

# Grid Operations at the Holland Computing Center

Derek Weitzel & Brian Bockelman



# Introduction – HCC

- University of Nebraska  
Holland Computing Center – 2009
- 162 Research groups
  - Not all are active every day or able to use the grid, but grid provides benefit to a sizeable group.



# Introduction – HCC Clusters

- Firefly
  - 5500 cores
  - Infiniband interconnect
- Prairiefire
  - 500 Cores
  - Infiniband interconnect
- CMS T2 – Red
  - 2100 Cores



# Introduction – HCC Grid

- 8,100 Cores, all opportunistic for grid
- HCC's Use of the Open Science Grid
  - 13.1 Million hours in 2010
- HTC jobs are based on Condor, and exported to OSG using GlideinWMS



# HCC Engagement

- Having cores isn't enough – having users is more important!
- 2 Fulltime Application specialists + 1 graduate student (ME).
- Actively seek out on-campus researchers
- Goal: Enable Nebraska researchers to run computational science.
  - Only one aspect is enabling grid usage.



# HCC Engagement

- Engagement starts by categorizing users – local or grid?
  - MPI or very large data ( $O(10\text{GB})+$ ) to local machines.
  - Single core to Condor and eventually the OSG.
- OSG acts as another, sometimes large, cluster in our arsenal.



# HCC Engagement - Jobs

- GlideinWMS has been key
  - Allows for smooth transition from local resources to OSG.
  - Grid problems are handled by OSG factory support.
  - Simple debugging for us.



# HCC Engagement - Data

Data movement is kept simple, 2 options:

## 1. Condor file transfer

- If input or output is less than 10MB, transfer files with the job.

## 2. Large files (10MB – 1GB)

- Stage files to Nebraska T2. Each job will pull/push the files using SRM.

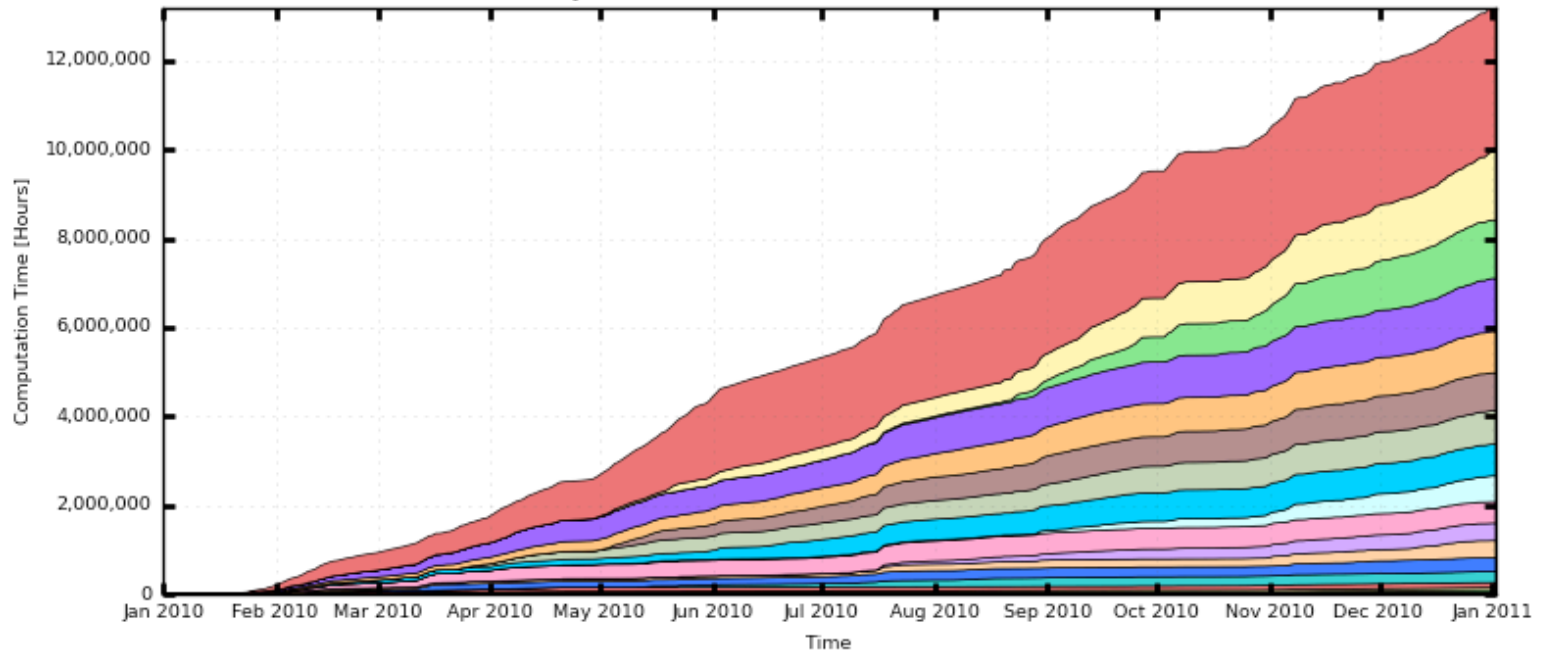
If neither work, the jobs probably should stay on campus.





# HCC Usage by Site

Cumulative Hours Spent on Jobs By Facility  
365 Days from Week 00 of 2010 to Week 00 of 2011



Firefly (3,235,650)	USCMS-FNAL-WC1 (1,537,372)	Clemson-Palmetto (1,312,756)	Nebraska (1,193,166)
prairiefire (932,404)	GridUNESP_CENTRAL (855,277)	UConn-OSG (752,406)	AGLT2 (711,738)
FNAL_FERMIGRID (583,710)	CIT_CMS_T2 (480,945)	UCSDT2 (390,851)	MIT_CMS (383,795)
GLOW (311,837)	Purdue-RCAC (246,352)	NYSGRID_CORNELL_NYS1 (111,597)	BNL-ATLAS (56,595)
Other (50,734)	RENCI-Engagement (19,050)	WQCG-Harvard-OSG (15,356)	SBBGrid-Harvard-East (6,198)

Total: 13,187,800 Hours, Average Rate: 0.42 Hours/s



# Building Relationships with Sites

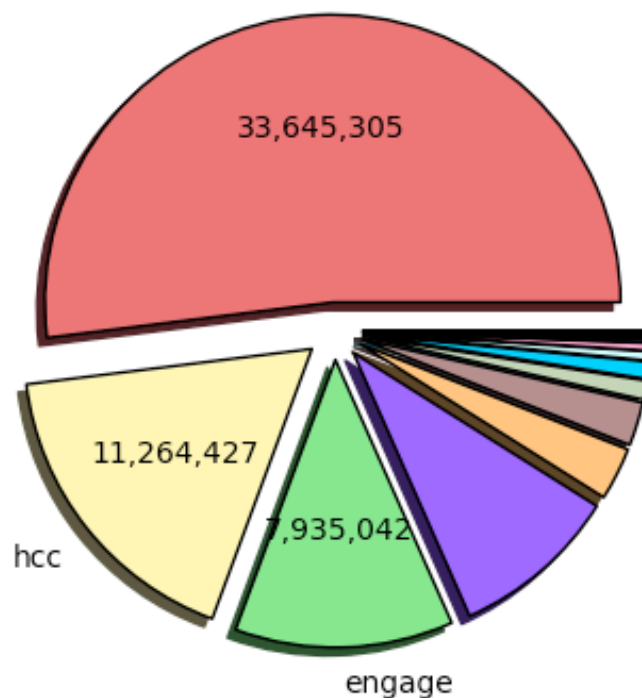
- Notice we run on many sites outside our own.
- Built partnerships with sites such as MWT2, NERSC, Oklahoma.
  - Goal was to enable opportunistic usage of the site.
  - We're the guinea pigs, but get cycles in return.



# Open Science Non-HEP Grid Usage

Wall Hours by non-HEP VOs (Sum: 64,741,636 Hours)

52 Weeks from Week 00 of 2010 to Week 00 of 2011



ligo (33,645,306)	hcc (11,264,427)	engage (7,935,043)	sbgrid (6,106,451)	gpn (1,923,371)
glow (1,580,933)	star (690,977)	osg (613,765)	c670 (325,633)	nysgrid (270,450)
fgstore (120,371)	nanohub (104,433)	auger (66,032)	loni_hep_09 (24,968)	Other (23,004)
ou2 (15,956)	compbiogrid (10,876)	pegrid (9,442)	along (6,740)	loni_binding04 (3,456)



# Take-Home Messages

- The grid is a great tool, but know when to not use it.
- Local Engagement.
- GlideinWMS.
- Build relationships with sites.

