E7 Lessons Learned

Albert Lazzarini
LIGO Laboratory

LIGO-G020013-03-E

Lessons Learned

- LDAS performed very well in all aspects
  - » There is still a lot to be done to get ready for S1 …
- System is robust and *invisible*
  - » Most of the LDAS analyses were being run, managed by remote users on client PCs throughout the country
  - » Kent was very ill throughout most of the run;
    - – Stuart, Peter and others were able to provide the needed support
    - – The programming team at Caltech provided critical support also
- Experience from 4 days of shifts at LLO
  - » Need Control Room visibility, monitoring tools for searches running in LDAS
    - – This is an LSC issue
  - » 3D visualization tools could help the operations and science teams
    - – f-t displays updated in realtime
    - – Dynamic f_1 - f_2 displays
  - » Many of the science shift tasks could be automated, reducing fatigue during late hours
    - – lots of repetitive tasks

E7 Lessons Learned

LDAS Software

Kent Blackburn

LIGO Laboratory

LIGO-G020013-03-E

LDAS Job Summary

	Hanford LDAS	Livingston LDAS	MIT LDAS	CIT-TEST LDAS	TOTAL
Total Jobs	63600	48775	280	915	113570
Database Rows	4188188	2789132	1062	2096	6980478

- LDAS for full E7 Run: Dec. 28th, 2001 - Jan. 14th, 2002
  - » Approximately one job every 10 seconds (averaged).
  - » Approximately five rows every second (averaged).
- Greater than 90% of jobs completed successfully
  - » LHO roughly 92%; LLO roughly 95%; Not checked elsewhere.
- Pre-Release testing revealed 0.3% failure rate!
  - » Pre-release dominated by dataConditionAPI thread problems.
  - » Fraction due to mpiAPI/wrapperAPI communications issues.

LDAS E7 Failure Modes

- Thread safety in dataConditionAPI cause multiple jobs to be lost at once.
  - Caused roughly half of all failures directly due to LDAS bugs.
- Number of files per directory used to store data products by the managerAPI filled up twice near end of run - *deleting pre-E7 files fixed each time*.
- Locked segments *wrapping* around from N^th^ frame directory to first frame storage directory not found by frameAPI - *this only occurred once*.
- Twice users killing jobs involving the metaDataAPI caused all jobs using database to fail - *restarting metaDataAPI fixed*.
- Known communication issues between mpiAPI and wrapperAPI caused node table to confuse available node list - *restarting mpiAPI fixed*.
- Use of stderr and stdout in LAL/LALwrapperAPI cause managerAPI to become unresponsive to requests - *new LAL code submitted to fix*.
- Bugs in LAL/LALwrapper cause jobs to fail - *subsequent LAL releases fixed*.
- *Pilot errors* in scripts caused job failures - *external user scripts corrected*.

LDAS E7 Failure Modes

- Time server drop out - *caused frames to not be available for requested time intervals*
- The dataConditionAPI failed to collect data from both the frameAPI and metaDataAPI in a timely manor to continue processing -- *This was only an issue for periodic search, which did not run on-line*

LDAS
Cost of getting ready for E7

- Dedicated ~ 6 weeks of schedule to getting ready
  - Beta release functionality had to be deferred
- More than 2 weeks of testing required to ensure high probability of robustness, success
- LSC LAL shared object C code has evolving in parallel
  - LDAS had to rely on *dated* LAL/LALwrapper code for integrated tests
- Significant increase time spent in meetings with LSC in order to integrate codes, planning for E7, etc.

Open Issues Raised Post-Run

- LDAS users from the LSC would like top level visibility into all search codes running on LDAS systems in the form of a GUI
  - Agreed in LSUG meeting that this is not be an LDAS function, but should reside in super-job control GUI/proxy outside of LDAS and visible in control rooms. Note: This type of interface is planned for GriPhyN/LDAS.
- Creators of driver scripts used to control job submittal would like more standardized error reporting
  - Functionality did in fact exist in LDAS, but only just appearing in the E7 release and was not widely known. Additionally, most driver scripts used a job-control library which didn't expose this information - plans are to extend the job-control library.
- LAL/LDAS Software Users Group not comfortable with responsibilities of allocation and scheduling resources
  - Requests made that new LSC Computing Committee manage these.

LDAS

High priority items needed to get back on track with LDAS (slipped)schedule.

- Rework configuration & build rules to support migration towards beta & final releases of LDAS.
- Create new diskCacheAPI; pull out this functionality from frameAPI.
- Improve reliability of dataConditionAPI (thread issues).
- Add common resampling library to LDAS for use in both the frameAPI & dataConditionAPI
- Extend system monitoring to track API shutdowns & restart, core files &debugging, job & database statistics and user account management.
- Add interpolation, Kalman filters, regression and replace intermediate() function.
- Reduce memory usage in dataConditionAPI by average of ~ 5x

LDAS

High priority items needed to get back on track with LDAS (slipped)schedule.

- Move LDAS API processes onto new dataserver
- Improve documentation, interfaces, and table designs per LSC recommendation
- Implement new TCL channel management interface to better control datasockets.
- Add new detector geometry metadata to LDAS pipeline to better support use of ALLEGRO bar data in stochastic search code
- Add job load monitoring commands to support GryPhyN integration
  - » Continue providing LSC and GriPhyN support on day-to-day bases.

E7 Lessons Learned

LDAS Databases

Peter Shawhan
LIGO Laboratory

LIGO-G020013-03-E

Database Insertion Statistics During the E7 Run

		LHO	LLO
Segments:	IFOLocked	17919	5899
GDS triggers:	BitTest	34640	17761
	ChannelReadOutError	26	–
	eqMon	28	–
	glitchMon	1790683	1056375
	Glitch	271430	201113
	Lock transition	140468	11328
	MC_F violin mode	11016	7156
	Rho2 [from CorrMon]	511	195
	TFCLUSTERS	290295	68551
	TimeSliceError	1755	23762
	TID	1663	–
LDAS inspiral:	template	428970	176655
	FCT	2970	24295
LDAS burst:	power	1082676	411127
	slope	17561	58044
	TFCLUSTERS	1700621	2519617

Over 10 million entries added!

Database Insertion Issues

- DMT segments and triggers
  - » Insertions handled by “SeqInsert” process to preserve sequence
  - » Occasional LDAS job failures caused SeqInsert to go into “error state”
    - – Required human intervention
    - – Delayed LDAS search codes which were checking for segments
  - » Midway through run, SeqInsert was modified to be “fault-tolerant”
- Events from LDAS search jobs
  - » No technical problems
  - » A few “learning curve” issues on the part of search code authors
- Overall insertion rate was easily handled by LDAS

Database Retrieval Issues

- Database contents can be retrieved
  - » Database queries can be submitted using guild or getMeta
  - » LIGO_LW files of database records can be converted to ASCII, read into Matlab, or parsed by a C program
- … but there are limitations
  - » LDAS limits query output to 10000 rows
    - – Action item: develop code to fetch more rows, concatenate files
  - » Have to direct query to the appropriate database
  - » Some data will need to be copied to multiple databases
    - – Action item: develop software tools to facilitate copying
  - » Sophisticated “event analysis tool” still under development
- The real challenge: how to make sense of all the information

E7 Lessons Learned

LDAS Hardware

Stuart Anderson
LIGO Laboratory

LIGO-G020013-03-E

LDAS Hardware + Data Archive

- Interface between LDAS and CDS
  - » Disk failure worked
  - » Shared QFS filesystem worked, but...
  - » Still improving coordination
    - – a) Control room monitoring of disk cache was not established until data was lost
    - – b) Differing versions of frame data are expensive to manage downstream.

LDAS Hardware + Data Archive

- Data archive
  - » Good choice of tape drive at observatories but poor robotics
  - » Good choice of robotics and tape drive for central archive
  - » Growing dissatisfaction with HPSS
    - – Need to accelerate evaluation of SAM-QFS alternative (SUN product)
- Software management
  - » Multiple test/development LDAS systems are essential
  - » Software mirroring tools are mature and useful
  - » Realtime development of software release policy is painful.
    - – Still need to differentiate between core LDAS/LAL software versus individual search codes

LDAS Hardware + Data Archive

- Data management
  - » Reduced data sets are needed now
  - » Replication or distributed access to metadata is needed
  - » *All the data cannot be available everywhere all the time*
  - » HPSS archive (pre-E1 through E7):
    - – 30 TB
    - – 300,000 files
    - – 10% of a 1 year 7x24 science run
      - one more order of magnitude to go

E7 Lessons Learned

DMT

John Zweizig
LIGO Laboratory

LIGO-G020013-03-E

DMT Use During E7

- Monitors used in c.r. to track IFO/PE status
  - » IFO performance (lock statistics, line tracking, servo stability checks)
  - » Environmental noise
- Many (>4M) triggers generated
  - » Transient searches (glitchMon, ZGlitch)
- Locked segments used to steer analysis jobs

DMT Lessons Learned

- Operational status visibility
  - » Lack clear top-level summaries, acoustic alarms
- Monitor status
  - » No record of which monitors running
  - » Performance statistics not readily available.
- DMT trends not available in control room
  - » History is lost when DMT is restarted.
  - » Need to display trends with data viewer

Multiple frame builders

- Control room tools, DMT and LDAS use different data paths.
  - » Control room, DMT use 1s for fast response
  - » Analysis uses 16s to reduce overhead
  - » DMT frame broadcaster reduces fb0 load.
- Data lost at LLO when fb2 not rebooted.
- Needs better monitoring and status visibility
  - » GDS tools on 16s frames where possible

E7 Lessons Learned
Search Group Rediness

Shawhan, Anderson
LIGO Laboratory

LIGO-G020013-03-E

Inspiral Searches:

Experience from the E7 Run (Shawhan)

- Template and FCT search codes were run on-line during E7
  - Both used “flat” (non-hierarchical) algorithms
  - Restricted to a high-mass region of parameter space
  - Used transfer function measured for H2 interferometer during E6
  - Jobs submitted by scripts to analyze data only for times when the interferometer was locked
  - Not all of the data was processed
- Lessons learned
  - Job scripts had to be babysat due to various failure modes
  - Job execution times include non-negligible overhead from time spent in data conditioning API, etc.
  - Need to get reasonably reliable calibration info in a timely fashion

Inspiral Searches
Plans (Shawhan)

- Near-term activities
  - » Detector characterization studies
    - – Investigate time variability of interferometer behavior
    - – Correlate inspiral candidates with environmental transients
  - » Exercise procedures for setting trigger thresholds, and study efficiencies, using simulated events buried in Gaussian noise
  - » Analyze the "playground" data
- Longer-term plans
  - » Develop veto conditions using "playground" data
  - » Study efficiencies using simulated events superimposed on E7 data
  - » Analyze all of the E7 data for each interferometer
  - » Combine results (using TBD statistical method) to set an upper limit
  - » Develop hierarchical search algorithms

Periodic Source
Status (Anderson)

3 source categories and 4 algorithms

» All sky unbiased
  – Sum short power spectra (no doppler correction)

» Known pulsar
  – Heterodyne narrow BW
  – Coherent frequency domain

» Wide area search
  – Hierarchical Hough transform

Periodic Source

Status (Anderson)

- No E7 real-time analysis was performed
- Post-run Time->Freq pre-processing (SFT) is proceeding and may be of more general use
- 1-of-4 algorithms is being developed for LDAS
  - » All 4 algorithms to use LAL
- Still developing core search code

Stochastic Background Search

Status (Fritschel)

- No on-line analysis run
- Before calculating the cross-correlation statistic (output of the search code), simply look at the coherence between H2-L1 ASQ
  - » for 1024 sec blocks, over a 2 hour coincident stretch, no statistically significant coherence seen
  - » weak coherence at power lines sometimes seen over this time scale, but …
- Coherence between power line monitors of the sites calculated for the full 17 day run: power lines are incoherent over this duration

Stochastic Background Search

Near term plans(Fritschel)

- Calculate possible sensitivity to stochastic background, given the posted strain sensitivities and duration of coincidence operation
- Look at the coherence between the two Hanford interferometers
- Continue analysis of ifo –ifo cross-correlation
  - » Determine frequency band(s) that are most significant for this search
  - » Look at variability of the interferometer noise in this band