Containers on OSG

Frank Würthwein
OSG Executive Director
Professor of Physics
UCSD/SDSC

The OSG Vision in a nutshell

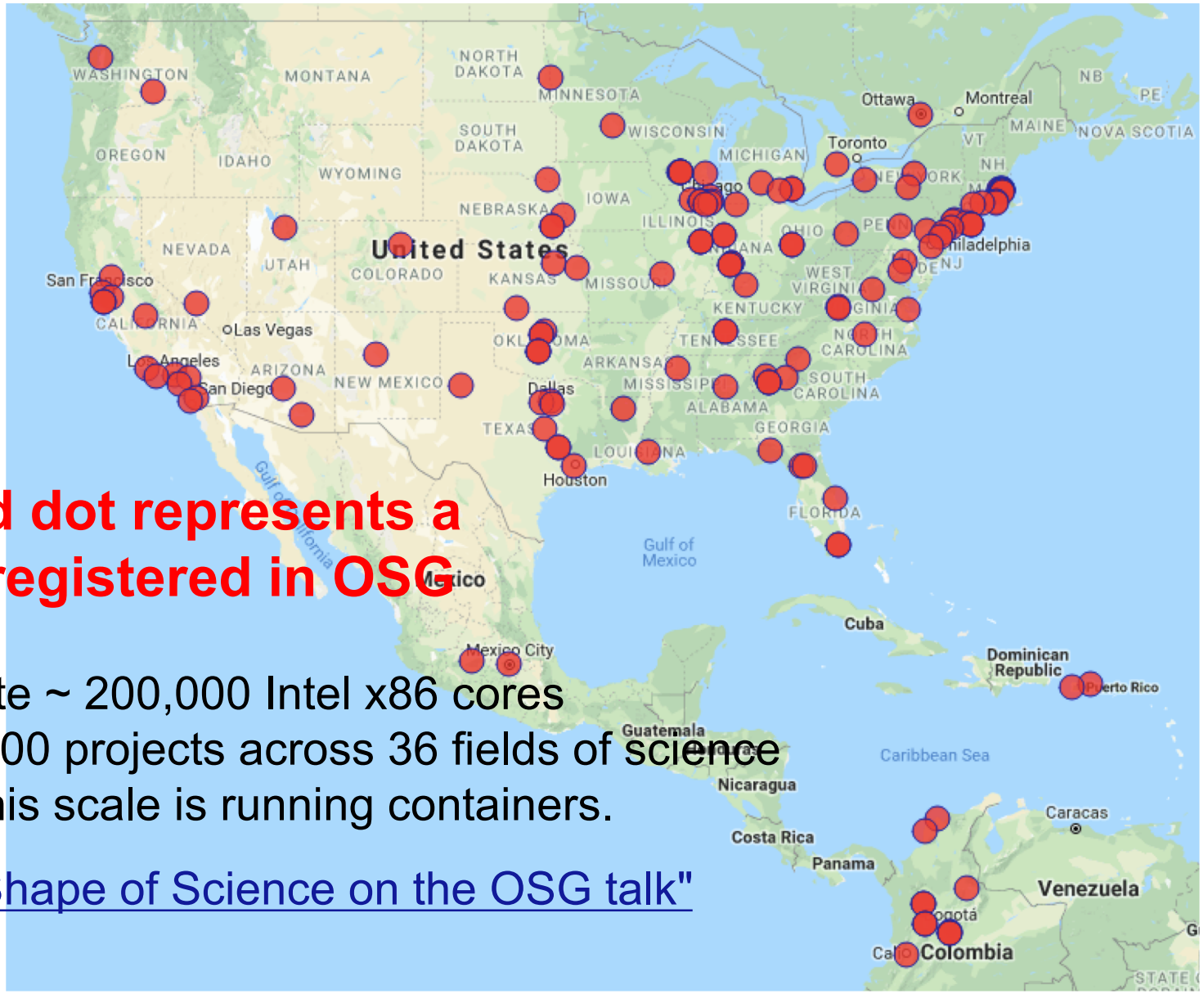
- Across the nation, institutions invest into research computing to remain competitive
- Science is a team sport, and institutions with an island mentality will underperform.
- Integration is key to success

OSG empowers researchers to use compute & data resources across institutional boundaries.

Commonality Across Science

- OSG's business model is to empower Scientists and their home institutions to work together for long term sustainability.
  - OSG provides knowledge & software infrastructure.
  - OSG can offer storage and service hosting to jump start projects, but prefers to enable institutions for growth and sustainability.
- OSG provides global integration across commercial and academic computing.
  - OSG respects local ownership and control.

Integrating Compute & Storage Clusters

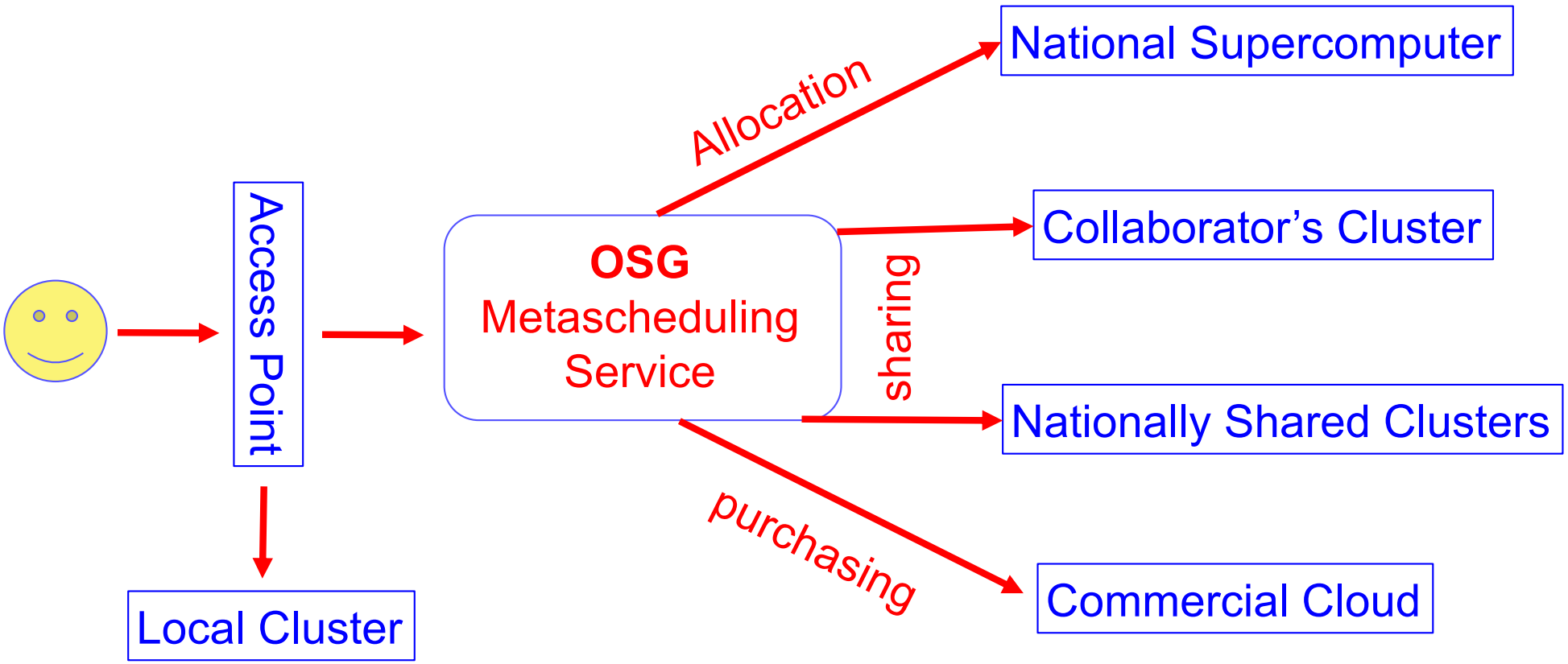


Each red dot represents a cluster registered in OSG

In aggregate ~ 200,000 Intel x86 cores used by ~400 projects across 36 fields of science
~ 80% of this scale is running containers.

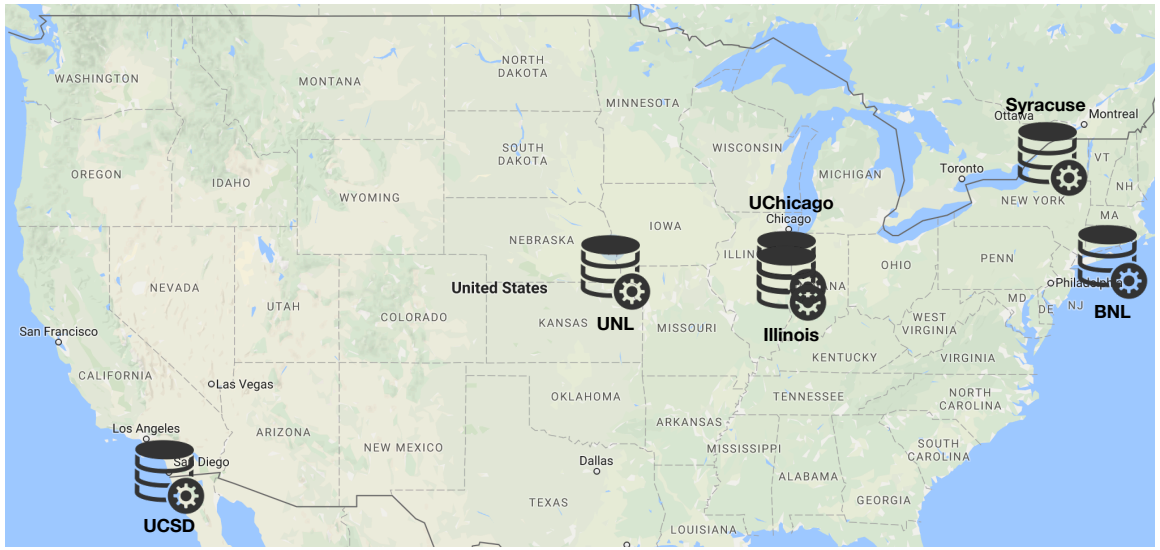
See "Shape of Science on the OSG talk"

Transparent Computing across different resource types

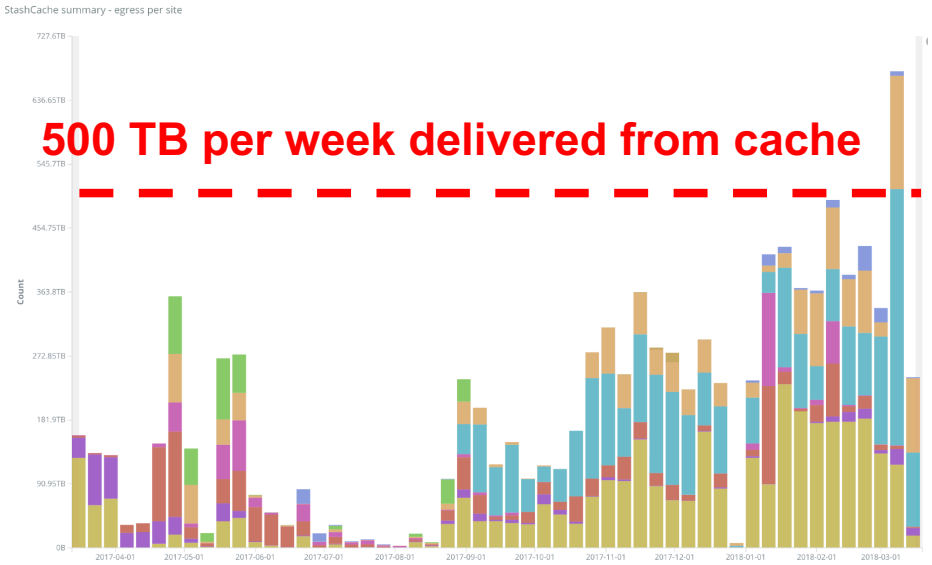


OSG integrates computing across different resource types and business models.

Caching Service



[See Derek Weitzel's talk at OSG AHM 2018 for more details]



500 TB per week delivered from cache

Among the users:

Minerva, Minos, DES, Xenon1T, LIGO, …

… but also single PIs and small groups especially from Bioinformatics analyzing genomics data.

Containers and OSG

- Containers are part of our DNA for longer than they existed, in a way.
  - Distributed High Throughput Computing for purists is all about making applications portable to be able to run anywhere.
    - In the old days this was a.out
    - … then it became elaborate user written scripts to package executables with all their dependencies as tgz.
    - … now it's containers

OSG Overlay batch system

- We are based on the late-binding principle, i.e. an overlay batch system gets submitted to the resources.
- Containers thus play multiple roles:
  - Containers for the glideins (i.e. the batch system we submit)
  - Containers for the payloads, i.e. user jobs
    - Different users get separated from each other to allow for clean sharing of resources inside the glidein.
- We may use containers to produce the appearance of homogeneity for the users in an environment that is fundamentally heterogeneous.
  - users may bring their own containers but are not required to do so.
  - Scientists are not sysadmins, and should not be forced to become such.

Example uses of Containers on OSG

- Providing singularity containers for others to build on.
- Hiding some of the GPU complexity from users.
- Machine Learning on the Grid.

Containers for others to build on

- OSG starts with Docker containers, and derives the singularity containers from there.
  - While most sites support singularity, some large places don't, or prefer shifter.
    - E.g. LIGO use of BlueWaters was done running OSG containers on shifter. The same container we run as singularity container elsewhere.
- At present, CMS, LIGO, Xenon1T, IceCube, SBGrid, and GlueX in addition to the OSG VO all build their operations on OSG on top of the OSG supported containers.

GPUs on the grid

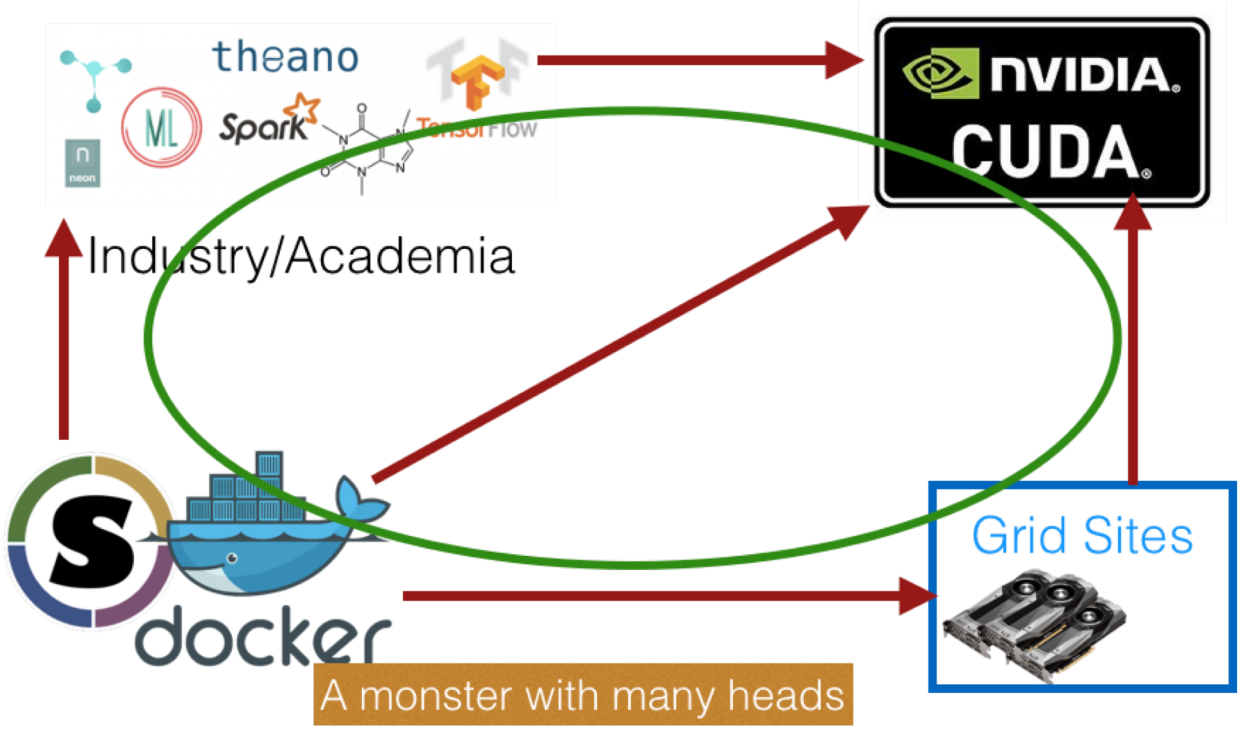
- Currently 4 institutions provide GPU access via OSG
  - Nebraska, UCSD, Syracuse, Vanderbilt
- The complexity and rate of change in GPUs introduces a heterogeneity explosion to the science community.
- We created 3 singularity images to soften the blow a little.
  - MultiPurpose RHEL6 + CUDA
  - MultiPurpose RHEL7 + CUDA
  - Tensorflow (Ubuntu) + CUDA

More details on ML on OSG

OSG establishes environment

- User jobs are run inside a wrapper script that checks first the “+SingularityImage” jobAd and then runs the job in the image chosen by the user.
- Wrapper script binds the GPU to the drivers
- Runs condor_gpu_discovery
- Updates the Startd Ad CUDACapability, CUDADeviceName, and CudaDeviceNumber(s) assigned to this overlay batch slot.
- Number of GPUs per slot is set by the site.
  - Most sites set it to 1 per slot.

Complexity …



… that we try to hide from scientists as much as possible.

opensciencegrid

osgusers

user-support@opensciencegrid.org

support.opensciencegrid.org

www.opensciencegrid.org/links

"Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program."

The Networking and Information Technology Research and Development (NITRD) Program

Mailing Address: NCO/NITRD, 2415 Eisenhower Avenue, Alexandria, VA 22314

Physical Address: 490 L'Enfant Plaza SW, Suite 8001, Washington, DC 20024, USA Tel: 202-459-9674, Fax: 202-459-9673, Email: nco@nitrd.gov, Website: https://www.nitrd.gov