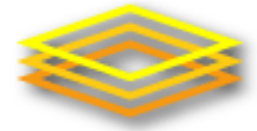




# Summary of OSG Activities by LIGO and LSC

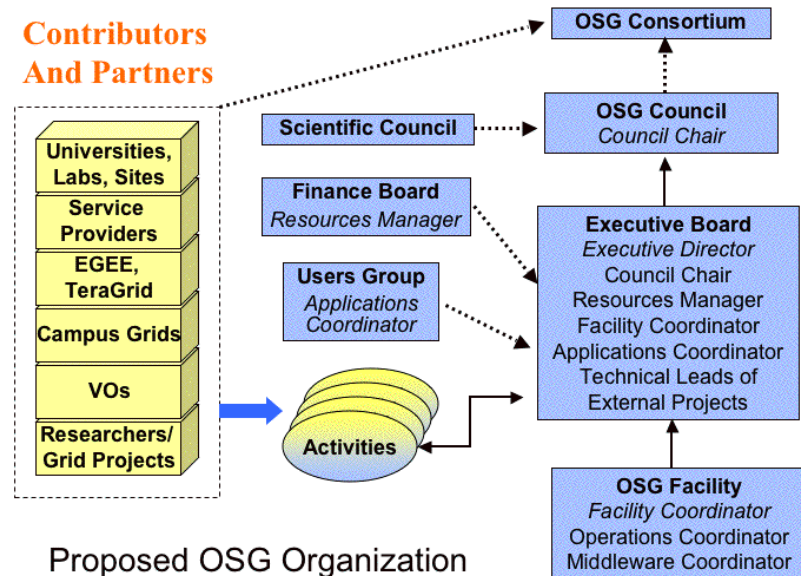
LIGO NSF Review  
November 9-11, 2005

*Kent Blackburn  
LIGO Laboratory  
California Institute of Technology*



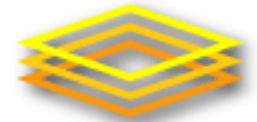
# OSG Organization

- Albert Lazzarini currently sits on Executive Board
- Kent Blackburn currently sits on OSG Council
- Kent Blackburn is LIGO liaison to OSG and participates in many OSG activity groups
- PSU, UWM, Caltech also providing effort to support OSG site administration and application development





# OSG Production Grid



- Evolved out of GRID3
- More than 50 sites on OSG Production Grid
  - *UWM & PSU from LIGO*
    - ~10% of these LDG compute cycles made available
- Over 15000 CPUs
- Used to support VO production runs for simulations and data analysis
- To date, only one release of software stack for use on production sites.

The screenshot shows the OSG Production Grid website in a Microsoft Internet Explorer browser window. The page features a map of the United States with various sites marked by colored dots. Below the map is a table listing site details.

Status	Site Name	Grid Version	Jobs ?	Disks ?	Service	Loc	Facility	CPUs ?
●	Purdue-Physics	osg 0.2.1	25,83	102,252	CS	IN	PURDUE	57
●	Purdue-ITaP	osg 0.2.1	21,122	25,146	CS	IN	PURDUE	1224
●	CIT_CMS_PG	osg 0.2.1	28,121	23,276	CS	CA	CALTECH	120
●	NERSC-PDSF	osg 0.2.1	22,012	10,1126	CS	CA	NERSC	812
●	GRASE-CCR-ACDC	osg 0.2.1	6,46	72,249	CS	NY	BUFFALO	68
●	FNAL_GPFARM	osg 0.2.1	20,134	11,191	CS	IL	FNAL	107
●	FNAL_FERMIGRID	osg 0.2.1	6,4	6,6	CS	IL	FNAL	4
●	UIOWA-OSG-PROD	osg 0.2.1	6,4	1,192	CS	IA	UIOWA	6
●	UCSandiegoOSG-Prod	osg 0.2.1	10,498	27,22	CS	CA	SDSC	396



# OSG Integration Testbed



- Roughly 30 sites currently in the OSG
  - Caltech from LSC
- Almost 5000 CPUs (some shared with production sites)
- Used to as test-bed to validate new OSG software stacks and develop new VO applications.

The screenshot shows the OSG Gridcat interface in a Microsoft Internet Explorer browser window. The address bar shows the URL: [http://osg-itb.ivdgl.org/gridcat/index.php?site\\_name=osgitb](http://osg-itb.ivdgl.org/gridcat/index.php?site_name=osgitb). The page features the Open Science Grid logo and a map of the United States with various sites marked. Below the map is a table of site details.

Status	Site Name	Grid Version	Jobs ?	Disks ?	Service	Loc	Facility	CPUs ?
●	UFlorida-IGT	osg 0.2.1	0.0	0.0	CS	FL	UFL	4
●	CIT_CMS_OSG	osg 0.1.6	0.0	0.0	CS	CA	CALTECH	24
●	FSU-OSGDEV	osg 0.3.0	0.0	0.0	CS	FL	FSU	4
●	CHANGE-site-name	osg unknown	0.0	0.0	CS	MI	UMICH	0
●	GRIDCAT_JM_TEST	osg unknown	0.0	0.0	CS	FL	UFL	102
●	CIT_CMS_OPT	osg 0.1.5-2	0.0	0.0	CS	CA	CALTECH	18
●	TACC-ITB	osg 0.1.6	0.0	0.0	CS	TX	UTEXAS	6
●	IU_umatlas	osg 0.1.6	0.0	0.0	CS	IN	INDIANA	1
●	BNL_OSG_Test1	osg 0.1.6	0.0	0.0	CS	NY	BNL	2696



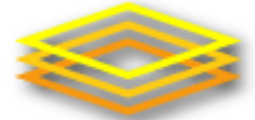
# LIGO's ITB Cluster

- Located on 6<sup>th</sup> floor of Millikan Library
- Eight dual CPU nodes
- 1.6 TB disk space
- Used to develop LIGO applications for the OSG Production Grid
- Used to test new releases of the OSG software stack (& propose future changes supporting LIGO's OSG application development)

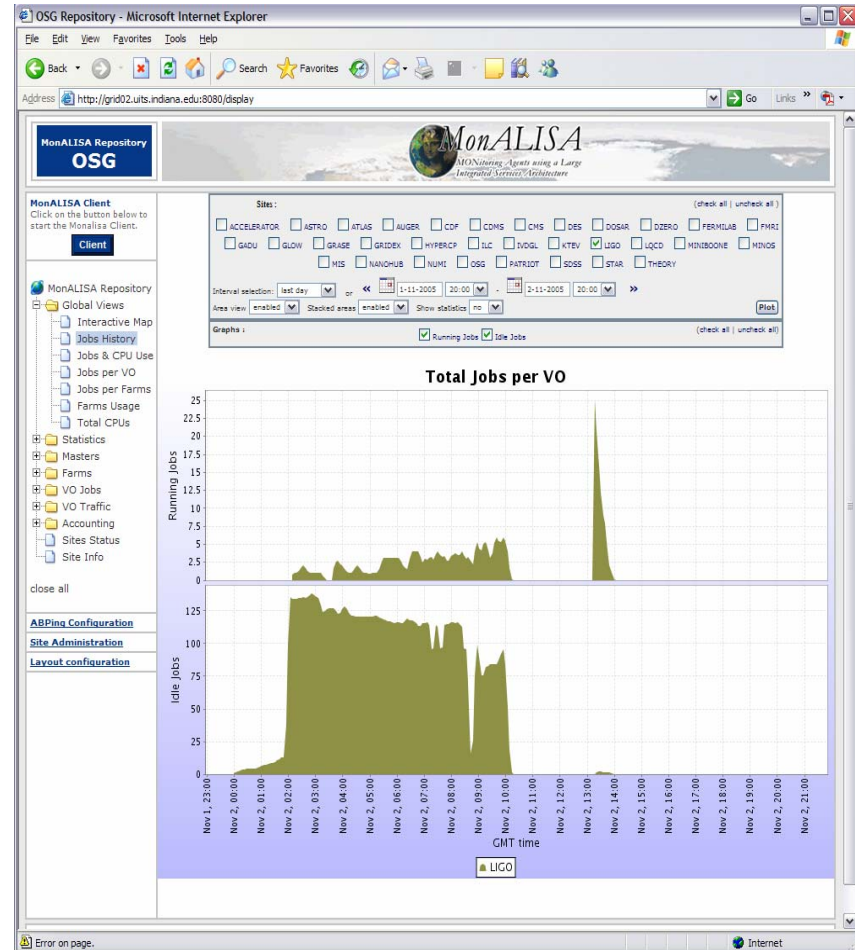




# LIGO Applications on OSG



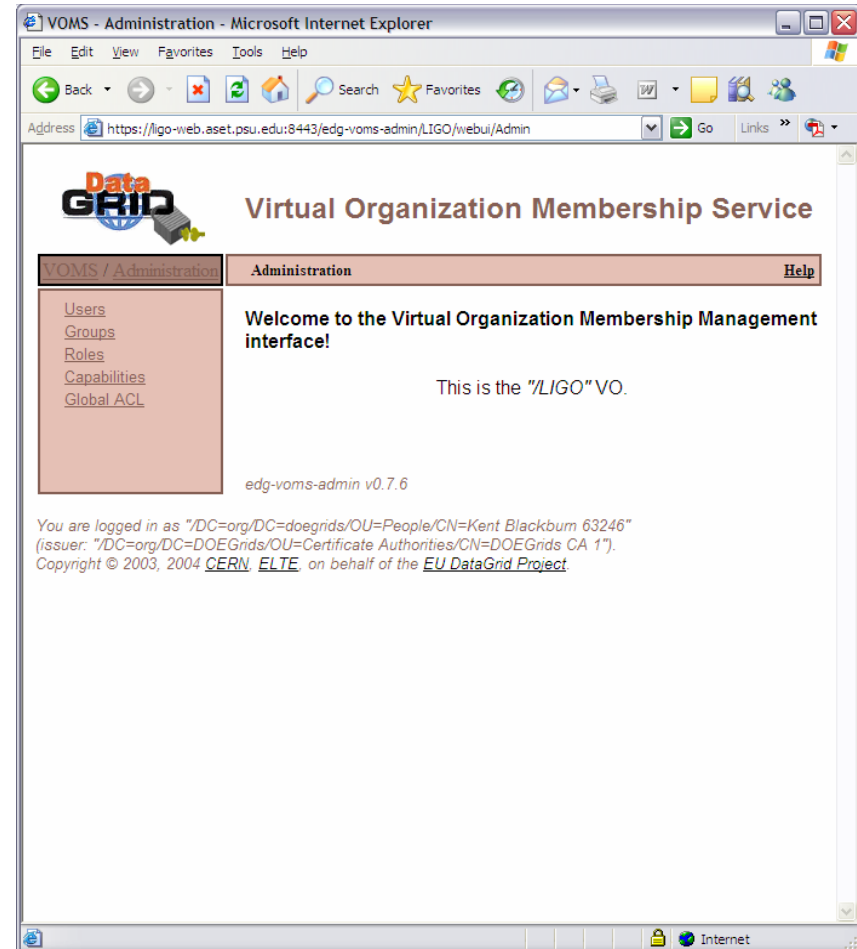
- Targeted binary inspiral searches
- Effort began in preparation for July OSG Consortium meeting in Milwaukee of this year
- October 1<sup>st</sup> succeeded in running binary inspiral application on LIGO's OSG integration testbed cluster using a test DAG of ~800 dag-nodes.
- Two weeks later, succeeded in running on LIGO's OSG production sites
- Have now run binary inspiral application at three CMS OSG production sites with mixed results:
  - 1.25 to 19 hours to transfer 38 GBs of LIGO data
  - 0.73 to 3.98 hours to compute
  - Requires customization of OSG software stack beyond current production "cache"
- Running jobs to fill queues for Super Computing 2005:
- Plan to run binary inspiral DAGs with 16,000 dag-nodes on OSG in near future!
  - Need to overcome data delivery and storage issues on non-LIGO sites.





# Virtual Organization

- PSU hosts LIGO's OSG VOMS server
  - *Until this July, LIGO was under iVDGL VO's VOMS*
- Approximately 100 LIGO certificates registered
- Two production sites:
  - *UW Milwaukee*
  - *PSU*
- One integration testbed site:
  - *Caltech*
  - *Some interest in having a second site*
- Currently only 3 non-LIGO OSG production sites recognize LIGO's VO for running on their sites
  - Pitfall is that new VOMS info being tied to OSG releases





# OSG Challenges

- Ease of porting LIGO data analysis applications from LDG to OSG: goal is to be transparent
- Efficiency of utilization of other OSG VO sites
  - *LDG benefits tremendously from co-locating data and computing*
  - *Additional computer cycles from OSG coupled to inefficiencies of other VOs at utilizing these cycles.*
- More frequent releases of software stack to better support VO and application needs
- Down the road issues:
  - *Funding*
  - *Reorganization*
  - *Storage Management*
  - *Interoperability with TeraGrid and EGEE*





# OSG Summary

- LIGO and the LSC have contributed significantly to the OSG process, providing personnel at all levels from oversight to administrators, developers and applications.
  - *LIGO and the LSC plan to continue to contribute*
  - *The LIGO Data Grid is critical to LIGO data analysis, migration and evolution of both the LDG and OSG will determine the degree of polymorphism*
- OSG has only released one production software stack
  - *LIGO's VOMS server didn't exist at the time of this release so most OSG sites are not LIGO aware due to software releases being so strongly coupled to the releases.*
  - *LIGO has worked with individual production sites to get the LIGO VO supported during the release interim so that LIGO applications can be developed and tested on the OSG*
- LIGO has now developed an application based on the binary inspiral search code for running on the OSG
  - *Requires patches beyond official OSG 0.2.1 software stack*
  - *Data transmission times significantly mask CPU benefits for LIGO application*