LIGO Scientific Collaboration Computing and Software

Manpower needs for operating and continuing to develop the production science infrastructure of the LSC

LIGO Scientific Collaboration Computing Committee
Meeting with Beverly Berger
16 August 2005

LIGO-G050365-01-Z

LIGO-G050668-00-Z

Required Activities

- 2 FTE: **End Users & Applications Support**
  - Operate a help desk to keep codes running and improve them
- 13.5 FTE: **Data Analysis Software Working Group (DASWG)**
  - Manage data analysis software development, releases, deployment
  - Responsible for DMT, GLUE, LAL, LALApps, MatApps, LDAS, LDR, Onasys, LIGOTools, and packaging maintenance of required software libraries
- 7.25 FTE: **LSC Data Grid (LDG) Client/Server**
  - R&D activities and movement to production (within DASWG)
  - Releases of LDG (within DASWG)
- 12 FTE: **Administrators/Data management/Hardware management**
  - Insure day to day operations of clusters and computers necessary for data analysis. Administer LDG, VOMS, and other services
  - Manage the LSC data archive

Computing Hardware

- LIGO Hanford/Livingston Observatories (per site)
  - » Cluster & Gateway machines
  - » Replication Services
  - » DMT machines
  - » Metadata Services
- LIGO Caltech
  - » Cluster & Gateway machines
  - » Replication Services and Data Archive
- LIGO MIT
  - » Cluster and Replication Services
- PSU
  - » Cluster & Gateway machines
  - » Replication Services
  - » VOMS (coming soon)
- UWM
  - » Cluster & Gateway machines
  - » Replication Services
  - » E@H

Personnel Requirements

	LIGO Data Grid Personnel Requirements -- Collaboration-wide													
								FTEs funded today				FTE deficit		
	Task	Admin	Help	Programming	Packaging	Architect	FTE total	By LIGO Lab Ops	By NSF GW program	By ITR 2000, 2001, 2003 programs	By university matching programs	FTE Deficit Today	FTE support expiring after ITRs	FTE deficit after ITRs
Applications	Help Desk		1	1			2					2		2
SW Deployment/R&D	System Architect					0.5	0.5		0.13		0.13	0.24		0.24
	System Documentation		0.5				0.5		0.25			0.25		0.25
	System Packaging				0.5		0.5		0.25			0.25		0.25
	GLUE/Onasys		0.25	0.5	0.25	0.25	1.25		0.5			0.75		0.75
	MatApps		0.2	0.5	0.2	0.1	1			0.25		0.75	0.25	1
	LDR		0.5	1.5		0.5	2.5			2.5			2.5	2.5
	LDAS/LIGOTools/Tcl Globus		0.5	1.25	2	0.75	4.5	3		1		0.5	1	1.5
	LAL/LALApps		0.2	0.6	0.2	0.25	1.25		0.2		0.2	0.85		0.85
	DMT		0.25	0.5	0.5	0.25	1.5	0.5		2		-1	2	1
LIGO Data Grid Clients/ Server/ R&D	System Architect					0.5	0.5					0.5		0.5
	Packaging/Help	0.5	0.5	0.5	0.5	0.25	2.25	0.25		1.5		0.5	1.5	2
	Middleware R&D	0.25	0.25	1.75		0.25	2.5			1		1.5	1	2.5
	Grid Services & integration	0.65	0.45	0.45	0.45		2			0.75		1.25	0.75	2
Sysadmin/ Conf. Mgmnt/ Cluster Mgmnt/ Archive Mgmnt	UWM	1.5	0.5			0.5	2.5			1	0.5	1	1	2
	PSU	1.5	0.5				2			2			2	2
	LLO	1	0.5				1.5	1.5						
	LHO	1	0.5				1.5	1.5						
	MIT	1					1	0.5				0.5		0.5
	CIT	2	1			0.5	3.5	3.5						
	Totals	9.4	7.6	8.55	4.6	4.6	34.75	10.75	1.33	12	0.83	9.84	12	21.84

The LIGO Data Grid

- NSF's support of the LSC through the ITR awards has enabled a production grid for gravitational wave data analysis
- The tiered grid operates, but needs continued effort for integrating grid tools into its infrastructure
  - » We do not efficiently use all our sites in a transparently seamless manner; still scientist - labor intensive.
  - » Need to continue effective collaboration with the Condor, Globus and LHC communities
  - » Need to continue to integrate LSC data analysis activities that are not yet capable of running within the OSG environment
- Data Grids' long-term support needs to be assured for the duration of LIGO I data analysis and beyond (Adv. LIGO)
  - » Present support of 12 FTEs based on one-time grid R&D awards is not a tenable solution because it cannot support the kind of technical people who are needed to run the complex
  - » In many cases $/FTE scaled to postdoctoral, GRA salaries. This is **NOT** the type of individual NSF should encourage to maintain IT infrastructure ….
  - » As a result faculty, senior scientific staff are acting as systems administrators, database managers, etc.
    - – **NOT** the best for LIGO science -- this is a brain drain on the collaboration
- The nascent OSG is ***not likely*** to fill the needs of the LSC
  - » Appropriate focus of OSG is on continued development of general infrastructure needed to enable the grid to grow robustly
    - – LIGO will participate to drive the technology along with the C/S groups -- but this cannot provide the operational support for the LSC virtual organization (VO).
  - » Support for use of the grid outside the scope of the OSG plan
    - – OSG model assumes each virtual organization (VO) brings to the table its own resources to enable it to join, utilize the grid

What does the LSC need?

- The LSC Datagrid (in the US) consists of 6 sites
  - » 4 are operated and managed by the LIGO Laboratory under its operations budget
    - – New-scope grid-based work provided by NSF gird projects.
  - » 2 are operated and managed by individual LSC institutions as a service to the collaboration under mainly ITR $ (to date) and use postdocs/faculty/students to perform needed development and maintenance work
- The LSC body of software has grown to a point where dedicated developers and maintenance is needed to support the libraries
  - » Outside of LIGO Laboratory this work is being borne by faculty, postdocs, students

AT THE PRESENT TIME, 12 FTEs WHO: (i) DEVELOP AND MAINTAIN PRODUCTION SOFTWARE PACKAGES, and (ii) MANAGE AND OPERATE THE PRODUCTION FACILTIIES AT UWM, PSU, MIT AND CIT ARE SUPPORTED BY NSF ITR FUNDS

Options

- OSG
  - » May be able to obtain support ~ 2 - 3 FTEs to work on OSG core software infrastructure
  - » This is critical work, but IT IS NOT specific to LSC Data Grid science analysis operations
  - » Leaves an ~10 FTE shortfall TODAY
  - » Does not address the chronic long-term shortfall of FTEs
- OCI
  - » The recent restructuring of CISE to address NSF's cyberinfrastructure needs includes in its scope/mission the support of professionals enabling scientific operations:

    "*OCI supports the preparation and training of current and future generations of researchers and educators to use cyberinfrastructure to further their research and education goals, while also supporting the scientific and engineering professionals who create and maintain these IT-based resources and systems and who provide essential customer services to the national science and engineering user community. Deborah Crawford will serve as the acting director of the OCI." (ref:* http://www.nsf.gov/od/oci/about.jsp)
  - » How can NSF's LIGO science program needs be matched to OCI's mission?