Me, My Friends and Everyone Else – Campuses and OSG

David Swanson, Director
Holland Computing Center
University of Nebraska-Lincoln
Holland Computing Center

10 PB, 35,000 cores, 100 Gbps
1045 active users, 237 research groups
NSF and Campus Resources

• Like other institutions across the nation, Nebraska invests into research computing to remain competitive

• From a funding perspective, the proportion spent on R & D by the Federal government with respect to Academic Institutions has decreased over the last several years

• It is critical to integrate our resources to maximize our investment

• OSG integrates HCC with more than resources, it connects us to science we’d otherwise not contribute to
Trends in Funding Academic R&D in Science and Engineering

Figure 5-1

- Federal Government
- Academic Institutions

* Academic institutions’ funds exclude research funds spent from multipurpose accounts.


Science and Engineering Indicators 2016
OSG appeal to Campus Resource Providers

• Better ROI for equipment
• Interact with world experts
• Improve local services
• Challenge and engage staff
• Support Science!

• Hosted CE’s
• OSG Connect
• CC* expects it!

3/20/19
CPU Hours Total: by Department

- [UNL] Chemistry: 35,652,030.5
- [UNL] Mechanical and Materials...: 19,334,709.5
- [UNL] Arts and Sciences: 16,540,471.2
- [UNL] Physics and Astronomy: 15,443,883.7
- [UNL] Center for Biotechnology: 11,642,751.0
- [Grid] Grid: 4,378,146.4
- [UNL] Computer Science and Eng...: 3,655,239.3
- [UNL] Educational Psychology: 2,849,746.4
- [UNMC] Pharmaceutical Science: 2,651,922.2
- [UNO] Physics: 2,085,532.3
- [UNO] Economics: 1,408,676.6
- All 64 others: 8,008,967.2
Who ran at HCC last year

- OSG
- ICECUBE
- NOVA
- GLOW
- GM2

Total:
- OSG: 2.971 Mil
- ICECUBE: 724 K
- NOVA: 567 K
- GLOW: 561 K
- GM2: 393 K
- /microboone: 252 K
- mu2e: 195 K
- /icecube/LocalGroup=tcarver: 95.0 K
- dune: 75.1 K
- seaquest: 41.4 K
- sbnd: 28.4 K
- /icecube/LocalGroup=hoshina: 27.6 K
- icarus: 17.4 K
- sbgrid: 16.5 K
- /gluex/LocalGroup=enp: 14.53 K
- minos: 10.94 K
- gluex: 8.81 K
- des: 8.27 K
- lariat: 7.67 K

3/20/19

Me, My Friends and Everyone Else, AHM 2019
Where HCC ran last year

- BNL: 474 K
- SU-ITS-CE2: 376 K
- SU-ITS-CE3: 341 K
- GPGGrid: 238 K
- UConn-OSG: 213 K
- GLOW: 205 K
- NUMEP_CE: 93.0 K
- SPRACE: 65.3 K
- MWT2: 60.5 K
- AGLT2: 38.9 K
- CIT_CMS_T2: 29.4 K
- SU-OG-CE1: 24.6 K
- GridUNESP_CENTRAL: 7.56 K
- N/A: 5.63 K
- Tusker: 3.94 K
- SU-OG-CE: 1.479 K
- SWT2_CPB: 954
- T3_US_NERSC: 703
- Crane: 326
- USCMS-FNAL-WC1: 78
Spheres of resources

- Me = Local
  - Departmental or research group
- My Friends = Campus
  - Same organization
  - Colleagues
- Everyone Else = Grid
  - All affiliated resources
  - Can’t expect any favors
  - Largest scale
• Me, My Friends, Everyone Else
• We know how to make a single cluster and how to integrate with the OSG
• But how to enable “small” sharing?
All a resource provider wants is ...

- To know who is using resources
- To prioritize this usage (for my friends)
- To do this in a way that maintains my ownership (i.e. I want to implement a new policy on Friday afternoon unilaterally)
- To be able to combine resources in a way that reflects the social aspects of collaboration
- Science remains a team sport – but some teams are small or short-lived
Long tail of science?
In-state collaboration

• Dr. Tessa Durham Brooks
• Doane University
• Root image datasets
• Ran via HCC to OSG initially, now goes directly
• Doane accounts at HCC

• Collaborated on CC*
• Working toward Statewide consortium
  • OneOCI
Multi-State Collaboration

Statistics, Data and Computational Science

HCC resources

Ecological Genomics

Beocat cluster
Multi-State Collaboration

- Jennifer Clarke from Nebraska and Brad Olson from Kansas State are involved with genome to phenome research
- “Let’s collaborate and better predict crop yield!”
- EPSCoR Multi-jurisdiction award submitted
  - Critical mass across state lines
  - There’s cash, let’s be friends!
- Includes teams of students
- Resource providers from each Institution for computing, data
Multi-state Collaboration wish list

• Share data without moving it all (StashCache)
• Share processing – esp. elevated priority or reserved resources
• Familiar environment (OASIS)
• IdM handled conveniently
  • Not have to add loginID, passwd, MFA to each cluster
  • And then their students, postdocs, etc.
• Lightweight (i.e. not GUMS)
• FIONAs ftw!
Regional Organizations

- Great Plains Network
  - VO
  - FIONAs
- PRP
- RMACC
- Proximity can facilitate training, meetings
- Efficiency of scale
OSG developments (see other talks)

• Hosted CE / OSG-Connect
• StashCache – RO from others’ storage – OSG Data Federation
• InCommon / SciTokens
• OASIS (Modules)
• Singularity
• Prioritization within Global Pool
• GRACC
• Numerous Personnel Resources ramping up

• Do I just want to enable ... Flocking?
  • Run SLURM at both HCC, KSU
  • HCC VO could be used
Philosophical Interlude

• Technology should mirror sociology
  • Each collaborator brings some ideas and expertise to the collaboration
  • Natural if cyber-resources similarly shared
  • Currently share access to global pool of resources with fair share
  • Want that _and_ ability to aggregate and/or prioritize among a few institutions
  • Agrees with OSG principle of Local Autonomy and Control
Summary

• OSG serves both Big Science and Individual Investigators extremely well
• Small, ad-hoc groups and long tail users are a “My Friends” scenario
• OSG has intentionally made great strides in both support and user features for long tail science
• Resource providers would find it attractive to have more fine-grained control readily available