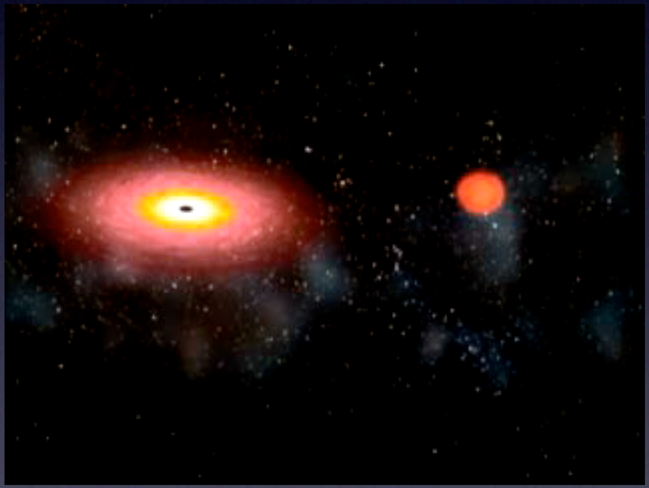
# LIGO Data Grid and Analysis

Patrick Brady for the LIGO Scientific Collaboration
University of Wisconsin-Milwaukee

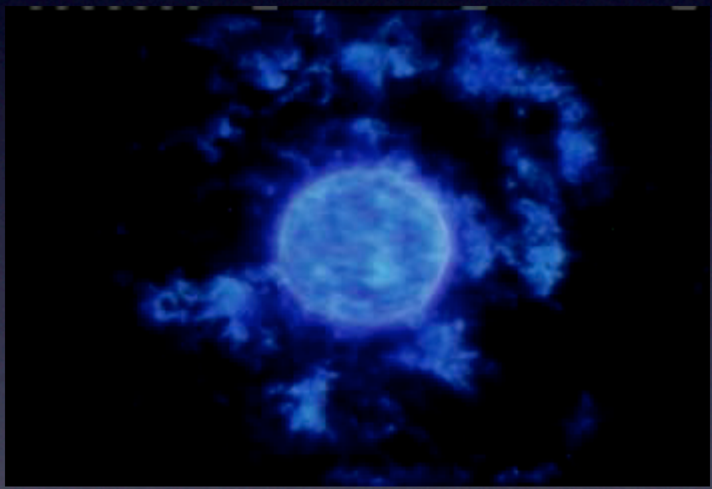
# Compact Binaries

- Pairs of black holes, neutron stars, or a black hole and neutron star
- As they orbit one another, they emit gravitational waves and the objects get closer together, eventually merging
- LIGO is sensitive to last few minutes of inspiral and merger



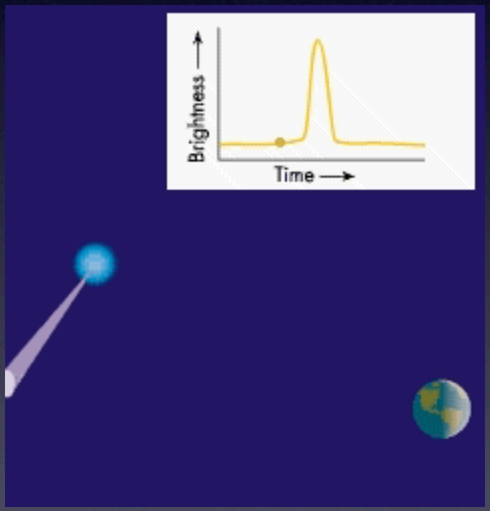
# Gravitational-wave bursts

- Very short (< 1 sec) bursts of waves from violent astrophysical events
- Examples include supernova explosions, mergers of compact binaries, and cosmic string kinks and cusps



# Continuous Signals

- Signals lasting as long as, or longer than, the observation time
- Known radio pulsars could also emit gravitational waves
- Unknown radio pulsars that are not beamed toward earth



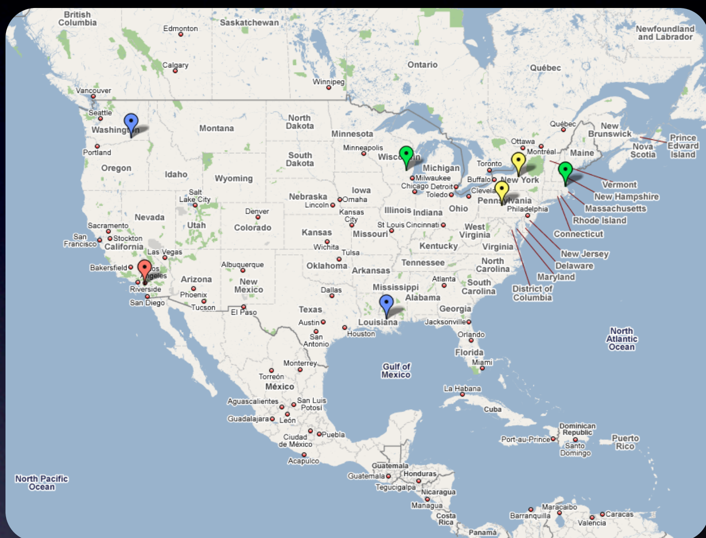
# LIGO Computing Drivers

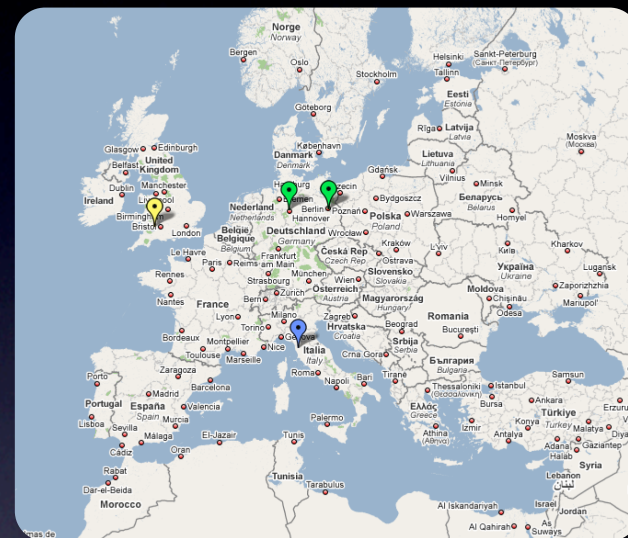
- Computing drivers:
  - Most of the analysis involves cross-correlating filter banks with data; dominated by Fourier transforms
  - Three instruments acquire 1TB/day of data; 1% in gravitational-wave channel
  - Filter banks with 10,000's filters for compact binaries; search needs $\sim 5 \times 10^7$ flops/byte.
  - Filter banks containing billions of templates for continuous-wave sources

# LIGO Computing Drivers

- Maximum scientific exploitation requires data analysis to proceed at the same rate as data acquisition
- Computational requirements for flagship searches:
  - Stochastic & bursts: 50 units (workstation day / day of data)
  - Compact binaries: up to 6000 units
  - All-sky pulsars = $10^9$ units, but longer delays tolerable
- LIGO's scientific pay-off is bounded by ability to perform computations on the data ..... leads to LIGO Data Grid

# LIGO Data Grid

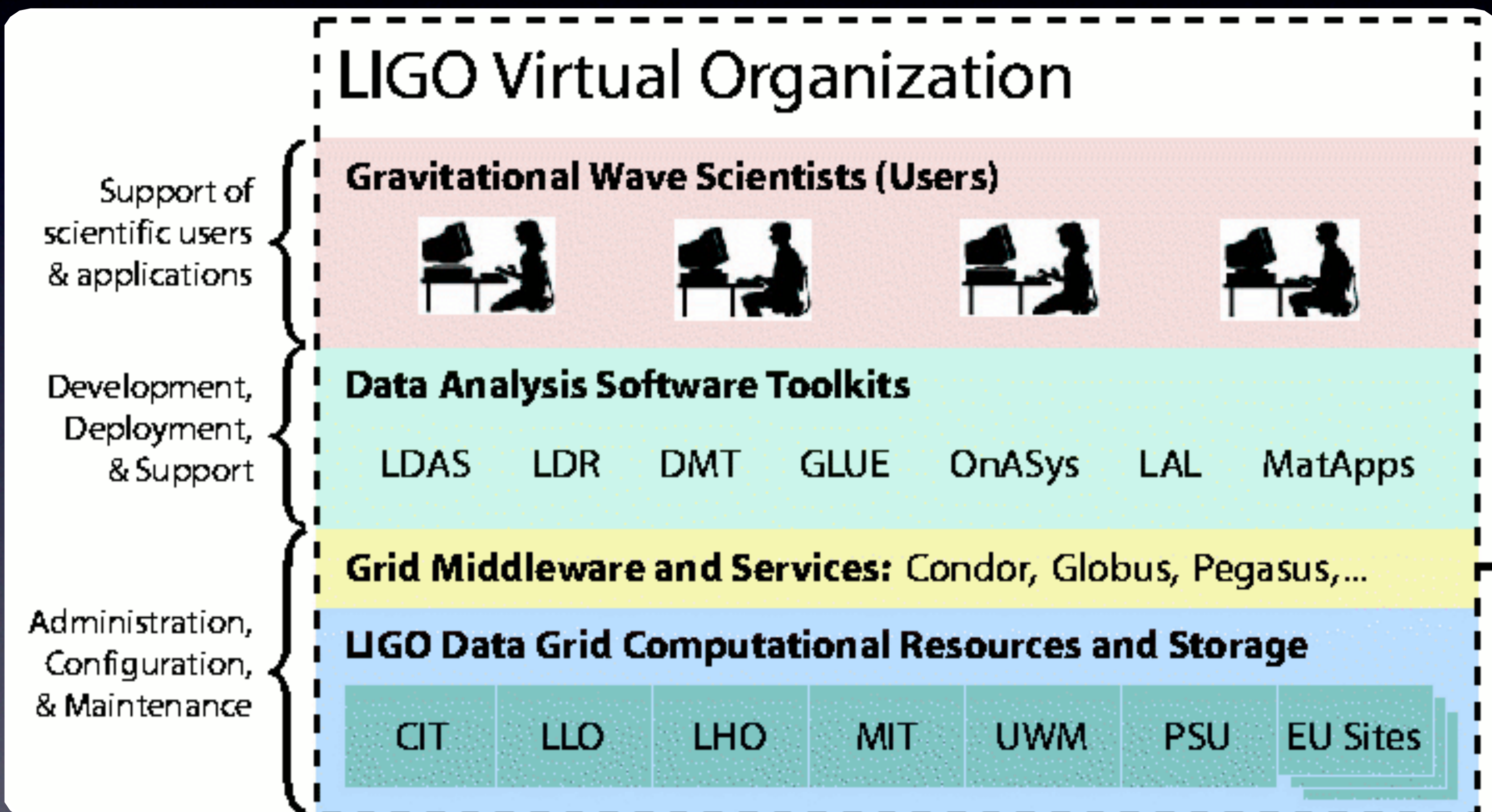




- Observatories
- Tier-1 Center
- Tier-2 Center
- Tier-3 Center

- LIGO Data Grid: combination of computational and data storage resources with grid-computing middleware to create a distributed gravitational-wave data analysis facility.

# Fabric of LIGO Data Grid



# LDG Operations

- Administration of storage and computational resources:
  - Heterogeneous hardware, but CentOS5 and Debian as reference platforms
- Grid Middleware:
  - Use the VDT releases to provide majority of middleware for LDG. Rely heavily on Condor and Globus
  - Rely on OSG to share our resources with other scientists on non-interference basis
- User Support:
  - About 200 users, many doing analysis prototyping and pushing limits of the facility - LIGO specific support

# Project Highlight

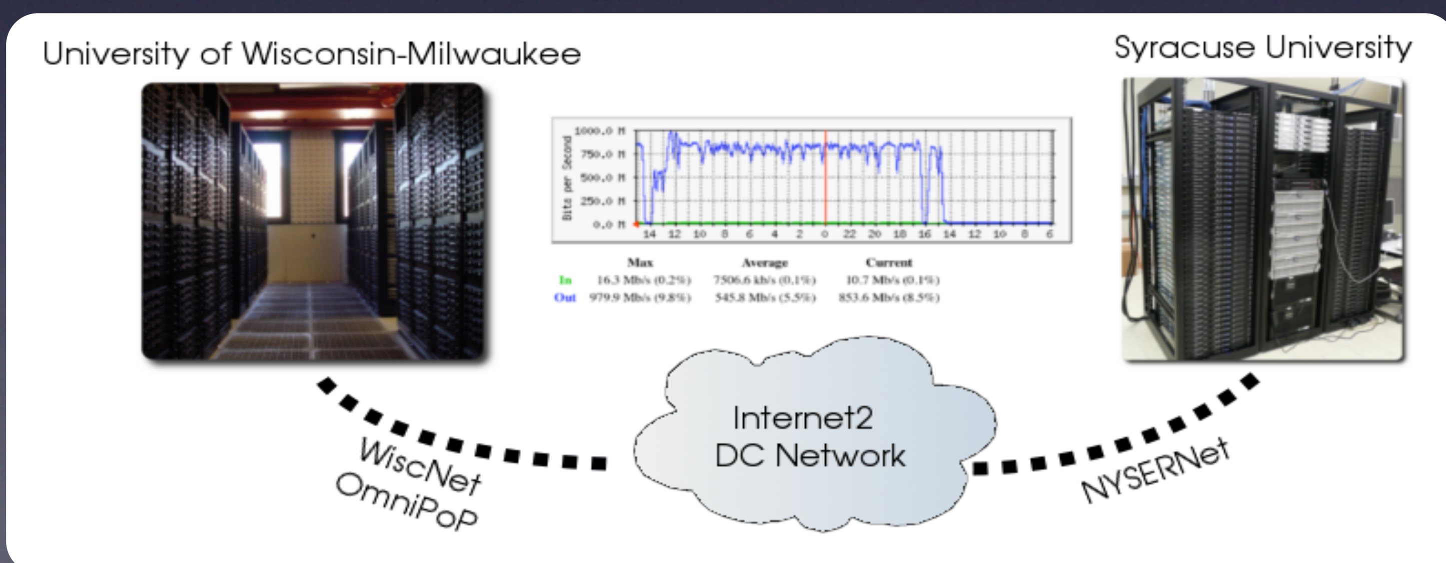
- LIGO Data Replicator (LDR; see Koranda):
  - moves data around the LDG before jobs are run
  - metadata catalog contains metadata information on more then 25 million files
  - each Replica Location Service (RLS) catalog between 1 and 50 million mappings
  - With eight LDR installations in the LIGO Data Grid the RLS network serves more than 300 million mappings
  - Datafind Servers: use LDR metadata to provide searchable index of data on each site using standard API

# Project Highlight

- Authentication and Authorization (see Anderson):
  - Address need to provide easy authorization of new collaboration members
  - Deliver combined mechanism to manage collaboration tasks, e.g. which group and what activity <u>and</u> LDG authentication/authorization
  - Use combination of Kerberos, LDAP, Grouper, and Shiboleth, to provide single sign and ability to manage the complex set of authorizations easily

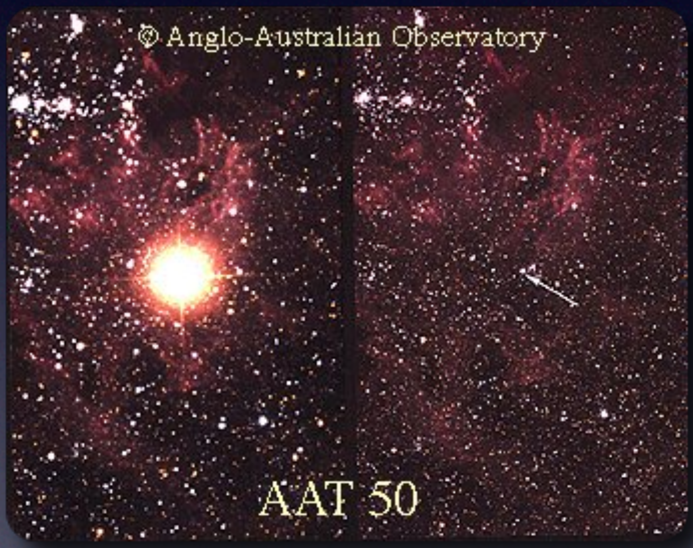
# Project Highlight

- NYSERNet, Syracuse University and University of Wisconsin-Milwaukee integrated LIGO Data Grid software with the Internet2 Dynamic Circuit Network (DCN) technology to achieve high-speed transfer.
  - 22 Tbytes of data from UWM to Syracuse in just 3 days with sustained transfer speeds of 800 Mbits per second
  - 8 times faster than the normal network connection



# Transient Astronomy in the 1980's

- Discovered by Shelton and Duhalde at the Las Campanas Observatory
- Independently discovered by Jones and Henshaw, amateur astronomers
- Astronomers learned about SN1987a via IAU telegram and phone calls to/from colleagues
- Follow up observations were made by almost every telescope in the southern hemisphere



# The New Millenium

- Instrumentation, networking, and computing has improved dramatically
- Automated follow-up observations of transients is now possible using many telescopes
- International Virtual Observatory is making strides to improve interoperability and access to information
- All-sky surveys will come of age during this decade


It's time to bring gravitational wave observatories to the party

# Where to next for LIGO computing and analysis

- Next production science run to start in spring 2009
- Scientific driver for analysis:
  - Perform analysis latency of seconds to hours to inform other astronomical observers
  - Requires data from LHO, LLO, Virgo (Pisa, Italy), GEO (Hannover, Germany) to be brought together for analysis
- Laundry list of scientific and computing issues:
  - Real-time generation of calibrated data, real-time distribution of data across the globe, analysis pipelines need to provide robust alerts, robust monitoring of data quality and environmental systems ........