Collaborative Analysis and Infrastructure Prototyping

TAMA Symposium, University of Tokyo, 2002

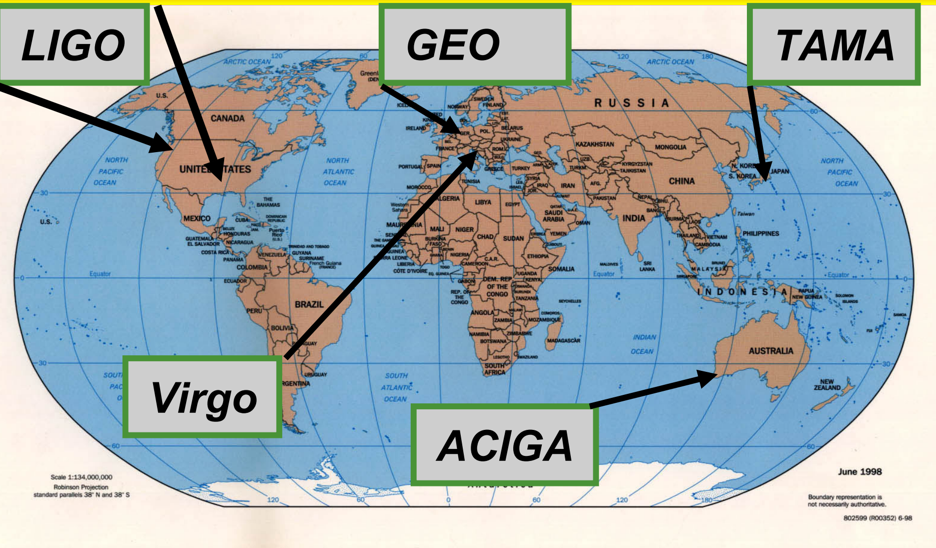
Szabolcs Márka, LIGO/Caltech

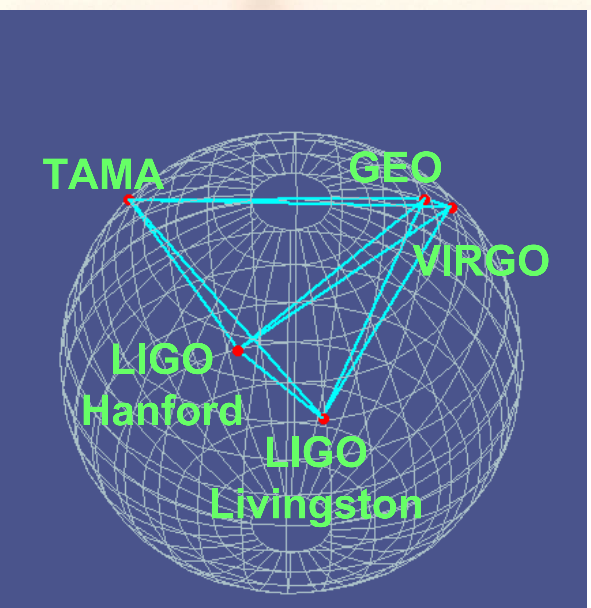
February 6, 2002

Outline

- Need for collaborative (network) data analysis
- Network data analysis group
  - Wide variety of tasks
    - Network Data Analysis Infrastructure
      - Topologies
      - Present emphasis on environmental data
      - Goals and experiences with the prototype system
      - Future plans

Interferometric Gravity Wave Detectors Worldwide



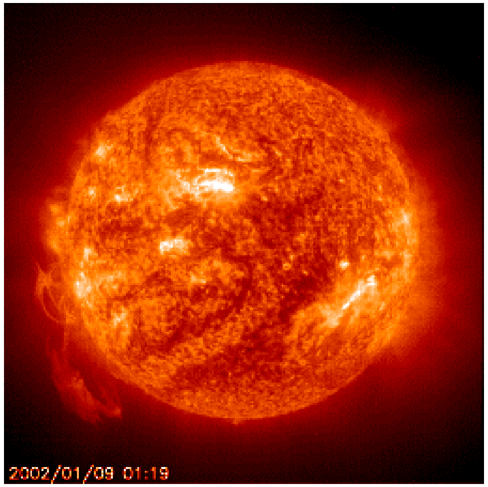


Some of the advantages of cooperation:

- Increased detection confidence
  - Redundancy
  - Various perspectives
  - Independent contributions
  - Independent observatories
  - Different technology, people and approaches
- Ability to locate the sources
- Decompose the polarization
- Redundancy
  - Downtime coordination

Collaborative (Network) Analysis Group

- Founded more than a year ago to address critical network analysis issues
- Well functioning working group
- Several members from every collaborations
- Hosted 2 international meetings to
    - Discuss progress
    - Resolve critical issues
    - Set standards
    - Set goals and milestones
    - Coordinate subgroups and taskforces



Contact person for each experiment:

Albert Lazzarini (Ligo)

Lee Samuel Finn (LSC)

Giancarlo Cella (Virgo)

Nobuyuke Kanda (Tama)

Susan Scott (Aciga)

Benno Willke, B. Sathyaprakash (Geo)

Task sampler

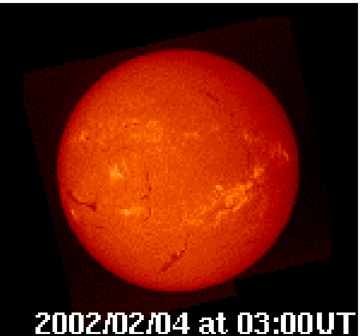
The motto is: Think ahead, be ready

- **Large number of tasks are addressed by subgroups**
  - » **Analysis approaches, requirement**
    - – inspiral, burst, continuous, stochastic
  - » **Pointing (event localization) issues**
  - » **Timing verification**
  - » **Technology interfacing**
  - » **Issues connected with widely varying sensitivities**
  - » **Standards for**
    - – Data containers
    - – Access tools
    - – Naming conventions
    - – Sampling rates
  - » **Multi-observatory data representation, access and storage**
    - – Network Data Analysis Server Prototype (NDAS)
      - Environmental data exchange and analysis
      - Timing verification
  - » **etc…**

Purpose of collaborative data analysis

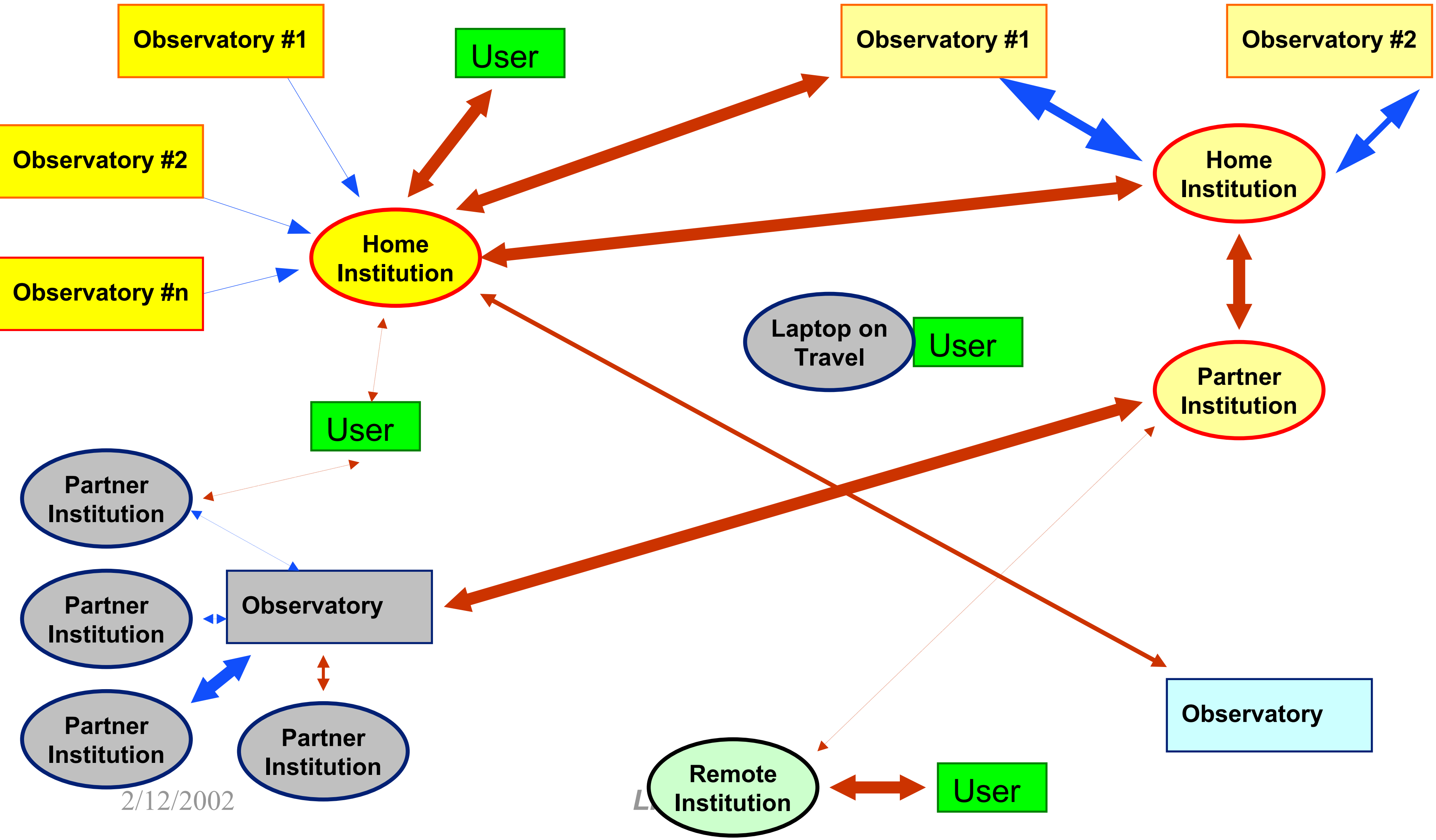
Quote from the 'Purpose of collaborative data analysis' section of the Collaborative Data Analysis White Paper

"…There is no too much time to prepare the required infrastructures, and it will be unavoidable that the implementation of a data analysis activity (code writing, software and hardware testing) will be necessarily intermixed with more theoretical activities (definition and improvement of data analysis algorithms, study of detection with non conventional noises and so on). This can be seen as a concrete opportunity for a proficuous exchange between theoretical activity and real experimentation. But it means also that the concrete implementation of the infrastructures must start as soon as possible and collaborative data analysis activities must start with simulated data. In this way: for data analysis: we will be able to accurately cross-validate data analysis codes and software tools. Also we will obtain real numbers for computational requests. for data exchange: we will be able to test and modify the data sets format and the data flow structure, and to make the required modifications to software and hardware tools for environmental monitoring: we could evaluate very soon, and with the final procedures, the correlation between noises on different detectors. This can be obtained in a first step (which does not require the creation of hardware infrastructures) by constructing simulated prototypes of network data analysis (for example, in Virgo, using Siesta). A second step could be the construction of a real prototype of simulated network data analysis, which will use the real infrastructures (network connections, computational facilities) of each experiment…"



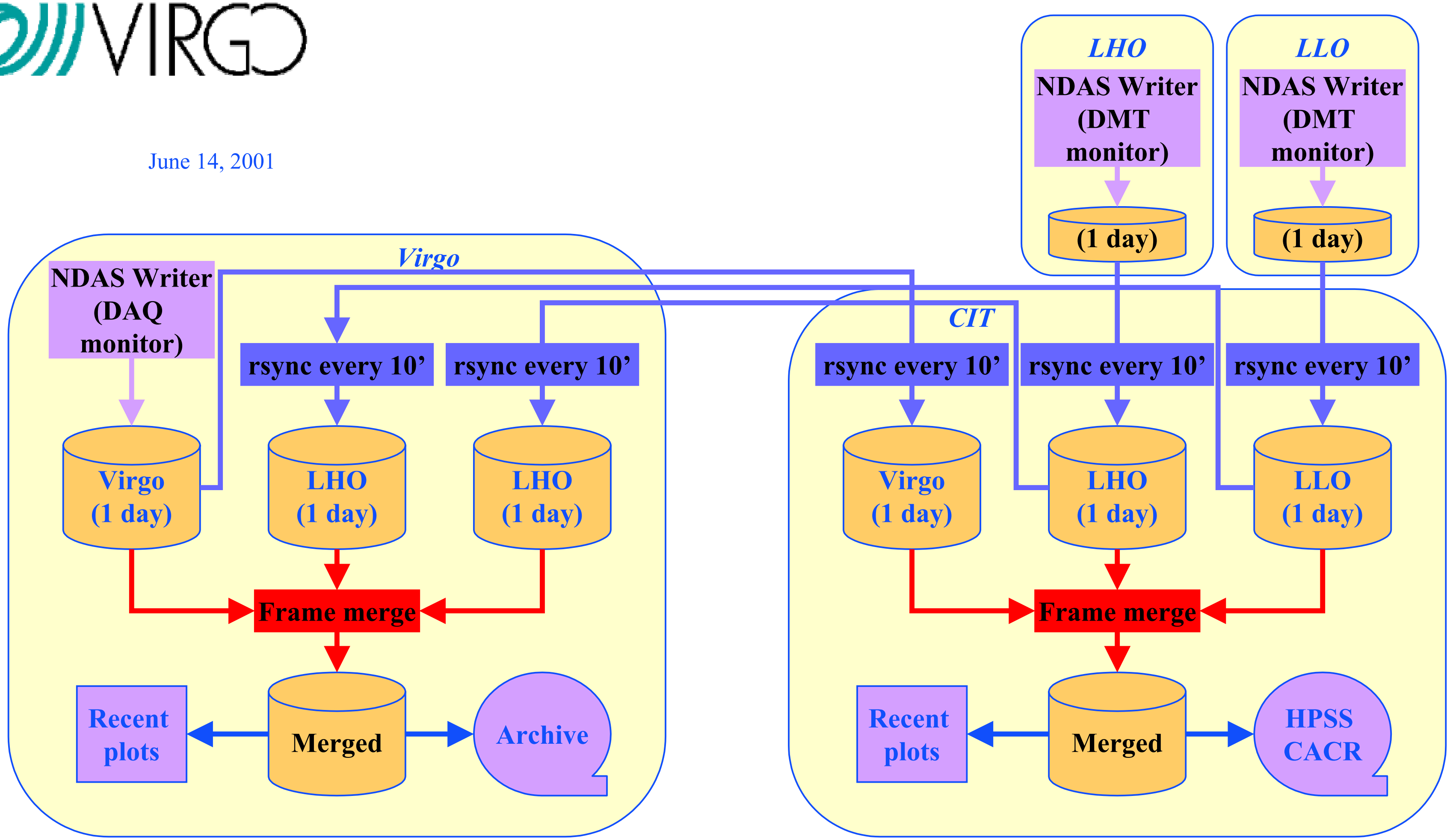
It is very clear that NDAS is a critical task to complete in time !

Wide variety of Network topologies (This slide was skipped at the talk)



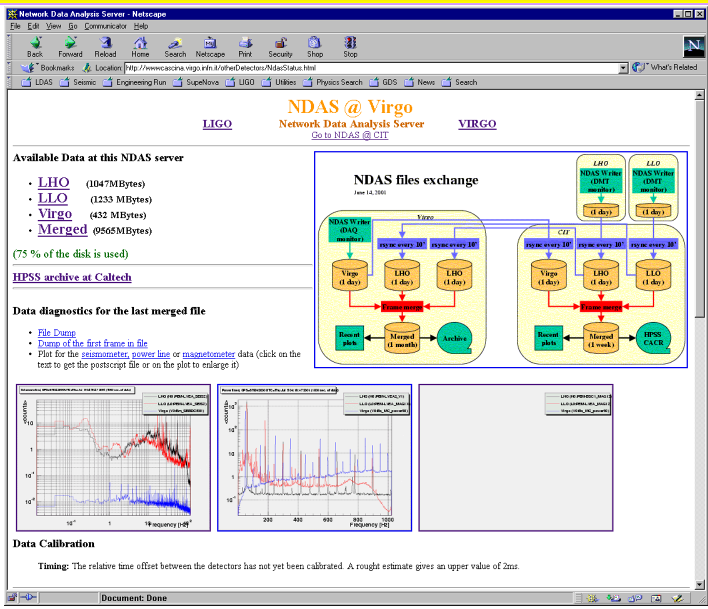
Early NDAS Topology

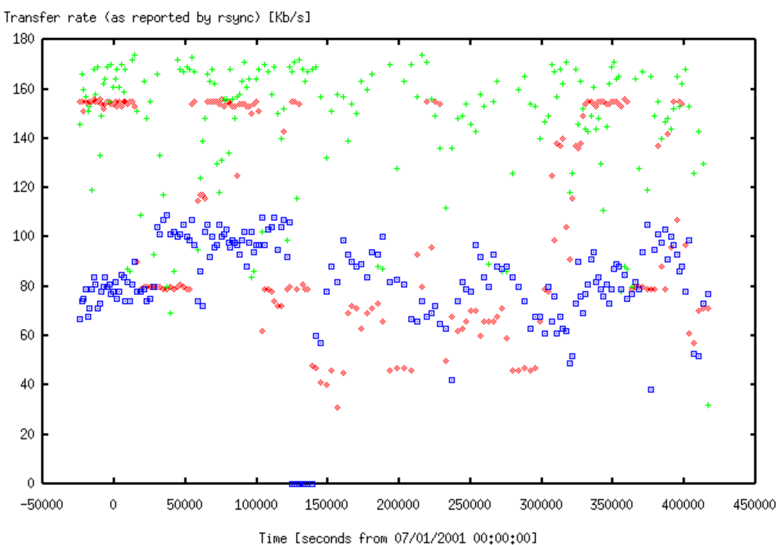
June 14, 2001



LIGO

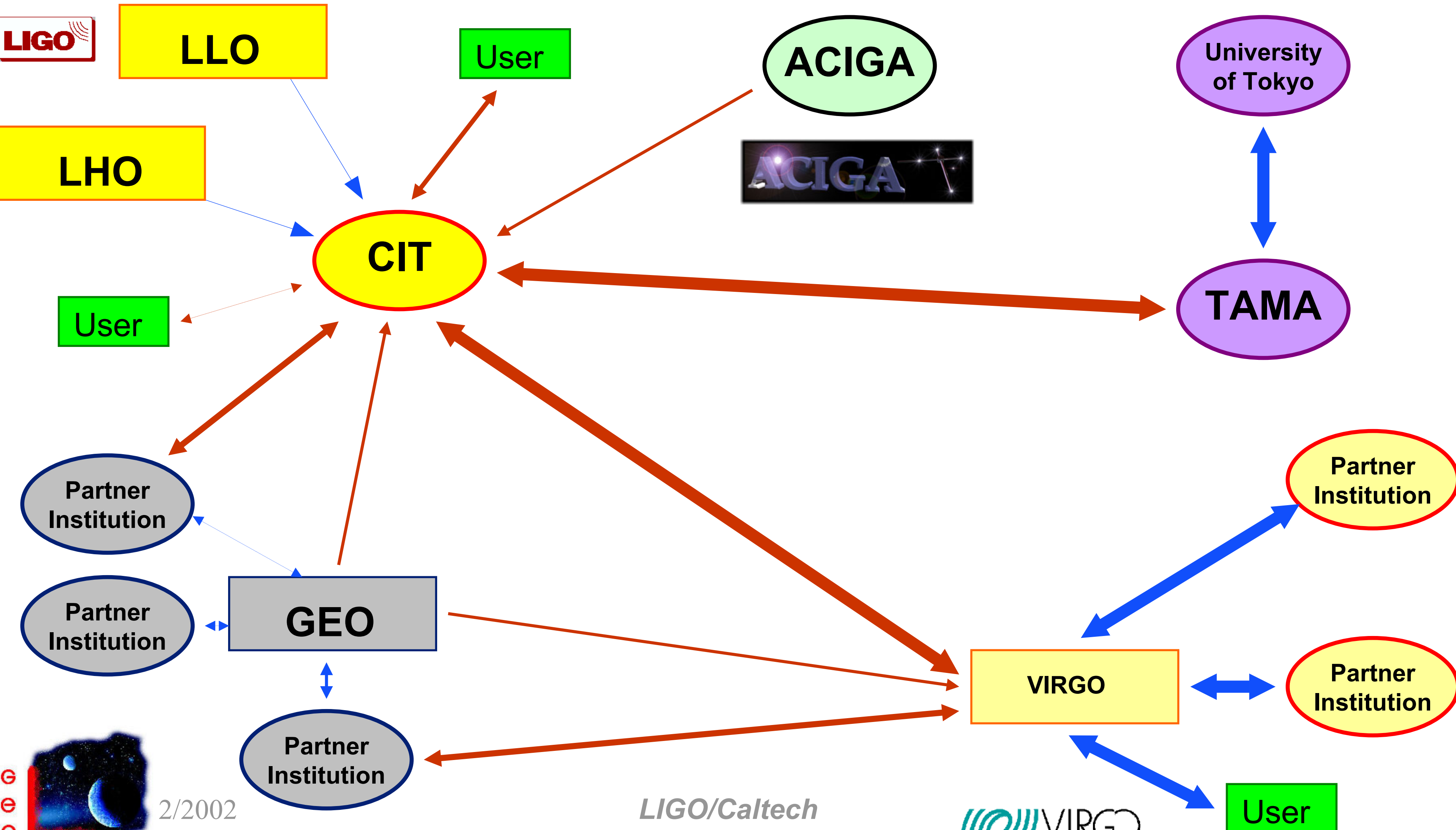
Some Early NDAS Experiences





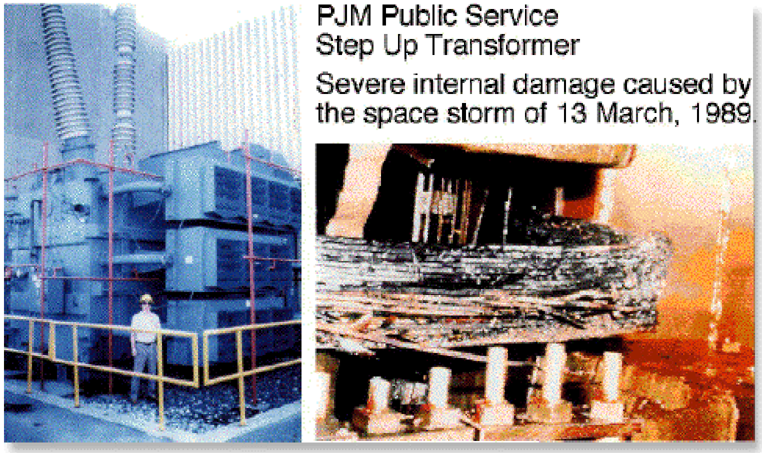
- NDAS is a modular, distributed and robust system utilizing standard unix tools and frame libraries to provide near real time environmental data exchange and merging between sites worldwide.
- NDAS gives the opportunity for participating members to share their data and download/access/analyze the full merged dataset containing all information from all participating sites.
- NDAS presently exchange only environmental data from a carefully selected set of few channels but it is freely extendable and scalable.
- Due to the modularity and simplicity of our solution NDAS requires absolutely minimal overhead to let participants concentrate on the joint data analysis.
- We developed a infrastructure and successfully exchanged and merged long stretches of data including the two LIGO and the VIRGO site.
- GEO and ACIGA joined our effort we will exchange and merge environmental data from five observatories in the following days
- TAMA agreed to join NDAS and we are surveying the implementation possibilities.
- We do actively encourage the participation of other existing or planned interferometric gravity wave observatories

Present NDAS topology



Only environmental and timing data for now… but it is still very interesting!

- Merits of environmental data
- Equivalent from technical viewpoint to h(t)
  - » Same DAQ
  - » Same transfer
  - » Same procedure
  - » Very similar problems and difficulties
  - » Other channels can be included any time
- Advantages for now
  - » Moderate data rates
  - » Very good practice
  - » Politically inert
  - » THERE IS GOOD PHYSICS TO LOOK AT!
    - Seismic
    - Magnetic
    - Lines
    - Timing
- Relatively easy and gives chance to
  - » Experiment
  - » Try alternative systems

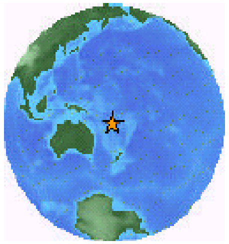


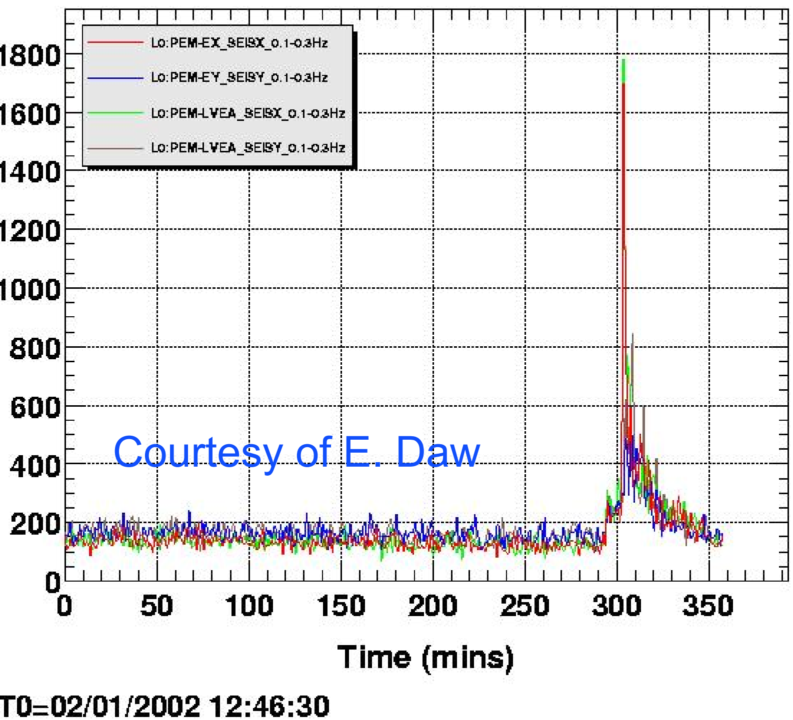
E7 Earthquake example

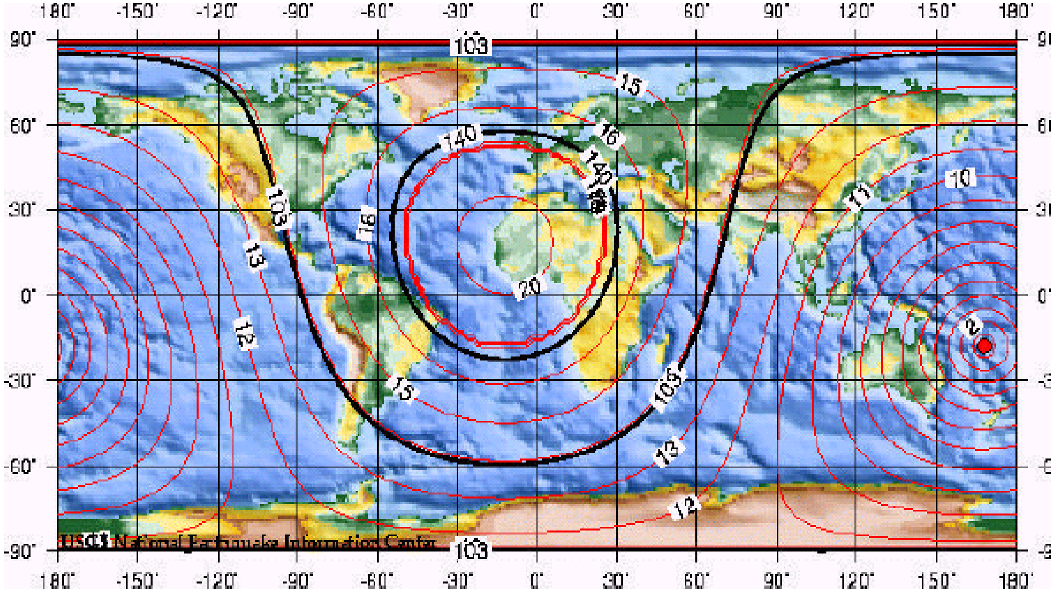
- **Predicted arrival times:**
- Knoxville, Tennessee 113.75 14:42.5 17:37:31.5 Pdiff
- Los Angeles, California 87.10 12:42.4 17:35:31.4 P
- Seattle, Washington 90.13 12:56.7 17:35:45.7 P
- Brownsville, Texas 101.59 13:48.5 17:36:37.5 Pdiff
- Knoxville, Tennessee 113.75 14:42.5 17:37:31.5 Pdiff
- Boston, Massachusetts 124.60 15:30.7 17:38:19.7 Pdiff

Theoretical P-Wave Travel Times

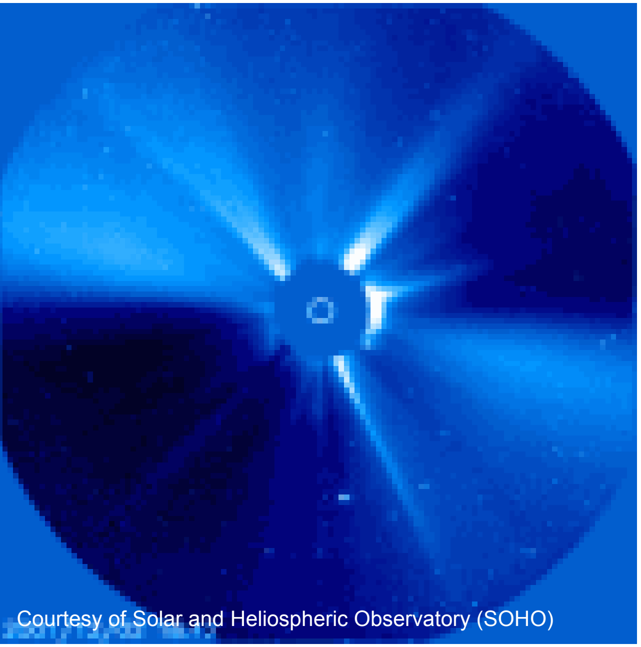
Date-Time	2002 01 02 17:22:49 UTC
Location	17.78S 167.88E
Depth	33.0 kilometers
Magnitude	7.2
Region	VANUATU ISLANDS
Reference	30 miles (45 km) W of PORT-VILA, Vanuatu



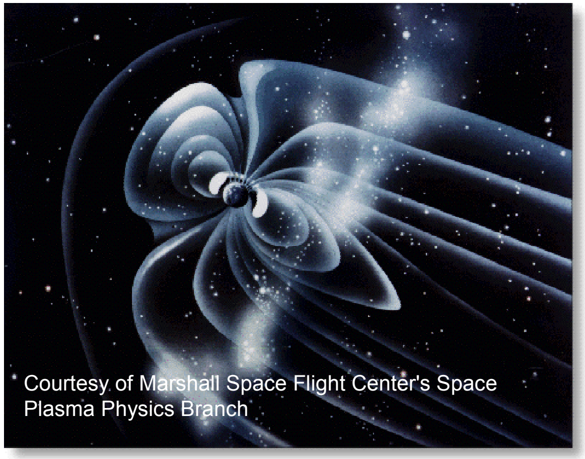




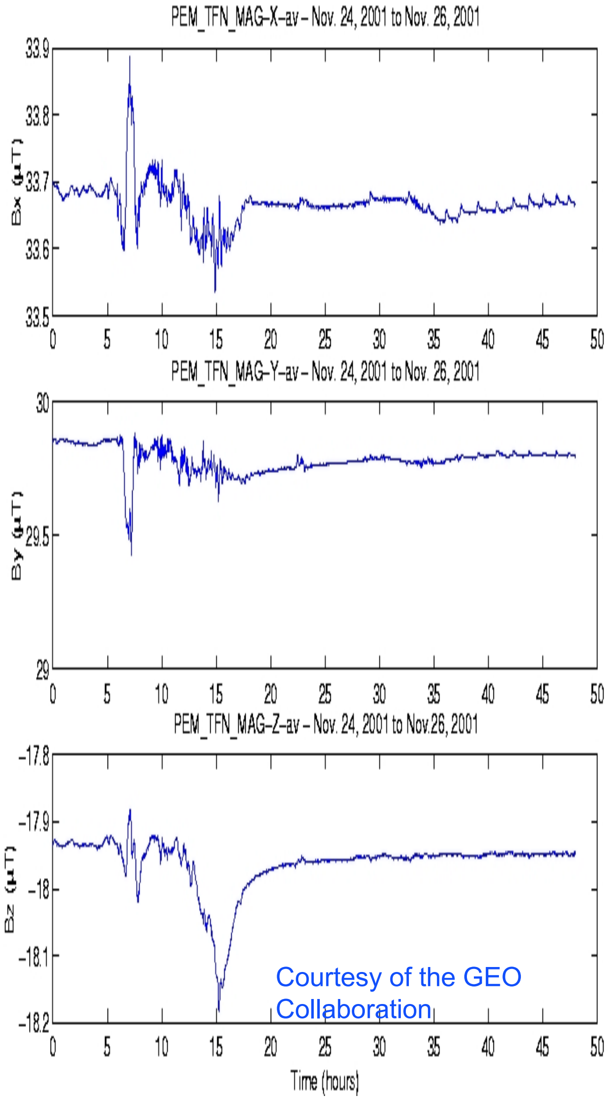
Geomagnetic Storms



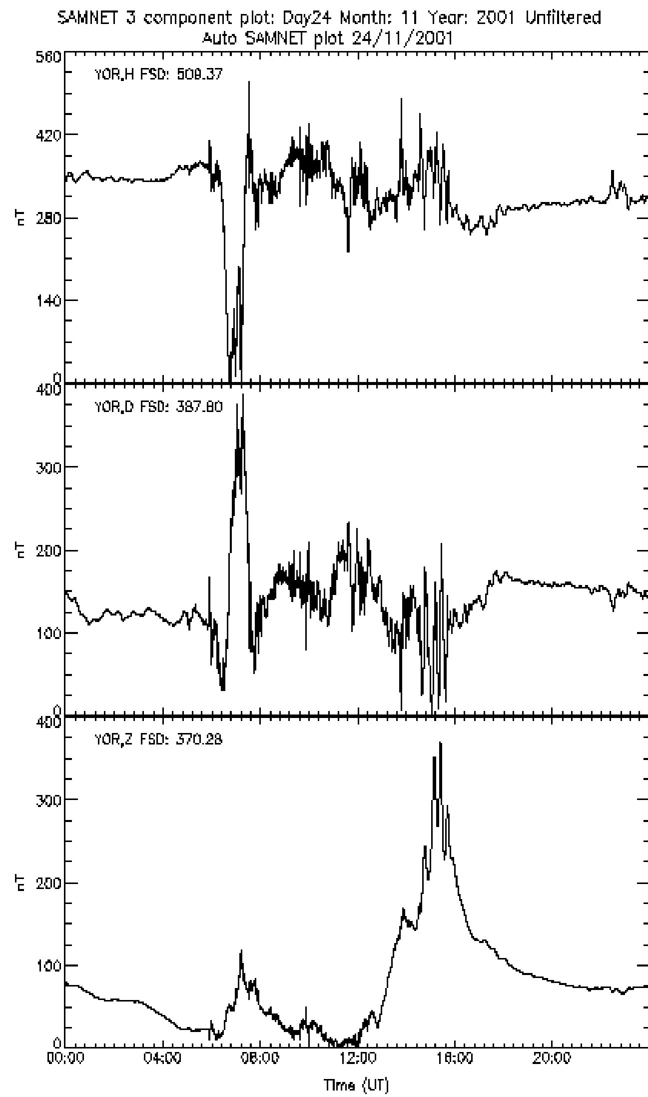
Courtesy of Solar and Heliospheric Observatory (SOHO)



Courtesy of Marshall Space Flight Center's Space Plasma Physics Branch



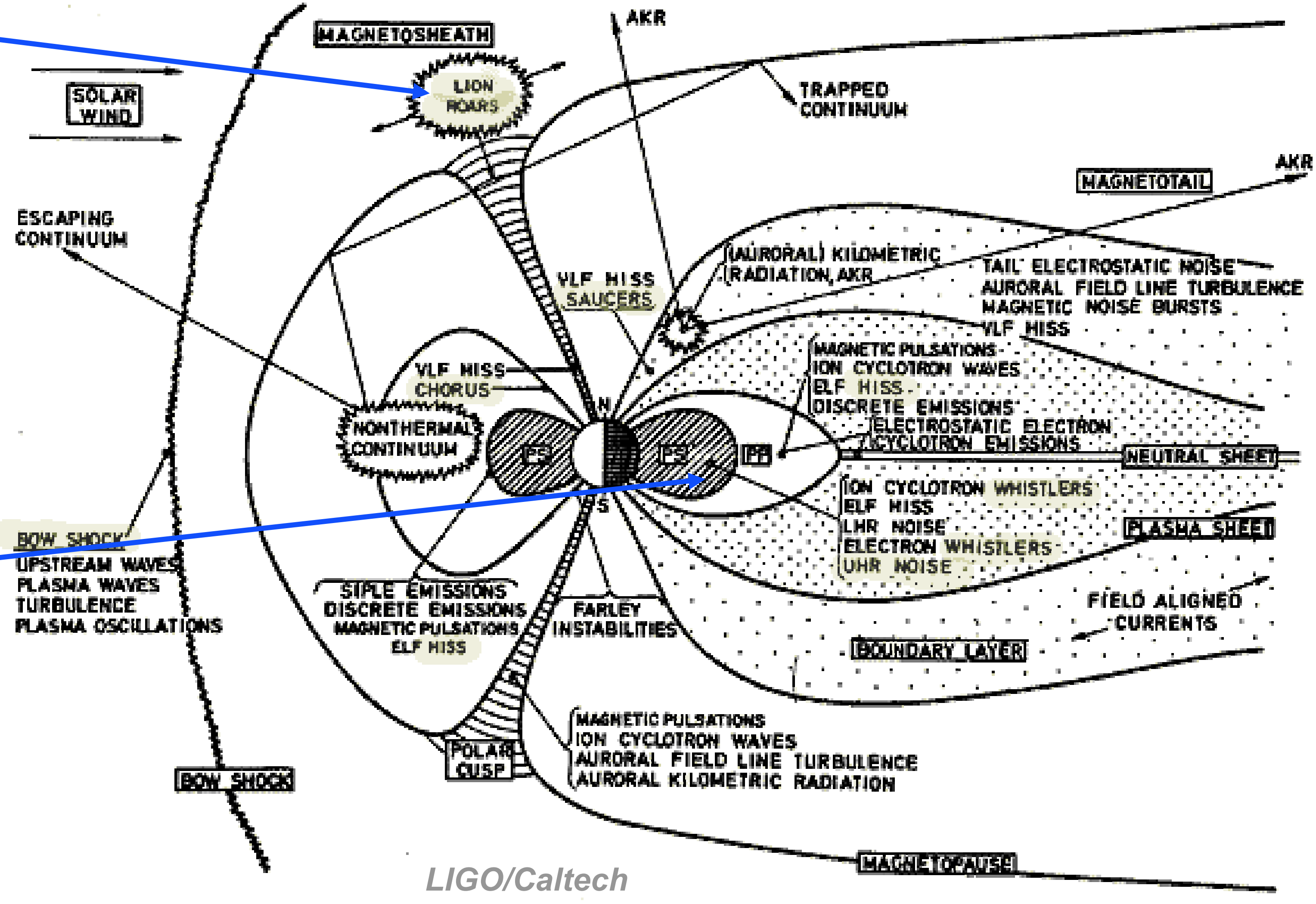
Courtesy of the GEO Collaboration



Roars and Whistlers of the Magnetosphere

Lion Roars

Whistlers



Courtesy of Marshall Space Flight Center's Space Plasma Physics Branch

Wide variety of seismic sensors

LIGO: Streckeisen STS-1 VBB Seismometer

VIRGO: EpiSensor triaxial FBA

LIGO: Guralp CMG-40T triaxial seismometers

ACIGA: Guralp CMG 3 single-axis seismometers

Magnetometer and Line Monitor Variety

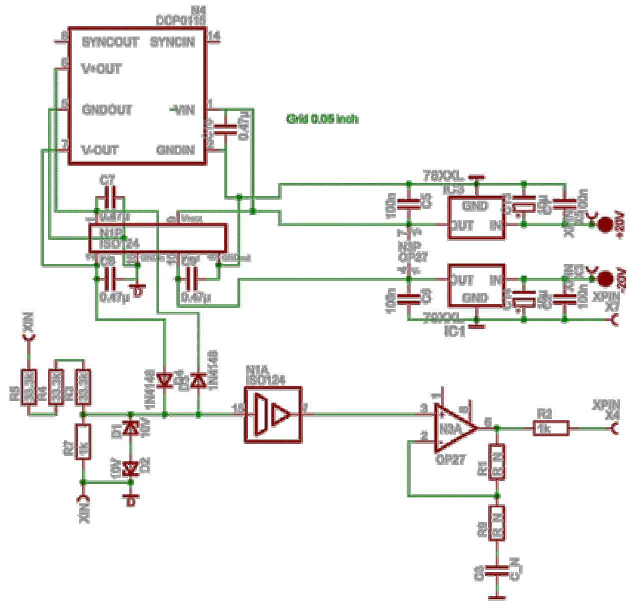
LIGO:

- **Triaxial Fluxgate Magnetometers**
  - » Bartington MAG-03MCES100-L7
    - – **Sensitivity: 7uT/V**
    - – 2048 Hz sampling rate
- **Triaxial Coil Magnetometers**
    - – Custom made
    - – 2048 Hz sampling rate

ACIGA

Stefan Mayer FL3-100
- triaxial fluxgate magnetometer
- sensitive for f < 2 kHz
- sample rate: 4096 Hz

Powerline Monitor Circuit
- voltage divider followed by isolating amp. (GEO design)
- sample rate: 2048 Hz



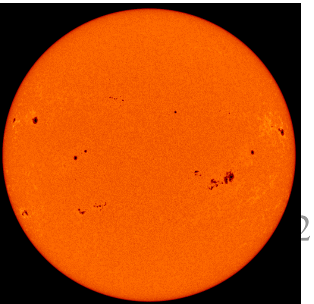
Future direction:
The Globus Data Grid Effort

"Access to distributed data is typically as important as access to distributed computational resources."

Distributed scientific and engineering applications often require acces to large amounts of data (terabytes or petabytes). Future applications envisioned by our team also require widely distributed access to data (For example, access in many places by many people, virtual collaborative environments, etc.)

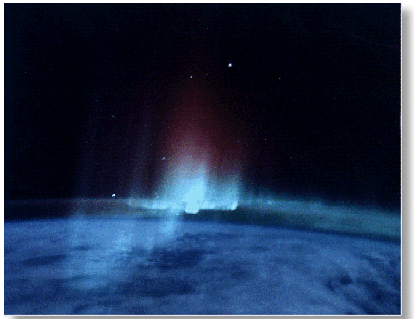
The Globus Project's data grid effort attempts to identify, prototype and evaluate the key technologies required to support data grids for scientific and engineering collaborations.

We face the same challenge, we have the same goal –
We should share the solution!



Conclusion

- The Cooperative (Network) analysis data group is well established
- Various critical tasks and bottle necks were identified and pursued
- One of the most matured task is the Network Data Analysis Server prototype development
  - First time that all of the interferometric gravity wave detectors shall collaborate
    - TAMA, ACIGA, GEO, LIGO, VIRGO
  - Environmental data exchange is sufficient for full infrastructure tests
  - There is relevant physics we can learn from the analysis of the merged global data
  - Large amount of data was already exchanged
  - The primary UNIX model is proven and routinely used
  - We do plan to be modern – We think about grid



It is beautiful to work on a collaborative effort of *all* interferometric gravity wave projects…