Open Science Grid

Intro to the Virtual School Pilot, HTC, and OSG

Monday, Aug 2

Tim Cartwright, Lauren Michael

This work was supported by NSF grants MPS-1148698, OAC-1836650, and OAC-2030508

Welcome to the
OSG Virtual School Pilot 2020!

Why You Are Here

- You need large-scale, HTC-style computing – or you support researchers who do
- **Do not let computing block your research!**
  - Computing is cheap and plentiful
  - Push the limits of what you can do
  - If you run out of science to do, transcend the boundaries of your science
  - When computing becomes a barrier, push us to fix the problems
- Help & encourage others:
  In your lab, in your department, in your field, friends, etc.

Intro to HTC and OSG

Overview

- What is *high throughput computing (HTC)* ?
- What is the Open Science Grid (OSG)?
- How do you get the most out of the above?
  - School content organization

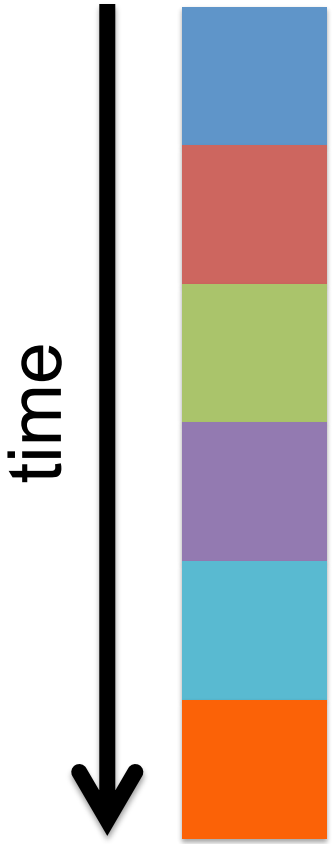
HTC: An Analogy

HTC: An Analogy

Serial Computing

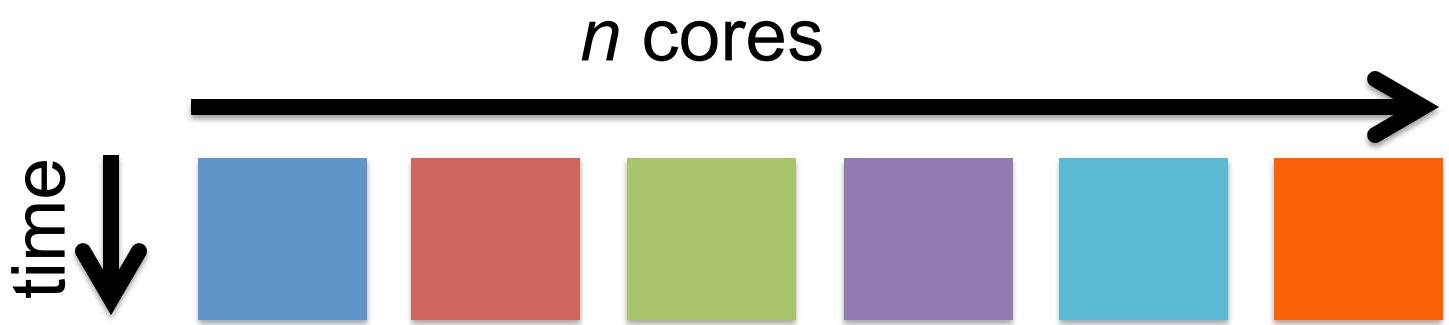
What many programs look like:

- *Serial execution*, running one task at a time
- Overall compute time grows significantly as individual tasks get more complicated (long) or if the number of tasks increases
- ***How can you speed things up?***



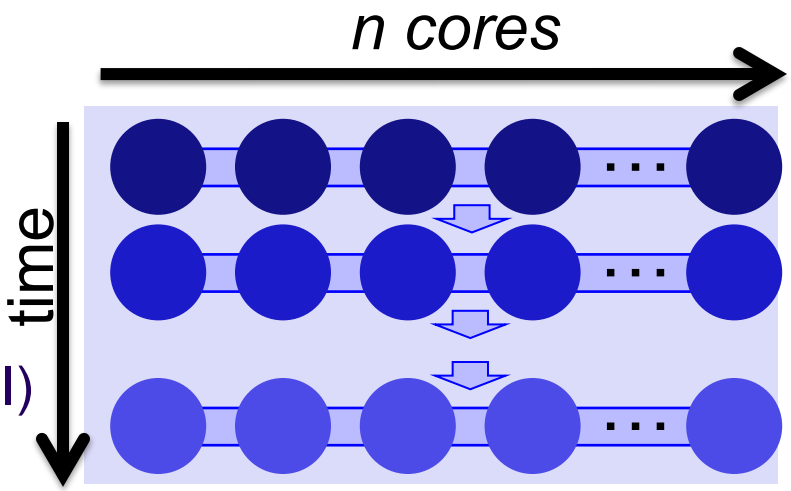
High Throughput Computing (HTC)

- Parallelize!
- Independent tasks run on different cores



High Performance Computing (HPC)

- Benefits greatly from:
  - CPU speed + homogeneity
  - shared filesystems
  - fast, expensive networking (e.g. Infiniband) and co-located servers
- Requires special programming (MP/MPI)
- Scheduling: **Must wait until all processors are available**, *at the same time* and *for the full duration*
- ***What happens if one core or server fails or runs slower than the others?***

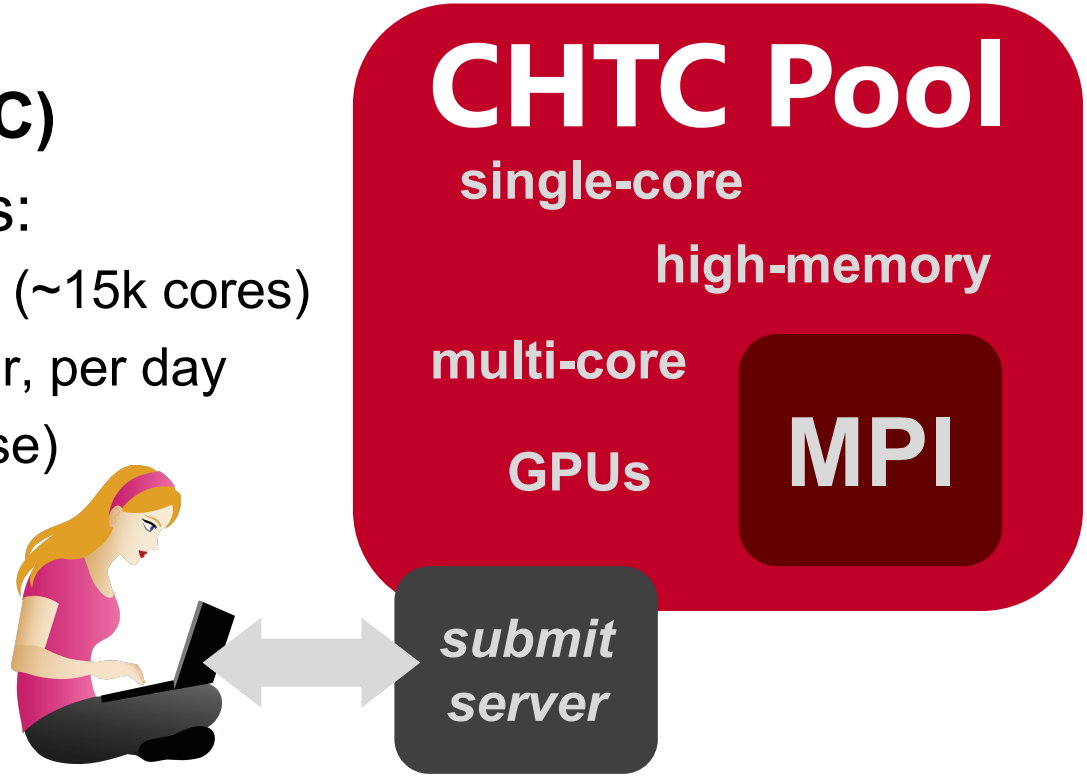


High Throughput Computing (HTC)

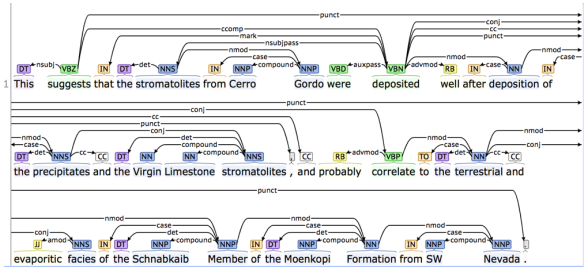
- Scheduling: only need **1 CPU core for each** (shorter wait)
- Easier recovery from failure
- No special programming required
- Number of concurrently running jobs is *more* important
- CPU speed and homogeneity are *less* important

Example Local Cluster

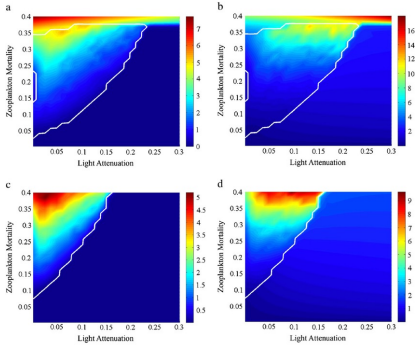
- UW-Madison's **Center for High Throughput Computing (CHTC)**
- Recent CPU hours:
  - ~120 million hrs/year (~15k cores)
  - Up to 15,000 per user, per day (~600 cores in use)



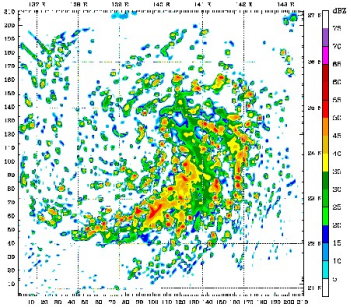
HTC Examples



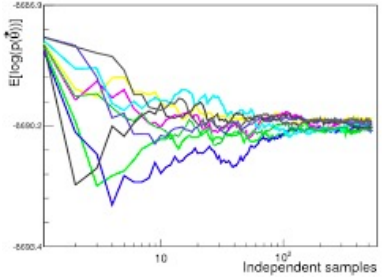
text analysis (most genomics …)



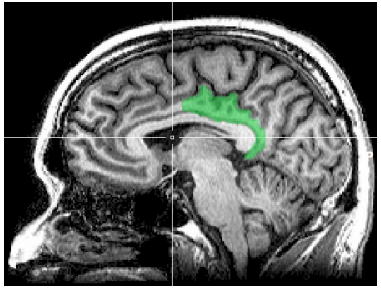
parameter sweeps



multi-start simulations



statistical model optimization (MCMC, numerical methods, etc.)

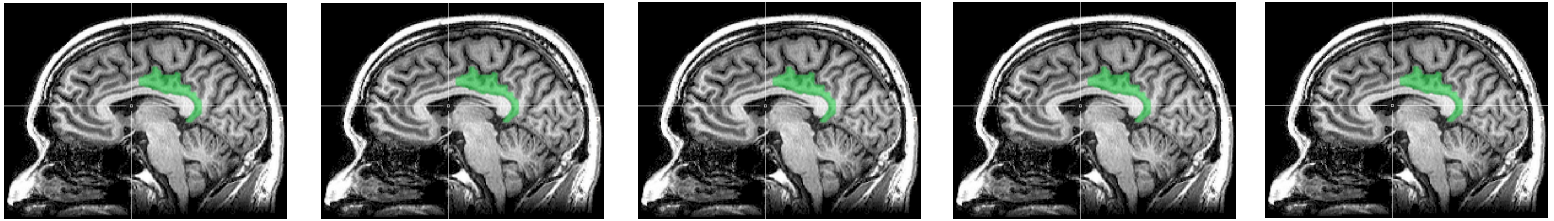


multi-image and multi-sample analysis

Signs of HTC-able work

- Any mention of **numerous** samples, images, models, parameters, etc.
- Nearly anything **written by the primary user** (e.g. c/fortran, Python, R)
  - Break out of loops!
  - Common internal parallelism could really be HTC (e.g. Matlab's 'parfor', 'distributed server', etc.)
- Some community softwares that use **multi-threading or multiprocessing** (e.g. OpenMP)
  - many are simply looping over data portions or independent tasks
  - HTC-able: break up input (or 'parameter' space), turn off multi-threading, combine results
- **Long-running** jobs (especially if non-MPI); see above explanations

Example Challenge



You need to process 72 brain images for each of 168 patients. **Each image takes ~1 hour of compute time.**

168 patients x 72 images = ~12000 tasks = ~12000 hrs

Conference is next week.

Distributed Computing

- Use many computers, each running one instance of our program
- Example:
  - **1 laptop (1 core) => 12,000 hrs = ~1.5 years**
  - 1 server (~40 cores) => 750 hrs = ~2 weeks
  - 1 MPI job (400 cores) => 30 hrs = ~1 days
  - **A whole cluster (10,000 cores) = ~1 hour**

What computing resources are available?

- A server?
- A local cluster?
  - Consider: Queue wait time? Can you program MP/MPI? Typical clusters tuned for HPC (large MPI) jobs may not be best for HTC workflows! Could you use even more than that?
- **OSG?**
- Other
  - EGI (European
  - Other national and regional grids
  - Commercial cloud systems (e.g. HTCondor on AWS)

OSG

- (d)**HTC for Open Science**
  - ~120 contributors
  - **Past year:**
    - >1.6 billion CPU hours
    - >200 petabytes transferred

- Can submit jobs locally, they backfill across the country
  - interrupted at any time (but not too frequent)
- https://www.opensciencegrid.org/

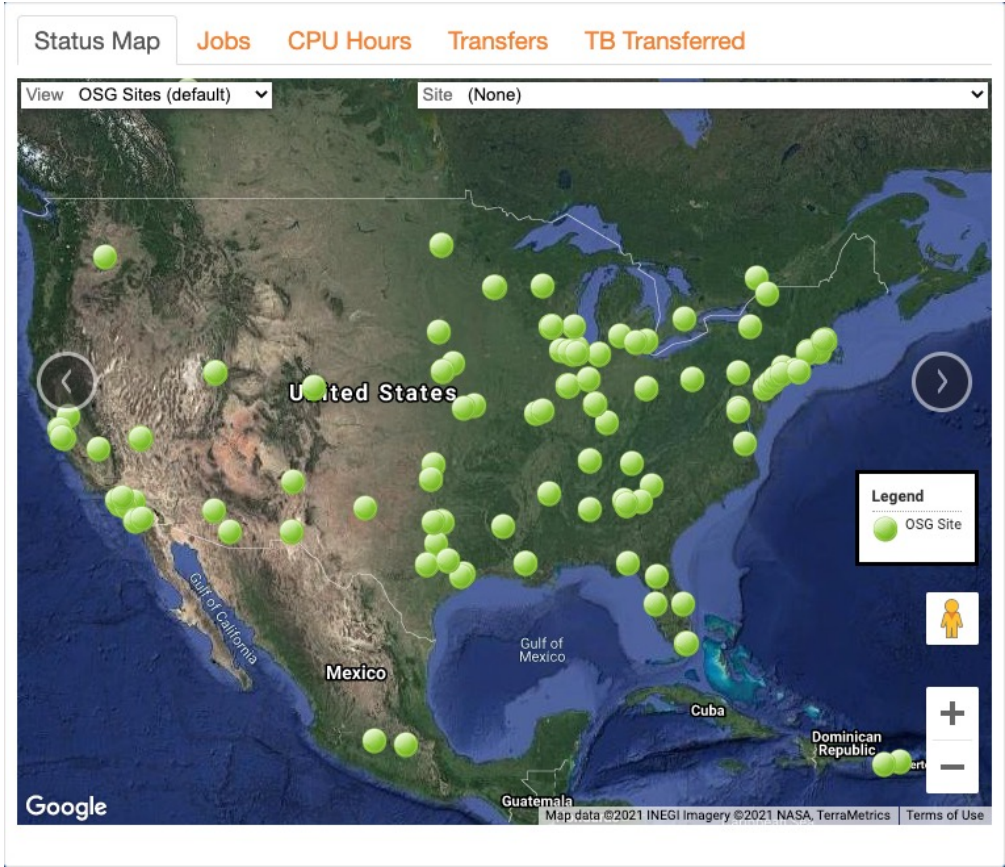
What is the OSG?

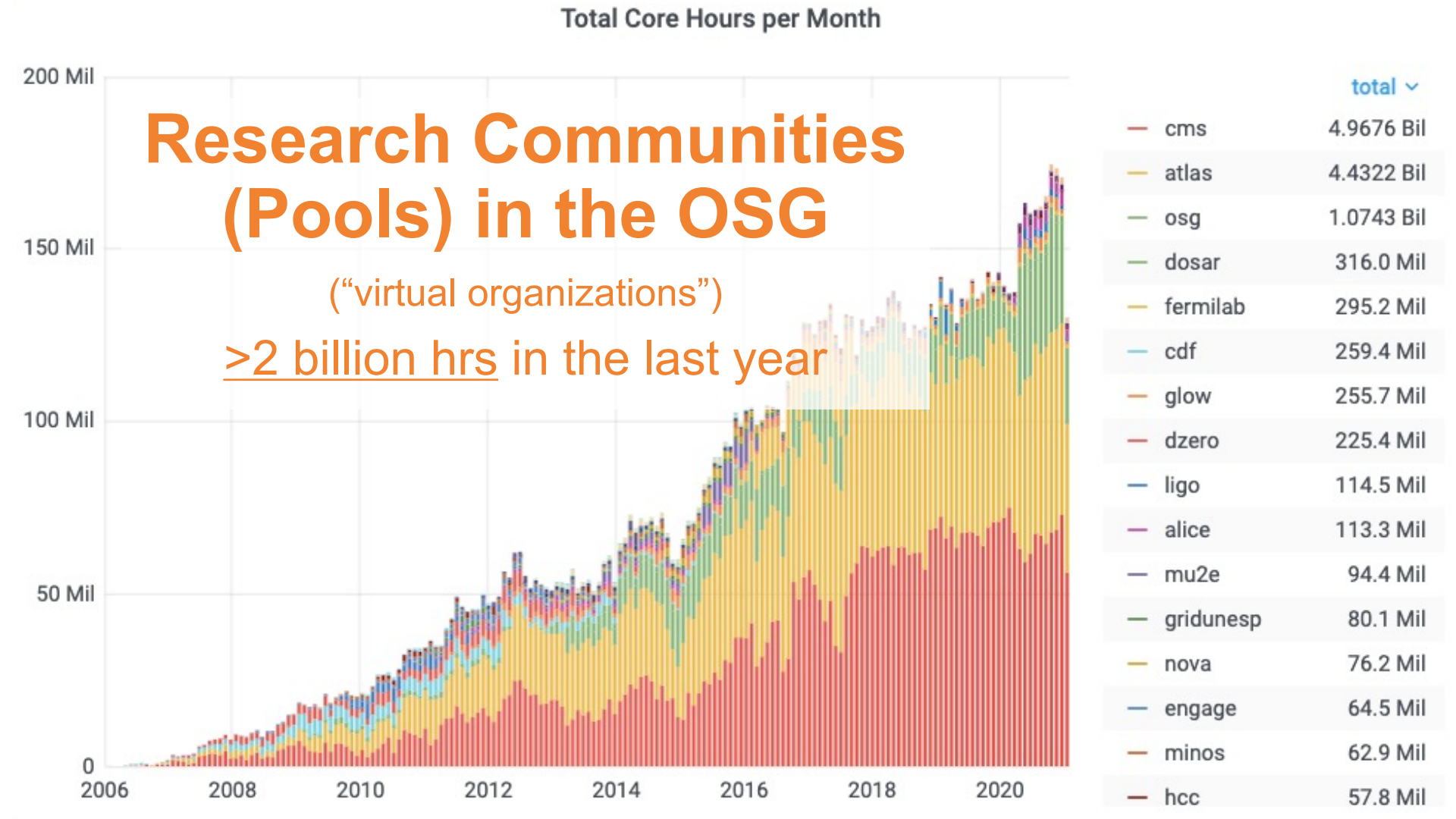
a consortium of researchers and institutions who *share* compute and data resources for ***distributed*** **high-throughput computing (dHTC)** in support of open science

Who Participates?

- Researchers
- Science Gateways
- Multi-Institution Collaborations
  - Atlas/CMS (Higg Boson), IceCube, South Pole Telescope, and others
- Academic Institutions and National Laboratories that support the above

Campuses are critical to OSG's ability to advance research.




Total Core Hours per Month
200 Mil
150 Mil
100 Mil
50 Mil
0
2006
2008
2010
2012
2014
2016
2018
2020
Research Communities (Pools) in the OSG
(“virtual organizations”)
>2 billion hrs in the last year
total
cms 4.9676 Bil
atlas 4.4322 Bil
osg 1.0743 Bil
dosar 316.0 Mil
fermilab 295.2 Mil
cdf 259.4 Mil
glow 255.7 Mil
dzero 225.4 Mil
ligo 114.5 Mil
alice 113.3 Mil
mu2e 94.4 Mil
gridunesp 80.1 Mil
nova 76.2 Mil
engage 64.5 Mil
minos 62.9 Mil
hcc 57.8 Mil


HOW IS CMS SEARCHING FOR
THE HIGGS BOSON?

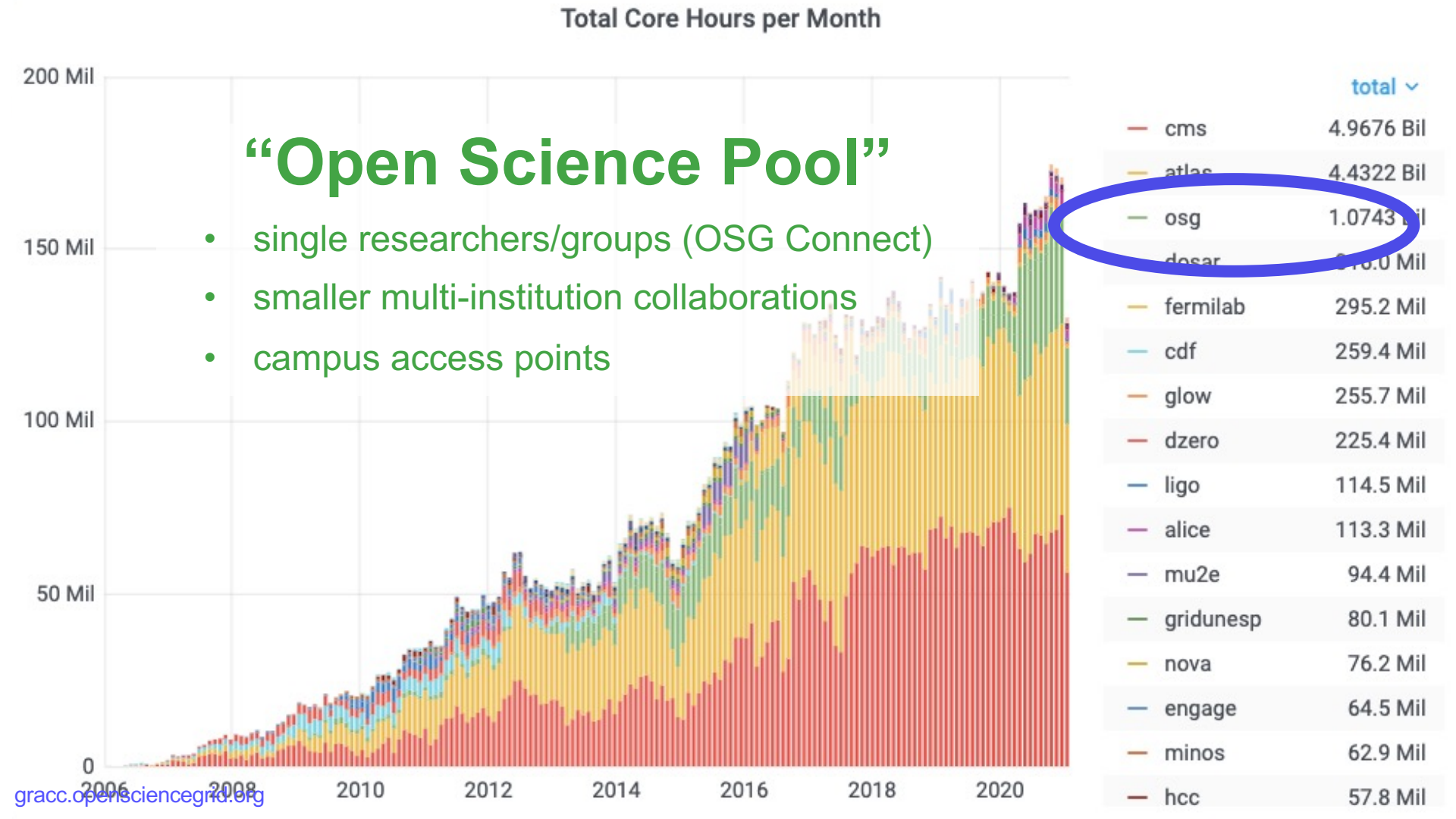
Previous

OSG Supports Multi-Messenger Astronomy.

OSG integrates global computing to support detection of colliding neutron stars by LIGO, VIRGO, and DECam.

Read more

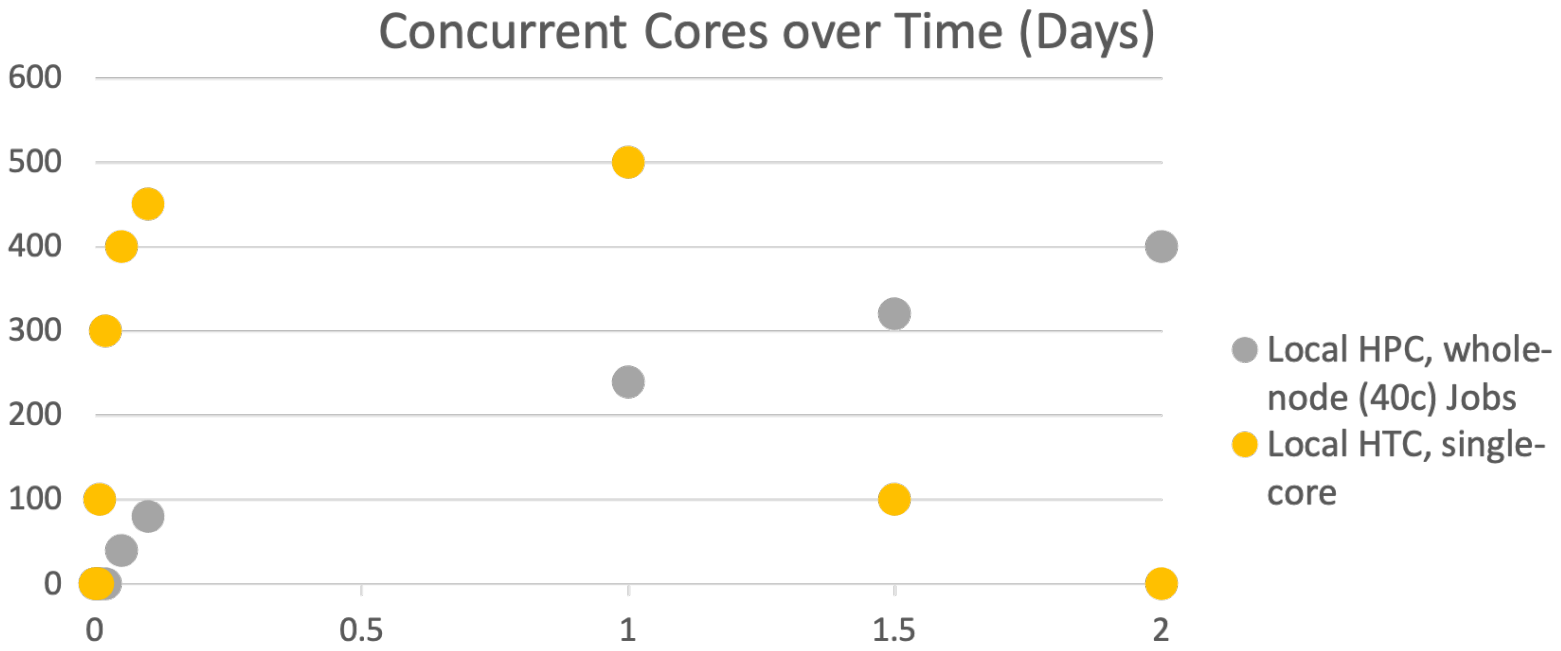
Next


Total Core Hours per Month
200 Mil
150 Mil
100 Mil
50 Mil
0
2006
2008
2010
2012
2014
2016
2018
2020
total
cms 4.9676 Bil
atlas 4.4322 Bil
osg 1.0743 Bil
dosar 316.0 Mil
fermilab 295.2 Mil
cdf 259.4 Mil
glow 255.7 Mil
dzero 225.4 Mil
ligo 114.5 Mil
alice 113.3 Mil
mu2e 94.4 Mil
gridunesp 80.1 Mil
nova 76.2 Mil
engage 64.5 Mil
minos 62.9 Mil
hcc 57.8 Mil
“Open Science Pool”
• single researchers/groups (OSG Connect)
• smaller multi-institution collaborations
• campus access points
gracc.opensciencegrid.org

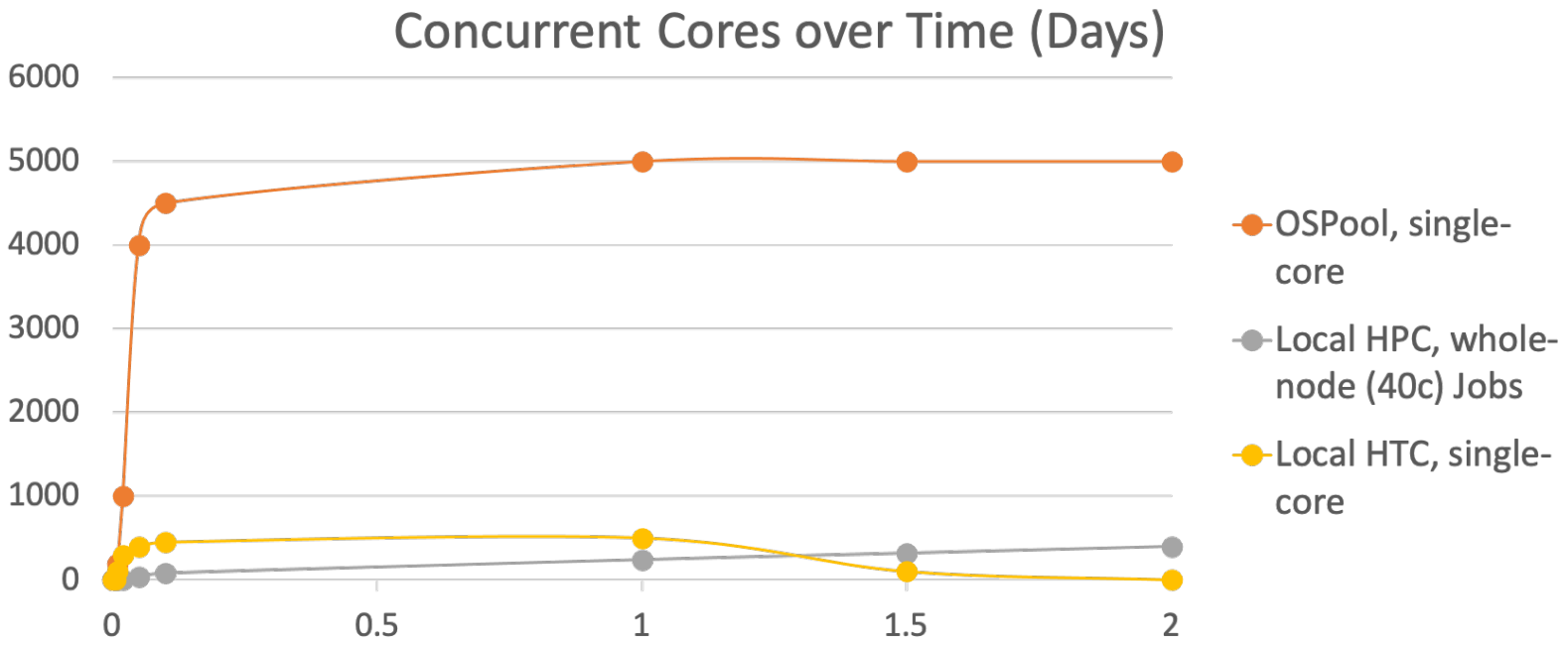
Can the OSPool Help?

Per-Job **Resources**	**Ideal Jobs!** (up to 10,000 cores, per user!)	**Still Very Advantageous!**	**Probably not…**
cores (GPUs)	**1** (1; non-specific)	**<8** (1; specific GPU type)	**>8 (or MPI)** (multiple)
Walltime *(per job)*	**<10 hrs*** *or checkpointable	**<20 hrs*** *or checkpointable	**>20 hrs**
RAM *(per job)*	**<few GB**	**<10 GB**	**>10 GB**
Input *(per job)*	**<500 MB**	**<10 GB**	**>10 GB**
Output *(per job)*	**<1 GB**	**<10 GB**	**>10 GB**
Software	*'portable' (pre-compiled binaries, transferable, containerizable, etc.)*	*most other than →→→*	*licensed software; non-Linux*

Hypothetical Throughput, 12k core hours



Hypothetical Throughput, ~~12k~~ core hours



OSG Virtual School Content

- **Lectures:** Tue-Fri, 10am CT & 2:30pm* CT
  - HTC via **HTCondor**
  - (d)HTC on the **OSG**
  - **Software** Portability for HTC
  - **Data** Portability for HTC
- **Bonus topics** (2nd Mon-Tue): optional
- **Showcase** (2nd Wed): science transformed by HTC
- **Lightning Talks+Close** (2nd Fri): chance to show work

**presented 'publicly', via registration; all others for selected participants*

For Researchers and Campuses

Proactive, personalized facilitation and support for:

- Individual researchers via **OSG Connect**
- Institutions and large collaborations
  - **Share local resources** via the OSG
  - Locally-supported **submit points**
    - data and identity federation
    - integration of cloud capacity
  - Local HTC Capacity
    - Learn from OSG's **Research Computing Facilitators**

- **Presentations/Training** in OSG compute execution, HTC Facilitation, and local HTC systems administration