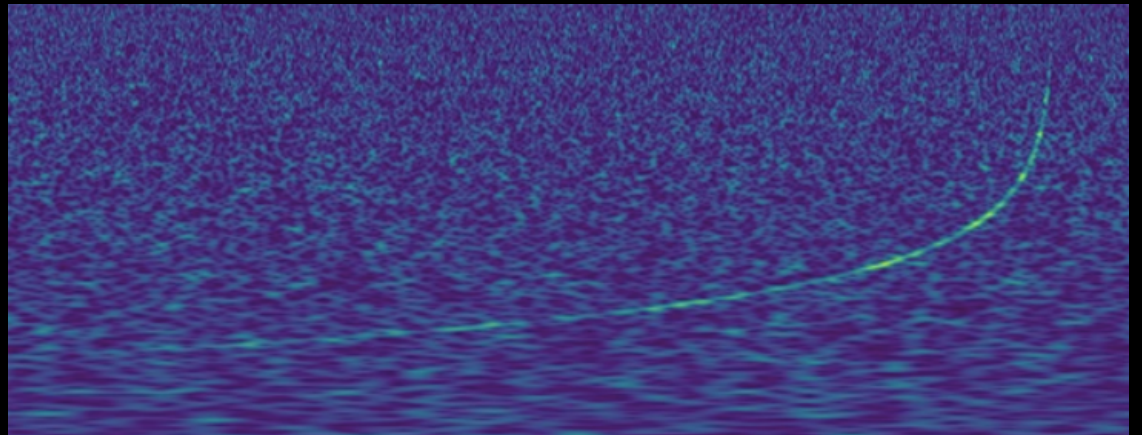# Open Science Grid: A Year in Review

Kyle Gross - OSG Operations
Pervasive Technology Institute - Indiana University

- Events of Importance

- OSG Area Successes

- *Mindless Musings* or *Down The Rabbit Hole*

LIGO

- LIGO "came home" and joined the OSG Council.

- LIGO Nobel Prize

- Multi-messenger astronomy

Hours, Hours, and More Hours

- Record week in April of 32 million CPU hours.

- Broke over 1,000,000,000 CPU hours per year.

- Currently have provided 1,588,941,000 CPU Hours in the last year.

More Science

- OSG Connect allowing great access to resources.

- Over 400 new users in 2017 alone.

- 36 fields of science in 346 fields of science.

- IceCube, Veritas, XENON1T, South Pole Telescope

Highlights by Functional Area

- Outreach - Kyle Gross

- Networking - Shawn McKee

- Operations - Rob Quick

- Production Support - Ken Herner

- Release - Tim Theisen

- Security - Susan Sons

- Software - Brian Lin

- Technology - Brian Bockelman

- User Support - Rob Gardner

Outreach

- Several high profile Science Highlights

  - LIGO Nobel Prize, LIGO/VIRGO/DECam, ScienceNode articles

- Merged with Education to become Outreach

- Twitter and website getting additional attention

Education

OSG Summer School

- July 17-21, Madison Wisconsin

- 56 participants from around the world

- Biological Sciences, Biomedical Sciences, Computer Sciences, Earth Sciences, Ecology, Economics, Engineering (most branches), Health Sciences, Physics, Plant Sciences, Psychology, Statistics, Zoology

- 8th annual school

- AHM Speakers who previously attended Summer Schools:

  - Hossein Saadatzi - Bipedal Walking Simulation

  - Ariella Gladstein - Ecology and Evolutionary Biology

  - Anthony Mills - Metabolomics

Registration for 2018 school now open:
https://www.opensciencegrid.org/UserSchool

CODATA/RDA School of Research Data Science

- Two week foundational Data Science school at ICTP in Trieste, Italy and ICTP/UNESP in Sao Paulo, Brazil

- ~40 students per class focused on students from Low and Middle Income Countries

- 2017 two instances in Trieste, July and Sao Paulo, December

- Students learn core Data Science Concepts, Data Management, Neural Networks and Machine Learning, and using International Cyberinfrastuctures

- Heavily leverage OSG-Connect and use DOSAR for support

- Very active student from Colombia (UPCDOSAR)

CODATA/RDA School of Research Data Science

- 2018 Scheduled Events

  - August 6-17 ICTP Trieste, Italy

  - October 22 - November 2 ICTP Kigali, Rwanda

  - December TBA - ICTP Sao Paulo, Brazil

  - Negotiating US based class in late summer

    - Plan to charge registration fee and use proceeds to help fund the LMIC events

- Target students from Low and Middle Income Countries worldwide for Trieste

- Target students from Africa for Kigali

- Target students from South and Central America for Sao Paulo

- Open registration for US Based Event, though some sponsorship may be available for students from underserved institutions

- Curriculum remains core Data Science Concepts, Open Science, Data Management, Neural Networks and Machine Learning, and using International Cyberinfrastuctures

  - Additional security module provided by CERN in Trieste

- Plan to continue to leverage OSG-Connect

Security

Zalak talks after me.

You should listen to him.

Networking

Shawn talks after Zalak.

You should listen to him, too.

Operations

- TWiki Turndown

- Single Sign-On

- Ansible management: Anyone on staff can rebuild a service instance; maintenance is a snap

- Configuration Management Database complete

- RSV elimination

Production Support

- ATLAS/CMS: running on more and more HPC machines

- FIFE: now on additional US and EGI sites, and adoption of StashCache by several experiments

- GlueX going again

- LIGO processing multiple triggers in parallel

Release

- Adopted flexible release policy. Still releasing about once per month.

  - Have been able to adjust our schedule to accommodate external development time-lines.

- In cooperation with Security Team, streamlined the CA certificate release process. Typical turn around is 2 days rather than 2 weeks.

- Testing HTCondor pre-releases on the Madison ITB.

Software

- Streamlined the OSG software stack with the release of OSG 3.4, which is less than half the size of OSG 3.3!

- Retired BDII and the osg-info-services in favor of the central CE collector

- Migrated the TWiki documentation to Github Pages and revamped the documentation update and review process

- Formed the GridC Community Forum, the organization dedicated to maintaining the Grid Community Toolkit (the open-source fork of the Globus Toolkit)

- Released Hadoop 2.6 into OSG upcoming; documentation and the move into the main OSG release incoming
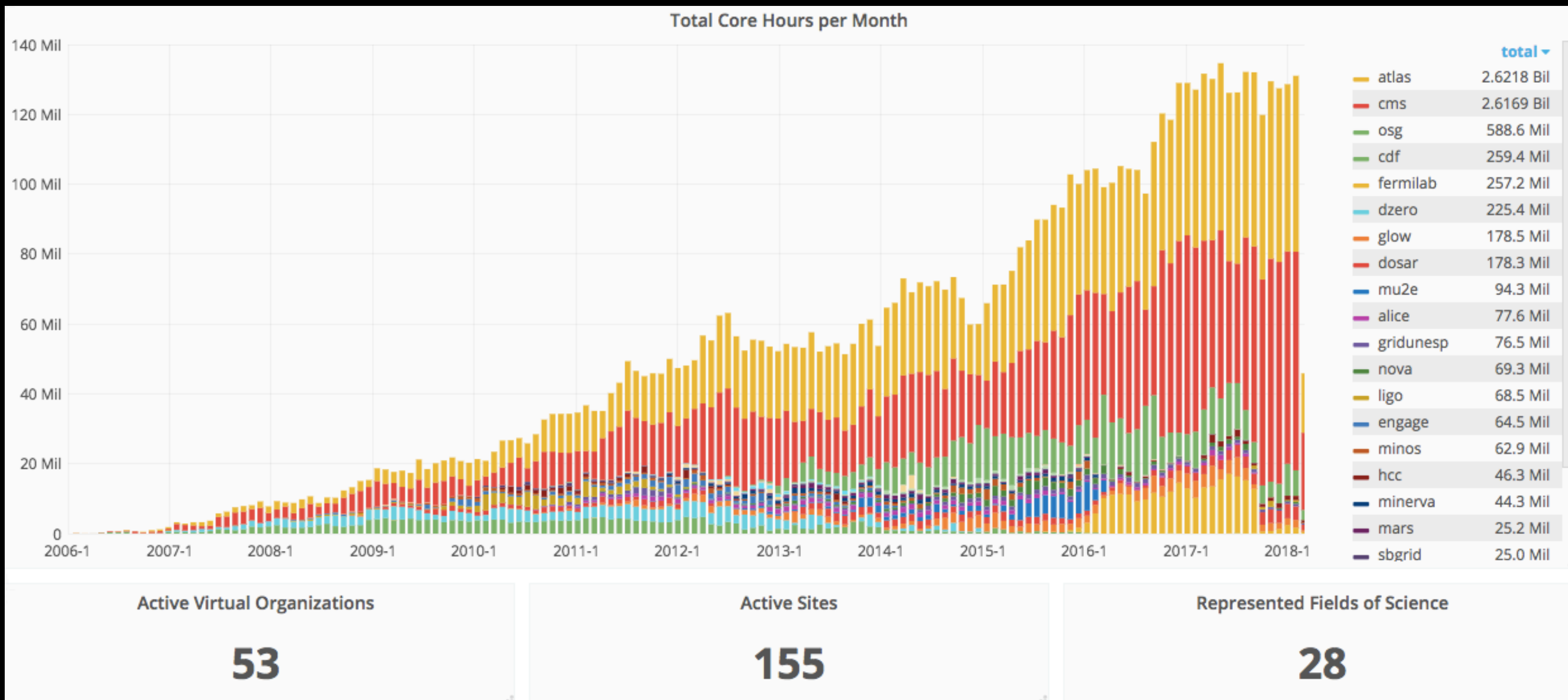
Technology

- Planning for end-of-Globus support and the post-Globus era.

- Singularity: Increasingly popular service with the user community.

- GRACC: Moving from "replacing prior accounting system" to managing more interesting accounting datasets beyond CPU usage (storage data, networking data, improving job annotations).

- StashCache: Explosion in traffic over the past year as it is better integrated with the remaining.

- With User Support, helping communities evaluate Rucio for data replication management.

- Continuing to retire software at the end-of-lifecycle to make room for exciting new projects.

User Support

- HostedCE, now providing ~100k hours/week, campus clusters, and XSEDE resources with allocations

- Support for larger communities/projects: SPT, XENON1T, GlueX, VERITAS, nEXO, FSurf

- Singularity work to support default and custom images + GPU

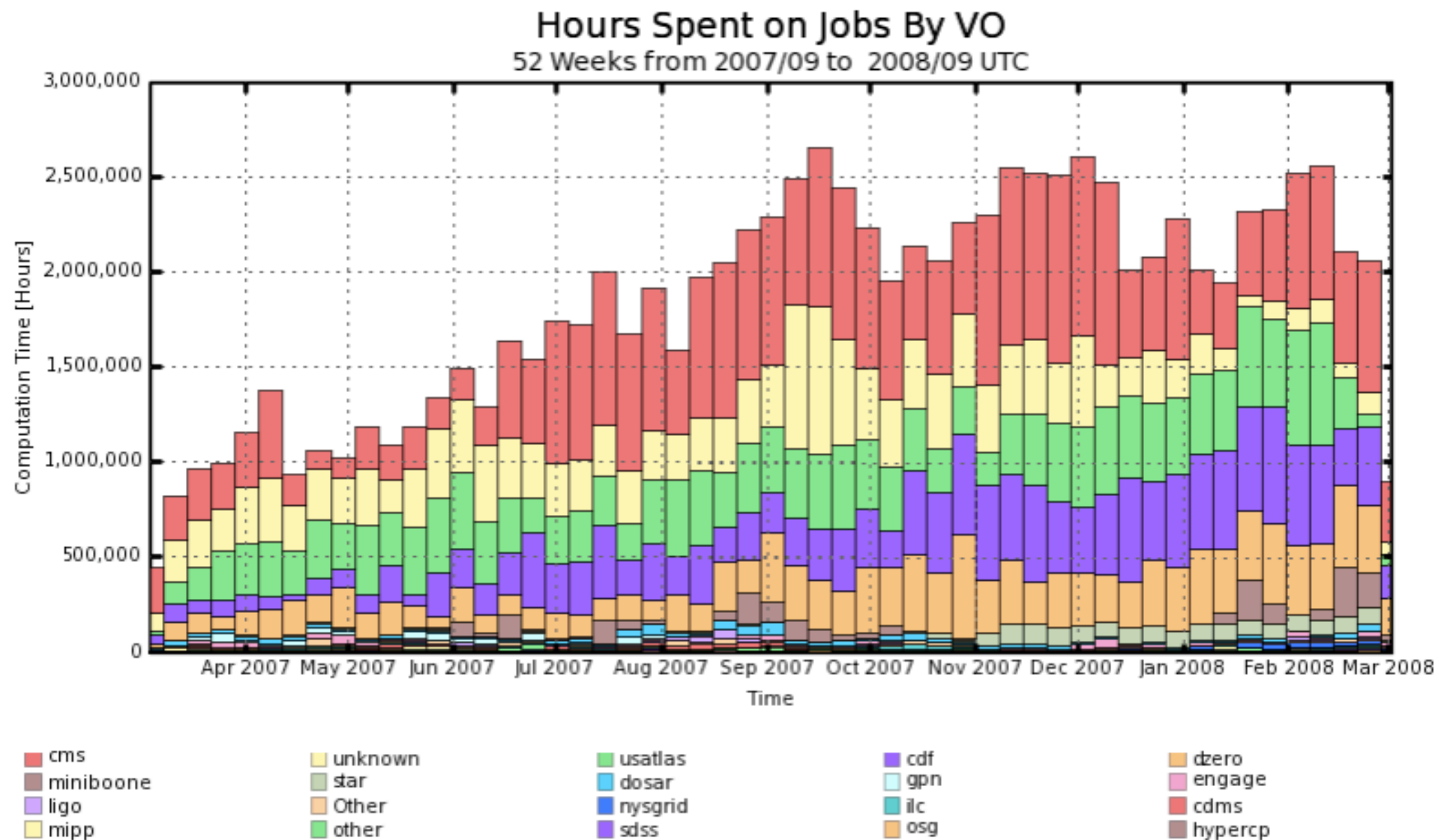- Training and workshops at PEARC'17, RMACC, OSG-AHM, JLab, CyVerse

The sum of these many successes

Total Core Hours per Month

total ▾	
atlas	2.6218 Bil
cms	2.6169 Bil
osg	588.6 Mil
cdf	259.4 Mil
fermilab	257.2 Mil
dzero	225.4 Mil
glow	178.5 Mil
dosar	178.3 Mil
mu2e	94.3 Mil
alice	77.6 Mil
gridunesp	76.5 Mil
nova	69.3 Mil
ligo	68.5 Mil
engage	64.5 Mil
minos	62.9 Mil
hcc	46.3 Mil
minerva	44.3 Mil
mars	25.2 Mil
sbgrid	25.0 Mil

Active Virtual Organizations	Active Sites	Represented Fields of Science
53	155	28

Why are we not backfilling ATLAS temporary reduction in use?

Open Science Grid

Hours Spent on Jobs By VO
52 Weeks from 2007/09 to 2008/09 UTC

Maximum: 2,659,041 Hours, Minimum: 449,054 Hours, Average: 1,833,457 Hours, Current: 898,228 Hours
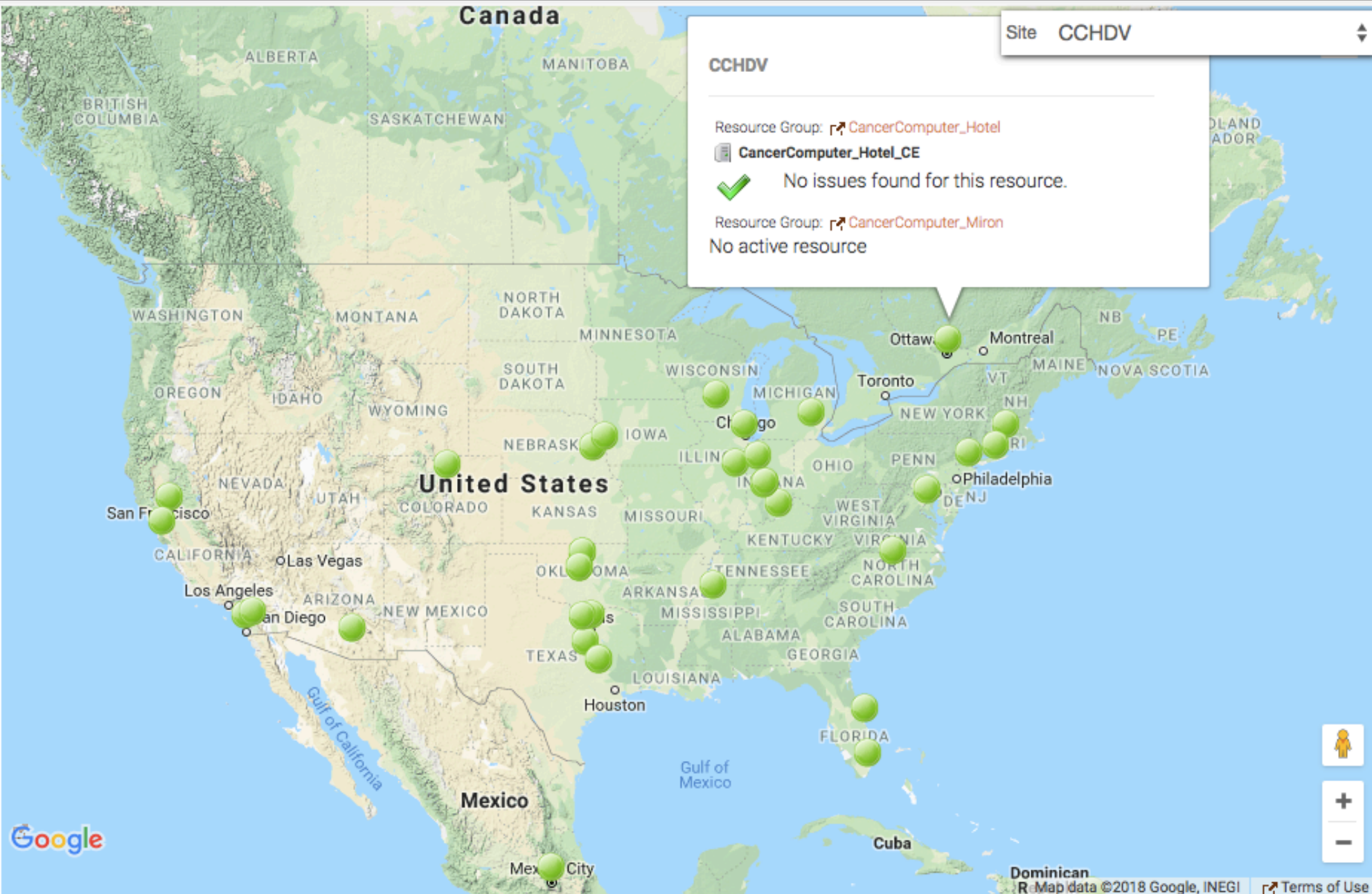
Site CCHDV

CCHDV

Resource Group: ↗ CancerComputer_Hotel
📄 **CancerComputer_Hotel_CE**

✓   No issues found for this resource.

Resource Group: ↗ CancerComputer_Miron
No active resource

Update Page

INFORMATION TO DISPLAY

☐ Show non-reporting sites
☑ Show Campus Grid

SITES TO DISPLAY

☑ All Sites
☐ Sites in Facility ❓
☐ Sites supported by Support Centers ❓

FILTER

Only show sites that matches following criteria
(conditions are and-ed).

☑ Has resource Group with grid type of
    ☑ OSG
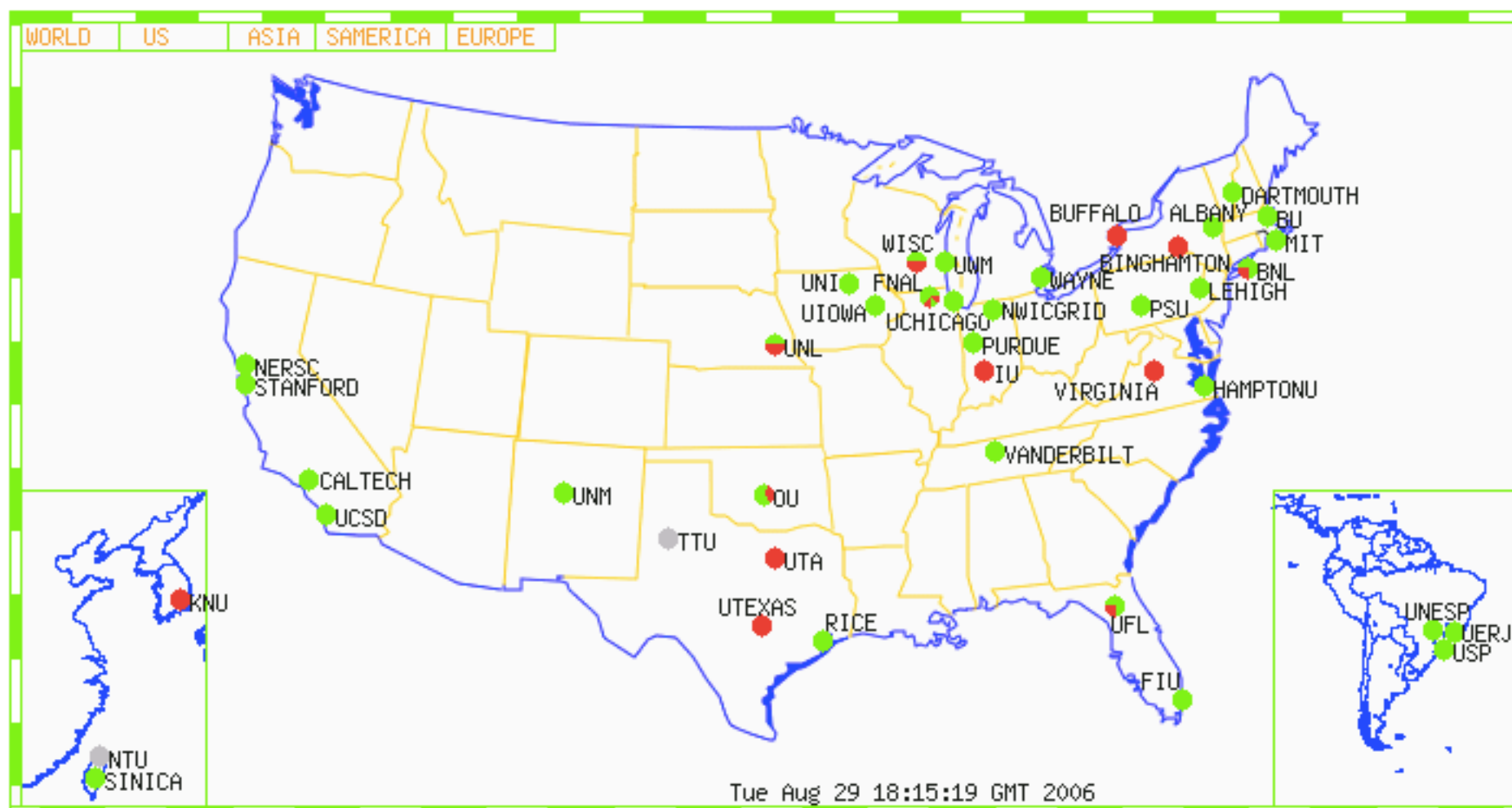   ☐ OSG-ITB

☑ Active Status
    ◯ Inactive
    ◉ Active

☐ Disable (Removed) Status

Update Page

Subscribe to the content of this page.

↗ Netvibes
↗ Mobile Content

GridCat

Open Science Grid

Service: CS = Compute Service, SS = Storage Service, WS = Web Service GRAM

WORLD | US | ASIA | SAMERICA | EUROPE

Tue Aug 29 18:15:19 GMT 2006

1 to 72 of 72 ◁ ▷ sort by: Service   entries per page: 100   map: U.S.+Asia+S.America   view: Summary

Status	Site Name	Grid Version	Jobs?	Disks?	Service	Loc	Facility
●	**Purdue-Physics**	**osg 0.4.0**	0/63	1179/1492	CS	IN	PURDUE
●	**ASGC_OSG**	**osg 0.4.0**	40/178	1566/1820	CS	TAIWAN	SINICA
●	**Purdue-ITaP**	osg 0.4.1	48/616	253/305	CS WS	IN	PURDUE
●	**NERSC-PDSF**	osg 0.4.1	452/1028	14805/1526	CS WS	CA	NERSC
●	**GRASE-ALBANY-NYS**	osg 0.4.1	0/0	dwn	CS WS	NY	ALBANY
●	**UFlorida-EO**	osg 0.4.1	0/20	569/748	CS WS	FL	UFL
●	**NWICG-NotreDame**	osg 0.4.1	dwn	dwn	CS WS	IN	NWICGRID
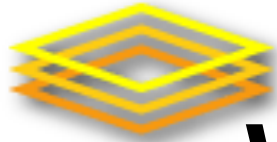
Why #2: OSG is an Experimental Tool for Computer Science

"The OSG is an exceptional experimental tool for computer scientists as well as a powerful resource for physicists, biologists and chemists."

– Miron Livny, U.Wisconsin-Madison
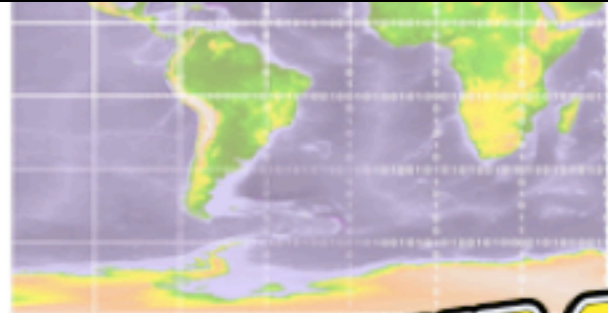
Core Operations and Common Support

1 2 3 - Join OSG

1. VO Registers with with Operations Center. User registers with VO.

2. Sites Register with the Operations Center.

3. VOs and Sites provide Support Center Contact and join Operations groups.

The OSG VO

1. A VO for individual researchers small groups.

2. Managed by the OSG itself.

3. Where one can learn how to use the Grid!
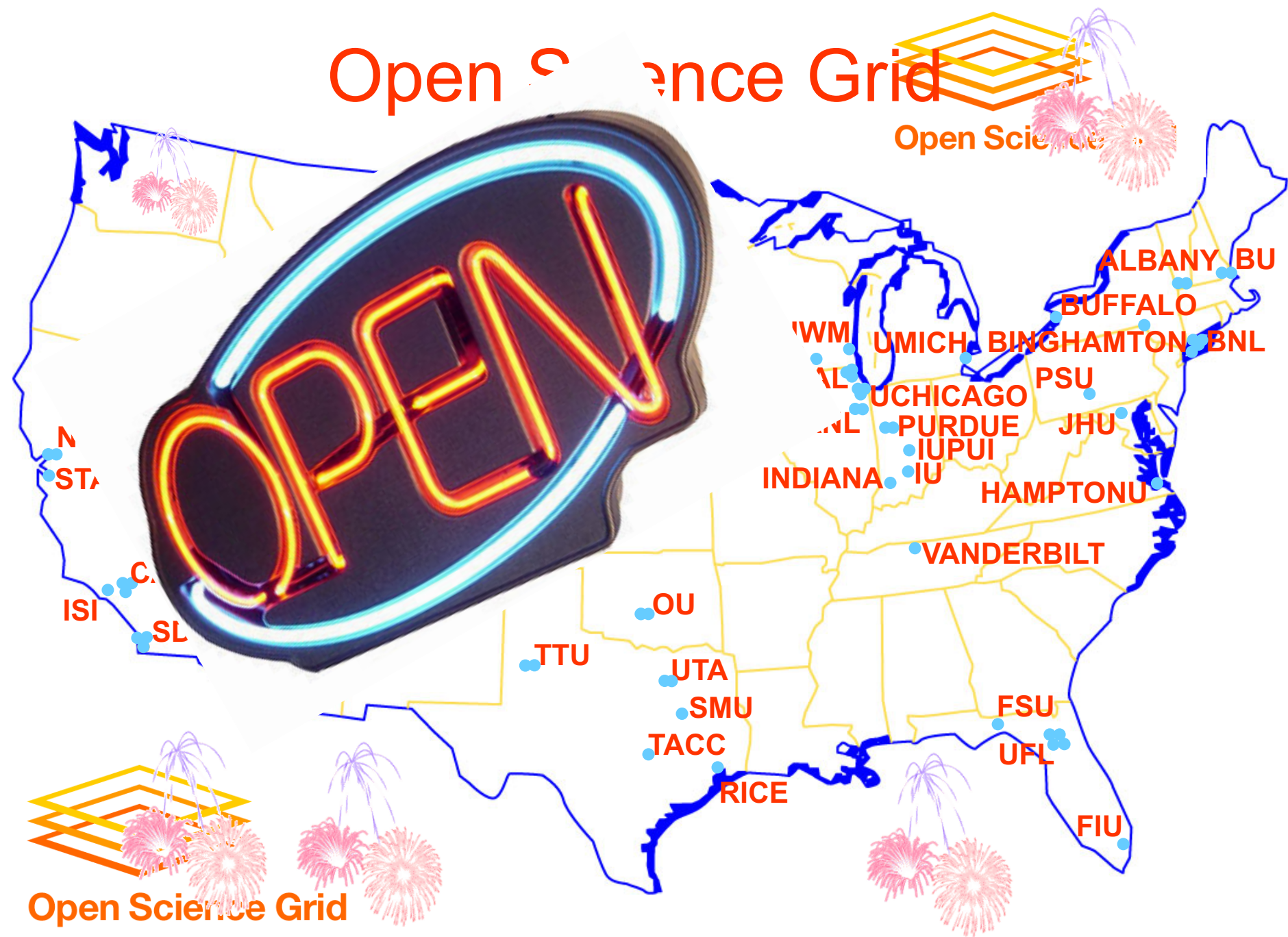
Open Science Grid

"The future of every community lies in capturing the passion, imagination, and resources of its people."

–Ernesto Sirolli