



Update on the LIGO Data Analysis System

*LIGO Scientific Collaboration Meeting
LIGO Hanford Observatory
August 19th, 2002*

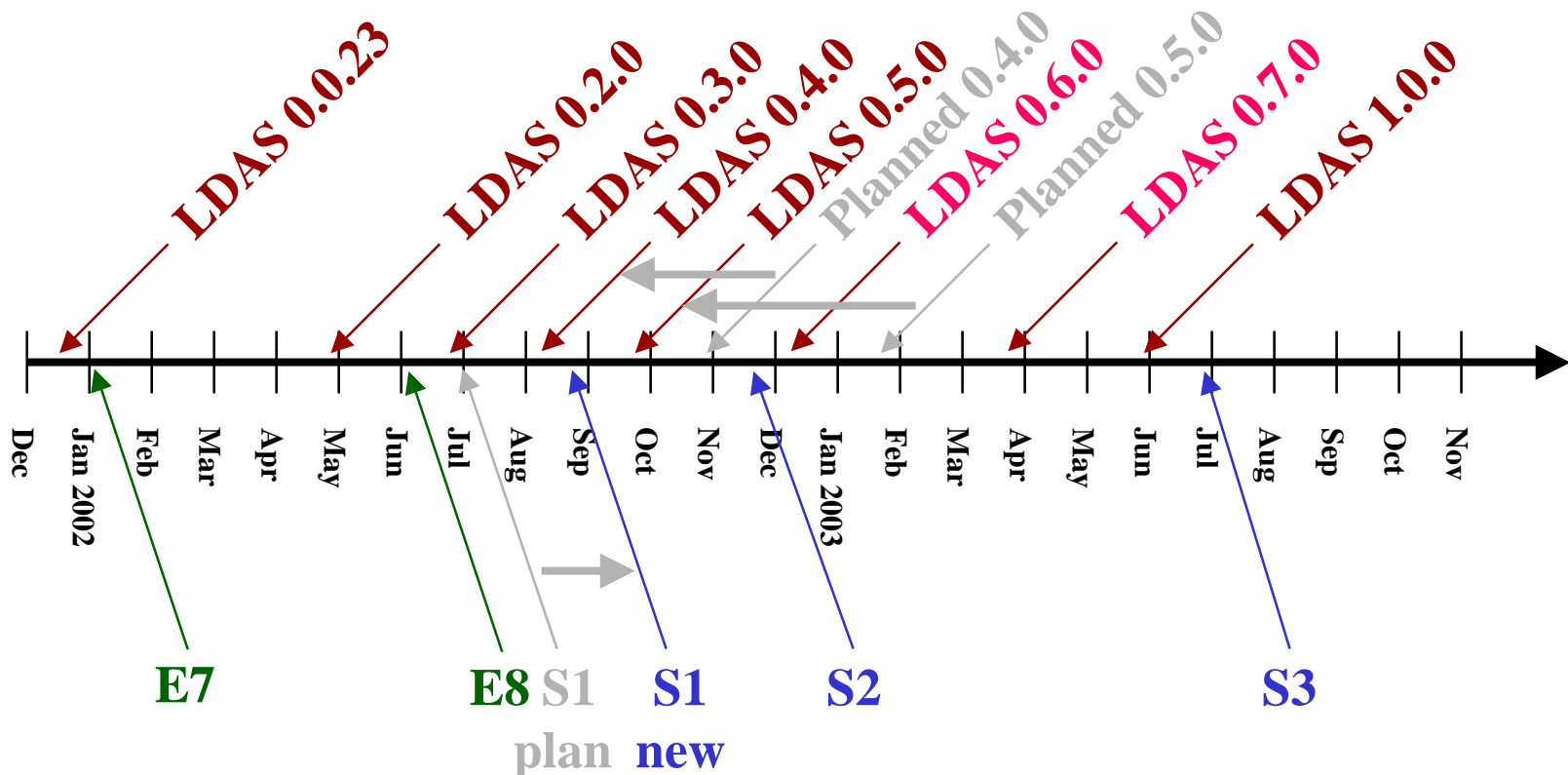
**Kent Blackburn
Albert Lazzarini & Stuart Anderson**



LDAS Software Update

Changes to Release Schedule

- Accelerated release of 0.4.0 to support rescheduled S1
- Accelerating release of 0.5.0 to support new Frame 6 Specification
- Most likely will have a 0.6.0 & 0.7.0 before 1.0.0



LDAS usernames/password Secure Web Server



LDAS Data Analysis System

Welcome to the LDAS CIT Web Site

SOFTWARE

- LDAS Software Index
- LDAS/ASIC Software Development
- LDAS Problem Reporting System
- User Access Tools
- LDAS Bulletin Board

HARDWARE

- LDAS System (Black Dagons)

GETTING STARTED

- How to Install and Configure the LDAS Database
- How to Build LDAS
- How to Configure LDAS
- How to Test LDAS
- LDAS Operator Commands
- LDAS User Commands
- Using controlMonitorAPI Client
- Build and Configuration of LDAS on a Single Host
- LDAS account application request
- LDAS FAQ

DATA

- Frame Archive
- LDAS Database

LOG TOOLS

- LDAS API Run Status
- Summary of Running Jobs
- API Status History Summary
- Electronic Logs

COMMUNICATIONS

- Videoconferencing

LDAS CONTACTS

- Software
- Sysadmin

LDAS CIT Job Files

	Current	Previous	Post
controlMonitor API	View	View	View
datacollection API	View	View	View
filterCache API	View	View	View
eventMonitor API	View	View	View
Frame API	View	View	View
lightweight API	View	View	View
Manager API	View	View	View
metadata API	View	View	View
pid API	View	View	View
wrapper API	View	View	View

LDAS Date: Fri Aug 2 10:04:09 PDT 2002 LDAS Version: 0.4.0

Skated relief map courtesy Johns Hopkins University Applied Physics Laboratory

Contact us

Last modified on March 26, 2002



LDAS Account Password Request Page

The LIGO Data Analysis System (LDAS) provides an interface to submit jobs for the purpose of accessing raw LIGO data, the LIGO database and to carry out computationally demanding analyses on LIGO data and data products. The system uses a job submission model described at <http://ligo-www.caltech.edu/submit/API.html>. In order to use this job submission system a user must provide a username, password and email address. If you are a member of the LIGO Scientific Collaboration and a participant in an active LIGO I Data Analysis MOU, then you are eligible to apply for access to the front of a LDAS job submission username and password. First select *Request new password* from the pull down menu below and click on next page to get started. You will be asked a couple of questions and then directed to an application form to be filled out in full. Once your application has been approved, you will be notified by email of the activation of your account. In addition, if you have forgotten your username and password, you may retrieve your existing username and password by selecting *Retrieve existing password* from the pull down menu and clicking on next page. Your forgotten username and password will be mailed to you immediately after you have correctly answered five questions based on your original request.

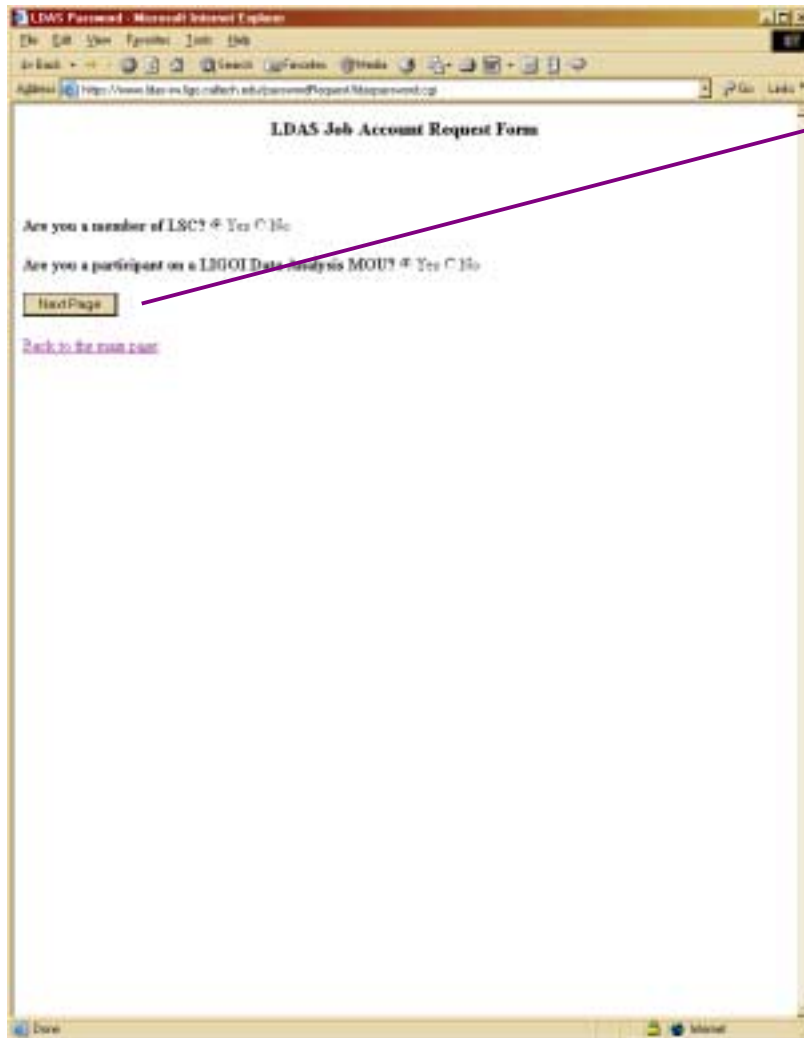
Remember, your LDAS username and password are for your use and your use alone. Do not share them with anyone else.

Request new password

Next Page

All shared account have been removed!

LDAS usernames/password Secure Application Form



LDAS Password - Microsoft Internet Explorer

Address: <http://www.ligo.caltech.edu/page/ldas/RequestManagement.cgi>

LDAS Job Account Request Form

Are you a member of LSC? Yes No

Are you a participant on a LIGO Data Analysis MOU? Yes No

[Back to the main page](#)



LDAS Password - Microsoft Internet Explorer

Address: <http://www.ligo.caltech.edu/page/ldas/RequestManagement.cgi>

LDAS Job Account Request Form

Please fill in all fields.

Full Name: Full Name must be at least 4 characters (at least 2 words space separated)

Classification: student postdoc researcher faculty other

Institution:

Office Phone Number:

Email:

Requested LDAS User Name:

Requested LDAS User Password: Password must be at least 8 characters
Allowed characters: -_a-zA-Z0-9@

Verify LDAS User Password:

Facsimile:

Proposal Purpose:

Number of Requested Beamforming Nodes per Job:

Number of Seconds of Frame Data per Job:

Number of Jobs Anticipated:

LDAS Account Duration: (you may)

LDAS@DEY LDAS@MIT LDAS@LHO
 LDAS@TEST LDAS@CTT LDAS@LLO

Lessons Learned: E7

“High priority items needed to get LDAS on track for Science Runs”

- Rework configuration & build rules - **DONE!**
- Create new diskCacheAPI; Remove from frameAPI - **DONE!**
- Improve reliability of dataConditionAPI (thread issues) - **DONE!**
- Create shared resampling library for use in both the frameAPI & dataConditionAPI - **DONE!**
- Extend system monitoring - **DONE!**
- Add interpolation, Kalman filters, regression and rework intermediate() function in dataConditionAPI - **MOSTLY DONE!**
- Reduce memory usage in dataConditionAPI below 5x data - **3x!**

Lessons Learned: E7

Continued

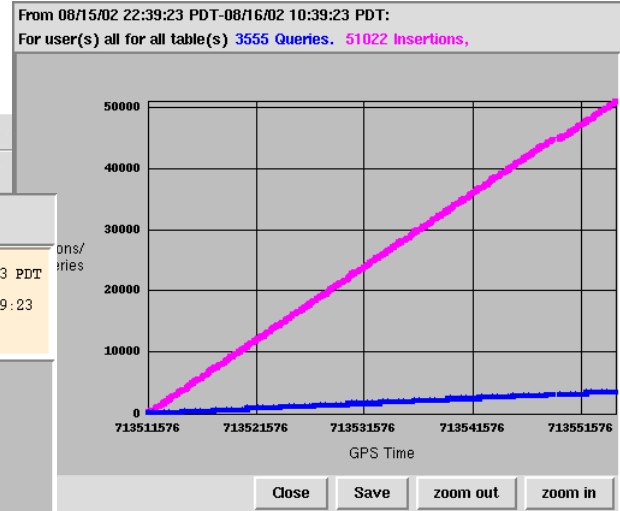
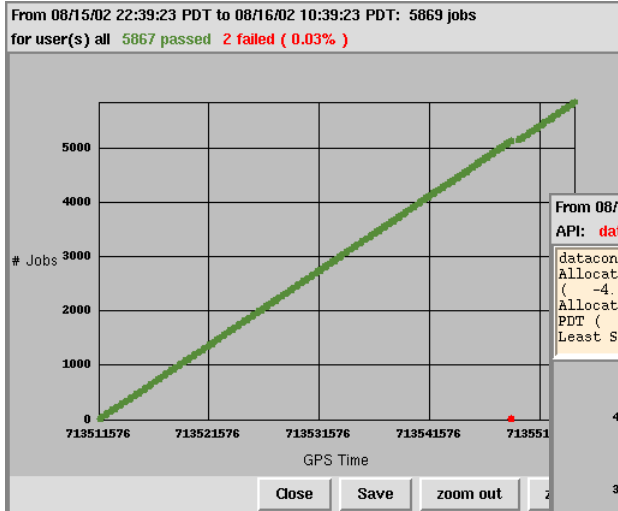
“High priority items needed to get LDAS on track for Science Runs”

- Move frameAPI / diskCacheAPI to new *dataserver* - **DONE!**
- Improve docs, interfaces and table design with LSUG - **DONE!**
- Implement new TCL channel management interface to better control data sockets - **DONE!**
- Add detector geometry metadata to dataPipeline for stochastic analyses - **DONE!**
- Add system load monitoring tools for GriPhyN - **DONE!**
- Determine archival technology (SAM-QFS vs HPSS) - **MUCH DONE - MORE TO DO!**
- Build up CIT LDAS System - **FULLY FUNCTIONAL!**

Performance Enhancements

- Increased concurrency of jobs (20 assistant managers now the default).
- Removed 30 seconds of startup overhead from “*mpirun wrapperAPI*” on LDAS beowulfs (down to ~3 seconds).
- Reduced latency to transmit data from one API to another by several seconds.
- Increased computational efficiency of *managerAPI* by order of magnitude.
- Improved overall compute performance by factor of two using C++ compiler optimization options.
- More extensive test scripts using more search DSOs.

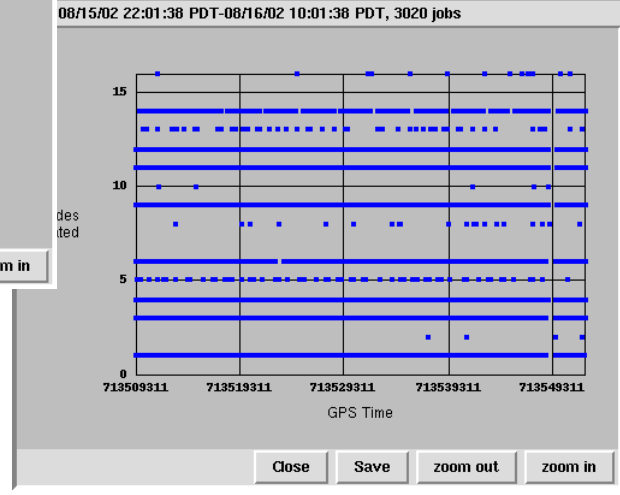
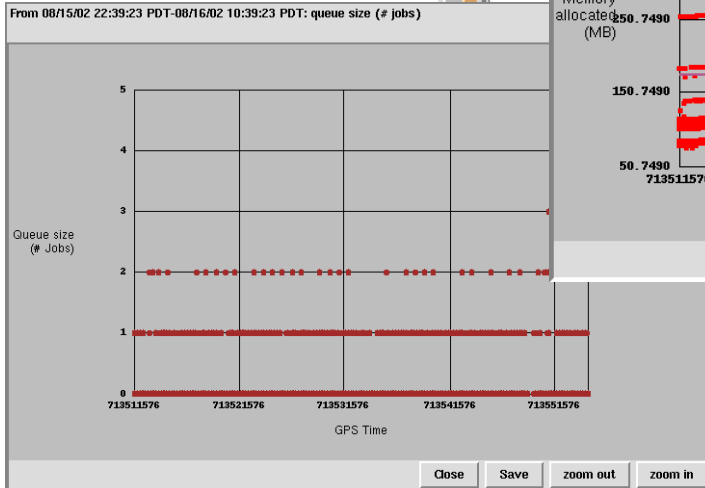
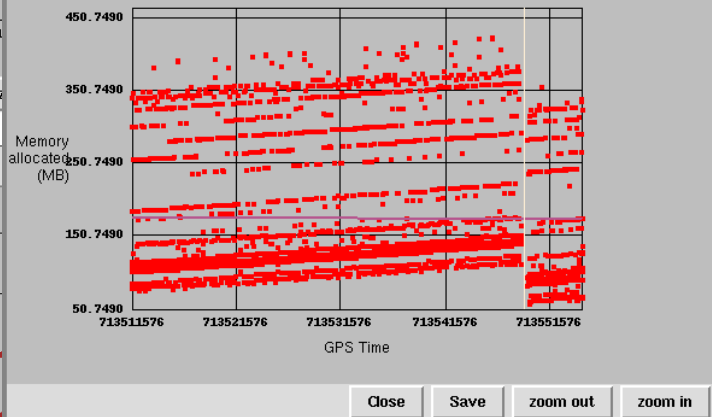
System Monitoring



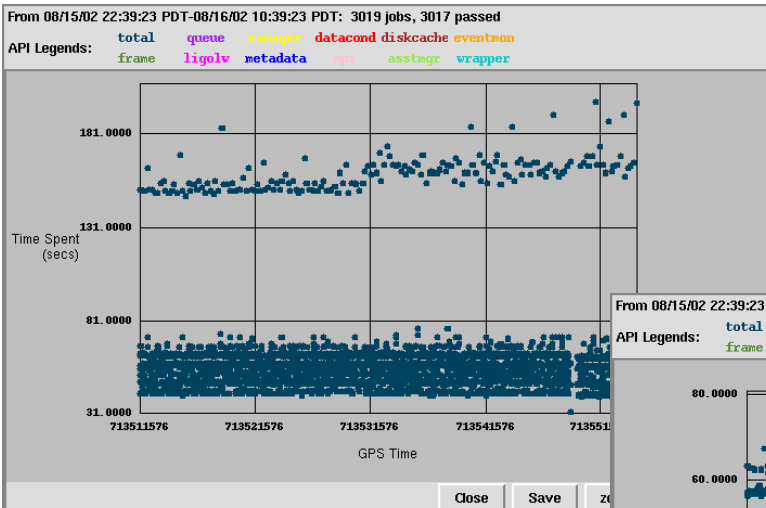
From 08/15/02 22:39:23 PDT-08/16/02 10:39:23 PDT: 3028 jobs
API: datacond

```

datacond memory allocation rates before each restart
Allocated -149.64 MB in 37500 secs before restart at 08/16/02 09:04:23 PDT
(-4.00 KB/sec, -50.00 KB/job)
Allocated 59.11 MB in 5700 secs from last restart to 08/16/02 10:39:23
PDT ( 10.00 KB/sec, 19.00 KB/job)
Least Squares fitted rate: -0.081 KB/sec, -1.155 KB/job
    
```



Job Monitoring

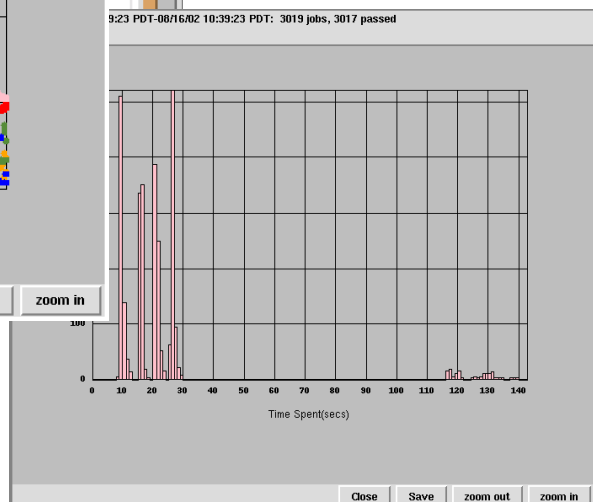
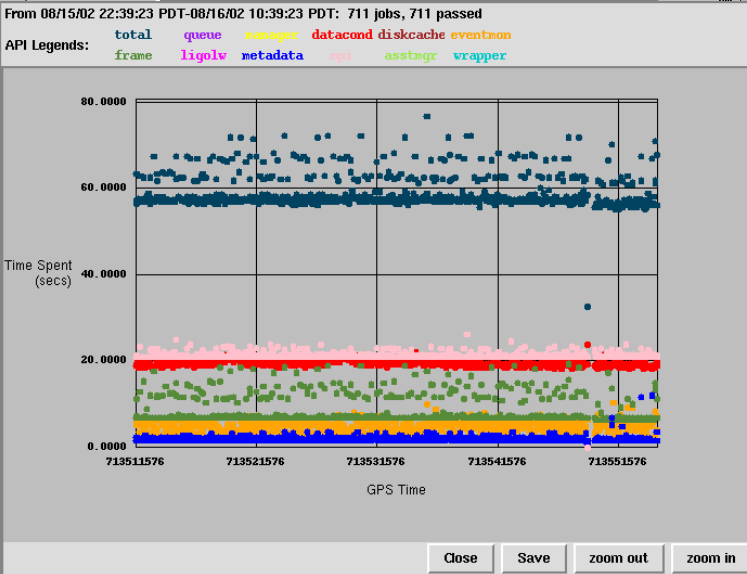
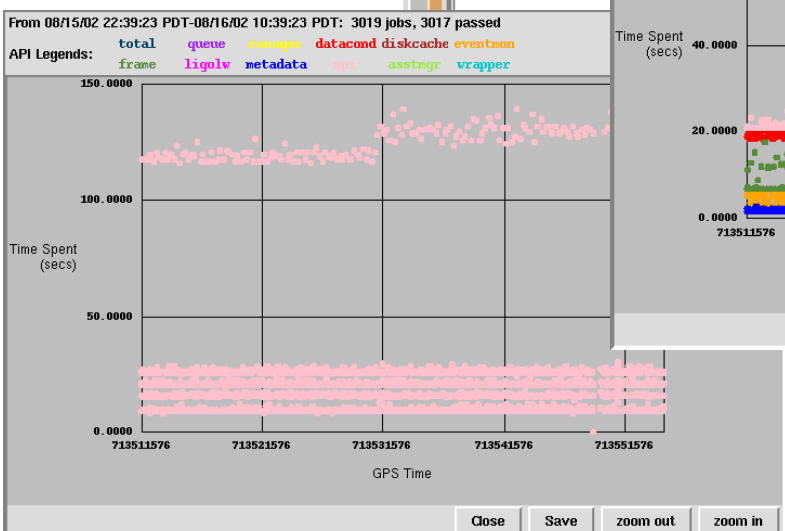
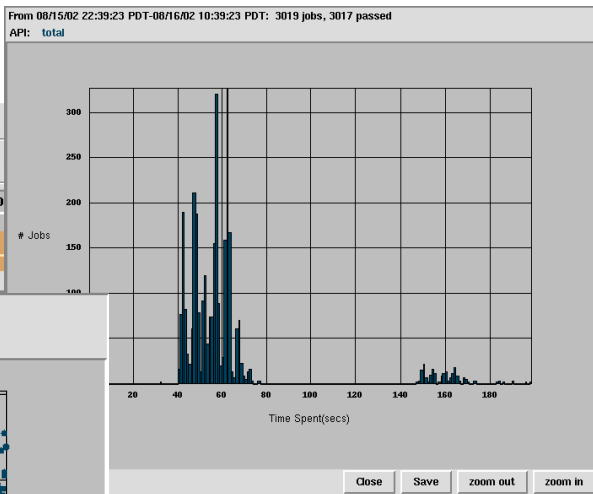


mas-dev.ligo.caltech.edu

Jobs LDAS Logs MPI Jobs Beowulf System Tape Co

Database Statistics Queue Statistics Rejected Job Statistics

itics



Other New Functionality

- **New User Commands**
 - **createRDS**: Creates the Reduced Data Set (RDS) Frames. Includes support for resampling and runs unique thread in frameAPI.
 - **dataStandAlone**: A dataPipeline job that stops short of issuing the mpirun command used to start the wrapperAPI on an LDAS beowulf cluster. Used to integrate GRID technology with LDAS.
 - **putStandAlone**: The second half of a dataPipeline that takes data products from a standalone wrapperAPI running on a GRID resource and re-integrates the data products into LDAS and the LIGO database.
- **New Web pages for monitoring job progress in system**
 - Monitors jobs in queue and jobs that are actively being processed.
 - Shows progress of active jobs through the system.
- **New encrypted LDAS password exchange when using ldasjobs package found in LIGO Tools.**

Software To Do: Short List

- All new *frameCPP* being developed based on Frame Specification Version 6 (previously labeled 5).
- The *frameAPI* is being completely rewritten to use multiple CPUs in threads to support increased demands placed by more complex searches and reduced frames.
- Further improve queue performance in the *managerAPI*.
- Adding option to *wrapperAPI* resource file to remove additional half a minute of overhead and improve efficiency of node usage.
- Implement new actions, improved metadata, resolve remaining memory leaks in *dataConditionAPI*.



LDAS Hardware Update

LDAS Hardware Update

- Initial archive analysis system is operating at Caltech.
- Beowulf upgrades at LHO, LLO, and MIT.
- Large IDE disk cache for raw S1 data at Caltech.
- All servers upgraded to Gigabit Ethernet.
- Shared SAN file system operational as data interface between LDAS and CDS.

Science Run Configurations

“Increased computational capacity over E7 and investigated advanced storage configurations but delay full compute farm deployment until S2+”

	SAN (TB)	IDE (TB)	CPU (GHz)	Tape (TB)
LHO	10	2	139	2
LLO	5	2	107	2
CIT	3	18	34	90
MIT	1	2	45	0
DEV	1	2	25	2
TEST	1	0	8	0

Archive Storage Solutions

SAM-QFS versus HPSS

- **SAM-QFS advantages**
 - **Simplicity/reliability**
 - **Media import/export**
 - **License cost allows for use at observatories**
 - **Disaster recovery (GNU TAR)**
 - **Metadata performance (x1000)**
 - **Single vendor solution (server, software and OEM storage)**
 - **Reduced dependency on CACR**
- **HPSS advantages**
 - **Few man year experience**
 - **Free at Caltech**
 - **40 TB successfully stored**

Hardware To Do: Short List

- Virtual Private Network (VPN) between LDAS systems to allow for database federation and replication
- Finalize HSM decision:
 - Fibre Channel tape drives.
 - ~500 tape slot library at observatories?
- Grow SAN at Observatories to allow high-speed access to raw frames from DMT and General Computing machines.
- Install full-scale Beowulf clusters in time for S2+.