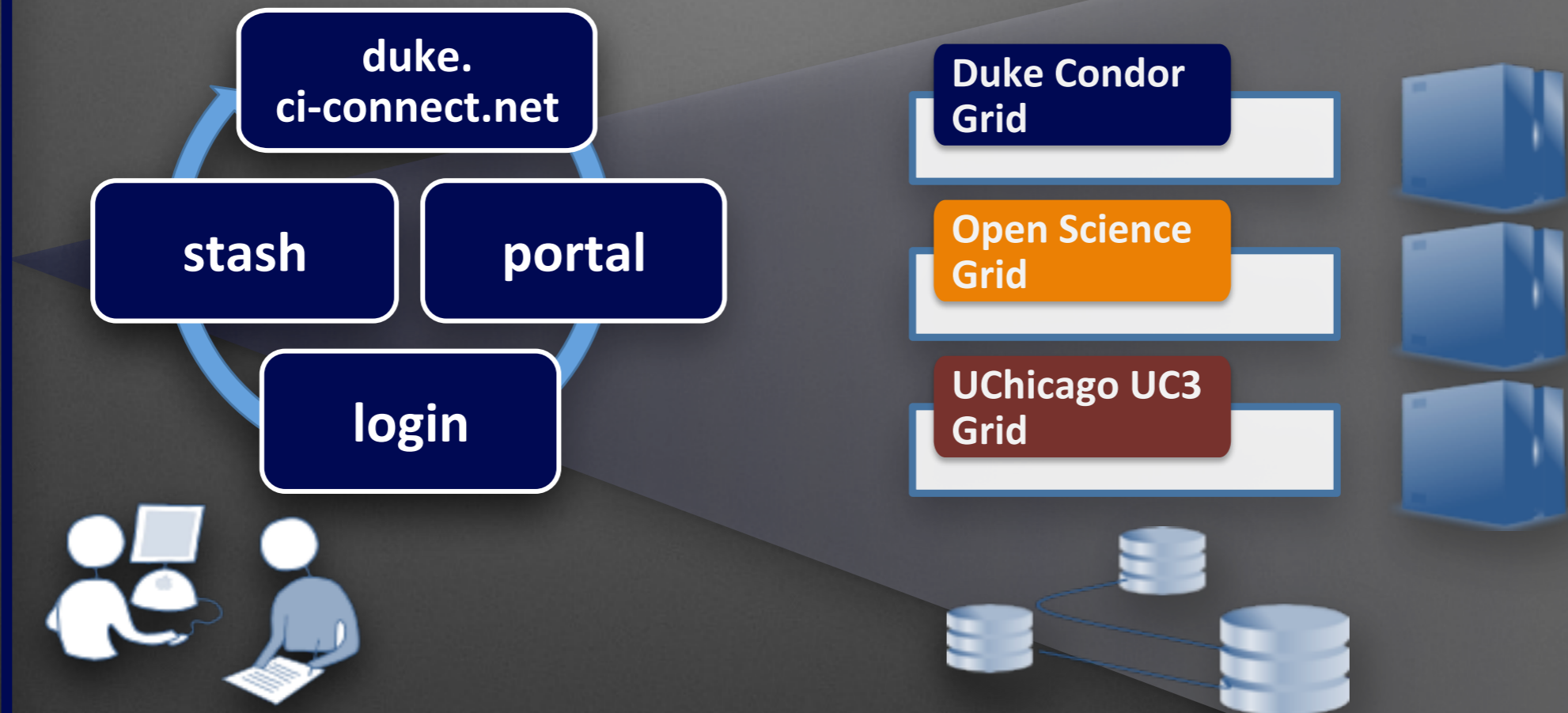
ci-connect.net built on Globus Platform and HTCondor

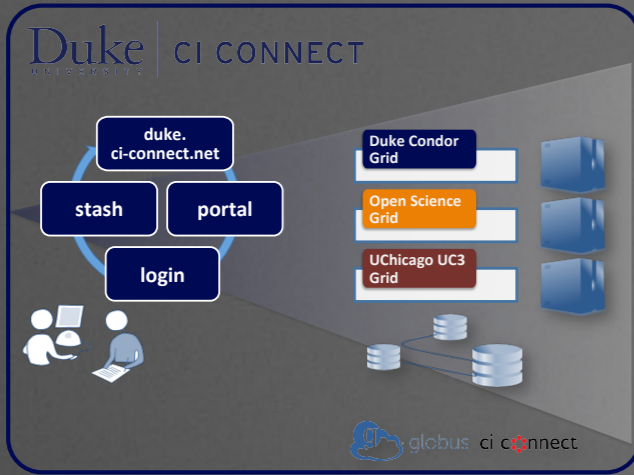
Services for Connecting



ci connect

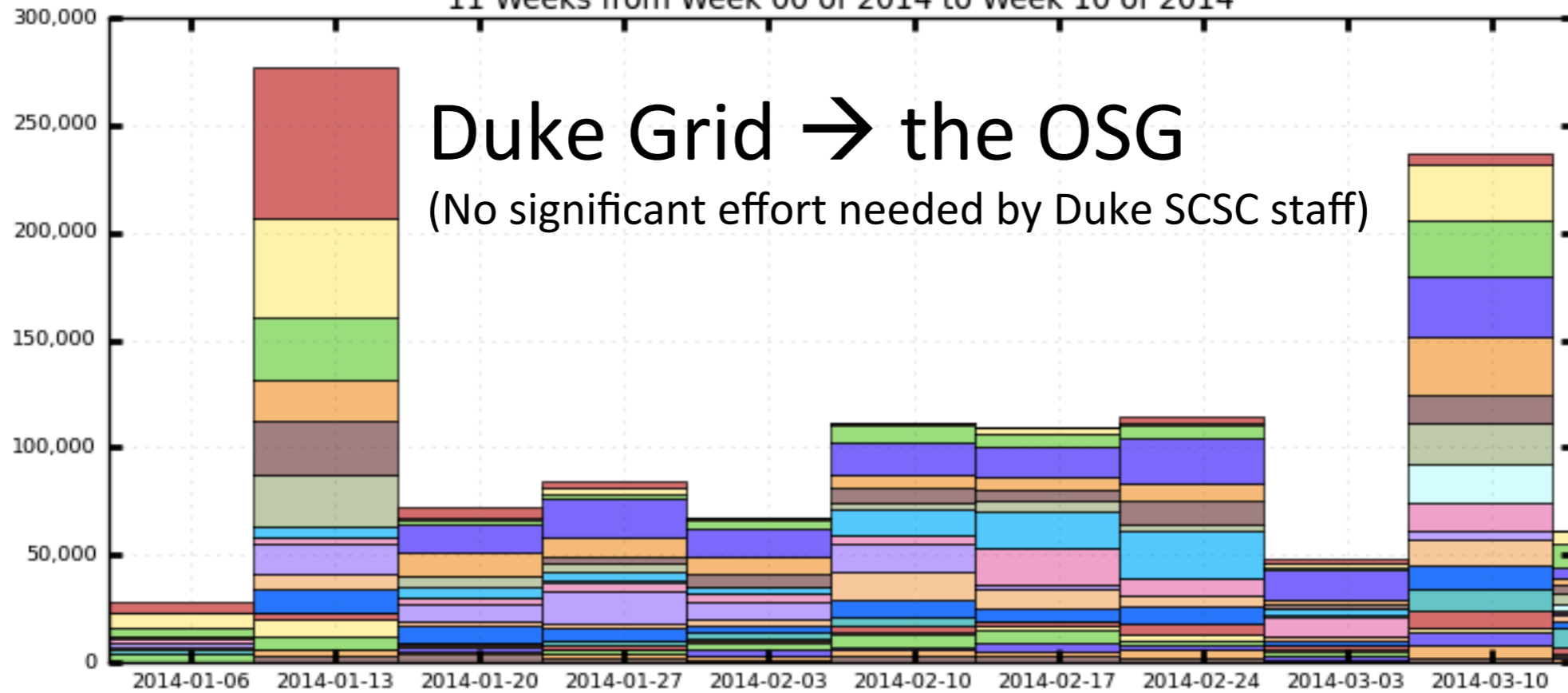






WMS Hours Spent on Jobs By Facility (Glidein)

11 Weeks from Week 00 of 2014 to Week 10 of 2014



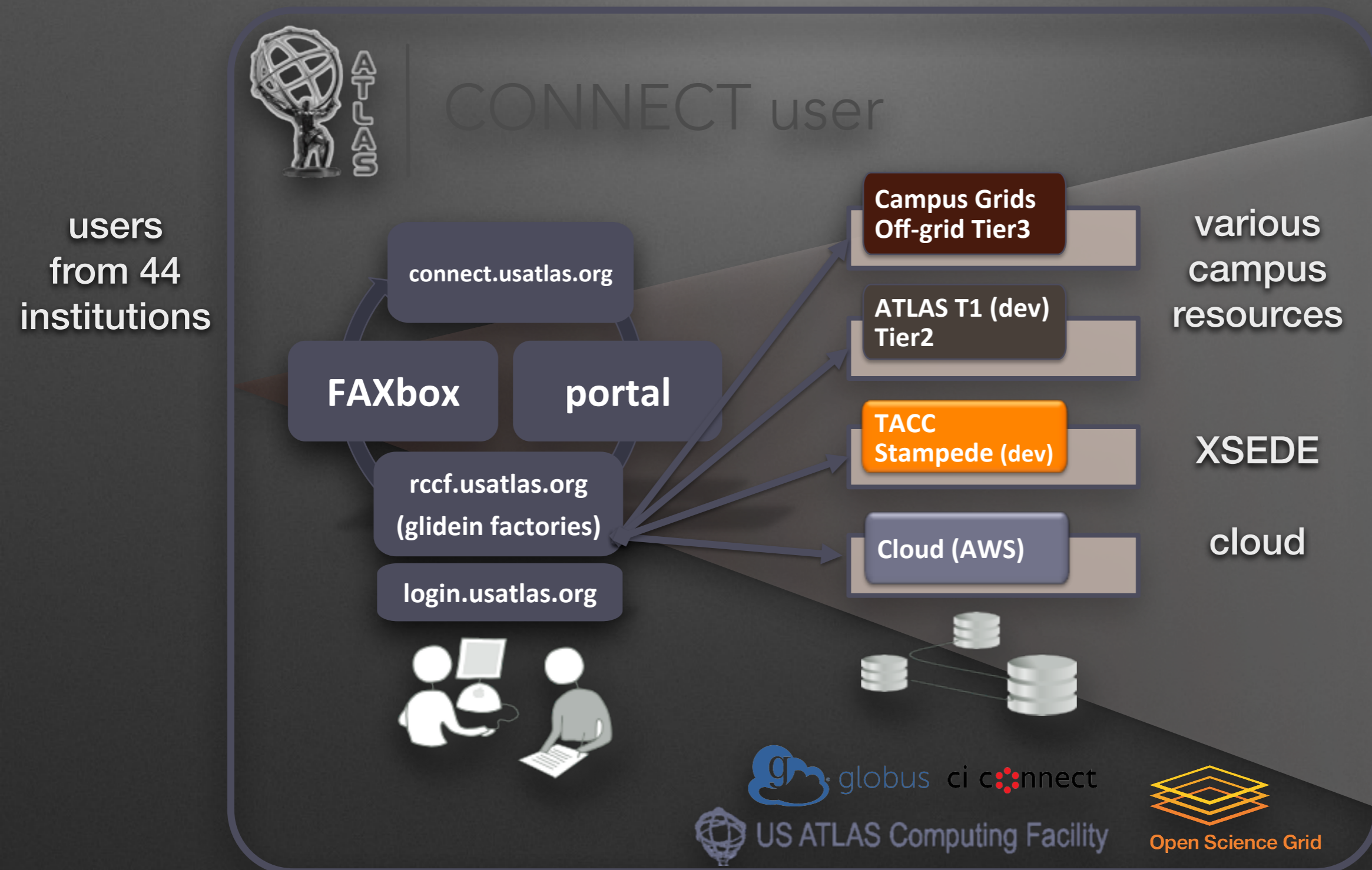
Duke Grid → the OSG

(No significant effort needed by Duke SCSC staff)

- | | | | |
|-------------------|-------------|--------|---------------------------|
| Crane-CE1 | Nebraska | UCSDT2 | login.duke.ci-connect.net |
| Sandhills | NWICG_NDCMS | Tusker | SMU_HPC |
| BU_ATLAS_Tier2 | Glux_VOMRS | MWT2 | Other |
| GridUNESP_CENTRAL | GLOW | AGLT2 | IU_OSG |
| CIT_CMS_T2 | Purdue-RCAC | UCD | cinvestav |

Maximum: 276,604 , Minimum: 28,237 , Average: 110,106 , Current: 61,593

Connecting communities to the national ecosystem



Strategy for XSEDE

TACC Stampede
<http://www.tacc.utexas.edu/stampede/>

P. Onysis, UTexas

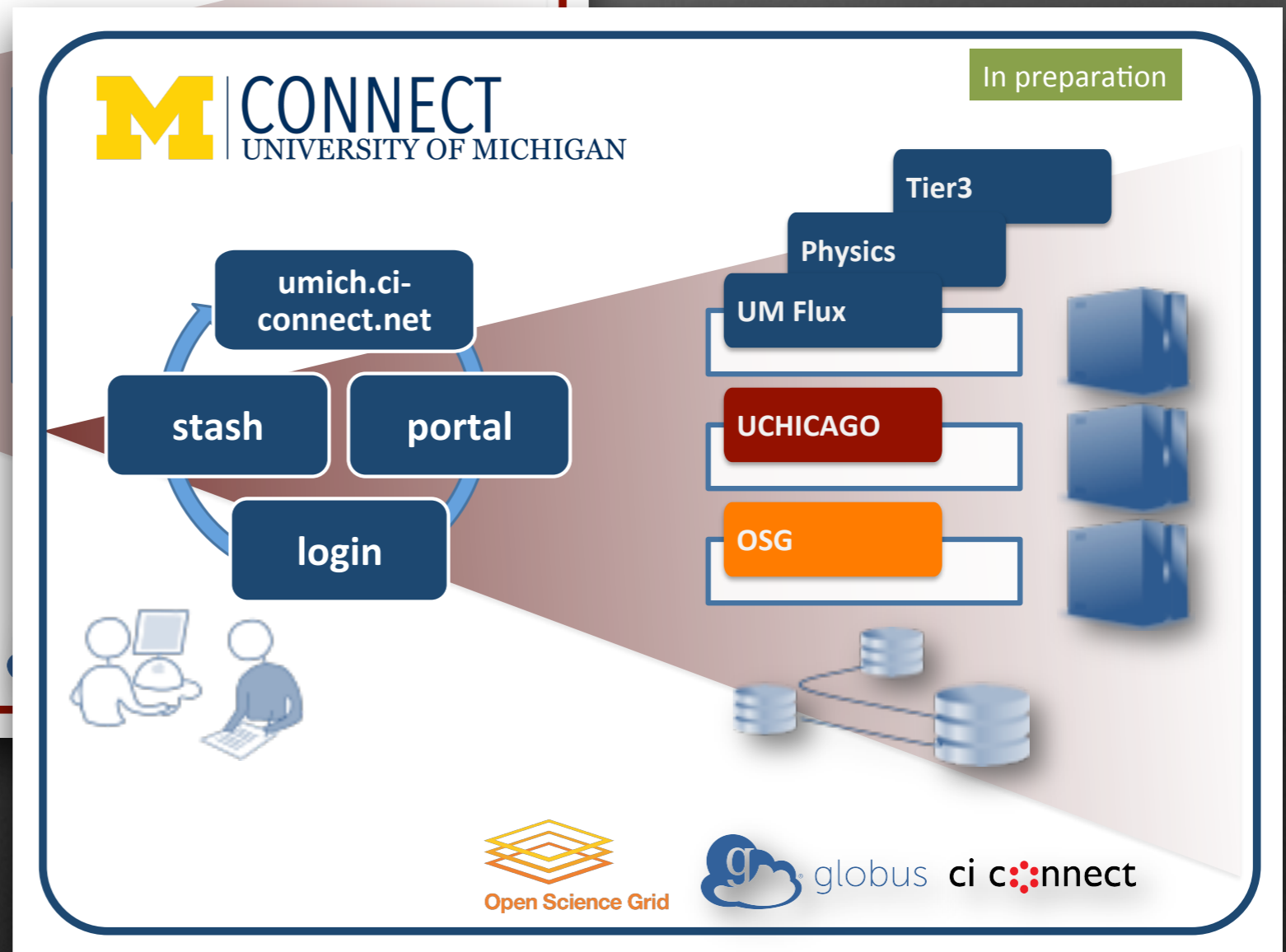
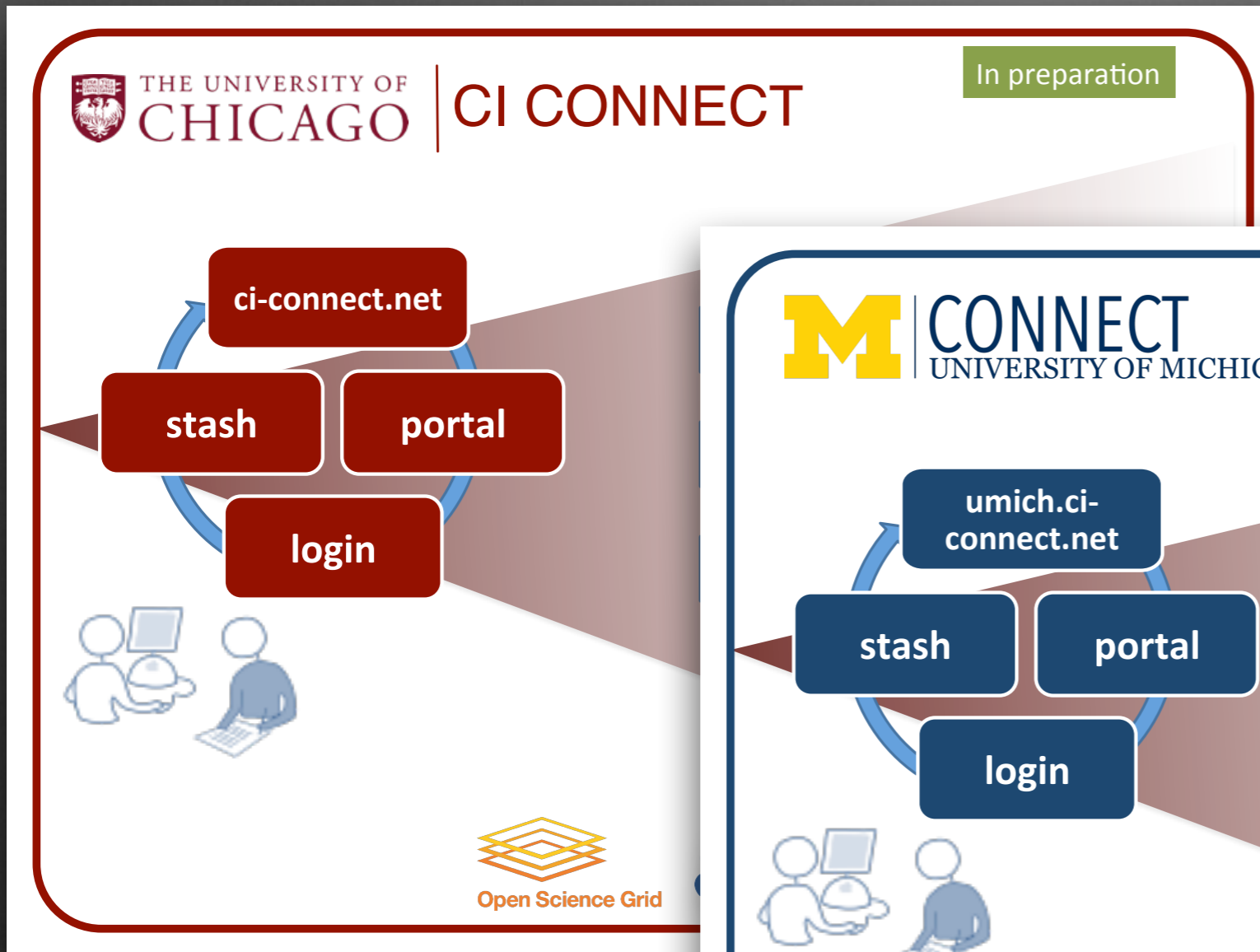
Science Gateway for ATLAS
gather lessons for campus bridging

6,400 nodes
102,400 Sandy Bridge cores
390,400 Xeon Phi cores
32 GB RAM/node
Special **highmem**, **NVIDIA** nodes
100 Gbps WAN
48 hr max job runtime

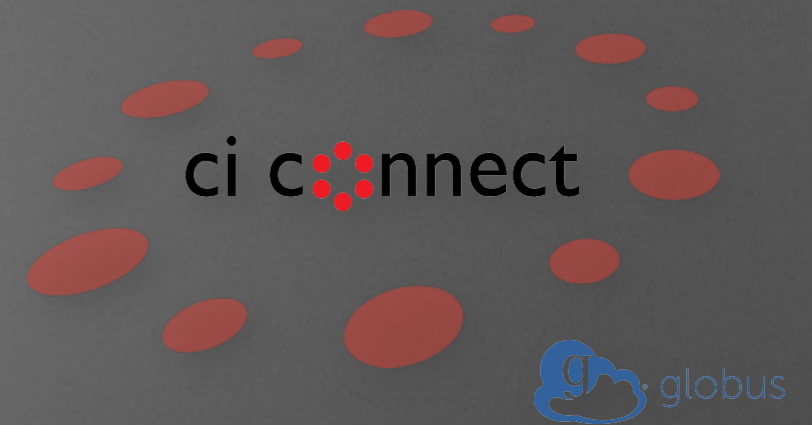
Minimize footprint @ TACC

- Key is minimizing Stampede admin involvement while hiding complexity for users
 - Simple SSH to Stampede SLURM submit node
 - ATLAS software mounted using CVMFS and Parrot
 - ATLAS squid cache configured nearby
 - Wide area federated storage access
- Leverage Globus, HTCondor, Glidein Factory, CCTools, OSG accounting
 - (CI Connect Services)

Campus Connect Services In Preparation



Summary



- Discovered we could leverage the Globus Platform to connect users to the Open Science Grid
- Suggested a model to couple users, data and distributed compute cycles as a service
- Easy path for campuses to “connect” or “bridge” to the national ecosystem

Acknowledgements

Steve Teucke, Rachana Ananthakrishnan + Globus Team!

Dave Lesny – UIUC (Midwest Tier2)

Lincoln Bryant, David Champion – UChicago (MWT2)

Suchandra Thapa (UChicago OSG)

Peter Onysis – UTexas

Jim Basney (CI-Logon) & InCommon Federation

XSEDE Science Gateway Team (Raminder Jeet, Suresh Maru, Marlon Pierce, Nancy Wilkins-Diehr)