

SuperComputing 2003 grid-distributed wide-area CW search

Bruce Allen, Marialessandra Papa, Scott Koranda,
and many CS Grid people...

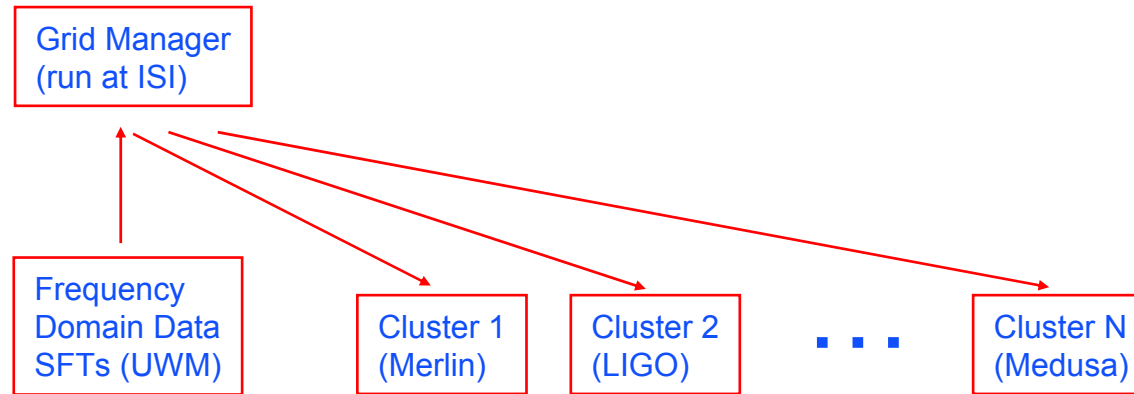
Why grid-enable Pulsar search?

- The existing validated algorithms are single-pass coherent methods, not hierarchical.
- To search over a broad frequency band, and even a limited area of the sky, requires enormous compute resources.
- Within the LSC, we currently have ~1500 CPUs that could be used – but of course they have other duties as well.
- A grid-enabled search code would allow us to leverage existing grid platforms and testbeds.

Why SC 2003?

- The GriPhyN and iVDGL Collaborations are LIGO's pathway into the grid computing community.
- The computer scientists in these collaborations work hard to demonstrate their work in the annual SuperComputing N meetings (SC2003 is next week in Phoenix, AZ).
- Past demos for SC2001 and SC2002 have not led directly to scientific results/papers. We'd like SC2003 to change this.

How does it work?

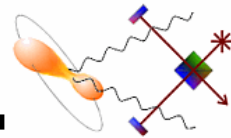


The grid manager has a list of search parameters (f , f_{dot} , sky positions). It sends the relevant part of the frequency-domain data to different clusters, along with a stand-alone executable, and runs these on the cluster. The executables that run on the clusters return lists of (clusters of) parameter sets and the corresponding values of the F statistic. They also return some simple statistics on the F-statistic distribution that enable one to quickly determine if the data has noise which is problematic.

The manager is “intelligent” and knows what data exists already at what clusters.

What search are we doing?

- Tiny area around Galactic center (approx e^{-4} radians)
- Wide band (150-350 Hz)
- Will search about 4×10^{10} points in parameter space
- Long observation time (all of S2) so high resolution in f , \dot{f} , and sky position
- Code should return about 10^6 parameter space points for follow-up studies
- Afterwards, we might search other interesting areas:
 - » galactic core
 - » first spiral arm
 - » Gould Belt (see Greg Mendell's talk). Claims exist that @100-200Hz there might be detectable S2 sources. Gould belt is 50-400 pc away.
 - » Unidentified X-ray sources
 - » SN remnants
 - » globular clusters



Current Compute Pool

- Currently about 2000 CPUs available in the pool
- Will run 800,004 jobs (1-4 hours each)
- Blue: LSC resources, Purple: non-LSC

Site	CPU count	login node	GridFTP server	batch jobmanager	RLS server	global scratch	GLOBUS_LOCATION	condor	private subnet
AEI	359	morgane.aei.mpg.de	morgane.aei.mpg.de	morgane.aei.mpg.de/jobmanager-condor	not yet	/home/NOBACKUP	/opt/ldg/globus	/opt/ldg/condor	yes
Birmingham	200	tsunami.sr.bham.ac.uk	tsunami.sr.bham.ac.uk	tsunami.sr.bham.ac.uk/jobmanager-condor	none	/raid/1/<login>	/opt/ldg/globus	/usr/local/condor	yes
Cardiff	160	mini.astro.cf.ac.uk	mini.astro.cf.ac.uk	mini.astro.cf.ac.uk/jobmanager-condor	not yet	none	/opt/ldg/globus	/opt/ldg/condor	yes
CIT	?	ldas-grid.ligo.caltech.edu	ldas-grid.ligo.caltech.edu	ldas-grid.ligo.caltech.edu/jobmanager-condor	not yet	?	/opt/ldg/globus	/opt/ldg/condor	yes
ISI	35	birdie.isi.edu	smarty.isi.edu sukhna.isi.edu sultan.isi.edu bindas.isi.edu pisa.isi.edu	columbus.isi.edu/jobmanager-condor birdie.isi.edu/jobmanager-condor skywalker.isi.edu/jobmanager-condor	rls://smarty.isi.edu rls://sukhna.isi.edu rls://skywalker.isi.edu	/nfs/cgt-scratch/griphyn	/nfs/v6/globus/GT2/linux/STABLE	<hostname>/columbus/condor	no
UTB	51	lobizon.utb.edu	lobizon.utb.edu	lobizon.utb.edu/jobmanager-condor	not yet	/home/<login>	/opt/ldg/globus	/opt/ldg/condor	yes
UWM	296	hydra.phys.uwm.edu	hydra.phys.uwm.edu	hydra.phys.uwm.edu/jobmanager-condor	rls://hydra.phys.uwm.edu:39281	/scratch	/opt/ldg/globus	/opt/ldg/condor	yes
AGT-Wisc	?	agt-login.cs.wisc.edu	agt-login.cs.wisc.edu	agt-login.cs.wisc.edu/jobmanager-pbs	none	/home/<login>	/vdt/globus	no	?
Grid2003	~1200	details	details	details	details	details	details	details	details
Madison Condor Pool	800	none	beak.cs.wisc.edu	beak.cs.wisc.edu/jobmanager-condor	none	/shared/scratch/griphyn-ligo /shared/scratch2/griphyn-ligo/	/scratch/vdt	/unsup/condor	yes/no



Ganglia Cluster Toolkit: LSC Data Grid Report - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address: <http://mothra.phys.uwm.edu/LSCDataGrid/?r=day&s=descending&c=>

LSC Data Grid (4 sources) [tree view](#)

CPU's Total: **553**

Hosts
Up: **279**
Down: **1**

LSC Data Grid CPU last day

LSC Data Grid Load last day

LSC Data Grid Memory last day

Avg Load (15, 5, 1m): 97%, 99%, 99% Localtime: 2003-11-13 11:04

AEI-LSC Merlin Grid [tree view](#)

CPU's Total: **365**

Hosts
Up: **183**
Down: **0**

Merlin at AEI Grid CPU last day

Merlin at AEI Grid Load last day

Merlin at AEI Grid Memory last day

Avg Load (15, 5, 1m): 128%, 130%, 130%

UWM-LSC Medusa [physical view](#)

CPU's Total: **302**

Hosts
Up: **296**
Down: **2**

medusa CPU last day

medusa Load last day

medusa Memory last day

Avg Load (15, 5, 1m): 44%, 56%, 66% Localtime: 2003-11-13 10:48

Birmingham Tsunami Grid [tree view](#)

CPU's Total: **154**

Hosts
Up: **76**
Down: **0**

Tsunami Grid CPU last day

Tsunami Grid Load last day

Tsunami Grid Memory last day

Avg Load (15, 5, 1m): 44%, 44%, 45% Localtime: 2003-11-13 11:04

CGT at ISI Grid [tree view](#)

CPU's Total: **34**

Hosts
Up: **20**
Down: **1**

CGT Grid CPU last day

CGT Grid Load last day

CGT Grid Memory last day

Internet