



LIGO Data Analysis System (LDAS) LIGO/LSC Algorithm Library (LAL) Mock Data Challenges

AMALDI 4

**The University of Western Australia
Nedlands, Western Australia, Australia**

July 8th – July 14th, 2001

Kent Blackburn

California Institute of Technology

LIGO Laboratory

LIGO G010249-00-E

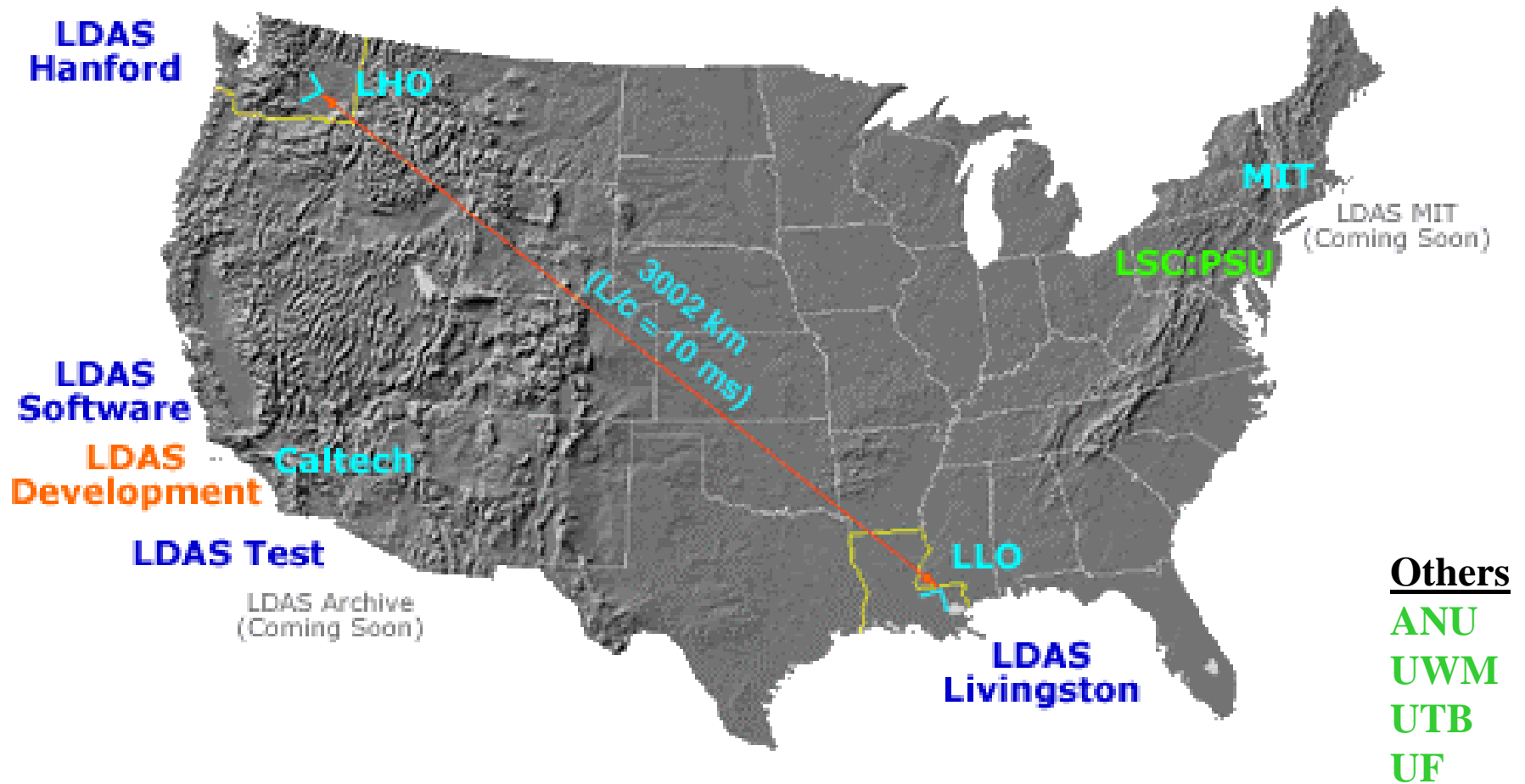


Overview

- **LDAS is a distributed computing environment mixing remote process control on servers with message passing on PC cluster.**
 - A TCL scripting layer is used to manage processing dynamics.
 - An underlying C++ extension to TCL provides the computational power.
- **LDAS provides a framework for conducting scientific studies of LIGO data primarily through the concept of a “data-pipeline”.**
 - LDAS user commands initiate a pipeline through a simple socket connection with username and password authentication.
- **LDAS uses dynamically loaded shared objects (dso) to uniquely customize the LAL analysis code used on LDAS PC cluster.**
 - Each node in a parallel cluster job loads search specific code at runtime.
- **LAL is a standardized library collection of routines used to carry out LIGO data analysis**
 - LALwrapper is a thin interface distributed along side LAL for dynamically loading LAL search codes into the LDAS wrapperAPI



Multiple LDAS Facilities





LDAS Hardware

- **Data Server** (*raw frame files & system*)
- **Metadata Server** (*database server & tables*)
- **Data Conditioning Server** (*pre-filtering*)
- **Linux PC Cluster** (*Message Passing Interface Standard Computation*)
- **Control Monitor Workstation** (*Configuration, Health & Status*)
- **Misc.** (*network switch, UPS, tape robot, guest workstations*)
- **NOTE:** *LDAS has successfully run on a single laptop computer!*



Hanford, WA Facility

Network Switch

Data Conditioning Server & Cluster Gateway

Parallel
Compute
Cluster
(Beowulf)

Frame data
& Metadata
Servers



Disk Storage
(Terabytes)

Control
& Monitor
Workstation

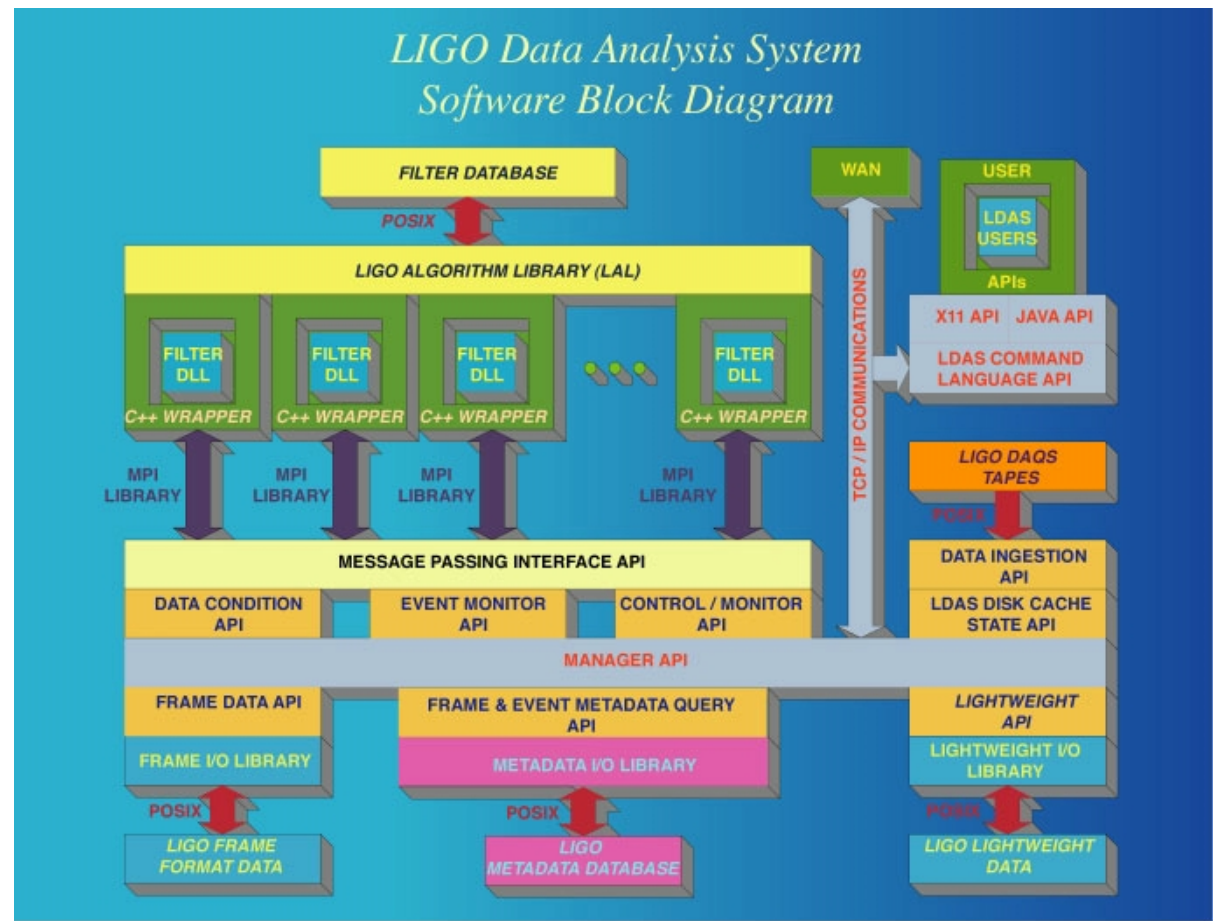
UPS



LDAS Software

Major Components:

- *12 Advanced Process Interface (API) modules*
- *12 libraries*
- *1 commercial database (DB2)*
- *1 internal & 2 external data formats*
- *Half dozen user interfaces already developed*





Software Update

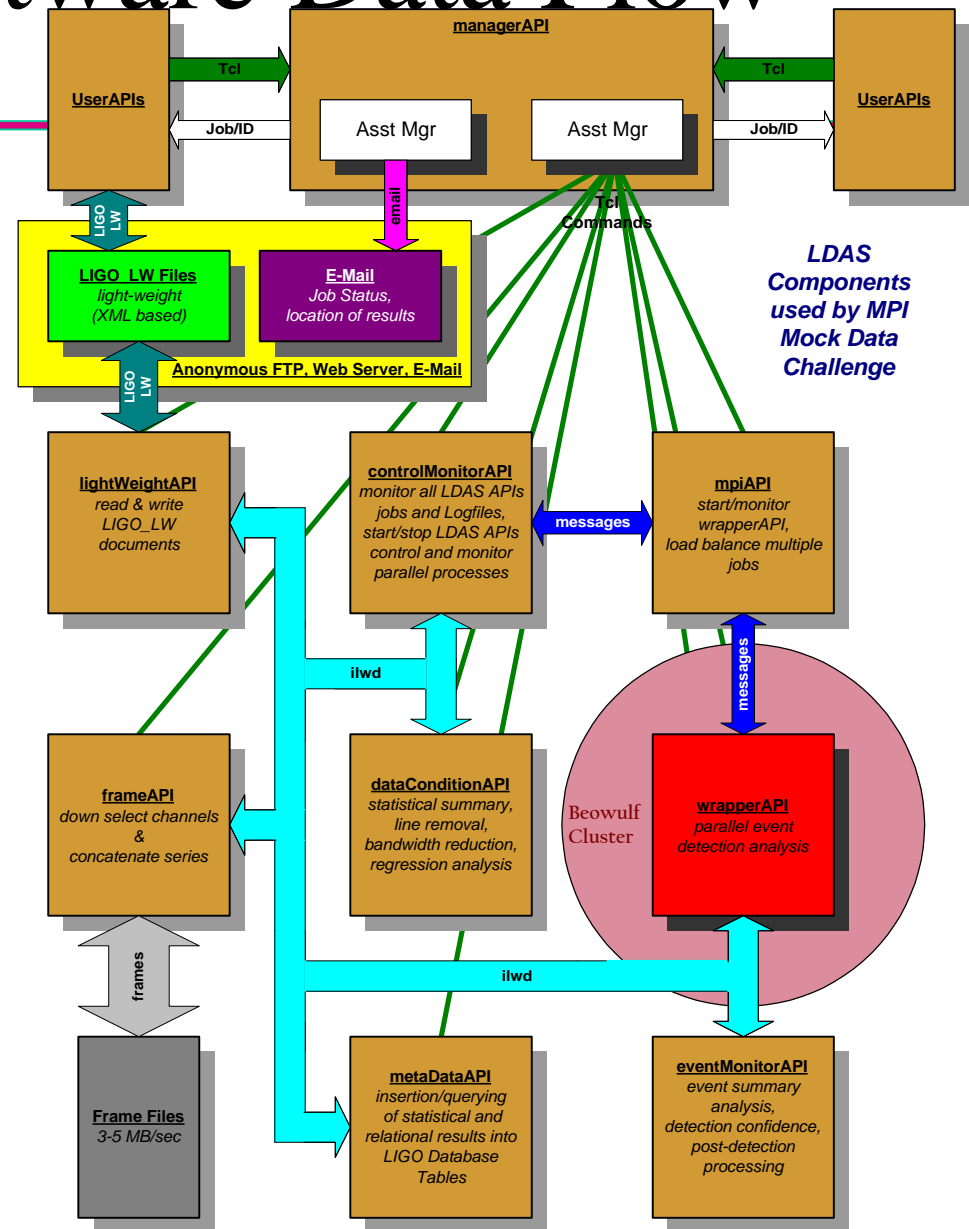
- CurrentLDAS Release 0.0.18 (last month).
- We are still developing at the **alpha** level.
 - Two major APIs between alpha and beta LDAS.
 - *diskCacheAPI (disk management / data availability).*
 - *dataIngestionAPI (tape control / inter-site data movement).*
- Approximately 600,000 lines of code in place.
- Represents ~80-90% completion of LDAS.
- Evolution of LDAS usage will uncover new flow control models requiring new user commands.



Flexible Software Data Flow

LDAS dataPipeline command:

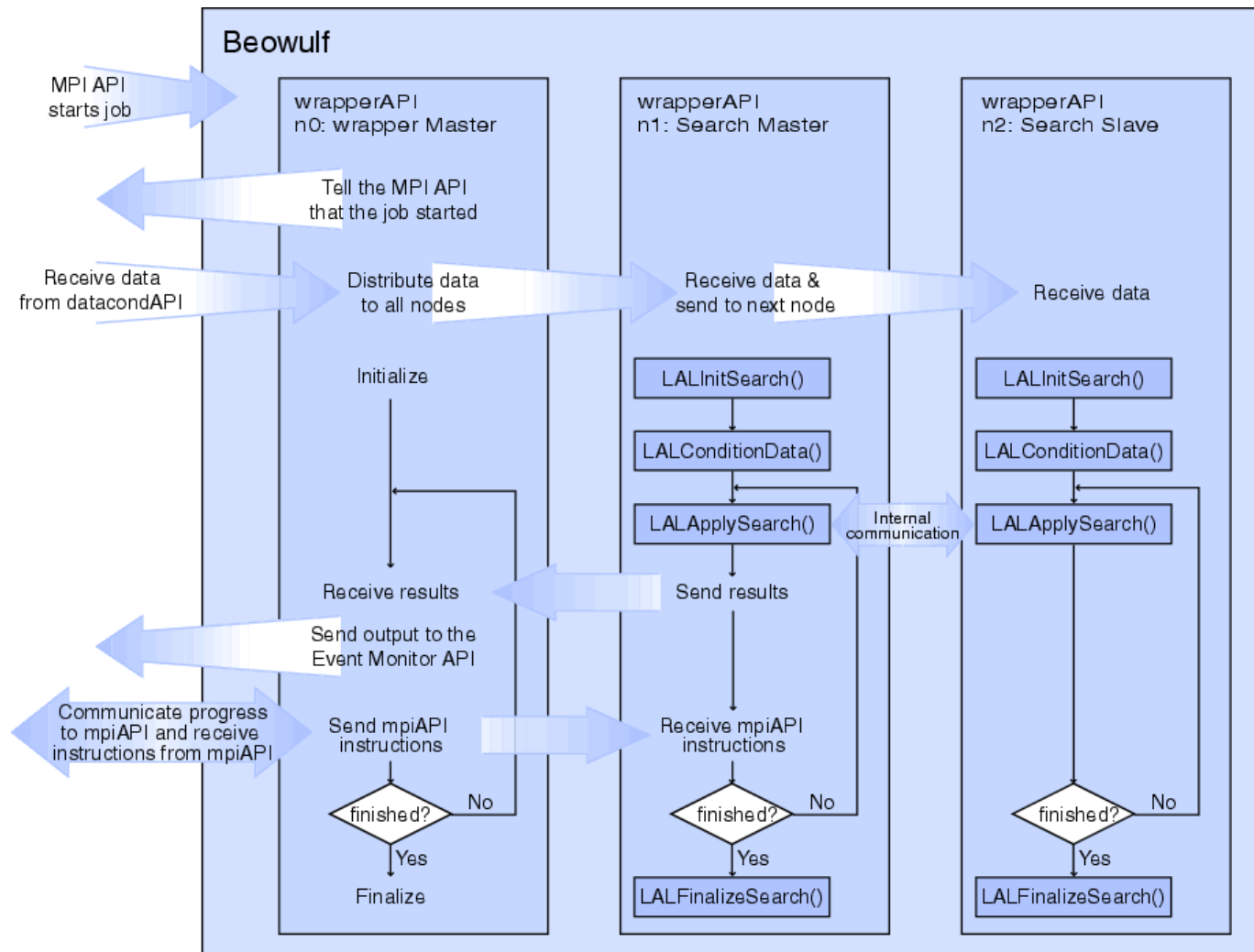
- UserAPI connects to **managerAPI** and starts pipeline.
- **managerAPI** assigns *assistant manager* to control job.
- **frameAPI** reads *frames* into the LDAS system, down-selects and sends data to the **dataConditionAPI**.
- **dataConditionAPI** pre-conditions data and forwards this to the **wrapperAPI** under the control of **mpiAPI**.
- **wrapperAPI** performs template based search using a dynamically loaded shared object.
- **eventMonitorAPI** receives search results and directs events to the **metaDataAPI** where they are placed in database.
- **eventMonitorAPI** sends user data products from search to **lightWeightAPI** where they are sent to user via FTP or URI commands.
- **managerAPI** notifies user via email of job completion.





WrapperAPI

(LDAS/LALwrapper/LAL Interface)





Mock Data Challenges

- Purpose is to test, verify and validate LDAS and LSC software under conditions which simulate the LIGO science run data analysis requirements.
- Series of increasing more complex challenges scheduled beginning August 2000 and spanning into early 2002.
- Four MDCs completed as of July 2001:
 - dataConditioningAPI (August 2000),
 - Message Passing Interface wrapperAPI (January, 2001),
 - Database (April, 2001),
 - Binary Inspiral Search (May, 2001).



Data Conditioning MDC

- July 31 – August 7th, 2001
- LSC Lead: Sam Finn
- 163 API & 40 performance for a total of **203 tests** exercised:
 - LDAS: CHEX, genericAPI, managerAPI, and dataConditionAPI
- Documented in “LDAS dataConditionAPI: Mock Data Challenge #1”
 - LIGO DCC: LIGO-T000124-00-E
- Test Scripts stored in LDAS CVS Repository and are repeated weekly
- Paved the way for signal conditioning to be used in future MDCs:
 - FFTs, Linear Filters, Mixers, Windowing, Power Spectral Densities, Cross Spectral Densities, Resampling, Statistics, Simple Mathematics...
- Algorithm results verified against Matlab.
- Provided first opportunity to distribute LDAS usage model beyond LDAS and into the LSC community



Message Passing Interface MDC

- January 15-22, 2001
- LSC Lead: Patrick Brady
- 4 MPI, 18 LALwrapper, 3 mpi-wrapper, 43 inspiral, 28 power, 8 long inspiral & 5 long power for a total of **109 tests** exercised:
 - LDAS: managerAPI, mpiAPI, wrapperAPI, and trivial eventMonitorAPI,
 - LSC: LAL, LALwrapper, inspiral shared object (dso), and power dso.
- Documented in “MPI Mock Data Challenge”,
 - LIGO DCC: LIGO-T010024-00-Z.
- Test Scripts stored in LAL CVS Repository and repeated weekly.
- Paved the way for inch-pebble search codes used in future MDCs:
 - Binary Inspiral, stochastic, burst, and periodic sources.
- Test waveforms injected “blindly” into 75 minute data streams and found!
- Distribute parallel analysis model beyond LDAS and into the LSC community.



Database MDC

- January - April, 2001
- LSC Lead: Peter Shawhan
- Wide range of evolving tests exercised:
 - LDAS: managerAPI, metaDataAPI, lightWeightAPI, Table Design, and GUILD.
- Documented in “Plan for the LDAS Database Mock Data Challenge”,
 - LIGO DCC: LIGO-T000089-00-E.
- Considered by Peter and Lab to be a tremendous success!
- Paved the way for inch-pebble MDCs and Upper Limits Engineering Run in Fall of 2001.
- Provided needed feedback in table schema to better support LIGO analyses.



Inspiral MDC

- May 15-18, 2001
- LSC Lead: Patrick Brady
- Repeated MPI MDC tests plus 11 new LDAS dataPipeline, 9 new inspiral dso, and 20 new FCT dso for a total of **40 new tests** exercised:
 - LDAS: managerAPI, frameAPI, dataConditionAPI, mpiAPI, wrapperAPI, eventMonitorAPI, metadataAPI, and lightweightAPI ;
(i.e., the complete LDAS dataPipeline command tests).
 - LSC: LAL, LALwrapper, inspiral dynamic shared object (dso), and Fast Chirp Transform (FCT) dso.
- Documented in “Inspirial Mock Data Challenge”.
- Test Scripts stored in LAL CVS Repository and repeated weekly.
- Discovered several frequency domain convention differences data conditioning and LAL which are presently being rationalized.
- Paved the way for later inch-pebble MDCs and Upper Limits Run.



What the Future Holds...

Software readiness, reliability, robustness and efficiency will continue to be tested, and challenged through the MDCs and Engineering Runs:

- Stochastic, Burst & CW Inch-Pebble MDCs (Sept 2001)
- **Beta Release LDAS 0.1.0 (Late September 2001)**
- Upper-Limit Engineering Run (October 2001)
- LDAS Archive MDC (Fall 2001)
- **LDAS Release 1.0.0 (December 2001)**
- Pilot Science Runs (First Half 2002)
- Start of LIGO I Science Run (July 2002)