LIGO Data Analysis

Peter Shawhan
(LIGO Laboratory / Caltech)

For the LIGO Scientific Collaboration

VIII International Workshop on Advanced Computing and Analysis Techniques in Physics Research

Moscow State University
June 26, 2002

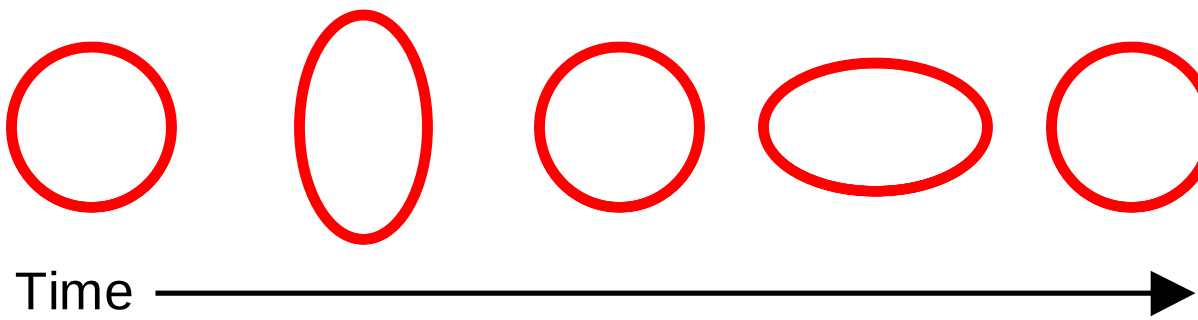
LIGO-G020284-00-E

Gravitational Waves

Predicted by Einstein's theory of General Relativity

Massive objects, moving at velocities near the speed of light, can distort the geometry of space-time

Far from source, appear as transverse quadrupole waves



Dimensionless strain: $h = \Delta L / L$

Typical strain amplitude at Earth: $< 10^{-21}$!

Gravitational waves have not been detected … yet

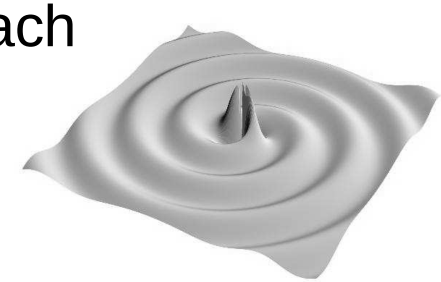
Sources of Gravitational Waves

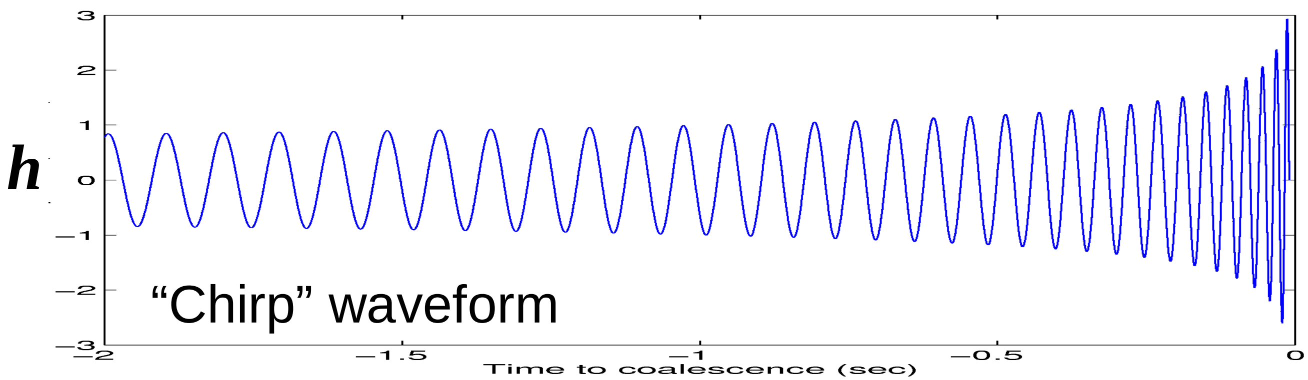
"Inspiral" (orbital decay) of a compact binary system

Two neutron stars, two black holes, or one of each

One of the most promising sources, since:

- Binary neutron-star systems are known to exist
- The waveform and source strength are fairly well known (until just before merging)





Note: the measured orbital decay rate of the binary pulsar PSR 1913+16 exactly matches the expected rate due to gravitational radiation !

Sources of Gravitational Waves

Supernova explosion

Wave emission depends on asymmetry of explosion

“Ringing” oscillations of a newly formed black hole

Stellar collapse to form a neutron star

Rapidly-spinning neutron star

Will radiate if slightly asymmetric

Stochastic radiation from the early universe

“Unexpected” sources ?

This is a new observational science !

Detection of gravitational waves will open a new window on astrophysics and will test the theory of General Relativity

Gravitational-Wave Detectors

First detectors: resonant aluminum "bars"

First built by Joseph Weber in the 1960s

Several cryogenic bars are currently in operation and achieve high sensitivity at their resonant frequencies

AURIGA detector

Several large *interferometric* detectors are now being commissioned: LIGO, VIRGO, TAMA, GEO

Use a laser to measure distance

Sensitive over a wide frequency range

The search for gravitational waves is an international effort

The LIGO Project

LIGO = Laser Interferometer Gravitational-wave Observatory

Funded by the U.S. National Science Foundation

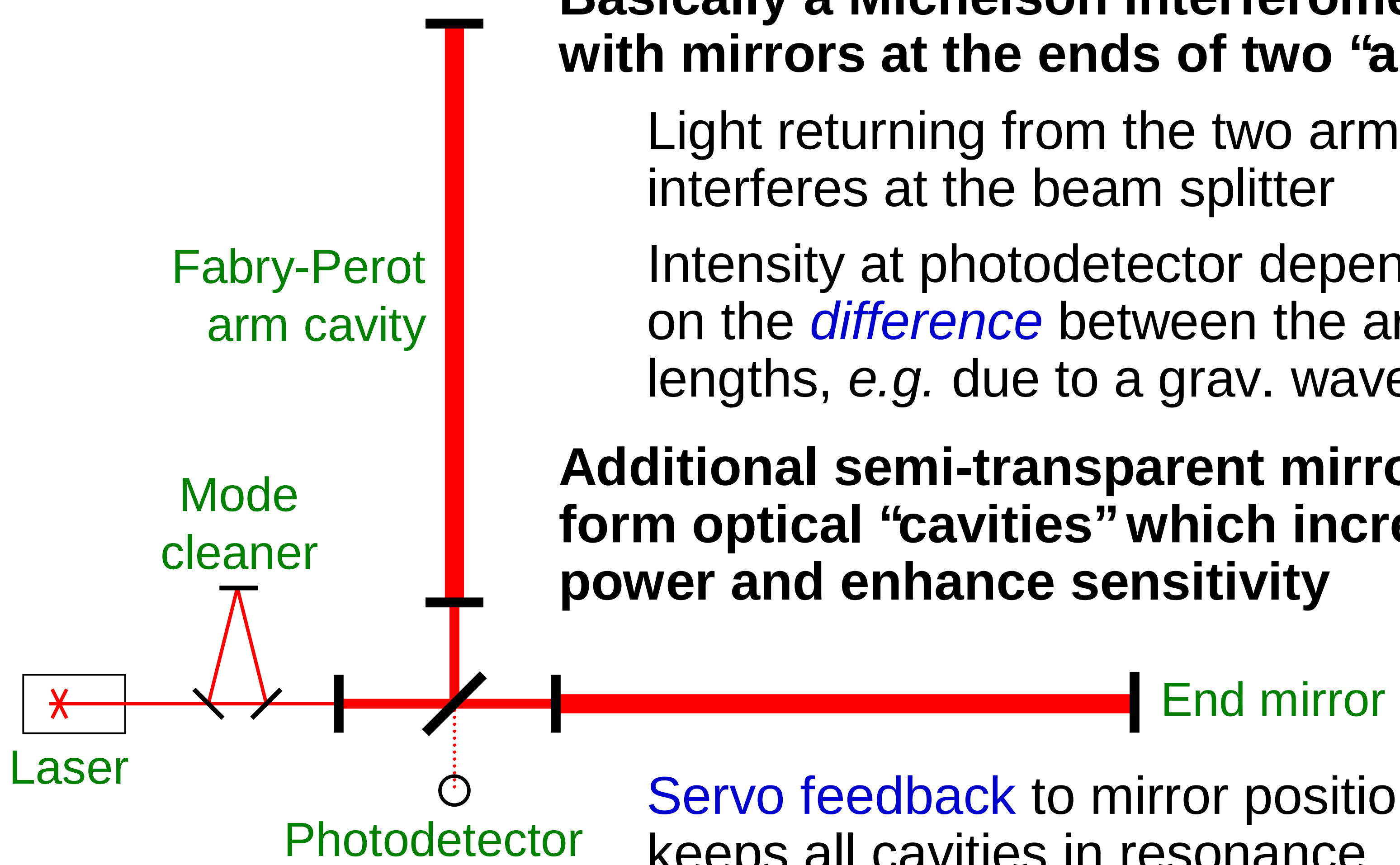
"LIGO Laboratory" consists of Caltech and M.I.T.

The broader "LIGO Scientific Collaboration" includes over 300 scientists from over 30 institutions worldwide

- Perhaps ~75 are actively involved in data analysis

LIGO Optical Layout

(Other large interferometric detectors are similar)



Basically a Michelson interferometer, with mirrors at the ends of two "arms"

Light returning from the two arms interferes at the beam splitter

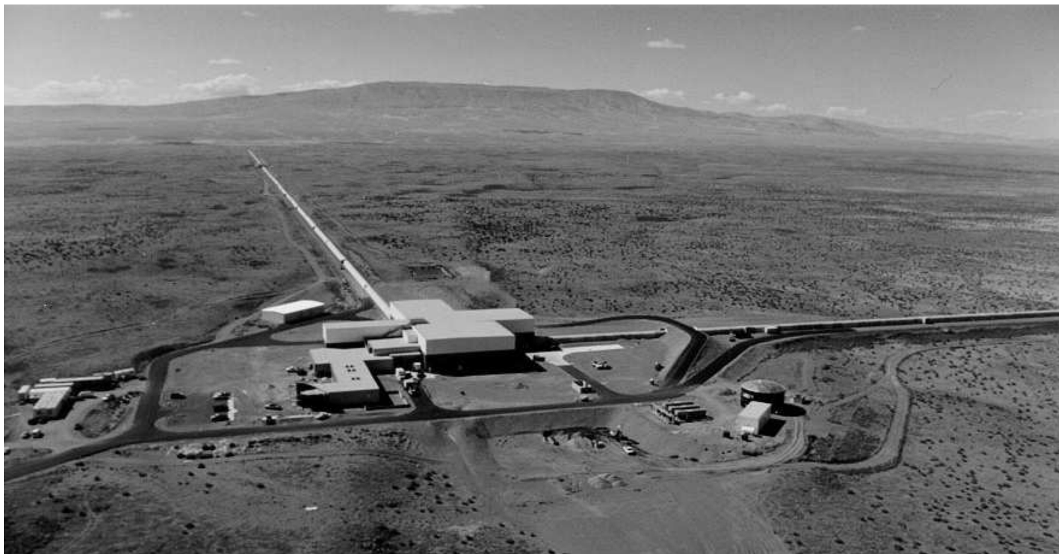
Intensity at photodetector depends on the *difference* between the arm lengths, *e.g.* due to a grav. wave

Additional semi-transparent mirrors form optical "cavities" which increase power and enhance sensitivity

Servo feedback to mirror positions keeps all cavities in resonance

A LIGO Mirror *in situ*

LIGO Observatories



Hanford Observatory
Washington
Two interferometers
(4 km and 2 km arms)

Livingston Observatory
Louisiana
One inteferometer (4 km)

LIGO Status

Construction and detector installation is complete

All three interferometers have been operated successfully in their complete optical configuration

- Some servo systems have not yet been commissioned
- Currently working to reduce noise and improve robustness

Have conducted several "engineering runs"

- To practice collecting and analyzing data

First "science run" begins this Saturday !

- Will collect data for 17 days

Sensitivity is still a few orders of magnitude away from the design goal, but is starting to be of scientific interest

LIGO Data

Main "gravitational wave channel" (servo signal which measures the arm length difference) is sampled at 16384 Hz

- Synchronized to GPS time reference

Data stream also includes many auxiliary channels

- Readback channels from the various servo systems
- Environmental sensors (seismometers, magnetometers, etc.)

There is no "trigger" — channels are sampled continuously and written first to disk, later to tape

Total data rate from each interferometer: ~3 MB/sec

- Gravitational wave channel is only ~2% of the data stream

Data to be archived: ~100-200 TB per year

Gravitational-Wave Data Analysis

Different scientific topics require different analysis methods

Searches for (short) transient signals

- Inspiral and other known waveforms: optimal filtering
- "Bursts" (transients with unknown waveform): several algorithms

Searches for (long) periodic signals

- Requires integrating over long periods

Search for stochastic gravitational-wave background

- Requires cross-correlating data from different detectors

Detector characterization

- Requires access to auxiliary channels

LIGO Data Analysis System (LDAS)

Follows a "computing center" model

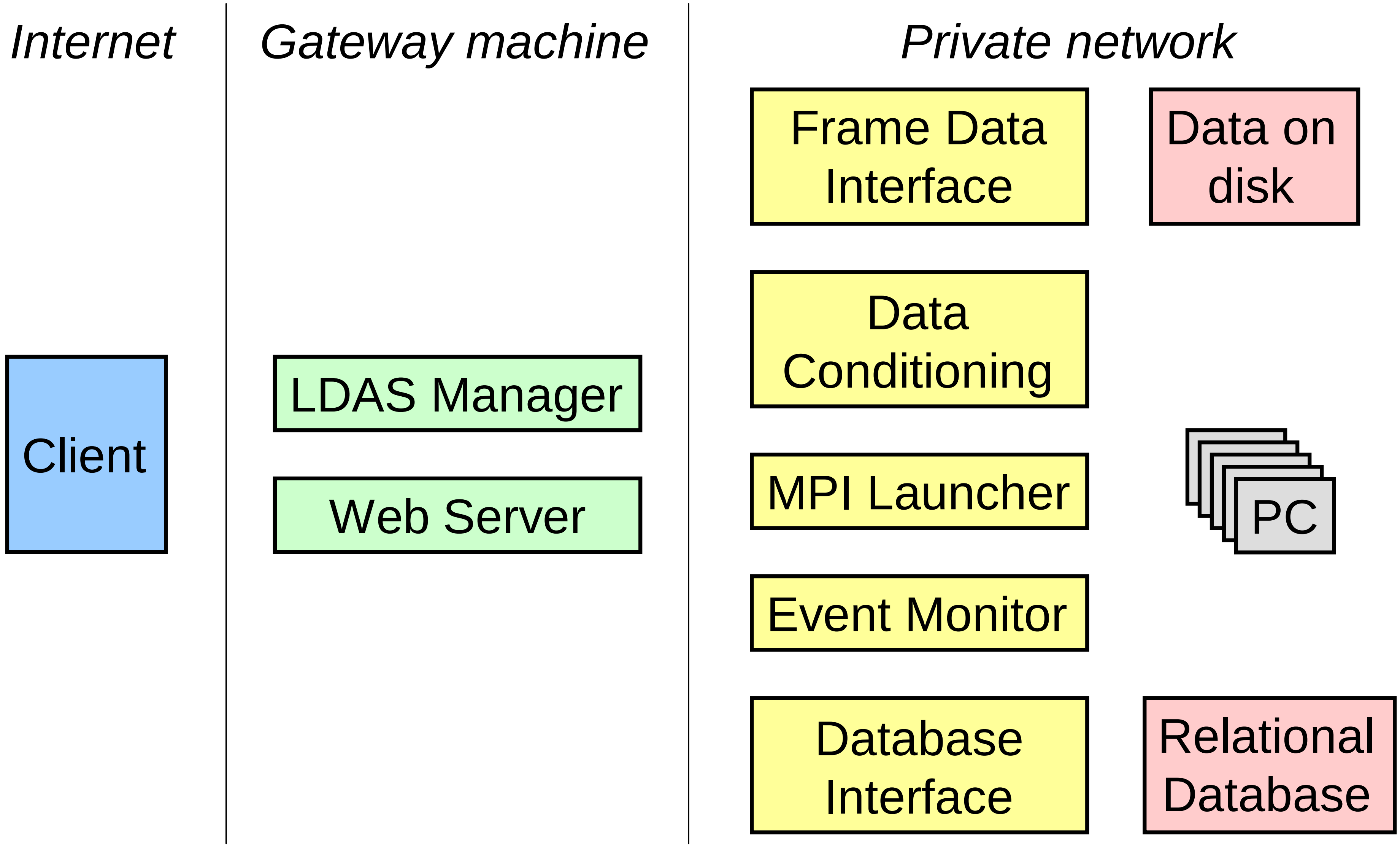
- Dedicated hardware
- Machines are on a "private" network; no access from Internet except to a single gateway machine
- Main data archive is at Caltech; data is copied to other sites as needed
- Software environment created specifically for LIGO

Remote job submission and result retrieval via gateway

- Socket-based job submission protocol; no unix login by users
- Access requires an LDAS username / password

LDAS systems at LIGO observatories, Caltech, M.I.T., and a few institutions in the LIGO Scientific Collaboration

Components of LDAS



Implementation of LDAS

LDAS components are separate unix processes

- Run on several different machines (normally)
- Socket-based job control and data transmission

The LDAS Manager controls job scheduling, as well as the sequence of component operations for each job

Components are written in Tcl and C++

- Tcl is used for job control, interprocess communication, and high-level operations
- CPU-intensive operations on data objects are implemented in C++, called from the Tcl layer

Parallel Processing in LDAS

Uses MPI running on a "Beowulf cluster" of PCs

LDAS delivers input data to memory on master node

User's analysis code is in the form of a shared-object library

- Must contain a standard set of entry points
- Loaded at run time, then called

Scientific users have control over the parallelization scheme

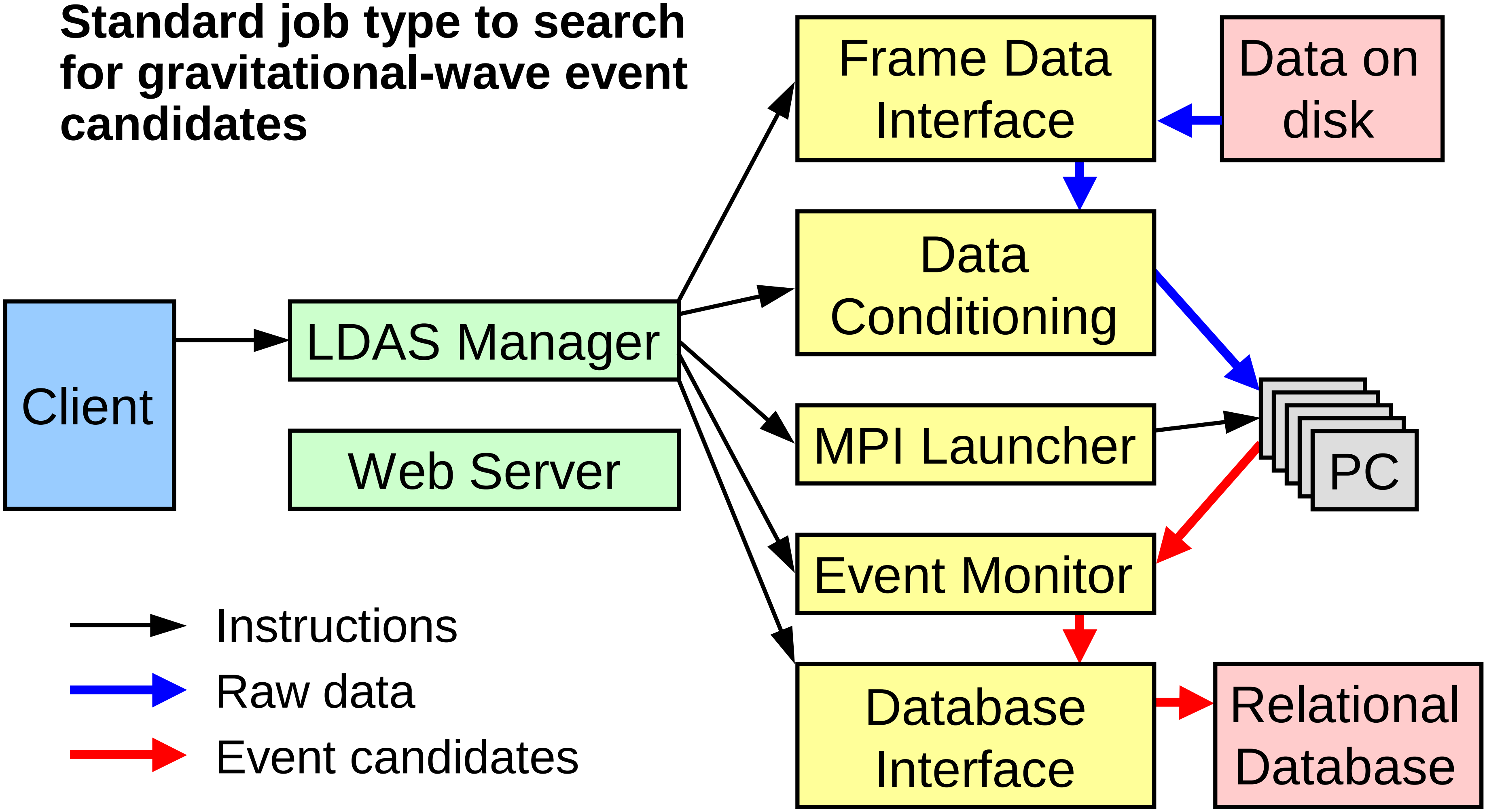
- Typically, the data is broadcast to all nodes, which process it differently (*e.g.* searching for different waveforms)

Results are collected on master node, then passed to "Event Monitor" component

- *e.g.* candidate event times and waveform parameters

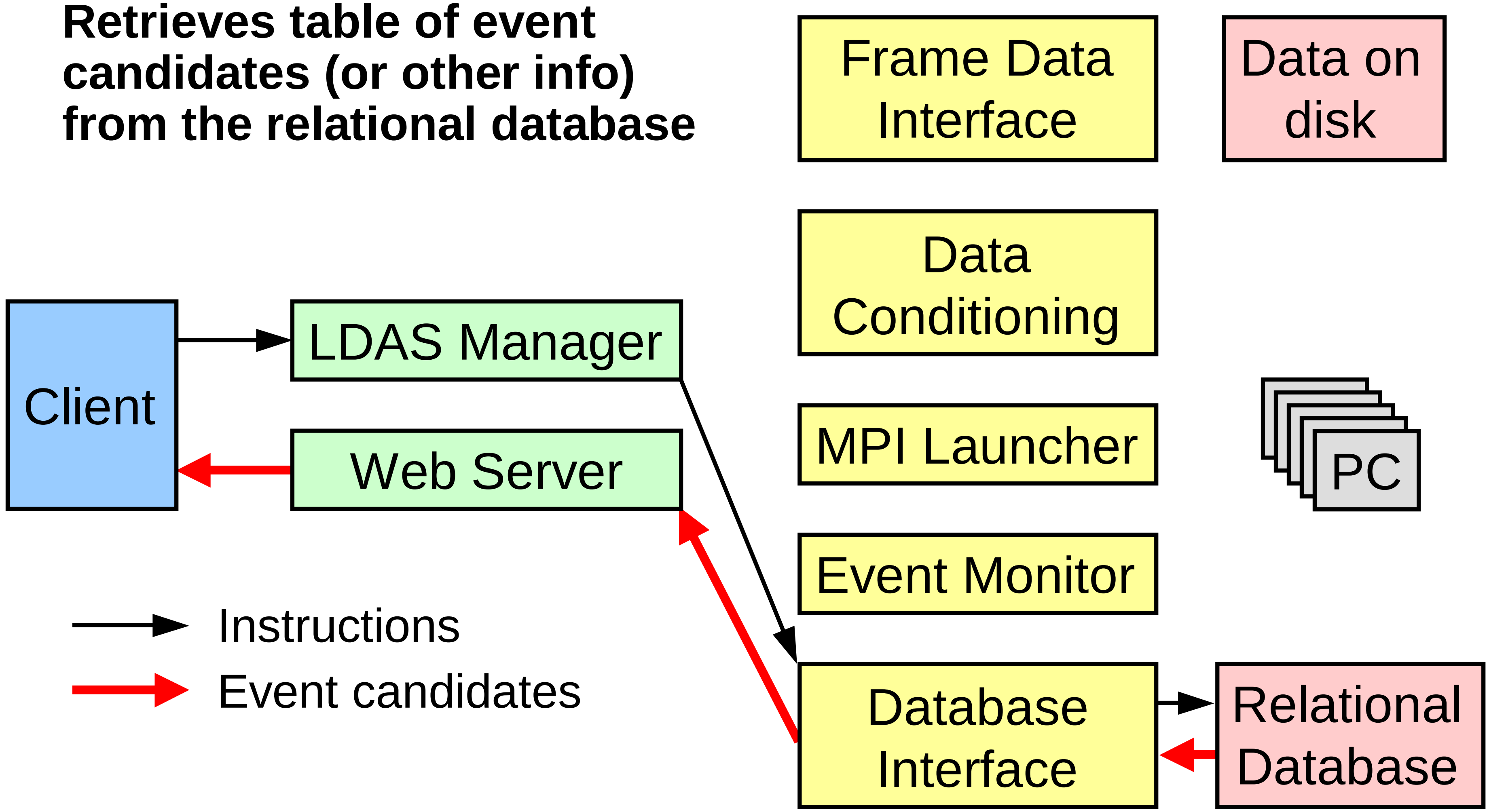
A “dataPipeline” Job

Standard job type to search for gravitational-wave event candidates



A "getMetaData" Job

Retrieves table of event candidates (or other info) from the relational database



The "LIGOtools" Software Suite (Client Utilities and More!)

Philosophy

Facilitate the sharing of useful software tools

Support diverse users in the LIGO Scientific Collaboration

Simple installation

Can be installed in any directory; doesn't require root privilege

Software is distributed as precompiled binaries [for Sun Solaris and Intel Linux] and scripts

⇒ No software prerequisites and no "build" step

Simple update procedure

Just type "ligotools_update" and select from the list of available new packages and/or new versions; the script downloads and installs them for you

Central web site with documentation, FAQs, links

LIGOtools Packages

Tcl library providing intuitive interface to execute LDAS jobs***

Graphical interface to LDAS relational database***

Utilities for remote retrieval of frame data***

Graphical user interfaces to view frame data

I/O libraries to read frame data and database table data into a C program or into Matlab

LIGO Algorithm Library (LAL)

… and more

***** These utilities can be (and are) "web-patched" to fix bugs, adapt to evolution of the LDAS communication protocol, or add new features without any action on the part of users

Lessons Learned

Creating a new data analysis system takes a lot of effort!

Integrating the parts is a large fraction of the work

It's hard to cast an analysis into a different paradigm

Analysis code written to run on LDAS's Beowulf cluster cannot trivially be adapted for execution on a grid

Right now, with science runs imminent, we can't afford to spend much time pursuing alternative approaches

Having multiple computing environments dilutes expertise

Number crunching is only one part of the whole analysis

Bookkeeping

Collecting, summarizing, and interpreting results

Summary

- LIGO is poised to begin collecting scientifically interesting data
- A customized environment has been developed to support various LIGO data analysis needs
- We are attempting to address issues of "usability" and support for a widely distributed community of users
- LIGO data analysis tools and algorithms will continue to evolve